International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(13s) 2026 pp. 868-883
https://doi.org/10.70917/ijcisim-2026-3095
Article

Diabetic Cardiomyopathy Staging using a Tri-branch Deep Learning Framework with Optimized Adaptive Loss Mechanism

Kavitha Bai A. S[1,*], J. Somasekar[2]

[1,*] Department of Computer Science and Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bangalore 562112, Karnataka, India
[2] Department of Computer Science and Engineering, Faculty of Engineering and Technology, JAIN (Deemed-to-be University), Bangalore 562112, Karnataka, India
Email: kavithabai.as@jainuniversity.ac.in [1,*], somasekar.j@jainuniversity.ac.in [2]

Abstract: Diabetic cardiomyopathy (DCM) is a progressive cardiac disorder characterized by stage-dependent structural, functional, and metabolic alterations. Most existing approaches treat DCM as a single risk prediction problem, ignoring the ordered and asymmetric nature of disease progression and resulting in poor discrimination of intermediate stages. To address these limitations, this study aims to propose a stage-aware tri-branch deep learning framework for DCM stage classification. A phenotype-informed tabular dataset is first constructed using a Conditional Tabular Generative Adversarial Network (CTGAN) to mitigate the scarcity of labeled stage-specific clinical data. Following preprocessing, the input data is transformed through a shared Neutral Attentive Feature Encoder (NAFE) to produce an unbiased latent representation. This representation is processed by tri-branch Bi-directional Long Short-Term Memory (BiLSTM) branches, each specialized in learning the characteristic temporal patterns of a specific disease stage. Final stage prediction is achieved using an ArgMax-based Winner Takes All (WTA) mechanism. The model is trained through a novel stage-adaptive competitive loss and a Stage-Aware Adam optimizer (SA-Adam). Experimental results show that the proposed framework performed better than the existing works with an accuracy of 99.2%, precision of 98.61%, and recall of 97.89%. This ensures the proposed framework enables DCM stage classification, improved interpretability, and supports stage-specific treatment planning.

Keywords: Diabetic Cardiomyopathy, Neutral Attentive Feature Encoder, SA-Adam, stage classification, tri-branch BiLSTM.

1. Introduction

Diabetes is a chronic metabolic disease that is growing in prevalence and causes several problems. In the year 2021, the global population of diabetes in 20 to 79 years old stands at 536.6 million people, and, predicted an increase to 783.2 million people in 2045 [1, 2]. Diabetic Cardiomyopathy (DCM) is a problem of Type 2 Diabetes Mellitus (T2DM). DCM is characterized by the prolonged dysregulation of glucose and lipid metabolism that triggers continual structural and functional heart damage [3, 4]. DCM affects around 70% of diabetic patients, which leads to coronary artery disease, heart failure, sudden cardiac arrest, and death [5]. Insulin resistance, glucotoxicity, impaired calcium handling, mitochondrial dysfunction, inflammation, and oxidative stress play a major role in the development of DCM [6, 7]. Existing DCM diagnostic faces challenges due to the lack of disease-specific biomarkers and depends on echocardiographic parameters. Early detection of DCM is vital for enhancing the survival rate of individuals [8, 9]. Figure 1 shows the stages of DCM. Stage I detects early pathological indicators such as mild Left Ventricular (LV) enlargement and initial reduction in myocardial contractility. Stage II shows moderate ventricular dilation by noticeable wall thinning and reduced systolic performance. This indicates worsening cardiac efficiency and increased physiological stress. Stage III shows severe LV dilation, thrombus formation, and critically impaired pumping capability. This reflects advanced heart failure conditions.

In recent years, Artificial Intelligence (AI) and Machine Learning (ML) models have been applied to automate diabetes screening. These models detect microvascular and macrovascular complications and allow multiomic phenotyping for prevention and treatment recommendations [10, 11]. ML models utilize large amounts of health data to detect DCM patterns and relationships that are missed by conventional statistical approaches. These ML models provide a nuanced view of disease phenotypes through supervised clustering [12, 13]. AI is extensively utilized for T2DM diagnosis with connections between diverse features, phenotypes, parameters, and body examination [14]. Deep Learning (DL) provides an appropriate algorithmic framework for cardiovascular diseases and diabetes. DL models provide improved results and less time-consuming preprocessing and feature extraction compared to traditional methods [15, 16]. These models demonstrate promising prediction accuracy in DCM detection. However, existing models fail to classify the stage-wise progression of DCM. Traditionally, loss functions and optimizers penalize all misclassifications equally and ignore the ordered nature of DCM progression. Existing models provide limited direction for DCM stage-specific treatment planning and monitoring [17, 18].

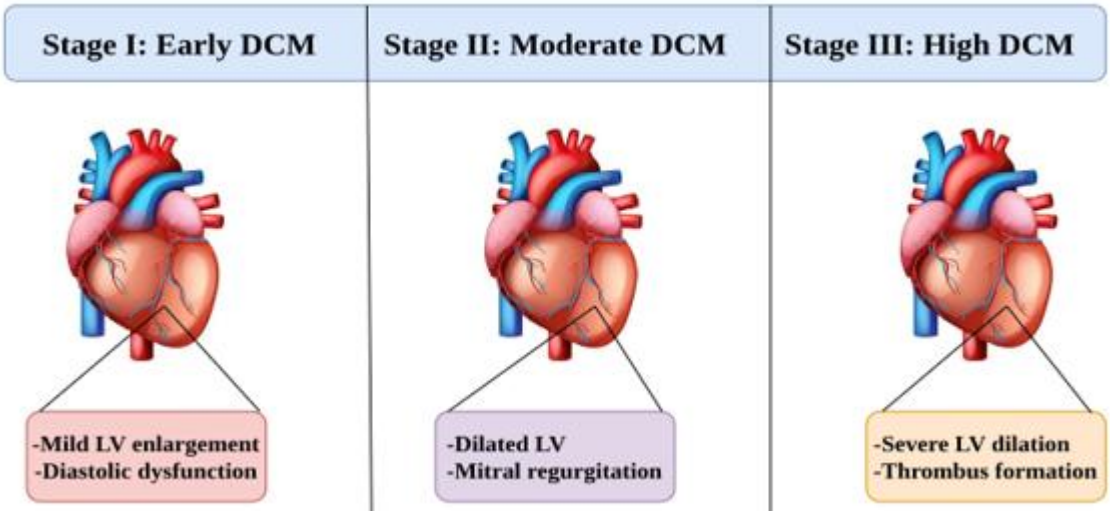

Figure 1: Stage-wise illustration of the DCM progression

To overcome these challenges, this study proposes a tri-branch DL framework for DCM stage classification. The study utilizes phenotype-informed synthetic data generation, a neutral attentive feature encoder, and tri-BiLSTM branches specialized for individual disease stages. To our knowledge, this is the first study to introduce a stage-wise classification of DCM disease. The major contributions of the proposed study are summarized as follows:

- **Phenotype-Informed Dataset Creation:** This study proposes a phenotype data generation strategy using Generative Adversarial Network (CTGAN) to construct a structured tabular dataset for DCM staging.
- **Stage-Aware Multi-Branch Deep Learning Architecture**: The study introduces a novel tri-branch Bidirectional Long Short-Term Memory (BiLSTM)-based architecture for DCM stage classification. Each branch independently learns characteristic temporal and clinical patterns of a DCM stage.
- **Severity-Aligned Learning:** The proposed framework incorporates a stage-adaptive competitive loss and a Stage-Aware Adam optimizer (SA-Adam) that explicitly balances the ordered and asymmetric nature of DCM progression.

The remaining part of the paper is organized as follows. Section 2 reviews the related works of DCM approaches, Section 3 explains the proposed methodology of the Tri-branch BiLSTM framework. Section 4 provides the results of implementation, analysis of the model and discussion. Section 5 concludes the paper.

2. Literature Survey

This section reviews the existing works related to DCM detection approaches. Zaidi et al. [19] developed a novel hybrid ensemble learning model called HeartEnsembleNet using ML Classifiers for Cardiovascular Disease (CVD) risk assessment. This model outperformed Conventional ML Classifiers and achieved higher accuracy along with high precision. However, this model failed to identify coronary artery disease and heart failure. Botros et al. [20] designed a multimodal data fusion approach using Electrocardiogram signals with selected blood test results for

improving heart failure detection. This multimodal fusion model enhanced heart failure detection, achieving higher accuracy with high sensitivity. However, this model relied on retrospective data (MIMIC-IV), which resulted in limited generalization. Nagavelli et al. [21] developed a ML-based Clinical Decision Support System (CDSS) using ML techniques such as Naive Bayes and Support Vector Machine (SVM) with XGBoost for heart disease diagnosis. This method achieved superior diagnostic performance when compared to Electrocardiography (ECGs). However, this approach resulted in low accuracy for heart disease detection.

Myat Oo et al. [22] introduced an integrated AI-driven heart failure identification framework using DL-based automated interpretation of Echocardiographic DICOM images for improving patient care through automated case detection and surveillance. This framework automatically interprets Echocardiographic DICOM images and helps to identify patients with heart failure. However, this framework failed to differentiate people affected with preserved ejection fraction and heart failure. Nabrdalik et al. [23] developed an AI-based DL technique with non-mydriatic 5 field color fundus imaging for detecting Cardiac Autonomic Neuropathy (CAN) using a ResNet-18 network. This framework improved non-invasive CAN screening and achieved high diagnostic performance. However, this framework required a longer duration for training the dataset.

Existing DCM diagnostic approaches depend on ML models that treat stage prediction as a single unified classification problem. These approaches struggle to capture heterogeneous medical data, including imaging-derived parameters, biomarkers, and clinical indicators. Additionally, traditional DL models tend to learn stage-biased representations and clinically unsafe severe misclassifications between different stages. Therefore, there is a need for an interpretable framework capable of learning neutral feature representations, distinguishing subtle stage-dependent patterns, and producing clinically reliable stage assignments. This work addresses these challenges by proposing a stage-aware DL architecture for DCM stage classification.

3. Materials and Methods

The study develops a DL framework for stage classification of DCM, as shown in Figure 2. The proposed framework creates a structured tabular dataset using CTGAN. The phenotype-informed synthetic samples are generated to reflect different disease stages based on clinically reported characteristics. The generated dataset is then preprocessed by removing physiologically implausible values, handling missing entries using stage-aware mean imputation, and normalizing continuous variables to ensure stable learning.

Then, a shared stage-agnostic attentive feature encoder Neutral Attentive Feature Encoder (NAFE) transforms heterogeneous clinical, biomarker, and echocardiographic inputs into a compact latent representation. This shared representation is fed into three independent BiLSTM-based branches. Each specialized in learning the characteristic pattern of a specific DCM stage, with the first branch focusing on early-stage mild ventricular remodeling. The second branch captures the intermediate-stage diastolic dysfunction and biomarker elevation.

The third branch is included to capture the advanced subclinical myocardial dysfunction. Then, the final stage prediction is obtained through a Winner Takes All (WTA) ArgMax selection of branch activation scores. The variables , , and are used to obtain the branch activation scores of stage I, stage II, and stage III, respectively. The model is trained using a stage-adaptive competitive loss that penalizes severe stage misclassifications and optimized using an SA-Adam optimizer. L1L2L3

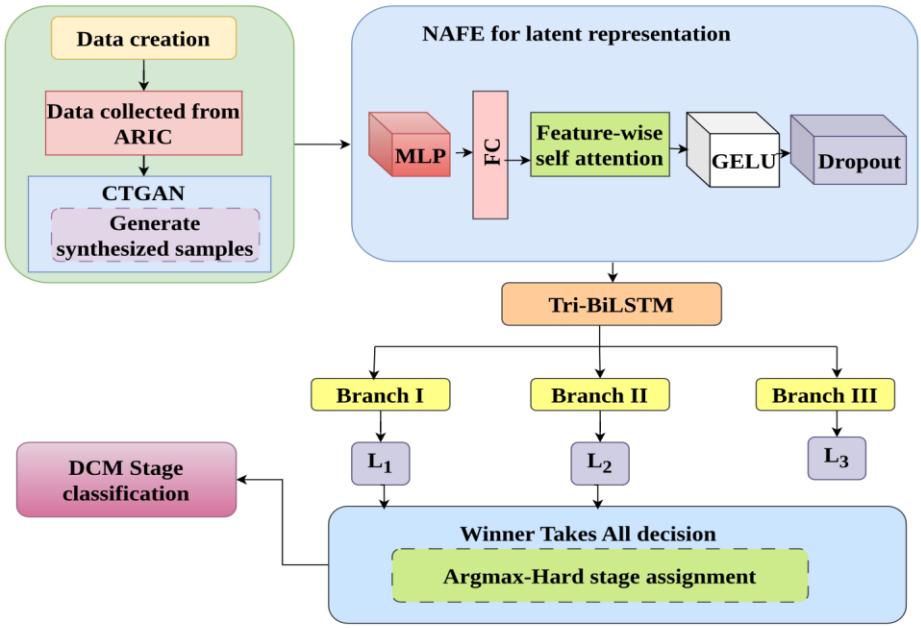

Figure 2: Schematic representation of the proposed tri-branch BiLSTM framework for DCM stage classification

3.1 Dataset Creation

The study uses the data from the Atherosclerosis Risk in Communities (ARIC) and Cardiovascular Health Study (CHS) at the University of Texas Southwestern medical center for the model training, testing and validation. ARIC and CHS data were obtained from the National Institute of Health Biological Specimen and Data Repository Information Coordinating Center (BioLINCC). In this study, the dataset is constructed to reflect stage-wise phenotypic characteristics of DCM from existing cohort-based studies [24]. A CTGAN is employed to generate realistic synthetic samples corresponding to different DCM stages and used to create a tabular dataset. The dataset includes demographic attributes, comorbidity indicators, medication usage, renal functional markers, cardiac biomarkers, and detailed echocardiographic measurements. Echocardiographic measurements capture cardiac structure, systolic and diastolic functions. The dataset comprises of the features including age, sex, blood pressure, body mass index, hypertension status, medications related to diabetes, renal indices, cardiac biomarkers, and echocardiographic variables left ventricular mass index, wall thickness, atrial size, ventricular volumes, global longitudinal strain, transmitral flow velocities, and derived ratios of diastolic echocardiography. The samples were classified into three phenogroups representing early, intermediate, and advanced subclinical DCM to ensure stage awareness. This reflects progressive myocardial remodeling patterns. These parameters merged into a single labeled dataset suitable for training and evaluating the proposed model.

3.2 Data Preprocessing

In this study, we utilize the mean imputation [25] and z-score normalization [26] preprocessing techniques to improve the quality of the dataset. Mean imputation technique handles the missing values in the dataset. The Z-sore normalization technique normalizes the continuous features in the dataset. This technique converts features, its mean becomes 0 and its deviation becomes 1 which ensures the feature follows a standard normal distribution.

3.3 Neural Attentive Feature Encoder

In this study, we introduce the NAFE encoder to transform heterogeneous clinical inputs into a latent representation. NAFE utilizes the Multi-Layer Perceptron (MLP) [27] as the backbone, which is augmented with a feature-wise self-attention mechanism. Every node of MLP includes two functions, such as summation and activation functions.

The product of input, weight, and bias values are obtained through the summation function S_f, as described in equation (1).

$$S_f = \sum_{u=1}^{n} w_{uv} I_u + b_v \quad (1)$$

Where, n represents the total number of inputs and I_u denotes the input variable u. b_v denotes the bias value and w_{uv} signifies the connection weight. Then, the preprocessed data $x \in \mathbb{R}^n$ fed to the Fully Connected (FC) dense projection layer. It maps the input into a latent space l_s and model nonlinear interactions among features. This transformation is expressed in equation (2).

$$l_s = W_1 x + b_v \quad (2)$$

Where, $W_1 \in \mathbb{R}^{d*n}$ denotes the learnable weight matrix and d represents the dimensionality of the latent feature space. Then, the feature-wise self-attention mechanism highlights clinically informative features and calculates attention scores, as denoted in equation (3).

$$\sigma = softmax(Z_a l_s + g_a) \quad (3)$$

Where, Z_a and g_a represents the attention parameters and σ denotes the normalized weights assigned to each latent feature dimension. The attended feature representation A_s is obtained through element-wise reweighting, as represented in equation (4).

$$A_s = \sigma \odot l_s \quad (4)$$

Where, $\odot$ signifies the element-wise dot product.

Then, a Gaussian Error Linear Unit (GELU) activation function, as mentioned in equation (5). Bernoulli dropout is used to mitigate overfitting by randomly deactivating a fraction of neurons during training.

The final latent representation f_s is expressed in equation (6).

$$G_s = GELU\ (A_s) \quad (5)$$

$$f_s = dropout\ (G_s), f_s \in \mathbb{R}^d \quad (6)$$

Where, latent vector f_s is neutral and stage-agnostic, shared across all stage-specific branches of the network. fs

3.4 Tri-Branch Bi-LSTM for DCM Stage-Wise Classification

After the NAFE phase, the shared latent representation is passed to three different pattern learning branches. Each designed to model the characteristic phenotypic patterns of a different DCM stage. Each branch utilizes the BiLSTM [28] to capture complex feature interactions within the latent space, as represented in Figure 3.

This Tri-branch BiLSTM employ to model a particular disease stage. BiLSTM transformation is denoted below for the k-th stage ($k \in \{1,2,3\}$).

$$H_k = BiLSTM_k(V) \quad (7)$$

In equation (7), H_k denotes the hidden representation learned by the k-th branch and V denotes the feature interactions. BiLSTM includes forward and backward LSTM to processed data. The forward hidden layer H_f, the backward hidden layer H_b, and the output sequence $BL_o(s)$ utilized to update the network.

The updated network is represented in equations (8)-(10).

$$H_f = \delta W_1 BL_o(s) + W_2 H_{f-1} + b_{H_f} \quad (8)$$

$$H_b = \delta W_3 BL_o(s) + W_5 H_{b-1} + b_{H_b} \quad (9)$$

$$BL_o(s) = W_4 H_f + W_6 L + b_{BL_o(s)} \quad (10)$$

Where, W represents the weight coefficients and b_{H_f}, b_{H_b}, and $b_{BL_o(s)}$ denotes the biases. In this study, each BiLSTM branch is specialized in learning different stage-relevant patterns. The first BiLSTM branch focuses on early-stage mild ventricular remodeling.

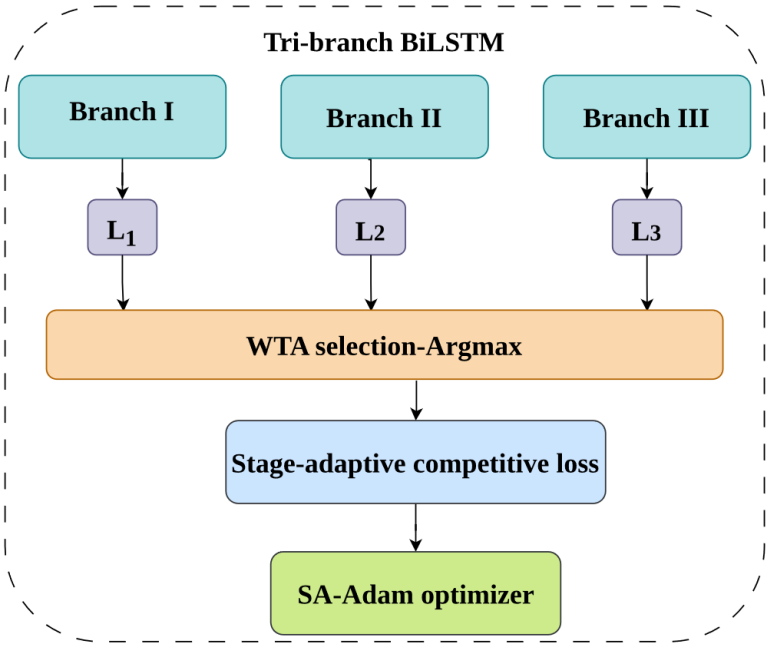

Figure 3: Architecture diagram of the proposed tri-branch BiLSTM framework. L_1, L_2, and L_3 represents activation outputs of stage I, stage II, and stage III branches. This shows competitive learning and WTA stage selection.

The second BiLSTM stage captures intermediate-stage diastolic dysfunction, and the third stage model subclinical myocardial dysfunction. The final stage prediction is obtained through a Winner Takes All (WTA) [29] ArgMax selection of branch activation scores, as denoted in equation (11).

$$A_k = \mathrm{w}_k^{\mathrm{T}} \mathrm{H}_k + b_k \tag{11}$$

Where w_k and b_k represents the learnable parameters with the k-th stage classifier. P_k represents the activation strength, indicating how well the patient's latent features align with that stage-specific pattern. The three activation scores are aggregated into a vector $\mathrm{L} = [L_1, L_2, L_3]$, and a WTA decision mechanism is applied using an ArgMax operation to assign the final DCM stage, as follows

$$\hat{c} = \arg \max_{k \in \{1,2,3\}} L_k \tag{12}$$

This competitive selection ensures clear and interpretable stage predictions while preventing feature dominance across branches. In equation (12), $\mathrm{L} = [L_1, L_2, L_3]$ denotes the three BiLSTM branches and L_k denotes the predicted score for the stage $k \in \{1,2,3\}$. After the WTA stage selection, the loss function is formulated to supervise the tri-branch framework to preserve competitive learning among branches. The predicted probabilities are computed using the softmax function, as represented in equation (13).

$$\widehat{L_k} = \frac{e^{L_k}}{\sum_{u=1}^{3} e^{L^u}} \tag{13}$$

The categorical cross-entropy is defined using equation (14).

$$L_{CE} = -\sum_{k=1}^{3} g_k \log(\widehat{L_k}) \tag{14}$$

Where, g denotes the ground-truth label. To reduce the misclassifications, a stage-adaptive penalty function is expressed in Equation (15),

$$L = L_{CE} + \rho * L_{severity} \tag{15}$$

Where, ρ represents the weighting coefficient controlling penalty strength. The severity term is defined based on-stage distance, as follows

$$L_{severity} = \sum_{k=1}^{3} |k - g| * \widehat{L_k} \tag{16}$$

In equation (16), penalizes prediction proportionally to how far they deviate from the true stage. This enables competitive branch specialization while maintaining clinically safe learning behavior.

3.4.1 Novel SA-Adam Optimizer

The proposed SA-Adam is a modified version of the standard Adam optimizer designed to improve convergence stability and stage-balanced learning in the tri-branch BiLSTM framework. The SA-Adam introduces a stage-adaptive scaling factor that adjusts parameter updates based on stage-specific error sensitivity. The Adam optimizer [30] efficiently updates model parameters during the training process. It reflects the correct bias ratings of the moving average gradient and the square gradient. In the Adam optimization, the training data θ is separated into θ^{ω},θ^{q} trainable parameters, hyperparameters β_1,β_2 controls the rates of moving averages, $\mathrm{G}_{\boldsymbol{z}}$ denotes the moving average exponential gradient, and $\mathrm{H}_{\boldsymbol{z}}$ denotes the squared gradient. The instatement tendency can be neutralized, taking about inclination redressed gauges $\widehat{G_z}$ and $\widehat{H_z}$. Initialize the exponential moving average as $\omega = 0$. The moving step of the updated time step is denoted as:

$$\mathrm{H}_{\boldsymbol{z}} = \alpha_{\boldsymbol{2}}.H_{\boldsymbol{z-1}} + (1 - \alpha_{\boldsymbol{2}}).A_z^2 \tag{17}$$

In equation (17), A_z^2 represents the component square of A_z. The time steps of the gradient function are computed as in equation (18).

$$\mathrm{H}_{\boldsymbol{z}} = (1 - \alpha_{\boldsymbol{2}}) \sum_{x=1}^{t} \alpha_2^{t-x} . A_x^2 \tag{18}$$

Take the exponential on both sides as represented below equations:

$$E[H_z^2] = E[(1 - \alpha_{\boldsymbol{2}}) \sum_{x=1}^{t} \alpha_2^{t-1} . A_x^2 \tag{19}$$

$$E[H_z^2] = E[A_z^2].(1 - \alpha_{\boldsymbol{2}}) \sum_{x=1}^{t} \alpha_2^{t-x} + \aleph \tag{20}$$

$$E[H_z^2] = E[A_z^2].(1 - \alpha_2^t) + \aleph \tag{21}$$

In equations (19)-(21) $\aleph = 0$ is fixed otherwise $\aleph$ is small. In SA-Adam, a stage-adaptive scaling factor β_s is introduced to highlight clinically critical stages.

$$\vartheta_{z+1} = \vartheta_z - \aleph\beta_s \frac{\widehat{\mathrm{H}_{\boldsymbol{z}}}}{\sqrt{\widehat{A_z} + \epsilon}} \tag{22}$$

$$\beta_s = 1 + \rho.L_{severity} \tag{23}$$

In Equations (22- 23), ρ controls the influence of severity-aware adaptation. The SA-Adam preserves the adaptive learning advantages of Adam. The integrated stage-aware optimization results in improved convergence stability, balanced stage learning, and reduced severe cross-stage misclassification in the proposed tri-branch BiLSTM framework.

4. Results and Discussion

4.1 Hyperparameter Settings and Complexity Analysis

This section presents the experimental results of the proposed tri-branch DL framework for DCM stage classification. The synthesized dataset is split into 70% for training, 20% for testing, and 10% for validation. The learning rate was set at 0.001. The proposed model finetunes a few hyperparameters to strike a balance between model performance and computational efficiency, as demonstrated in Table 1. The proposed DCM stage classification framework resulted in a balanced computational complexity. The NAFE phase has a complexity of $O(n * d)$, where n denotes the number of input features and d denotes the latent feature dimension. In the tri-branch BiLSTM, each BiLSTM has a time complexity of $O(T * h^2)$, results $3 * O(T * h^2)$. The WTA performs an Argmax operation, attains a complexity of $O(1)$. This makes it computationally efficient and suitable for real-time and clinical decision-support environments.

Table 1: Hyperparameter settings of the proposed framework

Component	Hyperparameter	Value
	Backbone	MLP

NAFE	Dense layer units	128
	Dense activation	GELU
	Attention mechanism	Feature-wise self-attention
	Dropout rate	0.3
Tri-branch BiLSTM	Number of branches	3
	BiLSTM hidden units	64
	Sequence length	10
	Latent dimension	64
Training	Batch size	64
	Epochs	20
	Learning rate	0.001
	Optimizer	SA-Adam

4.2 Performance Analysis of the Proposed Model

This section presents the performance analysis of the proposed model with the metrics of accuracy, precision, recall, f1-score, specificity, Matthews Correlation Coefficient (MCC), False Positive Rate (FPR), False Negative Rate (FNR), Area Under Curve (AUC), and Negative Predictive Value (NPV) values, as presented in Table 2. The proposed Tri-branch BiLSTM model for DCM stage classification achieves a highest accuracy of 99.2% and a lowest FNR value of 0.021. This highlights the effectiveness of the proposed model in DCM stage classification.

Table 2: Performance of the proposed tri-branch BiLSTM model for DCM stage classification

Metrics	Stage I	Stage II	Stage III	Average
Accuracy (%)	99.3	99.1	99.1	99.2
Precision (%)	99.1	98.4	98.3	98.61
Recall (%)	98.5	97.7	97.5	97.89
F1-score (%)	98.7	98.2	97.8	98.25
Specificity (%)	99.5	99.2	99.3	99.32
MCC (%)	97.7	96.2	96.1	96.78
FPR	0.005	0.008	0.007	0.008
FNR	0.014	0.023	0.026	0.021
AUC (%)	99.2	98.3	98.5	98.67
NPV (%)	99.4	98.7	99.2	99.13

4.3 Model Training Performance

The accuracy and loss graphs deliver a visualization of the learning progress of the proposed Tri-branch BiLSTM model over epochs. The accuracy of the proposed model shows reliable enhancement during the training process. This highlights its adaptability in learning from the synthesized data. The loss values follow a steady trend, decreasing during the training. This shows the effective convergence of the proposed model for DCM stage classification, as represented in Figure 4.

Figure 5 shows the confusion matrix of the Tri-branch BiLSTM model using three classes, namely Stage I, Stage II, and Stage III. It shows the true positive and true negative classification percentages for each class. Each row denotes the predicted labels, and each column denotes the true labels. This highlights where the model excels and where it struggles in the DCM stage classification.

Figure 6 shows the Receiver Operating Characteristic (ROC) curve and its respective AUC values to assess the performance of the proposed model. This marks the true positive rate as opposed to the false positive rate. The values

of the AUC range between 0 to 1. Higher values denote higher performance and lower values denotes poor performance. It highlights the estimation of how well a model can differentiate between positive and negative rates.

Figure 7 represents the original 34-dimensional latent feature space learned by the proposed model. Each point in the plot corresponds to a patient sample described by 34 extracted features. T-distributed Stochastic Neighbor Embedding (t-SNE) projects these high-dimensional embeddings on two axes. This dimensionality reduction shows that the acquired feature representations maintain high discriminative structure. The separation between clusters shows that the model effectively encodes stage-specific patterns in the high-dimensional space. This confirms that the proposed feature learning framework produces a structured and separable embedding space.

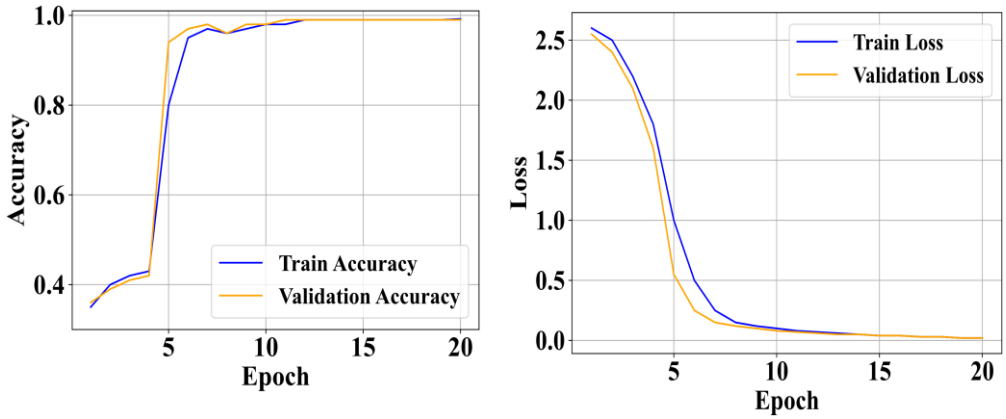

Figure 4: Accuracy and loss curve over epochs of the proposed model

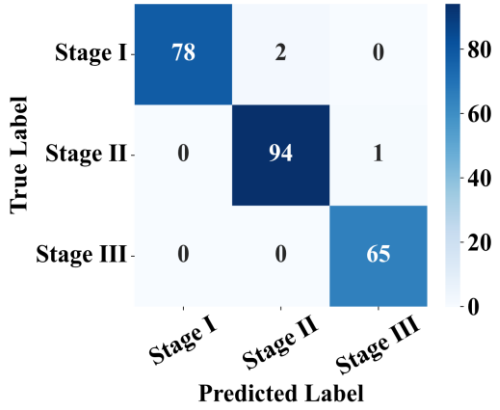

Figure 5: Confusion matrix of the proposed DCM stage classification model

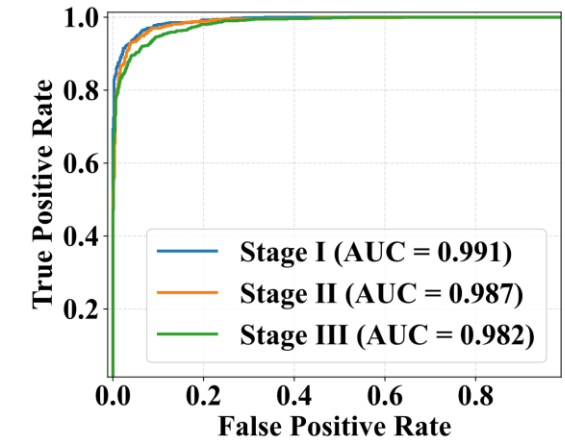

Figure 6: Receiver Operating Characteristic Area Under the Curve

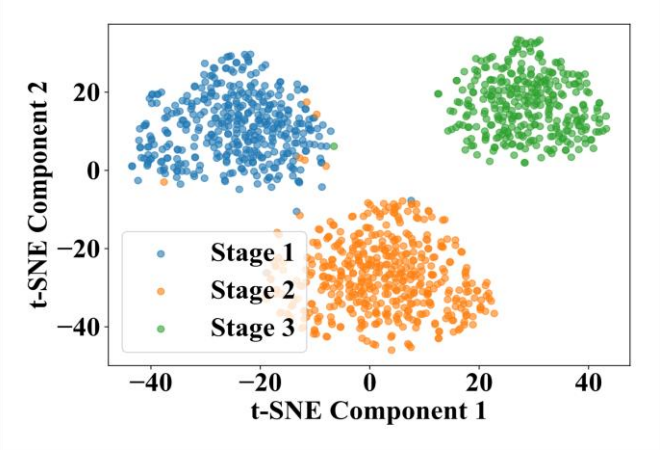

Figure 7: Visualization of latent feature space

Figure 8 presents density distribution analyses of key across three stages of DCM. It demonstrates the changes in the demographic, metabolic, structural, biomarker, and functional characteristics as the disease advances. The age structure presents a progressive change of stages. This confirms the validity of realistic demographic variation in the dataset that does not concentrate samples on a small age range. BMI distribution shows that there is a significant difference in stages. This indicates preserved cardiometabolic heterogeneity and suggests that metabolic burden contributes differently to disease progression. The LV mass index shows a rightward shift toward higher values in advanced stages. This highlights progressive left ventricular structural remodeling, which is a known hallmark of worsening DCM severity. The NT-proBNP distribution exhibits a highly skewed pattern. This confirms conservation of biologically realistic biomarker behavior. This confirms increased cardiac stress in advanced disease conditions. GLS distribution reveals the strain impairment that increases between Stage I and Stage III. These distributions confirm that the dataset captures clinically consistent progression patterns. This maintains stage-specific phenotypic characteristics essential for model training.

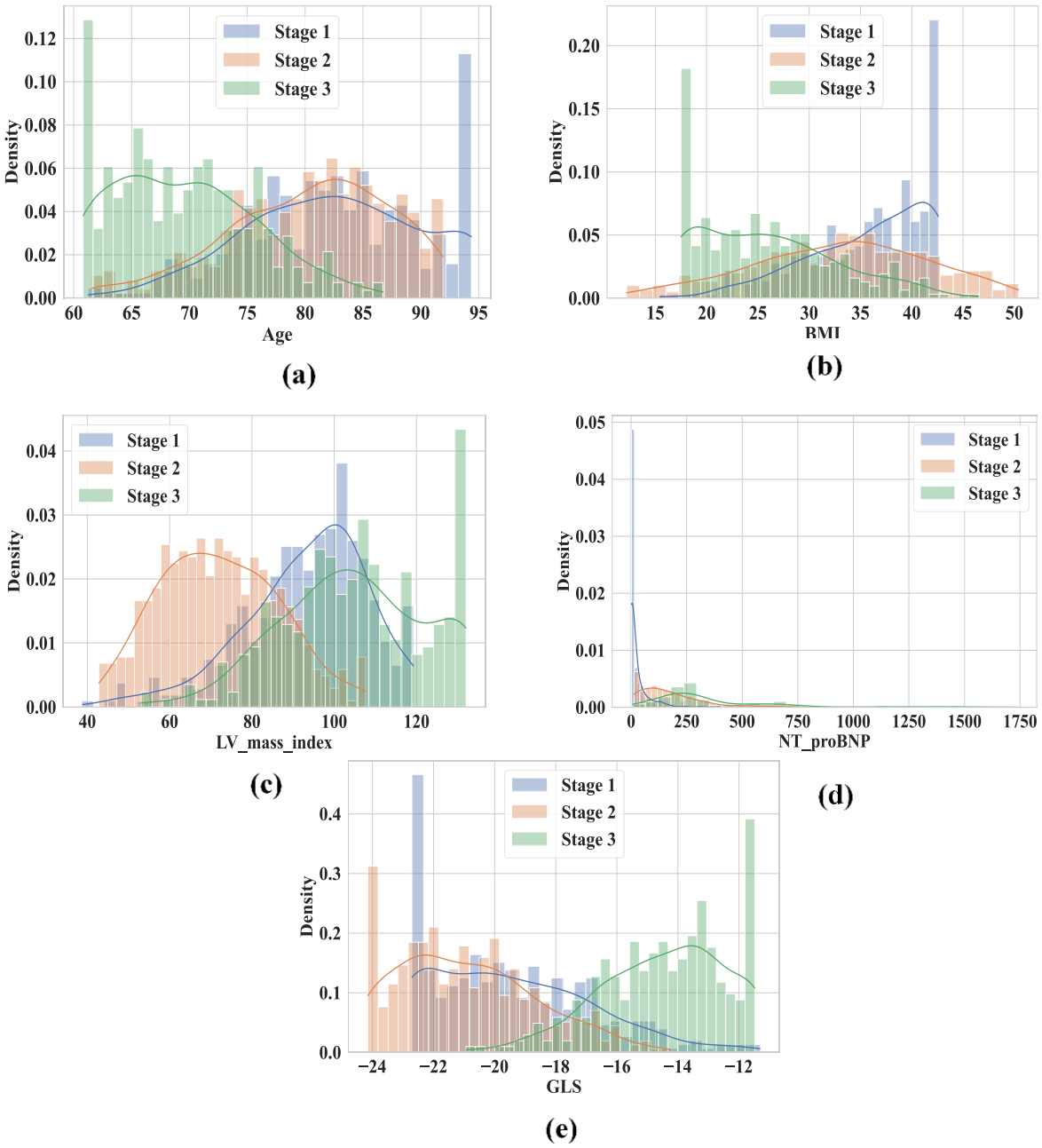

Figure 8: Stage-wise distribution analysis of clinical and cardiac indicators in the DCM dataset. (a) Age, (b) BMI, (c) LV mass index, (d) NT-proBNP, and (e) Global Longitudinal Strain (GLS).

Table 3: Generated activation scores for DCM stage classification

Samples	Stage I	Stage II	Stage III
20	-3.96016	-2.90078	2.744449
40	2.741279	-1.6854	-2.53251
60	-3.71251	-2.74973	2.576525

80	3.814747	-2.19219	-3.54689
100	-3.79933	-2.85987	2.664704
120	-1.16983	0.936135	-1.42654
140	-0.234	0.068921	-0.64975
160	-2.93448	-2.2872	2.042534
180	-1.48201	1.587552	-2.31969
200	-1.65459	1.341388	-1.76744
220	1.30117	-0.88372	-1.29856
240	-2.61548	2.216296	-2.62979

Table 3 shows the activation scores generated by the proposed model for DCM stage classification. It shows the stage-wise activation scores generated by the tri-branch BiLSTM model. Each corresponds to one branch of the model. This reflects how strongly that branch identifies its learned disease-stage characteristics. The fluctuations in activation values indicate the dynamic competition among branches. This validates the effectiveness of stage-specific pattern learning and supporting the WTA decision mechanism used for final classification. This confirms that the tri-branch produces competitive activation responses. This enables confident and interpretable DCM stage assignment.

4.4 Result Analysis of Proposed SA-Adam Optimizer

The training accuracy of the proposed model using the proposed SA-Adam optimizer is represented in Figure 9 across 20 epochs. The SA-Adam outperforms the Adam [30], AdamW [31], RAdam [32], Nadam [33], and AdaMax [34] optimizers. The proposed SA-Adam expands towards improved and stable accuracy faster. At the beginning of the epochs, SA-Adam shows a distinctively sharper rise. This shows a superior convergence speed due to the enhanced learning features. As training progresses, SA-Adam reaches an accuracy of 99.2%, while Adam achieves around 96% by the 20th epoch. This shows the SA-Adam's efficiency in stabilizing gradient updates and improving overall learning efficiency. The graph supports the use of SA-Adam as an efficient optimizer for the proposed DCM stage classification.

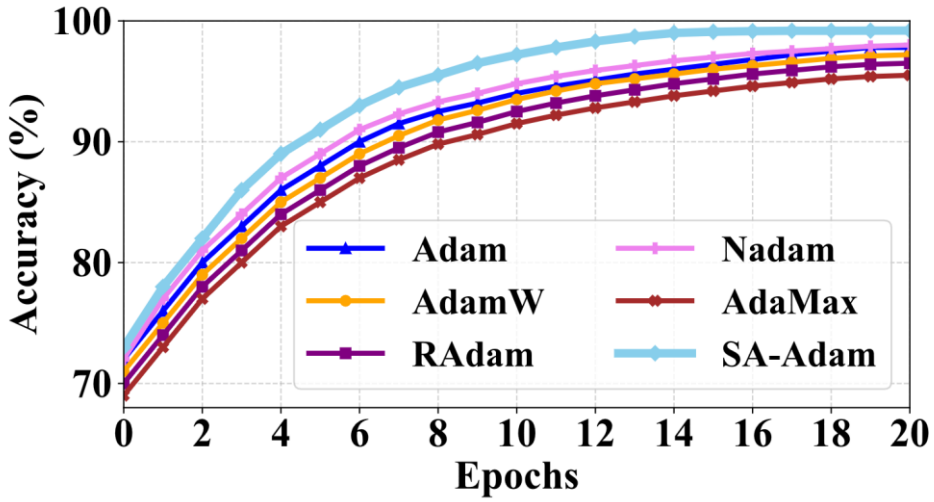

Figure 9: Analysis of the proposed SA-Adam optimizer's performance

Table 4: Convergence analysis of proposed SA-Adam optimizer

Optimizer	Epochs to Converge	Final Loss
Adam [30]	18	0.26
AdamW [31]	16	0.21
RAdam [32]	17	0.24
Nadam [33]	15	0.22
AdaMax [34]	19	0.28
SA-Adam	**13**	**0.12**

Table 4 highlights the convergence performance of the proposed SA-Adam optimizer over the existing optimizers Adam [30], AdamW [31], RAdam [32], Nadam [33], and AdaMax [34]. This table proves SA-Adam obtains faster convergence and lower final training loss. This highlights the contribution of the proposed stable optimization with the traditional Adam optimizers.

4.5 Ablation Analysis

This section presents the ablation analysis of the proposed tri-branch BiLSTM for DCM stage classification. This analyzes the configuration of the proposed tri-branch BiLSTM with different blocks. We also analyzed the significant performance of the individual NAFE and SA-Adam optimizer. This highlights that the proposed combination of modified NAFE, tri-branch BiLSTM, and SA-Adam attained a better performance, as represented in Table 5.

Table 5: Ablation analysis of the proposed framework

Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Specificity (%)
Full Model (Proposed)	**99.2**	**98.61**	**97.89**	**98.25**	**99.32**
Single-head BiLSTM	96.4	95.72	94.85	95.28	97.11
Without attention mechanism	97.3	96.12	95.43	95.74	97.95
Without loss function	97.8	96.85	96.09	96.47	98.23
Without NAFE	96.7	94.59	94.72	95.25	94.31
Replace SA-Adam with Standard Adam	98.1	97.02	96.45	96.72	98.41
Without CTGAN Augmentation	94.8	93.42	92.67	93.01	95.86
Without optimizer	96.7	95.59	94.92	95.25	97.31
With LSTM	97.4	96.42	95.78	96.05	98.01

4.6 Comparative Analysis

Table 6: Comparative analysis with the state-of-the-art approaches

Reference	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
[19]	HeartEnsembleNet	97.07	96.01	98.58	-
[20]	Multimodal fusion model	82.95	73.25	73.09	73.25
[21]	CDSS	95.9	97.1	94.67	95.35
[23]	ResNet-18	87	67	89	76
Proposed	**Tri-branch BiLSTM**	**99.2**	**98.61**	**97.89**	**98.25**

Table 6 compares the proposed DCM stage classification approach with the existing state-of-the-art approaches. The existing HeartEnsembleNet, Multimodal fusion model, CDSS, ResNet-18, and K-Means clustering algorithm approaches attained an accuracy of 97.07%, 82.95%, 95.9%, 87%, and 72.43%, respectively. While the proposed model attains an accuracy of 99.2%, a precision of 98.61%, a recall of 97.89%, and an F1-score of 98.25%. This shows that the proposed framework attained a better performance compared to the existing approaches.

4.7 Discussion

This study introduced a stage-aware learning framework for DCM stage classification. The study addressed the challenges with the early and subclinical disease staging. The NAFE plays a central role by transforming heterogeneous clinical, biomarker, and imaging features into a compact latent space. The attention-based reweighting learned by NAFE enhances interpretability by highlighting clinically relevant variables. This is supported by the observed structured clustering in the t-SNE visualization. The tri-branch BiLSTM architecture contributes to the core stage-discriminative component of the proposed framework. The independent BiLSTM branch is assigned to each DCM stage, enabled to learn nuanced and stage-specific temporal and structural patterns. This improves the recall for early and intermediate stages. The competitive WTA mechanism enforces hard stage assignment based on dominant activation scores. This ensures confident and definite predictions. This design reduces the risk of ambiguous outputs and minimizes severe misclassifications across non-adjacent stages. This enhances patient safety and decision reliability.

The experimental results validate the effectiveness of the proposed methodology. Stage-wise performance analysis shows balanced precision and recall across all disease stages. The lower error values confirm that most misclassifications are confined to adjacent stages. The WTA decision-making shows that the model consistently learns well-defined stage boundaries. This supports the effectiveness of competitive learning. These findings suggest that the proposed work is well aligned with the progressive nature of DCM. However, the study acknowledged some limitations. First, the limited dataset size to the number of clinical sources. It has an influence on its generalizability. The WTA mechanism does not directly measure the uncertainty in prediction. Also, the proposed structure is concerned with the categorization of phases and does not directly give prognostic prediction. Future work will focus on addressing these limitations and extending the clinical applicability of the proposed framework. The integration of multimodal imaging data may further strengthen temporal modeling predictions. Finally, extending the framework toward treatment recommendation systems represents a direction for real-world clinical decision support tools.

5. Conclusion

The study developed a Tri-branch BiLSTM framework for DCM stage classification. The proposed framework integrated NAFE, tri-branch BiLSTM, and competitive decision making. The NAFE efficiently transformed heterogeneous clinical, biomarker, and variables into a compact, noise-reduced latent representation. The tri-branch BiLSTM architecture independently learns characteristic patterns of early, intermediate, and advanced DCM stages. This enabled capture of subtle disease progression dynamics. The ArgMax-based WTA mechanism ensures confident and definite stage assignment. Experimental analysis, including stage-wise performance evaluation, Optimizer analysis, ablation analysis, and comparative analysis, demonstrates that the proposed approach achieves better performance. These confirm that the methodological design enhanced the accuracy, interpretability, and reliability of

the proposed framework. This makes the proposed framework highly suitable for real-world clinical decision support and early intervention planning.

Statements and Declarations

Authors' contributions: All authors contributed to the conception of the problem setting and overall design of the work. K.B built the conceptualization and methodology, J.S implemented the work, Validation and writing was done by both. This version was revised and improved by both authors, who also read and approved the final manuscript.

Funding: No funding was received for conducting this study.

Conflict of interest: The authors declare that they have no conflict of interest.

Availability of data and materials: Dataset will be provided as per the requested by the author.

Ethical approval: The research is original and all the figures and tables are created by the authors of this manuscript.

Consent to participate: Not applicable.

Consent for publication: Both authors agree with the submission of the manuscript to this journal and possible publication afterwards.

References

1. M. L. Peng, Y. Fu, C. W. Wu, Y. Zhang, H. Ren, S. S. Zhou. "Signaling pathways related to oxidative stress in diabetic cardiomyopathy", Frontiers in endocrinology, 13 (2022), pp. 907757, 2022.
2. B. T. Xu, S. R. Wan, Q. Wu, Y. H. Xing, Y. Q. He, W. Huang, Y. Long, C.X. Zhang, Y. Xu, Z. Z. Jiang. "BDH1 overexpression alleviates diabetic cardiomyopathy through inhibiting H3K9bhb-mediated transcriptional activation of LCN2", Cardiovascular Diabetology, 24(1), pp. 101, 2025.
3. V. Blumer, J. L. Januzzi Jr, Y. Liu, J. Butler, J. A. Ezekowitz, R. Perfetti, J. Rosenstock, S. Del Prato, W.W., Tang, A. Urbinati, C. S. Lam. "Sex differences in diabetic cardiomyopathy and treatment response to AT-001: insights From the ARISE-HF study", Heart Failure, 14(1), pp. 102433, 2026.
4. M. Bellemare, L. Bourcier, J. Iglesies-Grau, J. Boulet, E. O'Meara, N. Bouabdallaoui. "Mechanisms of diabetic cardiomyopathy: Focus on inflammation", Diabetes, Obesity and Metabolism, 27(5), pp. 2326-2338, 2025.
5. J. Wang, X. Huang, H. Liu, Y. Chen, P. Li, L. Liu, J. Li, Y. Ren, J. Huang, E. Xiong, X. Dai. "Empagliflozin ameliorates diabetic cardiomyopathy via attenuating oxidative stress and improving mitochondrial function", Oxidative medicine and cellular longevity, 2022(1), pp. 1122494, 2022.
6. P. Mone, A. Lombardi, U. Kansakar, F. Varzideh, S. S. Jankauskas, A. Pansini, S. Marzocco, S. De Gennaro, M. Famiglietti, G. Macina, G. Santulli. "Empagliflozin improves the MicroRNA signature of endothelial dysfunction in patients with heart failure with preserved ejection fraction and diabetes", The Journal of Pharmacology and Experimental Therapeutics, 384(1), pp. 116-122, 2023.
7. Z. Luo, D. Pi, T. Xi, W. Jiang, F. Qiu, J. Yang. "Predicting diabetic cardiomyopathy in type 2 diabetes: development and validation of a nomogram based on clinical and echocardiographic parameters", Frontiers in Endocrinology, 16(1), pp. 1641114, 2025.
8. B. Balakrishnan, D. M. Somashekara, V. Nayak, R. C. Hariharapura. "Diabetic cardiomyopathy: a comprehensive review of diagnosis, management, and future directions", Diabetology & Metabolic Syndrome, 17(1), pp. 422, 2025.
9. J. L. Huo, Q. Feng, S. Pan, W. J. Fu, Z. Liu, Z. Liu. "Diabetic cardiomyopathy: Early diagnostic biomarkers, pathogenetic mechanisms, and therapeutic interventions", Cell death discovery, 9(1), pp. 256, 2023.
10. E. K. Oikonomou, R. Kher. "Machine learning in precision diabetes care and cardiovascular risk prediction," Cardiovascular Diabetology, 22(1), pp. 259, 2023.
11. K. Irlik, H. Aldosari, M. Hendel, H. Kwiendacz, J. Piaśnik, J. Kulpa, P. Ignacy, S. Boczek, M. Herba, K. Kegler, K. Nabrdalik. "Artificial intelligence-enhanced electrocardiogram analysis for identifying cardiac autonomic neuropathy in patients with diabetes," Diabetes, Obesity and Metabolism, 26(7), pp. 2624-2633, 2024.
12. C. Luo, Y. Zhu, Z. Zhu, R. Li, G. Chen, Z. Wang. "A machine learning-based risk stratification tool for in-hospital mortality of intensive care unit patients with heart failure", Journal of translational medicine, 20(1), pp. 136, 2022.
13. Y. Liu, J. Zhang, Q. Han, Y. Li, Y. Xue, X. Liu. "Identification of biomarkers associated with macrophage polarization in diabetic cardiomyopathy based on bioinformatics and machine learning approaches," Life Sciences, 364(3), pp. 123443, 2025.
14. Y. Kumar, A. Koul, R. Singla, M. F. Ijaz. "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda," Journal of ambient intelligence and humanized computing, 14(7), pp. 8459-8486, 2023.

15. A. G. D. J. Raju, S. S. Manohar, T. C. J. Jeaunita, “Mitigating catastrophic forgetting in imitation learning for embodied AI using progressive neural networks,” Econom. Comput. Econom. Cybernet. Stud. Res., 59(3), pp. 97- 112, 2025.
16. E. H. Houssein, R. E. Mohamed, A. A. Ali. “Heart disease risk factors detection from electronic health records using advanced NLP and deep learning techniques,” Scientific Reports, 13(1), 7173, 2023.
17. S. Deng, F. Tayefi, Y. Jin. “Metabolic-stress-induced mitochondrial calcium dysregulation: a central hub in diabetic cardiomyopathy pathogenesis and treatment,” Frontiers in Endocrinology, 16(3), 1696344, 2025.
18. H. Tanaka. “Evolving role of echocardiography across the heart failure spectrum: from prevention to advanced management,” Heart Failure Reviews, 31(1), 6, 2026.
19. S. A. J. Zaidi, A. Ghafoor, J. Kim, Z. Abbas, S. W. Lee. “HeartEnsembleNet: An innovative hybrid ensemble learning approach for cardiovascular risk prediction,” In Healthcare, 13(5), 507, 2025.
20. J. Botros, F. Mourad-Chehade, D. Laplanche. “Explainable multimodal data fusion framework for heart failure detection: Integrating CNN and XGBoost,” Biomedical Signal Processing and Control, 100(2), 106997, 2025.
21. U. Nagavelli, D. Samanta, P. Chakraborty. “Machine learning technology-based heart disease detection models,” Journal of Healthcare Engineering, 2022(1), 7351061, 2022.
22. M. M. Oo, C. Gao, C. Cole, Y. Hummel, M. Guignard-Duff, E. Jefferson, J. Hare, A.A. Voors, R.A. de Boer, C.S. Lam, C. C. Lang. “Artificial intelligence-assisted automated heart failure detection and classification from electronic health records,” ESC heart failure, 11(5), 2769-2777, 2024.
23. K. Nabrdalik, K. Irlik, Y. Meng, H. Kwiendacz, J. Piaśnik, M. Hendel, P. Ignacy, J. Kulpa, K. Kegler, M. Herba, U. Alam. “Artificial intelligence-based classification of cardiac autonomic neuropathy from retinal fundus images in patients with diabetes: the Silesia Diabetes Heart Study,” Cardiovascular Diabetology, 23(1), 296, 2024.
24. M. W. Segar, M. S. Usman, K. V. Patel, M. S. Khan, J. Butler, L. Manjunath, C.S. Lam, S. Verma, D. Willett, D. Kao, A. Pandey, Development and validation of a machine learning-based approach to identify high-risk diabetic cardiomyopathy phenotype. European Journal of Heart Failure, 26(10), 2183-2192, 2024.
25. K. Mallikharjuna Rao, G. Saikrishna, K. Supriya. “Data preprocessing techniques: emergence and selection towards machine learning models-a practical review using HPA dataset,” Multimedia Tools and Applications, 82(24), 37177-37196, 2023.
26. D. Geem, D. Hercules, R. S. Pelia, S. Venkateswaran, A. Griffiths, J. D. Noe, J.L. Dotson, S. Snapper, S. Rabizadeh, J.R. Rosh, S. Kugathasa. “Progression of pediatric Crohn’s disease is associated with anti–tumor necrosis factor timing and body mass index Z-score normalization,” Clinical Gastroenterology and Hepatology, 22(2), 368-376, 2024.
27. S. Afzal, B. M. Ziapour, A. Shokri, H. Shakibi, B. Sobhani. “Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms,” Energy, 282(2), 128446, 2023.
28. P. Singla, M. Duhan, S. Saroha. “An ensemble method to forecast 24-h ahead solar irradiance using wavelet decomposition and BiLSTM deep learning network,” Earth Science Informatics, 15(1), 291-306, 2022.
29. K. A. Aziz, L. Elsonbaty. “Comparative study of different classification methods and winner takes all approach,” Civ Eng J, 10(10), 3370-85, 2024.
30. H. Sun, W. Zhou, J. Yang, Y. Shao, L. Xing, Q. Zhao, L. Zhang. “An improved medical image classification algorithm based on Adam optimizer,” Mathematics, 12(16), 2509, 2024.
31. A. M. Moraes, L. F. Pugliese, R. F. D. Santos, G. B. Vitor, R. A. D. S. Braga, F. R. D. Silva. “Effectiveness of YOLO architectures in tree detection: Impact of hyperparameter tuning and SGD, Adam, and AdamW optimizers,” Standards, 5(1), 9, 2025.
32. I. A. Alshalabi, T. Alrawashdeh, S. Abusaleh, M. Z. Alksasbeh, K. Alemerien, S. Al-Eidi, H. Alshamaseen. “Detection of COVID-19 using EfficientnetV2-XL and Radam Optimizer from Chest X-ray Images,” Journal of Applied Data Sciences, 6(2), 1143-1158, 2025.
33. Q. Zhang, Y. Zhang, Y. Shao, M. Liu, J. Li, J. Yuan, R. Wang. “Boosting adversarial attacks with nadam optimizer,” Electronics, 12(6), 1464, 2023.
34. M. Uppal, D. Gupta, S. Juneja, T. R. Gadekallu, I. El Bayoumy, J. Hussain, S. W. Lee. “Enhancing accuracy in brain stroke detection: Multi-layer perceptron with Adadelta, RMSProp and AdaMax optimizers,” Frontiers in Bioengineering and Biotechnology, 11(3), 1257591, 2023.