Mathematical Modelling of Engineering Problems
Vol. 12, No. 10, October, 2025, pp. 3361-3375
Journal homepage: http://iieta.org/journals/mmep

Enhanced Atrial Fibrosis Detection Using 2D Echocardiogram Images with an Advanced Deep Learning Framework and Weighted Multi-Feature Fusion

Pilli Sudheer[1*], Balasubramaniam Kirubagari[1], Ayyavoo Annamalai Giri[2]

[1] Department of Computer Science and Engineering, Annamalai University, Annamalainagar 608002, India
[2] Department of Computer Science and Engineering, Marri Laxman Reddy Institute of Technology and Management, Telangana 500043, India

Corresponding Author Email: p.sudheer@cvr.ac.in

https://doi.org/10.18280/mmep.121003

Received: 16 July 2025
Revised: 18 September 2025
Accepted: 23 September 2025
Available online: 31 October 2025

Keywords:
atrial fibrosis detection, HDF-RBMO, hybrid (2D, 3D) convolution based adaptive Trans-MobileUNet++, region-based vision transformer, CNNs

ABSTRACT

Atrial Fibrillation (AF) is a chronic condition characterized by structural variations within the atria. Fibrosis, identified based on the formation of collagen within the interstitial region of the heart is considered a result of arrhythmogenic structural changes. Recently, the implementation of deep learning approaches in medical science has become challenging because of their limited effectiveness in handling patterns or structural changes. Moreover, they demand large quantities of diverse and large-scale data for designing a robust detection framework. Considering the limited availability of resources, it is crucial to develop an effective approach to perform early detection of AF. Therefore, a deep learning-based AF detection framework is developed to track the progression of the disease. The required 2D Echocardiogram (ECG) images are collected from benchmark sources. Features including fuzzy entropy vectors, wavelet packet energy vectors, and hierarchical theory vectors are extracted from 2D images and concatenated to obtain Feature Set I. Next, the Region-based Vision Transformer (R-ViT) is utilized to retrieve Feature Set II from the input 2D images. Later, the deep features are extracted by employing Convolutional Neural Networks (CNNs). The weighted multi-feature fusion is performed using the suggested Hybrid Dark Forest and Red-Billed Blue Magpie Optimizer (HDF-RBMO) to tune the optimal features. These features from 2D ECG and the 3D ECG image series are input into the developed Hybrid (2D, 3D) Convolution-based Trans-MobileUNet++ (HC-TMU) for detecting AF. The proposed AF detection model may assist in analyzing high blood pressure and other heart diseases. The results of the developed model are compared with a previously developed detection model to ensure the presented approach's effectiveness.

1. INTRODUCTION

Atrial Fibrillation (AF) is considered one of the most frequently occurring chronic arrhythmia that remains a scientific issue that continues mysterious behavior even after several investigations [1]. The reason that leads to AF has not been identified, and suggesting an effective treatment plan is exceedingly complex in these cases [2]. The main symptom indicated the occurrence of AF which is the unusual contraction within the upper atrium of the heart and the Echocardiogram (ECG) signal denotes it by the reduction of sinus P wave [3]. The recognition of AF across multiple ECGs shows diverse quality and signal length. Ambiguity labels originated from various kinds of arrhythmia pulses within identical records, varying human anatomy, and issues in separating ECG signal features.

As a consequence, the strategy selected for AF must be capable of handling these scenarios while preserving system efficiency [4]. The existing detection techniques based on the computer technology introduced for enhancing the AF detection performance in traditional machine learning approaches created favorable outcomes. Traditional machine learning techniques depend upon classical approaches to perform feature extraction and selection that demand multiple phases to accomplish the categorization procedure [5]. Utilizing deep learning techniques for AF detection is easy and does not employ traditional hand-crafted engineering techniques to attain significant features. Yet, it is hard to determine a proper framework based on a deep learning approach, as it demands for huge quality of data for processing [6]. At present, only certain public datasets describing AF symptoms are available, including normal data more than AF problems or imbalanced information. Research concerning imbalanced circumstances primarily focuses on offering better classification outcomes for unique subclass [7].

However, classical machine learning methods for classifying AF are frequently employed for decreasing overall error rates instead of analyzing unique classes or unbalanced information [8]. Moreover, the rapid growth of deep learning approaches including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) Autoencoder (AE), and Deep Neural Networks (DNNs) offers better

efficiency in addressing the issues faced by the classical AF detection techniques [9]. Among the other classical approaches, CNN exceeds traditional models in processing 2D information including ECG image series. Yet, various researchers revealed that the CNN models provide more effective performance in processing ECG image series as 1D data than the approaches based on RNN and DNN [10]. Moreover, a 1-dimensional CNN (1D-CNN) is effective in processing long-term ECG information using a rapid and reliable method, as well as analyzing morphological features and gathering information [11]. Additionally, CNNs are capable of producing specific characteristics regarding the ECG signal series in order to recognize specific trends within the convolutional space. As every patch within the model experiences similar variations described by the convolution space, the structure acquired from a particular location is frequently identified by an alternate location, ensuring the transformation of 1D convolution systems is uniform [12].

In contrast, 1D-CNNs are often applied to temporal biomedical signals such as ECGs. However, since this study focuses on 2D echocardiographic imaging, we primarily adopt 2D- and 3D-convolutional strategies tailored for spatial and spatiotemporal cardiac structure analysis.

Building on these principles, we propose a novel deep learning–based AF detection framework with the following key contributions:

- **Hybrid feature fusion for early AF detection:** The design an innovative framework that integrates CNN-based spatial features, Region-based Vision Transformer (R-ViT) contextual features, and handcrafted descriptors. Weighted feature fusion ensures early and accurate AF detection while enabling longitudinal monitoring of disease progression.
- **Hybrid Dark Forest and Red-Billed Blue Magpie Optimizer (HDF-RBMO) optimization for feature selection:** The HDF-RBMO combines the strengths of differential forest algorithms (DFA) and RBMO techniques. This hybrid optimizer selects the most discriminative features and assigns adaptive weights based on correlation and relief scores, enhancing fusion quality and classification performance.
- **To develop the Hybrid Convolutional Trans-MobileUNet++ (HC-TMU) architecture for echocardiographic segmentation and classification:** To develop the HC-TMU, it integrates 2D convolutional layers for static echocardiographic images and 3D convolutional layers for echocardiographic sequences. This architecture simultaneously segments cardiac structures and performs AF classification, enabling a comprehensive analysis of atrial remodeling.

To address these limitations, this study makes the following key contributions:

- To develop a hybrid deep learning framework that integrates CNNs, an R-ViT, and handcrafted descriptors to capture both local and global echocardiographic features.
- To propose a novel HDF-RBMO for multi-feature fusion and feature selection, robust learning even with limited and imbalanced datasets are ensured.
- To introduce an HC-TMU architecture for joint 2D and 3D echocardiographic analysis, superior accuracy and dice score compared to state-of-the-art AF detection methods are achieved.

The outline of the designed deep learning-based AF detection framework is described here. Recent investigations executed to address the challenges of AF detection are explained in the 2nd Section. The architectural view of the developed AF detection model is described in the 3rd Section. The detailed explanation based on feature extraction and weight optimization along with the developed hybrid optimization is presented in the 4th Section. The development and implementation of the designed segmentation framework are discussed in the 5th Section. At last, the result and conclusion of the developed model are presented in the 6th and 7th Sections.

2. LITERATURE REVIEW

Research on AF detection using artificial intelligence has evolved from traditional machine learning toward deep learning–based frameworks. Existing studies can broadly be categorized into three groups: CNN-based approaches, RNN-based approaches, and hybrid models [13].

(1). CNN-based models

CNNs have been widely applied to extract spatial or spectral features from ECG and echocardiographic signals.

Pourbabaee et al. [1] proposed a deep convolutional neural network (DCNN) for AF detection from ECG time-series images, showing improvements over traditional classifiers. Cao et al. [2] developed Multi-Scale Decomposition Enhanced Residual CNNs (MSResNet) and FDResNets [14], combined via transfer learning, to improve F1-score and accuracy.

Nurmaini et al. [3] integrated Discrete Wavelet Transform (DWT) with 1D-CNNs for multiclass AF classification, demonstrating better generalization and early diagnostic capability. Chandra et al. [15] employed wavelet and Fourier transforms to generate 2D spectrograms fed into a DCNN, reporting enhanced AF detection reliability. Schwab et al. [7] applied a CNN-based framework validated on external datasets, demonstrating improved prediction of high-risk AF cases compared to classical models. Although CNNs can extract powerful features, they often require large balanced datasets, are sensitive to noise, and may lack interpretability in clinical contexts.

(2). RNN-based models

RNNs and their variants (e.g., LSTMs) are particularly suited for sequential biomedical signals.

Faust et al. [4] applied LSTMs to heart rate signals for AF detection, showing robustness to noisy and incomplete data. Andersen et al. [6] combined CNN and RNN layers for end-to-end AF classification from RR intervals, achieving strong performance even on unseen datasets. RNNs are effective for temporal dependencies but face challenges such as vanishing gradients, high computational cost, and limited scalability to large echocardiographic datasets [16].

(3). Hybrid models

Hybrid approaches combine CNNs with RNNs or other optimization strategies to address data imbalance and enhance generalization.

Petmezas et al. [5] proposed a CNN-LSTM model trained with focal loss to mitigate data imbalance, achieving higher sensitivity and specificity. While hybrids improve classification, they often increase architectural complexity [17], require fine-tuned hyperparameters, and lack efficient segmentation pipelines for echocardiographic imaging [18].

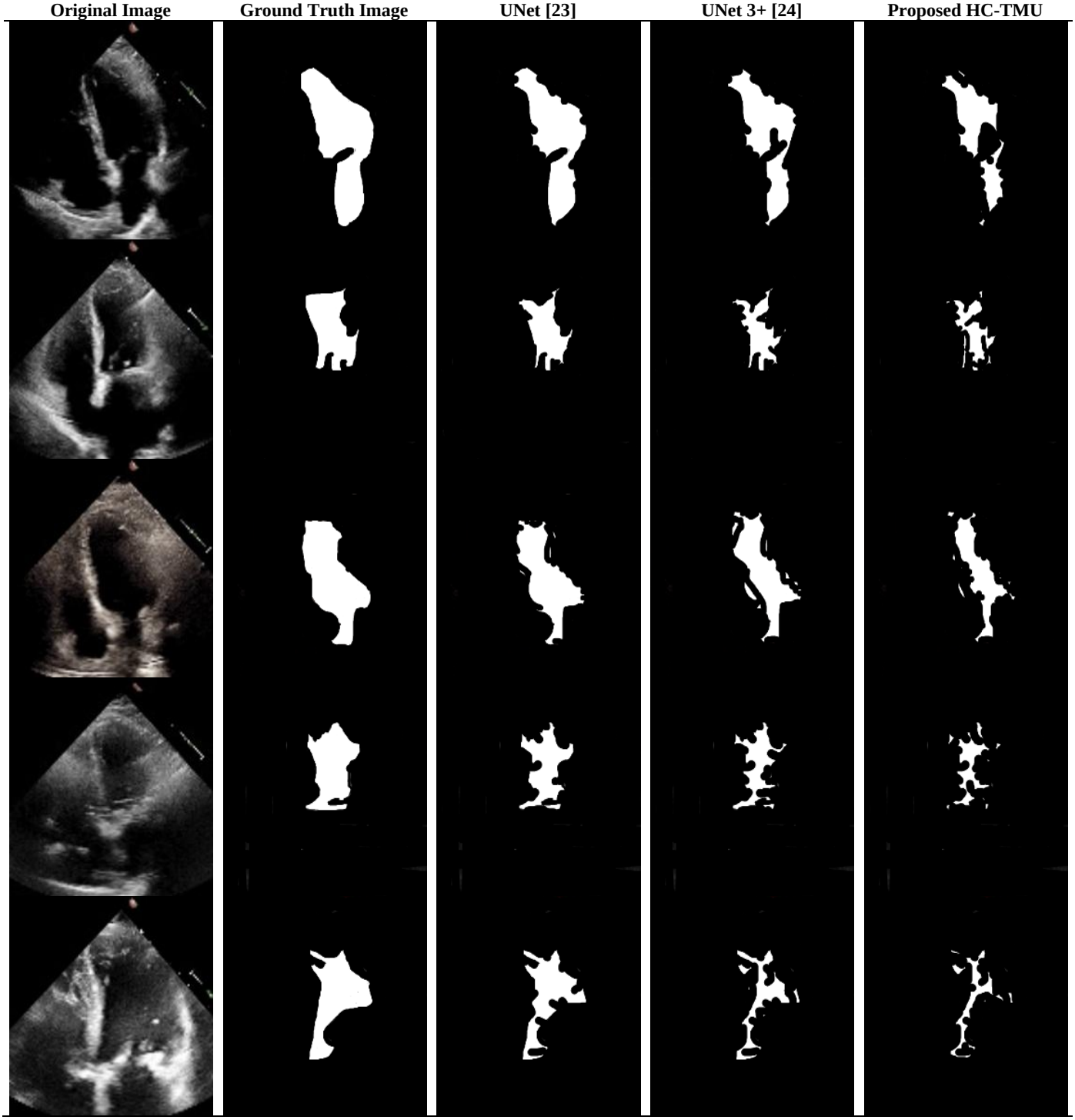
2.1 Problem statement

Atrial fibrosis detection methods have several limitations, including subjective interpretation, inability to detect early stages of fibrosis, variability in image quality [19], high computational cost, and inability to handle large datasets, etc. It may also face overfitting and underfitting issues. The features and challenges of existing atrial fibrosis detection techniques are given in Table 1 and the traditional model's experimental gap is given.

- In the existing techniques, high computational requirements and long processing times [20] hinder real-time diagnosis and clinical relevance. These issues can be overcome by using an optimization algorithm that minimizes computational requirements making the model more accurate, efficient, and clinically applicable.
- Existing atrial fibrosis detection methods [21] may suffer from limited data quality, availability, and high dimensionality. These problems can be overcome by using feature extraction methods, which extract relevant information from the data and reduce the data dimensionality.
- Existing atrial fibrosis diagnostic models may lack transparency, making it challenging to understand detection decisions. This problem can be overcome by combining two or more other algorithms to solve the identical issue, which improves computational efficiency and accuracy [22].

Table 1. Designed HC-TMU-based segmented images

Original Image	Ground Truth Image	UNet [23]	UNet 3+ [24]	Proposed HC-TMU



3. HYBRID META-HEURISTIC AIDED ATRIAL FIBROSIS DETECTION USING ADVANCED DEEP LEARNING NETWORK WITH WEIGHTED MULTI-FEATURE FUSION

3.1 Developed atrial fibrosis detection network

Generally, AF develops slowly and does not show visible signs until it reaches the final stage, which makes the detection process hard. Classical imagining approaches including ECG are not very effective in detecting AF in early fibrosis. In recent days, various approaches based on artificial intelligence have been employed to perform early detection of AF. Among these approaches, deep learning techniques including CNN have gained significant attention because of their capacity to examine the clinical images with better accuracy and to offer more exact outcomes in case of AF detection. Moreover, deep learning models can process huge quality of information effectively make them more applicable in the medical field, and assist them to offer highly generalized and robust outcomes. By considering these benefits, an advanced AF detection model is designed by leveraging the deep learning approaches. This model is developed by employing the 2D-ECG images [25] accumulated from the standard database. Following the data collection, the collected images are directly offered to the feature extraction. During feature extraction, three Feature Sets are extracted by employing certain techniques. At first, the features based on fuzzy entropy vectors [26], Hierarchical theory vectors, and Wavelet packet energy vectors are extracted from the ECG images and concatenated together to attain the final Feature Set, i.e., Feature Set 1. Then, the second set of features is extracted from the images by employing R-ViT, and the deep features (third set of features) within the images are extracted via CNN. Later, three features are chosen optimally from each Feature Set by utilizing the hybrid tuning approach named as HDF-RBMO strategy. Then, the optimally selected features from each Feature Set are multiplied with their corresponding weights, which are tuned by the same HDF-RBMO strategy to attain the weighted features. Followed by weight optimization, the resultant weighted features are concatenated together to form the fused weighted feature for enhancing the relief score and correlation coefficient on the segmentation performance. Further, the obtained fused weighted feature [27] is then fed into the developed HC-TMU model to perform AF segmentation. The HC-TMU framework was constructed by integrating a hybrid 2D and 3D convolution to the Trans-MobileUNet++. Moreover, in the designed HC-TMU, the input is fed into both 2D and 3D format, i.e., the fused weighted feature is taken as 2D input and the collection of 2D ECG image series is considered as 3D input. After further processing, the segmented image is obtained from the Trans-MobileUNet++. In addition, the AF detection outcome is determined based on the resultant segmented images [28]. At last, the performance of the developed technique is estimated by analyzing the outcomes of the developed model with a few classical models. The diagrammatic view of the developed deep learning-based AF detection framework is given in Figure 1.

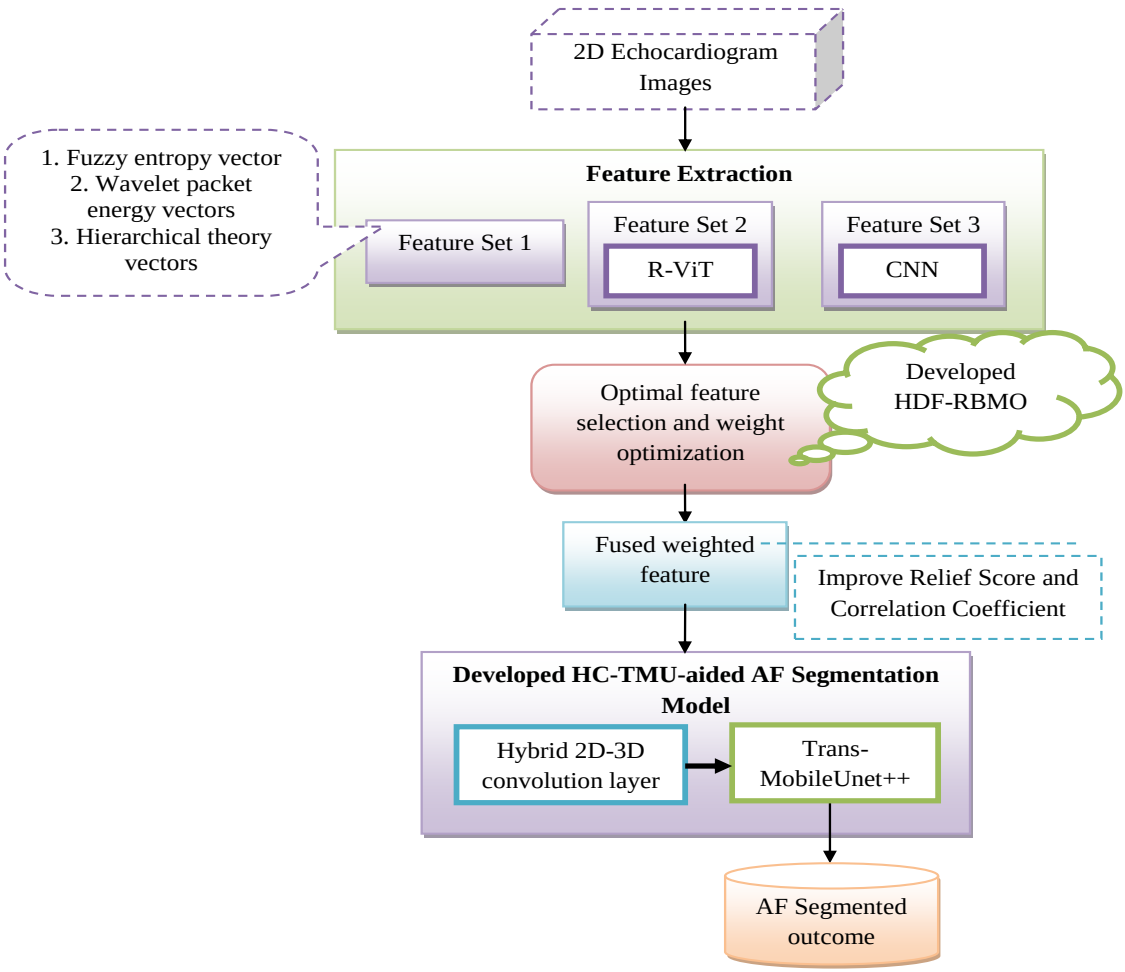


Figure 1. Developed deep learning-aided AF detection framework

Figure 2. Sample images taken from the dataset

3.2 Atrial fibrosis Echocardiogram signal dataset for model analysis

The implementation of this designed AF detection model is executed by employing the 2D ECG images obtained from the Echo Net-Dynamic dataset, which is available at https://stanfordaimi.azurewebsites.net/datasets/834e1cd1-92f7-4268-9daa-d359198b310a- access date: 07-11-2024. The Echo Net-Dynamic dataset is a huge collection of ECG video datasets that comprise nearly 10030 labeled ECG videos and explanations from human experts based on the tracing, measurements, and evaluation that is essential to offer a better understanding of cardiac motion and chamber dimensions. Among the 10030 records, 7523 are employed to train the detection model and 2508 are taken as testing data. The accumulated 2D ECG images are indicated as Cl_y^{ECG}. The sample images taken from the Dataset are given in Figure 2.

4. FEATURE EXTRACTION AND WEIGHTED FEATURE FUSION FOR ATRIAL FIBROSIS DETECTION USING HDF-RBMO

4.1 Echocardiogram 2D image feature extraction

Feature extraction is the major phase in this designed framework, the main intent behind this feature extraction is to extract the significant features from the accumulated 2D ECG images Cl_y^{ECG} that promote better detection outcomes. During feature extraction, three sets of features are extracted from the data. A detailed description of each Feature Set is given below.

Feature Set 1: The first set of features is attained by concatenating three diverse features including fuzzy entropy vector, wavelet packet energy vectors, and hierarchical theory vectors are extracted from the given input data Cl_y^{ECG}, and the obtained vectors are combined to form the Feature Set 1.

Fuzzy entropy vector [29]: Fuzzy entropy (FE) is employed to evaluate the time series data's irregularity and complexity and the execution procedure of FE is discussed below.

Let us consider, a time series data $Z = \{z(i): 1 \le i \le B\}$, in which time series data's length is denoted as B. Later, the mean value $z_0(i)$ is evaluated via Eq. (1):

$$z_0(i) = \frac{1}{n}\sum_{k=0}^{n-1} z(i+k) \tag{1}$$

Here, the embedding dimension is denoted as n, and the n-dimension vector $Z_i^n (i = 1,2,\ldots,B-n)$ is modified as in Eq. (2):

$$Z_i^n = \{z(i), z(i+1,\ldots,z(i+n-1))\} - z_0(i) \tag{2}$$

The fuzzy function $\mu(s_{ik}^n, b, e)$ is described in Eq. (3).

$$\mu(s_{ik}^n, b, e) = \exp\left(-\frac{(s_{ik}^n)^b}{e}\right) \tag{3}$$

Here, the function of exponential is denoted as $exp(\cdot)$, and the boundary gradient and width are signified as b and e. Later, the degree of similarity S_{ik}^n among Z_i^n and Z_j^n is described in Eq. (4):

$$S_{ik}^n(b,e) = \mu(s_{ik}^n, b, e) \tag{4}$$

Finally, the FE time series $\{z(i): 1 \le i \le B\}$ is evaluated via Eq. (5):

$$FE(Z,n,b,e) = \lim_{B\to\infty}\left(\ln\Phi^n(b,e) - \ln\Phi^{n+1}(b,e)\right) \tag{5}$$

Here, the natural logarithm operation is represented as $ln(\cdot)$. If B is finite, then the FE time series $FE(Z,n,b,e)$ is represented as in Eq. (6):

$$FE(Z,n,b,e,B) = \ln\Phi^n(b,e) - \ln\Phi^{n+1}(b,e) \tag{6}$$

The extracted fuzzy entropy vector is expressed as FE_x^{Vt}.

Wavelet packet energy vectors [30]: It is defined as the energy in diverse frequency bands that are evaluated from the wavelet packet decomposition outcome. The wavelet packet decomposition energy is estimated by Eq. (7) and the overall energy of the signal is derived by Eq. (8):

$$W_i = \int_{-\infty}^{+\infty} z_k^i(t)\,dt \tag{7}$$

$$W_{Tol} = \sum_{i=1}^{2^k} W_i \tag{8}$$

Here, the energy within each sub-band is indicated as W_i. Further, the value of normalized energy that is equivalent to each wavelet packet's energy is defined by Eq. (9):

$$Q_l = \frac{W_i}{W_{Tol}} \tag{9}$$

Here, each sub-band's probability distribution is expressed as Q_l. The extracted wavelet packet energy vectors are represented as Wp_v^{st}.

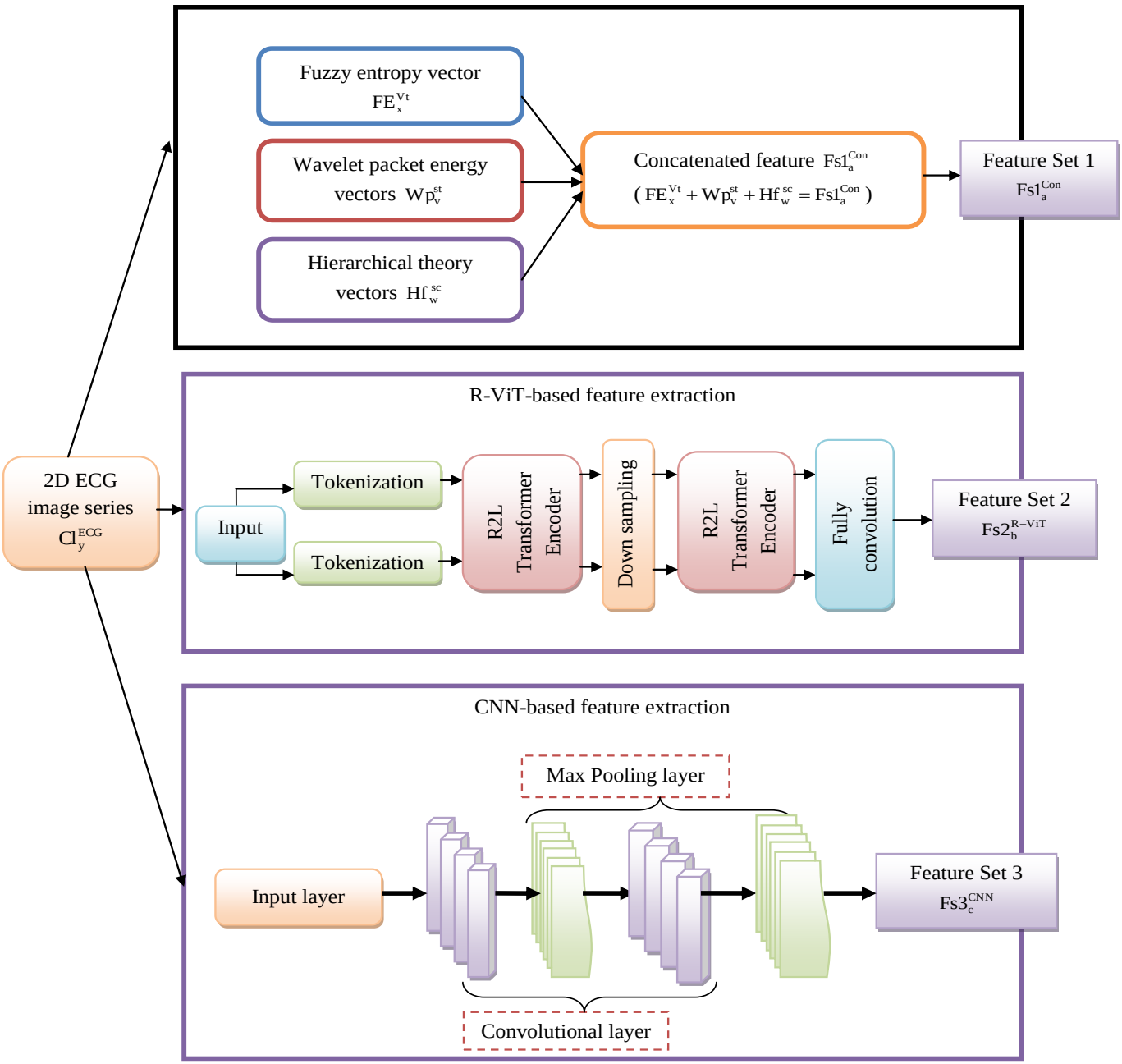


Figure 3. Echocardiogram 2D image feature extraction

Hierarchical theory vectors [31]: Hierarchical theory is described as the procedure of arranging or expressing the features within the given input in the hierarchical format. During hierarchical feature extraction, it extracts multiple features from the input to enhance the efficiency of the model in employing global context data. During feature extraction, each feature extracted from the image is represented in the vector format that helps to monitor the connection among the diverse features. Moreover, employing hierarchical theory in feature extraction helps in determining complicated patterns or connections within the features and accelerates the performance of the model during segmentation. The extracted hierarchical theory vectors are indicated as Hf_w^{sc}.

Feature concatenation: In order to attain the final Feature Set 1. The obtained fuzzy entropy vector FE_x^{Vt}, wavelet packet energy vectors Wp_v^{st}, and hierarchical theory vectors Hf_w^{sc} are concatenated together and are denoted as $Fs1_a^{Con}$ and it is expressed in Eq. (10):

$$FE_x^{Vt} + Wp_v^{st} + Hf_w^{sc} = Fs1_a^{Con} \tag{10}$$

Here, the first Feature Set is indicated as $Fs1_a^{Con}$.

Feature Set 2: The second Feature Set is obtained by employing R-ViT as a tool to extract the features from the input data Cl_y^{ECG}. The R-ViT [32] is a variant of classical ViT that comprises a regional-to-local attention module that helps to understand both global and local characteristics from the input data. The R-ViT model comprises two tokenization procedures that transfer input data into local and regional tokens, which assist in extracting significant features from the image. Moreover, this tokenization procedure is considered a convolution function with diverse path sizes. In R-ViT, the path size of the local and regional tokens is 4^2 and 28^2. Moreover, by employing the regional-to-local attention module the R-ViT processes both the tokens. Additionally, down-sampling procedures were carried out to separate the token's spatial resolution. At last, the features are obtained by taking an average value among the regional tokens. Similarly,

the fine-grained location data are also obtained from the local token. The features extracted by employing R-ViT are represented as $Fs2_b^{R-ViT}$.

Feature Set 3: The third set of features is extracted by employing the CNN [33]. CNN is employed to extract the deep features from the given input information. The key layers of the CNN model that are responsible for feature extraction are the convolution and max pooling layers. Each layer embedded in the CNN model has its own tasks to perform.

Convolutional layer: This layer comprises filters and feature maps. The filers within the layer are viewed as the neuron and the outcome from the filter utilized by the prior level is referred to as the feature maps.

Max pooling layer: The key function of this layer is to carry out the down-sampling on the feature maps with an aim to minimize the issues associated with overfitting. The deep features extracted from the input images by utilizing CNN are denoted as $Fs3_c^{CNN}$.

The pictorial view to describe the feature extraction phase is represented in Figure 3.

4.2 Proposed HDF-RBMO-based weighted feature fusion process

Followed by feature extraction, weighted feature fusion is the next phase that assists in enhancing detection performance by optimizing weight along with feature fusion. At this phase, the significant features from each extracted features set $Fs1_a^{Con}$, $Fs2_b^{R-ViT}$ and $Fs3_c^{CNN}$ are chosen optimally by utilizing the designed HDF-RBMO. During feature selection, the feature $OF1_t^{Con}$ from Feature Set 1 is selected from the range [1,10], the feature $OF2_u^{R\text{-}ViT}$ from Feature Set 2 is chosen from the limit [11,20] and the feature $OF3_r^{CNN}$ from Feature Set 3 is taken from the limit [21,30]. Additionally, the optimized weights Wt_t^{op}, Wt_u^{op} and Wt_r^{op} that lie within the range [0.01,0.99] are multiple with the corresponding feature, which results in weighted Feature Sets as expressed in the following equations:

$$Wf_t^{f1} = OF1_t^{Con} * Wt_t^{op} \tag{11}$$

$$Wf_u^{f2} = OF2_u^{R-ViT} * Wt_u^{op} \tag{12}$$

$$Wf_r^{f3} = OF3_r^{CNN} * Wt_r^{op} \tag{13}$$

Later, the resultant weighted Feature Sets are fused together as shown in Eq. (14) and the final fused weighted feature Fu_g^{Fea} is taken for further processing:

$$Fu_g^{Fea} = Wt_t^{f1} + Wt_u^{f2} + Wt_r^{f3} \tag{14}$$

The key objective behind weighted feature fusion is to increase the relief score as well as the correlation coefficient and it is numerically derived in Eq. (15):

$$Obj_1 = \underset{\left\{\substack{OF1_t^{Con}, OF2_u^{R-ViT}, OF3_r^{CNN}, \\ Wt_t^{op}, Wt_u^{op}, Wt_r^{op}}\right\}}{\arg\min} \left(\frac{1}{RF_{sc} + Cor_{cof}} \right) \tag{15}$$

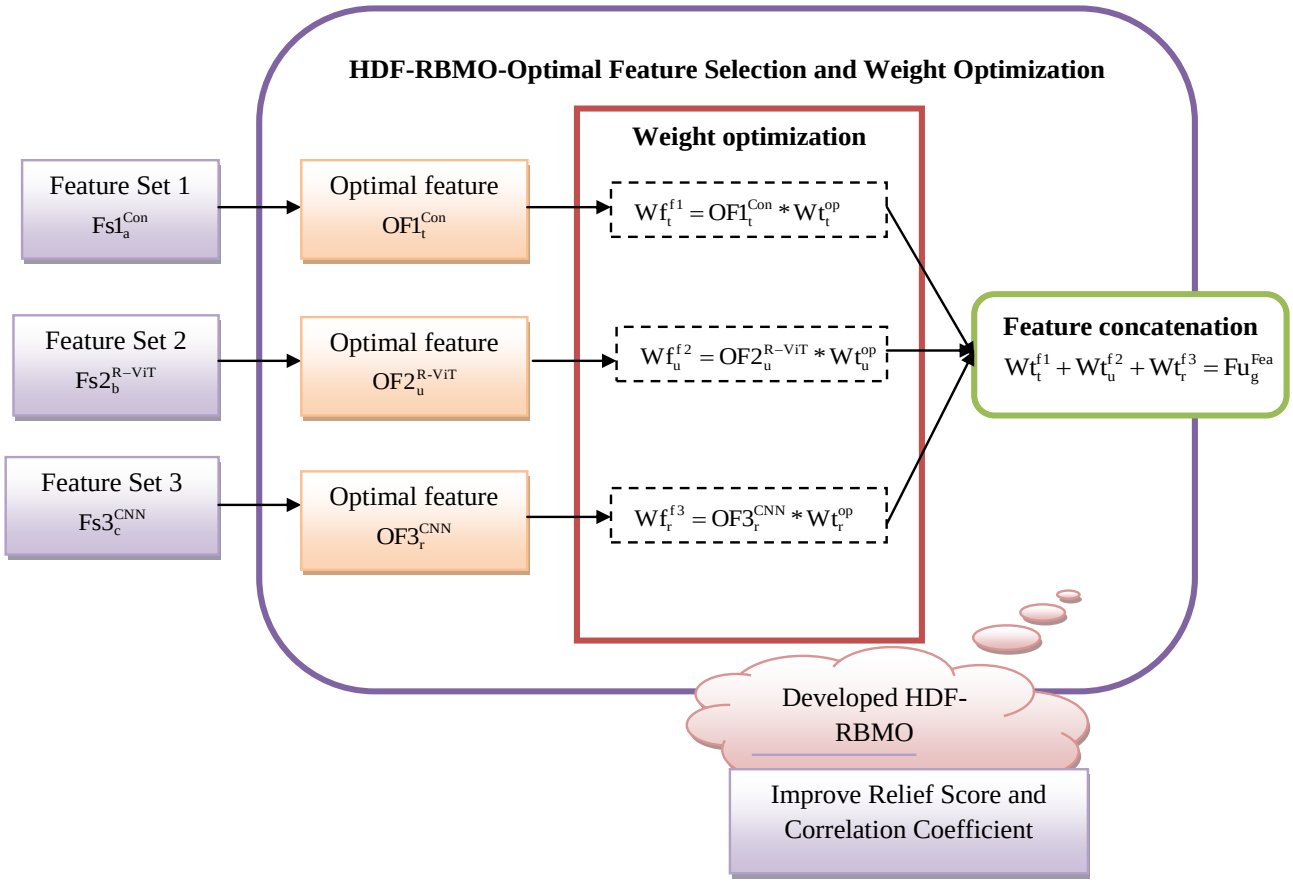


Figure 4. Developed HDF-RBMO-based weighted feature fusion process

Relief score RF_{sc} is described as the method to evaluate the features with multiple classes from the given input ECG images and they are capable of evaluating how the features within the images distinguish its instances among multiple identical classes. Moreover, it determines the weight T of the feature D, i.e., $T[D]$ by employing Eq. (16):

$$T[D] = Q\left(D_{SB} \mid S^{ni}\right) - Q\left(D_{SB} \mid F^{ni}\right) \tag{16}$$

From the above equation, the livelihood of the diverse values of the feature D on different instances is represented as Q, different values of the feature D are indicated as D_{SB}, the closest instance from a different class is denoted as S^{ni}, and the closest instance from a similar class is represented as F^{ni}.

Correlation coefficient Cor_{cof} is defined as the statistical metrics that evaluate the robustness of the linear connection among two variables as well as it is the normalized evaluation of the covariance among the variables and their values lying between the limit [-1,1] and it is numerically defined in Eq. (17):

$$Cor_{cof} = \frac{n\sum SG - \sum S \sum G}{\left(n\sum S^2 - \left(\sum S\right)^2\right)\cdot\left(n\sum G^2 - \left(\sum G\right)^2\right)} \tag{17}$$

The total count of data points within the given data is n and the summation of the product of Sth and G^{th} value for every data point is represented as $\sum SG$. The diagrammatic representation of HDF-RBMO-based weighted feature fusion is provided in Figure 4.

4.3 Proposed HDF-RBMO

In this designed framework, a hybrid optimization approach is developed by integrating the functions of classical DFA and RBMO and named HDF-RBMO. The prime function of the developed HDF-RBMO is to select the optimal features within the extracted Feature Sets and to optimize the weight for multiplying with optimal features to enhance the relief score and correlation coefficient, which helps to improve the detection performance. In this hybrid strategy, DFA [27] is employed by considering its ability to manage huge-scale optimization issues, which makes it more applicable in complicated situations. Moreover, they offer better performance in various tuning issues like multi-model sand non-linear issues, which results in robust performance. However, the DFA is highly complicated to execute and learn. The efficiency of DFA is vulnerable to parameter optimization, selecting the most significant parameter results in suboptimal outcomes. Additionally, in certain cases, the DFA faces difficulties in converging effectively for the optimal outcome, particularly in dynamic surroundings that also result in sub-optimal outcomes. Therefore, the RBMO [24] strategy is integrated with the DFA to address these shortcomings since they present effective performance in complicated regions and avoid local minimum as well as offer better convergence than other classical tuning approaches. In this developed HDF-RBMO, a random value r within the range [0,1] is upgraded according to the developed concept as shown in Eq. (18).

$$r = \frac{C_{fit}}{\left(W_{fit} + M_{fit} + B_{fit}\right)} \tag{18}$$

In the above equation, the best, worst, mean, and current fitness values are indicated as B_{fit}, W_{fit}, M_{fit} and C_{fit}. The pseudocode of the proposed HDF-RBMO is presented in Algorithm 1.

Algorithm 1: Implemented HDF-RBMO

Input: Extracted Feature Sets $Fs1_a^{Con}$, $Fs2_b^{R-ViT}$, $Fs3_c^{CNN}$, and initial weight.

Output: Optimal features $OF1_t^{\text{Con}}$, $OF2_u^{\text{R-ViT}}$, and $OF3_r^{\text{CNN}}$, optimized weights Wt_t^{op}, Wt_u^{op} and Wt_r^{op}.

Initialize the parameters present iteration t, highest iteration M_{itr}, and population count N_{pop}.

```
While (condition satisfied)
    For t = 1 to M_itr
        For i = 1 to N_pop
            Upgrade the random value r as shown
          in Eq. (19)
            If r > 0.5
                Update by employing DFA
            else
                Update by employing RBMO
            end
        end
    end
    Obtain the optimal solution
end
```

5. ATRIAL FIBROSIS DETECTION USING HYBRID (2D, 3D) CONVOLUTION BASED TRANS-MOBILEUNET++

5.1 Trans-MobileUNet++

Trans-MobileUNet++ is designed by integrating a transformer encoder into the MobileUNet++. The MobileUNet++ model implements the MobileNet on the UNet++ model. It is an effective variant of classical DNN that is introduced to execute the semantic segmentation procedure by employing the skip connections within the network. In addition, the UNet++ model holds a symmetric expanding path that enables the model to perform accurate localization as well as the contracting path to accumulate the context. The skip pathways of the UNet++ hold various skip connections, which are employed to boost the gradient flow within the network. The outcome attained from the bottom and top layers are combined to attain a final outcome as presented in Eq. (19).

$$z^{q,w} = \begin{cases} V\left(z^{q-1,w}\right), & w = 0 \\ V\left(\left[z^{q-w}\right], S\left(s^{q+1,w-1}\right)\right), & w > 0 \end{cases} \tag{19}$$

Here, the feature map is represented as $z^{q,w}$, and the convolution and activation function are indicated as $V(.)$ and $S(.)$. Moreover, integrating the MobileNet and UNet++ enables the framework to process effectively in case of minimum training data.

Furthermore, a transformer encoder is added to the MobileUNet++. In the transformer encoder, the input data is offered to the layer normalization that helps to normalize the activation. Later, the normalized images are offered to the multi-head self-attention layer. Following this, a residual connection is utilized, which is connected to a layer

normalization and feed-forward block. Feed-forward block is a combination of two linear layers, a Gaussian Error Linear Unit (GELU) and a dropout layer. The size of the input images is extended by the initial linear layer and later minimized by another linear layer and the final outcome attained from the transformer encoder is presented in the following equations.

$$O' = Linear\left(L \in T^{M\times f}\right) \in T^{M\times 4f} \tag{20}$$

$$O'' = Dropout\left(GELU\left(O'\right)\right) \in T^{M\times 4f} \tag{21}$$

$$O' = Linear\left(O'' \in T^{M\times 4f}\right) \in T^{M\times f} \tag{22}$$

From the above equations, the result obtained from the transformer encoder is indicated as O', the size and the sequence count are expressed as f and M. The pictorial representation of Trans-MobileUNet++ is offered in Figure 5.

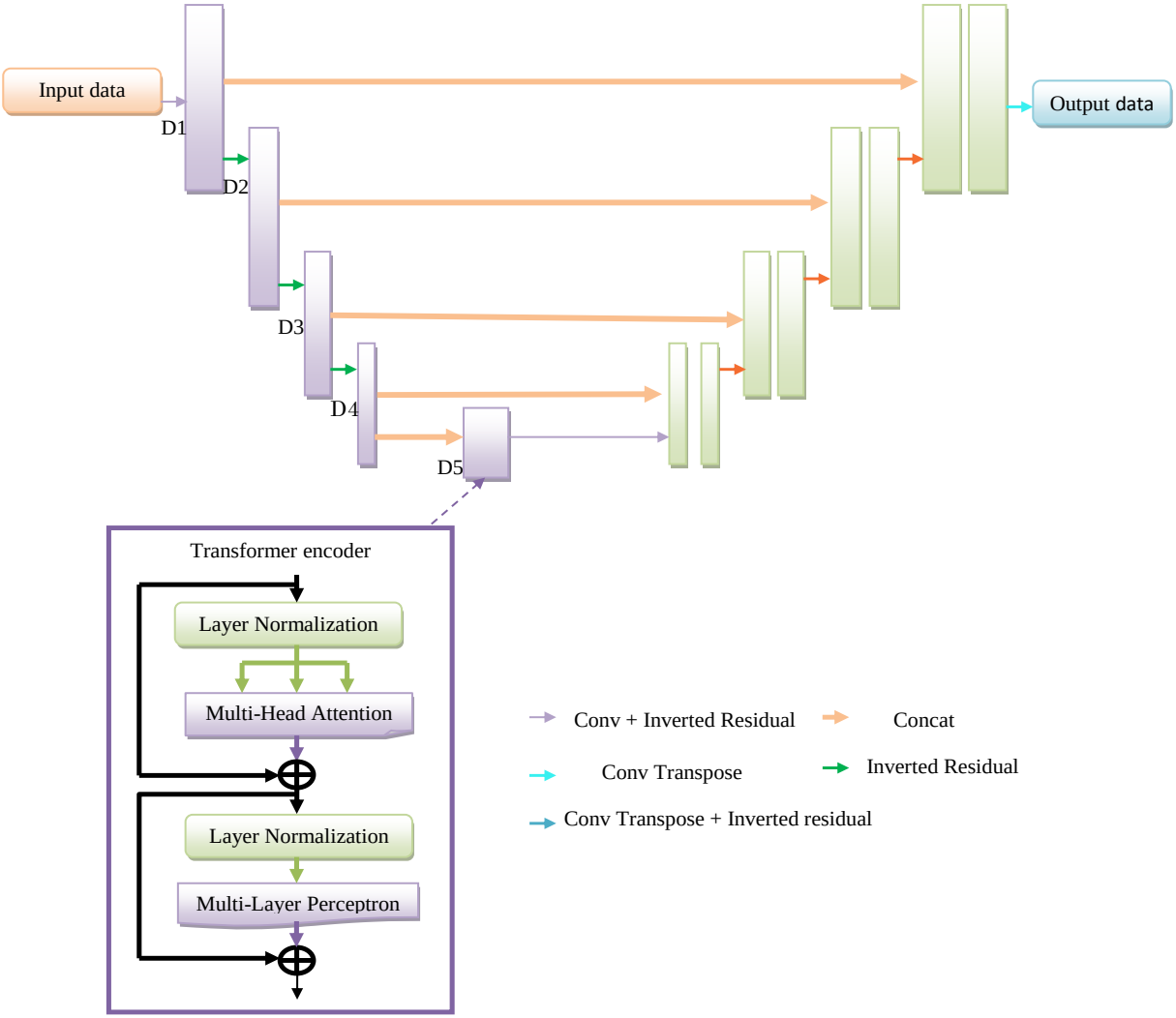


Figure 5. Pictorial view of Trans-MobileUNet++

5.2 Introduced HC-TMU for atrial fibrosis detection

The designed HC-TMU model is employed to carry out the segmentation procedure. This model utilized the fused weighted feature Fu_g^{Fea} and the collected 2D ECG image series as the input data. The designed HC-TMU is a combination of a hybrid 2D-3D convolution layer and Trans-MobileUNet++. The hybrid convolutional layer is a combination of two convolutional layers namely 2DCNN and 3DCNN. Here, the feature maps are created by concatenating the outcome from both the 2DCNN and 3DCNN equivalently. Additionally, a cross-domain transfer is established with the help of a data interface procedure. Later, the cross-domain concatenation function is employed to establish the unique 2D and 3D features. Let us consider, the input offered to the 2DCNN and 3DCNN as y_2 and y_3. Further, the output attained from both 2DCNN and 3DCNN are expressed as $t_2(c)$ and $t_3(c)$, the hybrid convolutional layer's data interaction is denoted as W, and the convolution operations performed on the 2DCNN and 3DCNN are indicated as $r_2(c)$ and $r_3(c)$. In addition, the cross-entropy loss function within the hybrid convolutional layer is represented in the below equations:

$$K(c,e) = G(e, g(r_2(c) + r_3(c))) \tag{23}$$

$$(r_2(c) + r_3(c)) = C(t_2(c), t_3(c)) \tag{24}$$

$$t_2(c) = \mathrm{M}_2 \otimes y_2 \tag{25}$$

$$t_3(c) = \mathrm{M}_3 \otimes y_3 \tag{26}$$

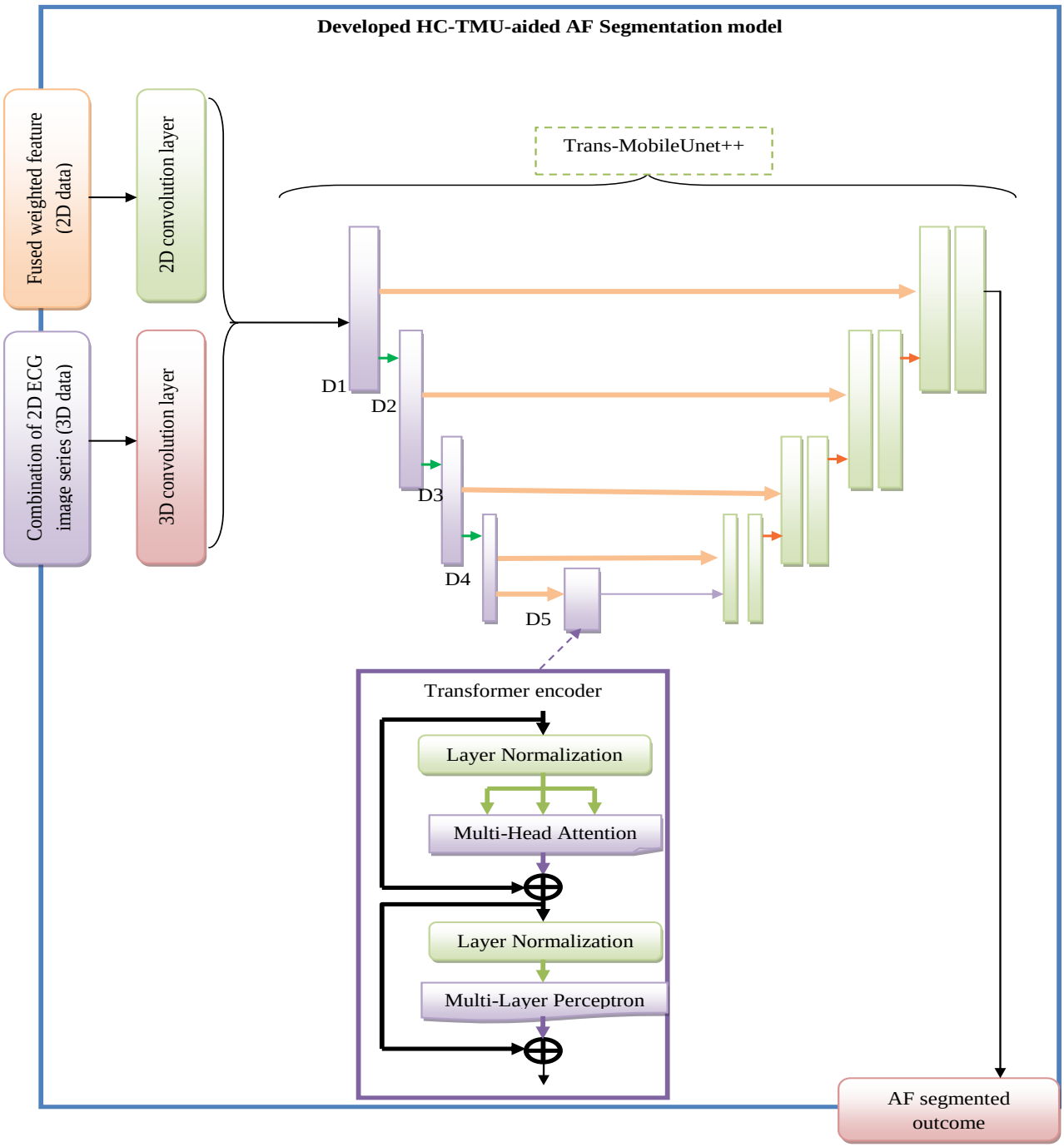


Figure 6. Developed HC-TMU-based AF detection model

In these equations, the cross-entropy loss function is denoted as $G(\cdot)$, and the actual label for c is represented as e. Initially, the input data are offered in two different formats namely 2D and 3D. Here, the fused weighted feature is taken as the 2D data for processing and the collection of ECG image series is considered as the 3D data. Followed by the convolutional operation, the processed input is fed to the Trans-MobileUNet++ to attain the segmented outcome. The diagrammatic illustration of the developed HC-TMU-based AF detection framework is offered in Figure 6.

6. RESULTS AND DISCUSSION

6.1 Simulation setup

The introduced AF detection framework was built and trained on the Python platform. The setup for this detection framework was established by considering the population counts as 10, the highest count of iteration as 50, and the chromosome length as 33. Additionally, the performance analysis was carried out by analyzing the outcome of the designed mode with a set of classical optimization strategies including Reptile Search Algorithm (RSA), Mud Ring Algorithm (MRA), Dark Forest Algorithm (DFA) and RBMO and the segmentation performance is analyzed with DCNN, MSResNet, LSTM, and RNN.

6.2 Evaluation measures

The effectiveness of the designed AF detection framework over diverse optimization strategies and segmentation approaches is demonstrated based on the following measures.

(a) Accuracy of the approach is estimated by Eq. (27).

$$Ay = \frac{Q^T + M^T}{Q^T + M^T + Q^F + M^F} \quad (27)$$

(b) The dice coefficient of the model is evaluated using Eq. (28).

$$Df = 2 \times \frac{\left| Mt_c^{ig} \cap Di_t^{ig} \right|}{\left| Mt_c^{ig} \right| + \left| Di_t^{ig} \right|} \quad (28)$$

(c) Jaccard is determined by employing Eq. (29).

$$J_{cf} = \frac{\left| Mt_c^{ig} \cap Di_t^{ig} \right|}{\left| Mt_c^{ig} \right| + \left| Di_t^{ig} \right| - \left| Mt_c^{ig} \cap Di_t^{ig} \right|} \quad (29)$$

(d) Sensitivity of the model is derived using Eq. (30).

$$Sensitivity = \frac{Q^T}{Q^T + M^F} \quad (30)$$

(e) Specificity is evaluated by Eq. (31).

$$Specificity = \frac{M^T}{M^T + Q^F} \quad (31)$$

Here, the false positive, false negative, true positive, and true negative are represented as Q^F, M^F, Q^T, and M^T. The ground truth and segmented image are described as Mt_c^{ig} and Di_t^{ig}.

6.3 Designed HC-TMU-based segmented images

The segmented images obtained by the developed HC-TMU model are presented in Table 2.

Table 2. Overall performance validation with classical techniques

Terms	DCNN [1]	MSResNet [2]	LSTM [5]	RNN [6]	HC-TMU
Dice Coefficient					
Median	0.8066764	0.8216339	0.8174575	0.8451191	0.9275383
Worst	0.7797717	0.7878065	0.7901691	0.8120909	0.8892463
Best	0.838269	0.8338267	0.8829492	0.8866828	0.943818
Mean	0.8041712	0.8134676	0.8244332	0.8447935	0.9226687
Jaccard					
Median	0.6759927	0.6972831	0.691279	0.7317839	0.864961
Worst	0.6390375	0.6499016	0.6531236	0.6836305	0.8005792
Best	0.7215689	0.715011	0.7904289	0.7964331	0.893613
Mean	0.6728436	0.6859655	0.7020573	0.7320043	0.8570105
Accuracy					
Median	0.8036957	0.816658	0.8187332	0.8472366	0.9242783
Worst	0.7724762	0.7945251	0.8025055	0.8234711	0.8907318
Best	0.8394775	0.8349457	0.8743744	0.8810425	0.9443207
Mean	0.804715	0.8134933	0.8244705	0.8458298	0.9227249
Peak Signal-to-Noise Ratio (PSNR)					
Median	55.201737	55.498199	55.547641	56.291896	59.36967
Worst	54.560535	55.003217	55.175253	55.662645	57.745866
Best	56.075446	55.954535	57.140022	57.376885	60.673864
Mean	55.245987	55.436212	55.720633	56.286607	59.353584
Mean Squared Error (MSE)					
Median	0.1963043	0.183342	0.1812668	0.1527634	0.0757217
Worst	0.1605225	0.1650543	0.1256256	0.1189575	0.0556793
Best	0.2275238	0.2054749	0.1974945	0.1765289	0.1092682
Mean	0.195285	0.1865067	0.1755295	0.1541702	0.0772751

6.4 Detection performance estimation of the designed model

The introduced HC-TMU's segmentation performance is monitored by considering epoch count and the obtained graphs are presented in Figure 7. This evaluation estimates how effectively this designed segmentation model determines the infected area with the given input data. This evaluation helps to determine the quality of resources employed by the segmentation models during the implementation phase. It is necessary to analyze that the introduced framework has the capacity to process the input data more effectively even in the presence of unwanted noise or artifacts. Moreover, this evaluation is performed with the standard framework to monitor the performance in terms of accuracy, Jaccard, and dice coefficient by varying the epoch count from 50–250. From Figure 7(a), the accuracy of the introduced HC-TMU framework is 14.45%, 11.76%, 9.19%, and 5.55% more advanced than DCNN, MSResNet, LSTM, and RNN, when considering the epoch count as 50. Hence, the segmentation efficiency of the introduced model is superior to other standard frameworks.

6.5 Convergence analysis for the developed model

The convergence graph attained by the developed HDF-RBMO model is shown in Figure 8. This evaluation is essential for evaluating the convergence of the employed tuning strategy throughout the iteration. From the given convergence graph, the straight line describes that the best fitness value identified by the tuning strategy is stable and reveals that the optimization technique is stuck within the local minimum. According to the obtained graph, the convergence of the proposed HDF-RBMO model has the ability to offer more advanced performance, when compared to the traditional techniques.

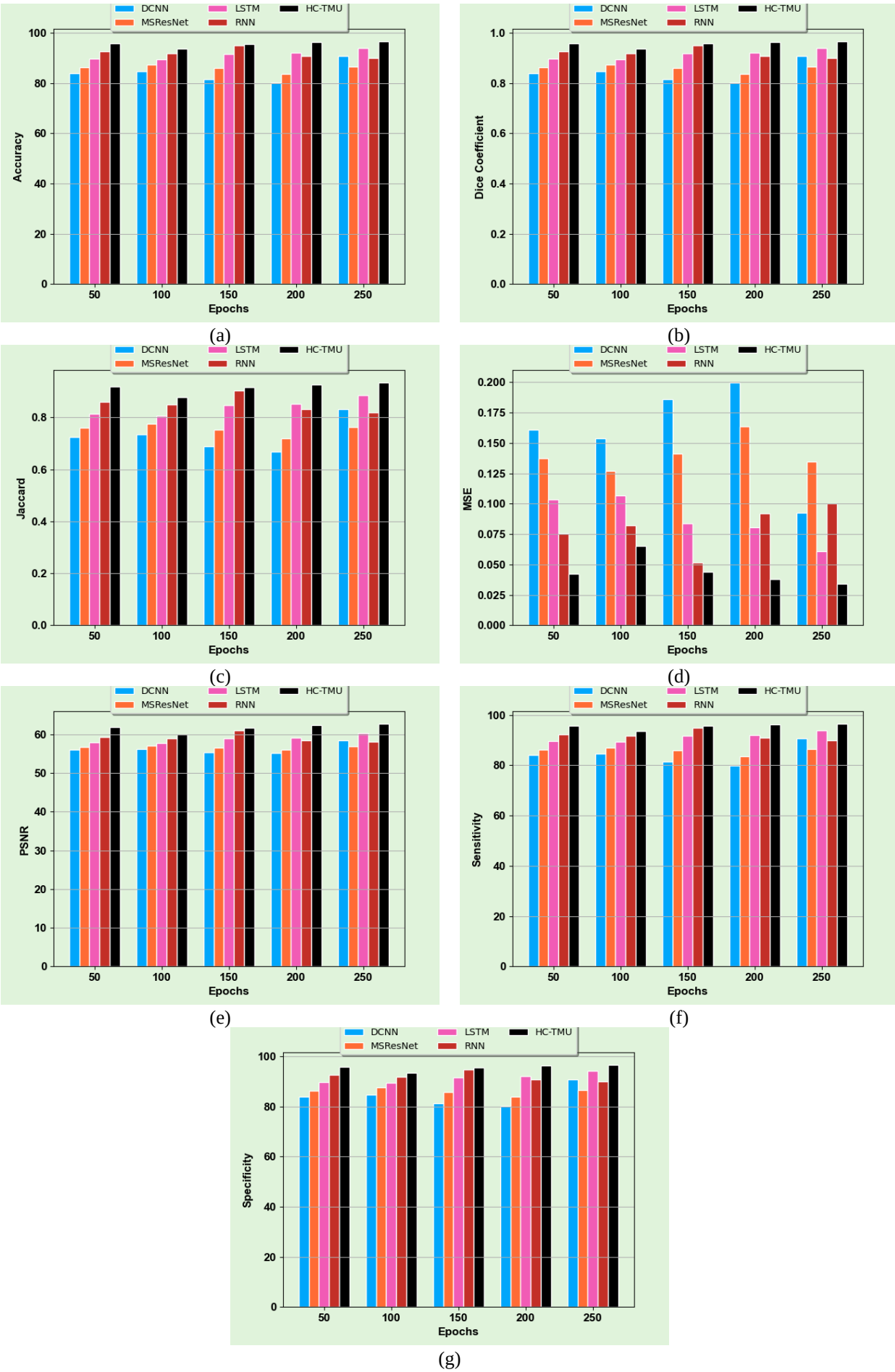


Figure 7. Detection performance evaluation of the proposed framework with traditional approaches with respect to (a) Accuracy, (b) Dice-coefficient, (c) Jaccaed, (d) MSE, (e) PSNR, (f) Sensitivity, (g) Specificity

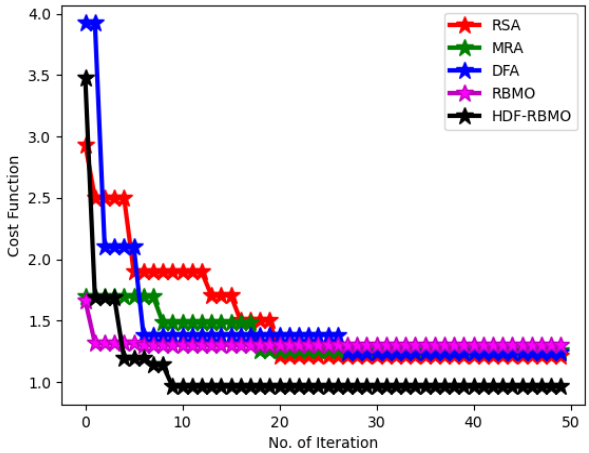


Figure 8. Convergence analysis for the proposed framework

6.6 Overall detection performance validation with classical models

The overall segmentation performance of the designed model is presented in Table 3. The segmentation performance is determined by considering the mean, best, median, and worst values. By varying the statistical measures, a constant threshold value is handled during the segmentation procedure. Moreover, the developed model's flexibility is improved which is necessary for determining the most significant features from the given input. From the Table, the dice coefficient of the suggested framework, when analyzing the mean performance is 14.73%, 13.42%, 11.91%, and 9.21% more effective than other classical techniques such as DCNN, MSResNet, LSTM, and RNN, and provided improved performance throughout AF segmentation.

6.7 Statistical evaluation for the suggested framework

Statistical analysis is essential to identify the effectiveness of the developed HDF-RBMO technique based on weighted feature fusion. The statistical analysis of the designed technique with standard approaches is given in Table 3. From Table 3, the best performance of suggested HDF-RBMO is 20.58%, 23.40%, 22.09%, and 26.11% better than the traditional techniques such as RSA, MRA, DFA, and RBMO. Thus, the analysis outcome revealed that the effectiveness of the designed optimization model is more effective than other standard techniques.

Table 3. Statistical evaluation for the developed framework

Terms	RSA	MRA	DFA	RBMO	HDF-RBMO
Worst	2.9343999	1.6979305	3.9218194	1.6634846	3.4747119
Best	1.21243	1.2571479	1.2358679	1.3031325	0.9628628
Mean	1.5119658	1.3736648	1.4749031	1.3117682	1.0772629
Standard deviation	0.4452767	0.1666598	0.5482389	0.0504273	0.385869
Median	1.21243	1.2582486	1.3844285	1.3031325	0.9628628

7. CONCLUSIONS

In this work, we proposed an innovative AF detection framework that integrates deep learning with hybrid optimization. The model combines HDF-RBMO for optimal feature selection and weighting with the HC-TMU segmentation architecture, achieving superior performance compared to baseline models. Experimental results demonstrated that HC-TMU attained a mean dice coefficient of 0.9226 and outperformed DCNN, MSResNet, LSTM, and RNN by 12.48%, 13.09%, 7.99%, and 7.18%, respectively. These results highlight the potential of the proposed method for early and accurate AF detection. Despite these promising outcomes, several limitations remain. First, the framework risks modality confusion since ECG-derived features and echo-inspired segmentation approaches are combined. Second, the model has not yet undergone clinical validation, limiting its immediate applicability in healthcare settings. Third, the reliance on a single dataset (EchoNet-Dynamic) may restrict generalizability across diverse patient populations. Finally, the interpretability of the fused weighted features remains limited, which may hinder clinical trust and adoption. Future research will focus on addressing these challenges by incorporating multimodal data sources (e.g., combining ECG with Echocardiogram sequences), exploring real-time deployment in clinical workflows, and developing explainable AI mechanisms to better interpret fibrosis-related patterns. These enhancements could significantly improve the reliability, clinical relevance, and translational impact of AF detection models.

REFERENCES

[1] Pourbabaee, B., Roshtkhari, M.J., Khorasani, K. (2017). Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 48(12): 2095-2104. https://doi.org/10.1109/TSMC.2017.2705582

[2] Cao, X.C., Yao, B., Chen, B.Q. (2019). Atrial fibrillation detection using an improved multi-scale decomposition enhanced residual convolutional neural network. IEEE Access, 7: 89152-89161. https://doi.org/10.1109/ACCESS.2019.2926749

[3] Nurmaini, S., Tondas, A.E., Darmawahyuni, A., Rachmatullah, M.N., Partan, R.U., Firdaus, F., Khoirani, R. (2020). Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. Future Generation Computer Systems, 113: 304-317. https://doi.org/10.1016/j.future.2020.07.021

[4] Faust, O., Shenfield, A., Kareem, M., San, T.R., Fujita, H., Acharya, U.R. (2018). Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. Computers in Biology and Medicine, 102: 327-335. https://doi.org/10.1016/j.compbiomed.2018.07.001

[5] Petmezas, G., Haris, K., Stefanopoulos, L., Kilintzis, V., Tzavelis, A., Rogers, J.A., Maglaveras, N. (2021). Automated atrial fibrillation detection using a hybrid CNN-LSTM network on imbalanced ECG datasets. Biomedical Signal Processing and Control, 63: 102194.

https://doi.org/10.1016/j.bspc.2020.102194
[6] Andersen, R.S., Peimankar, A., Puthusserypady, S. (2019). A deep learning approach for real-time detection of atrial fibrillation. Expert Systems with Applications, 115: 465-473. https://doi.org/10.1016/j.eswa.2018.08.011
[7] Schwab, P., Scebba, G.C., Zhang, J., Delai, M., Karlen, W. (2017). Beat by beat: Classifying cardiac arrhythmias with recurrent neural networks. In 2017 Computing in Cardiology (CinC), Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.363-223
[8] Yildirim, Ö. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in Biology and Medicine, 96: 189-202. https://doi.org/10.1016/j.compbiomed.2018.03.016
[9] Zihlmann, M., Perekrestenko, D., Tschannen, M. (2017). Convolutional recurrent neural networks for electrocardiogram classification. In 2017 Computing in Cardiology (CinC), Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.070-060
[10] Andreotti, F., Carr, O., Pimentel, M.A., Mahdi, A., De Vos, M. (2017). Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ECG. In 2017 Computing in Cardiology (CinC), Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.360-239
[11] Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A.Y. (2019). Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature Medicine, 25(1): 65-69. https://doi.org/10.1038/s41591-018-0268-3
[12] He, R.N., Wang, K.Q., Zhao, N., Sun, Q., Li, Y.C., Li, Q.C., Zhang, H.G. (2020). Automatic classification of arrhythmias by residual network and BiGRU with attention mechanism. 2020 Computing in Cardiology, 1-4. https://doi.org/10.22489/CinC.2020.044
[13] He, R., Liu, Y., Wang, K., Zhao, N., Yuan, Y., Li, Q., Zhang, H. (2019). Automatic cardiac arrhythmia classification using combination of deep residual network and bidirectional LSTM. IEEE Access, 7: 102119-102135. https://doi.org/10.1109/ACCESS.2019.2931500
[14] Warrick, P., Homsi, M.N. (2017). Cardiac arrhythmia detection from ECG combining convolutional and long short-term memory networks. In 2017 Computing in Cardiology (CinC), Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.161-460
[15] Chandra, B.S., Sastry, C.S., Jana, S., Patidar, S. (2017). Atrial fibrillation detection using convolutional neural networks. In 2017 Computing in Cardiology (CinC), Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.163-226
[16] Król-Józaga, B. (2022). Atrial fibrillation detection using convolutional neural networks on 2-dimensional representation of ECG signal. Biomedical Signal Processing and Control, 74: 103470. https://doi.org/10.1016/j.bspc.2021.103470
[17] Xia, Y., Wulan, N., Wang, K., Zhang, H. (2018). Detecting atrial fibrillation by deep convolutional neural networks. Computers in Biology and Medicine, 93: 84-92. https://doi.org/10.1016/j.compbiomed.2017.12.007
[18] Alfaras, M., Soriano, M.C., Ortín, S. (2019). A fast machine learning model for ECG-based heartbeat classification and arrhythmia detection. Frontiers in Physics, 7: 103. https://doi.org/10.3389/fphy.2019.00103
[19] Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., San Tan, R. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89: 389-396. https://doi.org/10.1016/j.compbiomed.2017.08.022
[20] Udawat, A.S., Singh, P. (2022). An automated detection of atrial fibrillation from single-lead ECG using HRV features and machine learning. Journal of Electrocardiology, 75: 70-81. https://doi.org/10.1016/j.jelectrocard.2022.07.069
[21] Asgari, S., Mehrnia, A., Moussavi, M. (2015). Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. Computers in Biology and Medicine, 60: 132-142. https://doi.org/10.1016/j.compbiomed.2015.03.005
[22] Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., Tan, R.S. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89: 389-396. https://doi.org/10.1016/j.compbiomed.2017.08.022
[23] Clifford, G.D., Liu, C., Moody, B., Lehman, L.W.H., Silva, I., Li, Q., Mark, R.G. (2017). AF classification from a short single lead ECG recording: The PhysioNet/computing in cardiology challenge 2017. In 2017 Computing in Cardiology, Rennes, France, pp. 1-4. https://doi.org/10.22489/CinC.2017.065-469
[24] Rashied, N., Jeribi, A. (2024). Enhancing image quality through a novel multiscale fractal dimension formulated by the characteristic function. Mathematical Modelling of Engineering Problems, 11(1): 107-113. https://doi.org/10.18280/mmep.110111
[25] Cheng, J., Zou, Q., Zhao, Y. (2021). ECG signal classification based on deep CNN and BiLSTM. BMC Medical Informatics and Decision Making, 21: 365. https://doi.org/10.1186/s12911-021-01736-y
[26] Ansari, Y., Mourad, O., Qaraqe, K., Serpedin, E. (2023). Deep learning for ECG arrhythmia detection and classification: An overview of progress for period 2017-2023. Frontiers in Physiology, 14: 1246746. https://doi.org/10.3389/fphys.2023.1246746
[27] Kiranyaz, S., Ince, T., Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering, 63(3): 664-675. https://doi.org/10.1109/TBME.2015.2468589
[28] Niu, J.H., Tang, Y.Q., Sun, Z.Y., Zhang, W.S. (2020). Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks. IEEE Journal of Biomedical and Health Informatics, 24(5): 1321-1332. https://doi.org/10.1109/JBHI.2019.2942938
[29] Swapna, G., Soman, K.P., Vinayakumar, R. (2018). Automated detection of cardiac arrhythmia using deep learning techniques. Procedia Computer Science, 132: 1192-1201. https://doi.org/10.1016/j.procs.2018.05.034
[30] Bhattacharyya, S., Majumder, S., Debnath, P., Chanda, M. (2021). Arrhythmic heartbeat classification using ensemble of random forest and support vector machine algorithm. IEEE Transactions on Artificial Intelligence, 2(3): 260-268.

https://doi.org/10.1109/TAI.2021.3083689
[31] Swaroop, P., Badolia, N., Ranjan, R., Kumar, M. (2024). Arrhythmia classification using hybrid CNN-LSTM model. In 2024 First International Conference on Electronics, Communication and Signal Processing (ICECSP), New Delhi, India, pp. 1-6. https://doi.org/10.1109/ICECSP61809.2024.10698222
[32] Van Zaen, J., Delgado-Gonzalo, R., Ferrario, D., Lemay, M. (2020). Cardiac arrhythmia detection from ECG with convolutional recurrent neural networks. In Biomedical Engineering Systems and Technologies. BIOSTEC 2019. Communications in Computer and Information Science, pp. 311-327. https://doi.org/10.1007/978-3-030-46970-2_15
[33] Salau, A.O., Markus, E.D., Assegie, T.A., Omeje, C.O., Eneh, J.N. (2023). Influence of class imbalance and resampling on classification accuracy of chronic kidney disease detection. Mathematical Modelling of Engineering Problems, 10(1): 48-54. https://doi.org/10.18280/mmep.100106