Synergising Principal Component Analysis with Pythagorean Neutrosophic Bonferroni Mean Approach for Arrhythmia Detection using Cardiovascular Signals

Majed Balkheer[1,*], Reda Salama[2], Mahmoud Ragab[2], Ashis Kumer Biswas[1]

[1]Department of Computer Science and Engineering, University of Colorado Denver, Denver, Colorado, USA

[2]Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

Emails: majed.balkheer@ucdenver.edu; ashis.biswas@ucdenver.edu; rkhalifa@kau.edu.sa; mragab@kau.edu.sa

Abstract

The neutrosophic set (NS) concept from the philosophical perspective extends and simplifies the principles of fuzzy set (FS) and intuitionistic FS (IFS). A NS is defined by truth, indeterminacy, and falsity membership functions, with each value belonging to the non-standard intervals (−0, 1+). In contrast to IFSs, there is no limitation in the membership function in NS, and the hesitancy degree is incorporated in NS. Arrhythmia is a medical illness wherein the regular pumping mechanism of the human heart becomes abnormal. The arrhythmia detection is one of the most essential steps to identify the disorder that can play a significant role in helping cardiologist with their decision. The initial identification of abnormal heart disease is critical for patients with heart disorders. Computer-aided diagnosis (CAD) has gained popularity in the arrhythmia domain recently, as artificial intelligence (AI) technology has matured. Still, the AI-based deep learning (DL) techniques are applied frequently to classify and detect arrhythmia. This paper presents an Enhanced Diagnostic Model for Cardiac Arrhythmia using Principal Component Analysis and Pythagorean Neutrosophic Bonferroni Mean (DMCA-PCAPNBM) technique in Cardiovascular Signal Processing. The objective is in the automated arrhythmia detection using advanced techniques. Initially, the DMCA-PCAPNBM model applies the min-max scaler-based data pre-processing technique for transforming input data into an appropriate format. In addition, the principal component analysis (PCA) method is applied for the feature subset selection model to pick out the optimal attributes from the dataset. For the procedure of arrhythmia detection, the PNBM model is utilized. Finally, the improved dung beetle optimization (IDBO) approach is applied for parameter tuning, resulting in enhanced classification performance. A comprehensive experimentation is implemented to verify the superior outcome of the DMCA-PCAPNBM model on the ECG arrhythmia classification dataset. The experimental validation of the DMCA-PCAPNBM approach illustrated an improved accuracy value of 99.06% over recent techniques.

Keywords: Neutrosophic Sets; Interval Neutrosophic Sets; Arrhythmia; Cardiovascular; Improved Dung Beetle Optimization; Fuzzy Set

1. Introduction

Neutrosophic logic is comparatively new domain, where all propositions are analyzed to have the percentage (proportion) of truth in subset T, the proportion of indeterminacy in subset I, and the proportion of falsity in subset F [1]. NS is successfully applied for indeterminate data handling, and provides assistance in coping with data indeterminacy, and is even a method endorsed for data analysis and classifier uses [2]. As stated by the WHO, cardiovascular diseases (CVDs) are the primary reason for death. 17.9 million person were lost to CVD, accounting for 31% of global fatalities [3]. Two-thirds of these fatalities occurred in low-income nations, where healthcare infrastructure for prompt disease detection and comprehensive patient care for CVD is widely considered insufficient [4]. Arrhythmia is a condition that arises when an individual's heart's pacemaker fails or is disrupted

DOI: https://doi.org/10.54216/IJNS.270208

by ectopic foci. Numerous diagnostic techniques exist to diagnose CVDs. Electrocardiography (ECG) is a widely used technique for identifying multiple heart diseases. ECG is a widely accepted and reliable method for detecting acute coronary syndromes, intraventricular conduction disturbances, and arrhythmias. ECG lead signifies the electrical view of the heart generated by measuring the voltage difference between two electrodes. A standard 12-lead ECG utilizes 12 different viewpoints of the heart's electrical activity. Six of these viewpoints, called limb leads, are obtained from electrodes found on the legs and arms. Remaining six, termed precordial mains, are recorded from electrodes positioned on the chest [5].

There are several kinds of arrhythmias, every kind is linked with a pattern, and as such, it is possible to detect and classify its type. Arrhythmias are categorized into two kinds: morphological and rhythmic arrhythmias [6]. The 1st one refers to the heart rhythm disturbances caused by isolated, abnormal heartbeats, as opposed to those stemming from a series of irregular beats [7]. The 2nd one is characterized by a sequence of heartbeats that deviates from a consistent, predictable pattern. These deviations can involve variations in the timing or regularity of the heartbeats, causing them to be too fast, too slow, or simply erratic [8]. The procedure for detecting and classifying arrhythmias is complicated for a human, as sometimes it is crucial to examine every heartbeat of the ECG records, attained by a Holter monitor, for instance, for hours or even days. An alternative is to employ computational methods for automated classification [9]. Even though many of the traditional pattern recognition methods are previously used effectively to the ECG arrhythmia recognition process, current advanced performances attained by deep learning (DL) methodologies, specifically convolutional neural networks (CNNs), in popular pattern or object detection challenges [10], fortified scholars to apply such approaches to the domain of medical image and signal processing [11]. The use of DL techniques in medical pattern recognition tasks is auspicious, attaining superior outcomes [12]. Unlike traditional classification techniques, CNNs possess the proficiency to automatically derive meaningful characteristics straight from raw data, negating the necessity for manually crafted transfer learning techniques [13].

1.1. Need for Advanced Techniques in Arrhythmia Detection

Cardiovascular diseases remain a primary reason of mortality worldwide, emphasizing the serious requirement for accurate and timely identification. Conventional methods mostly face difficulty with uncertainties and inconsistencies in physiological signal data, limiting their efficiency. Addressing these challenges requires advanced techniques capable of managing ambiguous and incomplete data while extracting meaningful patterns. By incorporating advanced data reduction and uncertainty modelling approaches, it becomes possible to improve the reliability and precision of arrhythmia detection. This advancement has the potential to enhance medical decision-making and patient results in cardiovascular care.

1.2. Study Contribution and Novelty

This paper presents an Enhanced Diagnostic Model for Cardiac Arrhythmia using Principal Component Analysis and Pythagorean Neutrosophic Bonferroni Mean (DMCA-PCAPNBM) technique in Cardiovascular Signal Processing. The objective is in the automated arrhythmia detection using advanced techniques. Initially, the DMCA-PCAPNBM model applies the min-max scaler-based data pre-processing technique for transforming input data into an appropriate format. In addition, the principal component analysis (PCA) method is used for the feature subset selection system to select the optimal attributes from the dataset. For the arrhythmia detection model, the PNBM method is utilized. Finally, the improved dung beetle optimization (IDBO) approach is applied for parameter tuning, resulting in enhanced classification performance. A comprehensive experimentation is implemented to verify the superior outcome of the DMCA-PCAPNBM model on the ECG arrhythmia classification dataset. The main contribution of the DMCA-PCAPNBM approach is summarized as shown.

- The DMCA-PCAPNBM method applies min-max scaler-based data preprocessing to normalize the input ECG signals, which improves data consistency and training efficiency of the classifier. This step confirms that overall features equally contribute during learning, mitigating bias caused by scale differences and ultimately resulting in improved classification performance.
- The DMCA-PCAPNBM methodology employs the PCA-based feature subset selection to effectively mitigate the ECG data dimension, eliminating redundant and irrelevant attributes. This mitigation improves computational efficiency and helps prevent overfitting by focusing on the most informative components. As a result, the model achieves improved accuracy and faster processing times during arrhythmia classification.
- The DMCA-PCAPNBM approach incorporates the PNBM-based arrhythmia detection process that utilizes the PNBM model for improved handling of uncertainty and imprecision in ECG signals. This methodology enhances decision-making accuracy by effectively aggregating ambiguous information from diverse sources. As a result, the detection process becomes more reliable and robust against noisy or incomplete data.
- The DMCA-PCAPNBM technique utilizes IDBO-based parameter tuning for optimizing hyperparameters, ensuring the best possible performance of the arrhythmia classification system. This enhanced dung beetle optimization algorithm efficiently searches the parameter space to avoid local minima and to improve

convergence speed. Consequently, the model attains higher accuracy and stability in detecting cardiac arrhythmias.

- The innovation of the DMCA-PCAPNBM technique is in incorporating the PNBM with DL and advanced optimization techniques for handling ambiguity and imprecision in ECG data. This fusion enhances the capability of the model in processing ambiguous data, resulting in more accurate arrhythmia classification. The combined approach improves robustness and adaptability across diverse cardiac conditions. It represents a unique framework that bridges neutrosophic theory with practical DL applications for improved healthcare outcomes.

2. Literature Review on Cardiac Arrhythmia Detection

Wang et al. [14] introduced a multi-teacher self-distillation (MTSD) architecture to enhance the outcome of arrhythmia techniques in wearable ECG observation. Particularly, the initially employed model from teacher representations through connection of patterns in middle layers, to extract inter-channel and inter-category relations, integrates the model to guarantee the precision and suitability. In [15], a polymorphic representation-aware network (PRA-Net) is projected for cardiac arrhythmia identification from the signals of ECG, in which the adaptive wavelet-aware (AWA) mechanism is designed to recognize ECG signals at diverse scales by extracting multiscale features, whereas the adaptive temporal-space-aware (ATA) mechanism is created for modelling the temporal and positional relationships in the sequence, and the collaborative semantic-aware (CSA) technique is developed for encoding semantic data from waveforms, lastly, the adaptive feature attention fusion (AFAF) mechanism is designed for selecting useful information for cardiac arrhythmia classification. Daduvy et al. [16] proposed the MultiScale Convolutional LSTM Dense Network (MS-CLDNet) framework, a sophisticated DL system specially tailored for addressing the problems and increase arrhythmia classification precision. This study focuses on developing an effective DL approach. Tao et al. [17] presented an interpretable multileader ECG recognition method that solves those restrictions in 2 central regions. Initially, a flexible DKR block (DKR-block) is offered for proficiently removing inter- and intra-lead features. Furthermore, a 2D ECG classification method is created depending on the DKR-block for precise arrhythmia recognition.

In [18], a Binarised spiking NN (BSNN) enhanced with a Momentum search model for Fetal Arrhythmia Identification based on ECG signals to classify the result as usual or arrhythmia. The output is captured by the hexadecimal local adaptive binary pattern (HLABP). In general, BSNN does not implement some optimization techniques to define the optimum parameters and to ensure a precise classification. Al Fahoum [19] suggested a new methodology for automatically diagnosing arrhythmia. This methodology comprises a multiscale filter bank with scalograms that utilizes pre-processed ECG information and non-weighted, pretrained CNNs. Temporal frequency textures offer 2D representations of basic features from single-lead ECG recordings. Then, DL-NNs are specially created to classify arrhythmia. This investigation examines the usefulness of DL systems in cardiac arrhythmia classification from ECG data. In [20], an EPO-driven deep quantum NN (QNN) is designed to classify the heartbeat to detect arrhythmias. Now, the ECG signal is preprocessed to eliminate noise. Additionally, Deep QNN performs the arrhythmia diagnosis, where EPO trains Deep QNN's weights. Prakash and Atef [21] developed a lightweight DL model integrating CNN and Bi-directional LSTM (BiLSTM) for accurate and efficient ECG beat classification. Panigrahi et al. [22] proposed a DL technique by utilizing deep TL (DTL) for feature extraction and the African Vulture Optimization Algorithm (AVOA) for FS, combined with ML classifiers to improve cardiovascular disease (CVD) diagnosis from electrocardiogram (ECG) signals.

2.1. Limitations and Research Gap in Existing DL-Based Arrhythmia Detection

Despite improvements, existing DL models illustrate limited generalization due to overfitting on specific ECG datasets. MTSD-based models depend heavily on mid-layer alignment, yet lack robustness across varying data distributions. PRA-Net introduces complex adaptive modules (AWA, ATA, CSA, AFAF), but their computational overhead affects real-time deployment. MS-CLDNet improves precision but overlooks interpretability. BSNN lacks effective optimization strategies, affecting classification accuracy. The research gap is in unifying interpretability, computational efficiency, and accurate classification in a single scalable architecture.

3. Proposed System Architecture

This paper designs an enhanced DMCA-PCAPNBM model in Cardiovascular Signal Processing. The DMCA-PCAPNBM paper aims to implement a precise and successful technique for automated arrhythmia detection using advanced algorithms. It comprises data pre-processing, feature selection, arrhythmia detection, and parameter tuning. Fig. 1 describes the complete flow of the DMCA-PCAPNBM method.

DOI: https://doi.org/10.54216/IJNS.270208

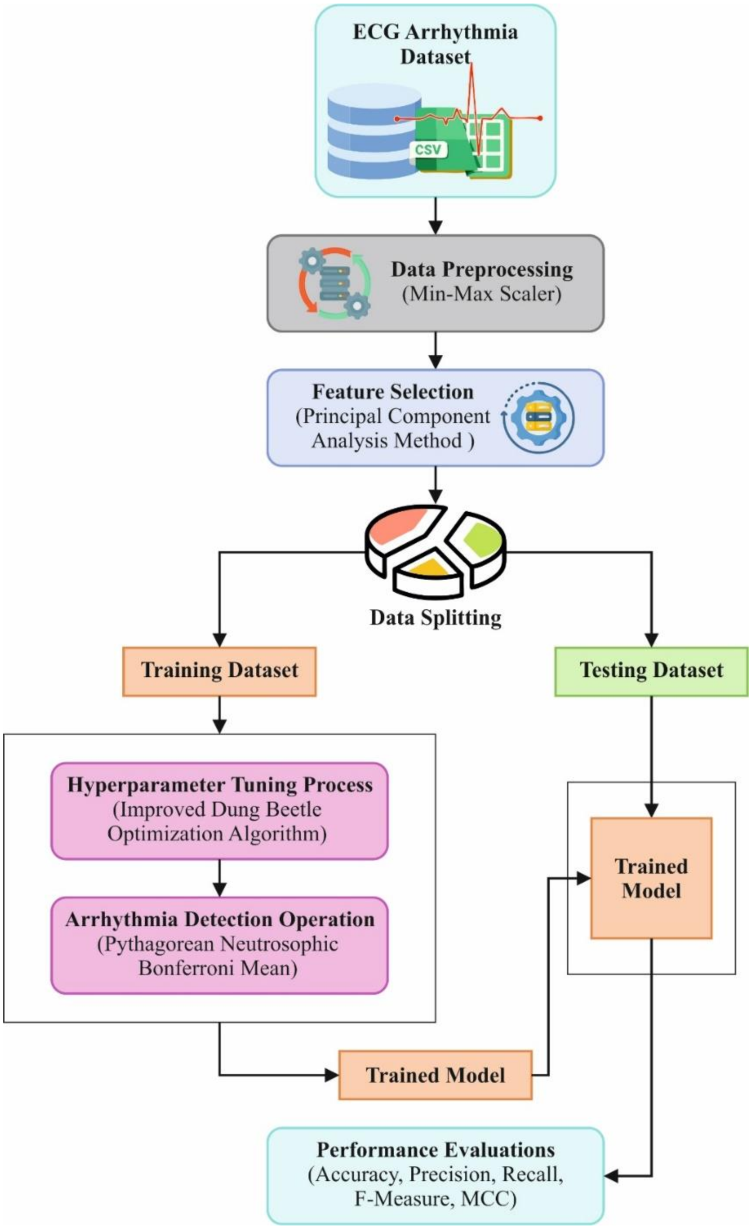


Figure 1. Complete procedure of DMCA-PCAPNBM model

A. Min-max Scaler-based Data Pre-processing Model

Initially, a min-max scaler-based data pre-processing approach is employed to transform the input data into the most appropriate design [23]. This model is selected for its simplicity and effectiveness in normalizing data within a fixed range [0-1]. This technique preserves the original distribution shape of the data, which is significant for algorithms sensitive to data scale, like neural networks. This approach confirms that overall features contribute equally to the learning process, preventing bias toward features with larger numeric ranges. Additionally, min-max scaling is computationally efficient and easy to implement, making it appropriate for large ECG datasets. Its ability to maintain the relative relationships between data points assists in improving model convergence and overall classification accuracy. At the initial phase of pre-processing, the signal is scaled to be accepted within (0, 1).

$$M_i = \frac{X_i - \min(X_i)}{\max(X_i) - \min(X_i)} \quad (1)$$

Whereas M_i and X_i signify the normalized and clear signals. Here, the variance between the maximum and minimum X values was divided. Formerly, the minimal value X was removed from the signal X_i that is considered.

DOI: https://doi.org/10.54216/IJNS.270208

The values of the array were converted into the standardized range of 0 to 1. Indeed, specific problems have been experienced with the dataset pre-processing, such as selecting the successful normalization range, as negative ranges failed to provide effective outcomes, and resolving data overlap in the database.

B. Feature Subset Selection using PCA Method

Additionally, the PCA technique is applied for the feature subset selection system to choose the optimal attributes from a dataset [24]. This method is selected for its efficiency in mitigating dimensionality while preserving the most crucial data variance. This approach effectually transforms the original features into a new set of uncorrelated components, minimizing redundancy and noise, unlike other methods that may rely on feature importance scores or heuristics. This reduction improves computational efficiency and prevents overfitting, specifically with high-dimensional ECG data. The capability of the model in capturing the most significant patterns in fewer components improves model performance and generalization. Additionally, it is a well-established, mathematically robust technique that works well with continuous data, making it appropriate for pre-processing physiological signals. Fig. 2 depicts the PCA method.

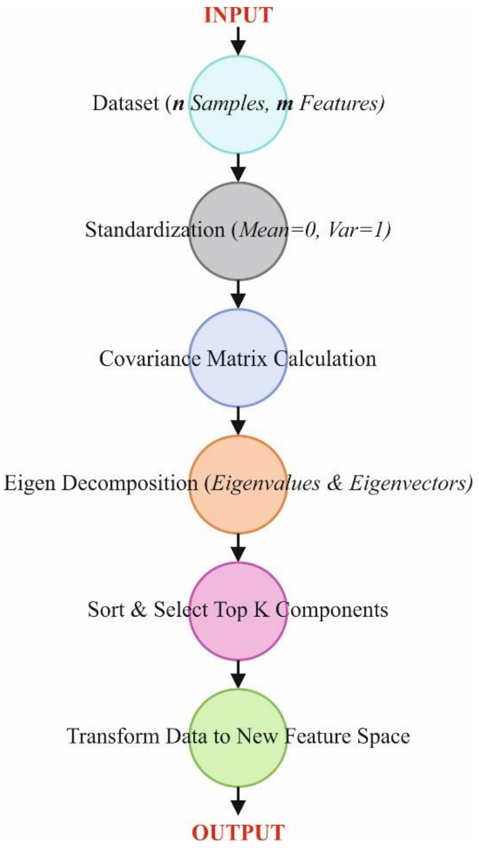


Figure 2. Complete procedure of DMCA-PCAPNBM model

Afterwards, data pre-processing is performed, and the batch impacts are used to decrease the size of the extracted data to utilize PCA. It reduces dimensionality to transform high-dimensional data towards a space with fewer dimensions, while preserving maximal variance. The model comprises standardizing the data to hold unit discrepancy and mean, and then calculating the covariance matrix C of features. PCA discovers the eigenvectors v_i of the covariance matrix and eigenvalues λ_i; the core dynamics are linear blends of initial features.

$$C = \frac{1}{n-1} X_t Y \tag{2}$$

Utilizing Eq. (3), eigenvalues exhibit the magnitude of variance beside the lines of maximum variation, while eigenvectors reflect the direction of variance.

$$Cv_i = \lambda_i v_i \tag{3}$$

DOI: https://doi.org/10.54216/IJNS.270208

Here, every key component depicts a linear integration of original features, which decreases and streamlines the data while maintaining the most crucial variance computed by the succeeding equation:

$$Y_{reduced} = Y \cdot W \tag{4}$$

Now W indicates the medium of eigenvectors equivalent to the leading eigenvalues, $Y_{reduced}$ indicates the reduced data, and Y signifies the standardized feature matrix. PCA is crucial for identifying images, because it facilitates the intricate data by emphasizing the core element that accounts for the variance, thus enhancing computation efficacy and performance of the model by concentrating on the most informative aspects and decreasing redundancy.

C. PNBM-based Arrhythmia Detection

For the arrhythmia detection operation, the PNBM model is employed [25]. This method is selected for its superior ability in handling uncertainty, indeterminacy, and inconsistent data in ECG signals compared to conventional methods. This technique effectually captures the intrinsic relationships among multiple criteria by incorporating truth, indeterminacy, and falsity components. These results in more accurate and reliable decision-making in noisy and ambiguous environments, such as those encountered in medical data. Its mathematical flexibility improves robustness against data imperfections, enhancing classification accuracy. Furthermore, the capability of the model in aggregating diverse data sources makes it an ideal tool for precise arrhythmia detection in complex cardiovascular datasets. Fig. 3 specifies the PNBM architecture.

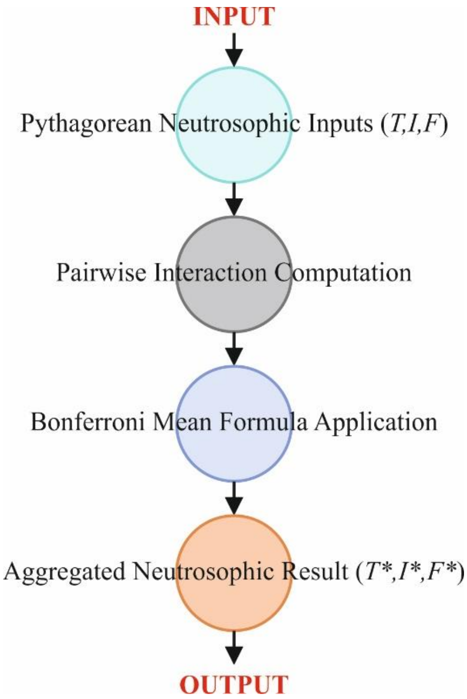


Figure 3. Complete procedure of DMCA-PCAPNBM model

The main idea related to the Pythagorean NS (PNS) is reviewed in this section.

Assume X refers to set of non-empty. The PNS with ψ and κ as dependent membership

$$A = \{\langle x_t \psi_A \sigma_A(x), \varsigma_A(x), \kappa_A(x)\rangle | x \in X\} \tag{5}$$

During Eq. (5), κ_A, ψ_A, and ς_A represent false, truth, and indeterminacy memberships, respectively.

$$0 \le \psi^2 + \kappa^2 \le 1 \tag{6}$$

$$0 \le \psi^2 + \varsigma^2 + \kappa^2 \le 2 \tag{7}$$

Considering $x_1 = \left(\psi_{x_1}, \sigma_{x_1}, \kappa_{x_1}\right), x_2 = \left(\psi_{x_2}, \sigma_{x_2}, \kappa_{x_2}\right)$ and $x = (\psi_x, \sigma_x, \kappa_x)$ are PNSs, formerly the operational instructions are provided as shown:

$$x_1 \oplus x_2 = \left(\sqrt{\psi_{x_1}^2 + \psi_{x_2}^2 - \psi_{x_1}^2\psi_{x_2}^2}, \varsigma_{x_1}\varsigma_{x_2}\kappa_{x_1}\kappa_{x_2}\right) \tag{8}$$

$$x_1 \otimes x_2 = \left(\psi_{x_1}\psi_{x_2}, \varsigma_{x_1} + \varsigma_{x_2} - \varsigma_{x_1}\varsigma_{x_2}\sqrt{\kappa_{x_1}^2 + \kappa_{x_2}^2 - \kappa_{x_1}^2\kappa_{x_2}^2}\right) \tag{9}$$

$$\mu x = \left(\sqrt{1-(1-\psi_x^2)^{\mu}}, \varsigma_x^{\mu}, \kappa_x^{\mu}\right) \; where \; \mu \in \Re \; and \; \mu \geq 0 \tag{10}$$

$$x^{\mu} = (\psi_x^{\mu} 1 - \left(1-\varsigma_x)^{\mu t}\sqrt{1-(1-\kappa_x^2)^{\mu}}\right) \; where \; \mu \in \Re \; and \; \mu \geq 0. \tag{11}$$

Let $p, q \geq 0$ with $x_i = \left(\psi_i(x), \varsigma_i(x), \kappa_i(x)\right)$ while $(i = 1,2,3, \dots, n)$ refers to PNS, then PNBM is provided

$$PNBM\ (x_1, x_2, \dots, x_n)^{p,q} = \left(\frac{1}{n(n-1)} \oplus_{i,j=1\ \ i\neq j}^{n} \left(x_i^p \oplus x_j^p\right)\right)^{\frac{1}{p+q}} \tag{12}$$

Stage1: Create the Matrix of Direct-Relation, X^k

PNS produced the matrix to distinguish the direct relationship depending on the decision-making selections and evaluate the choice as a non-negative matrix. $X^k = [x_{ij}^k]_{n\times n}$, whereas $1 \leq k \leq m$. The representation of $x_{ij} = \left(\psi_{ij}, \varsigma_{ij}, \kappa_{ij}\right)$ states the range of the decision-makers' beliefs that states i impact states j, with the diagonal element being 0. Seven semantic processes varying from ‘no effect’ to ‘high effect’ based on the PNS’s linguistic variable calculate the mark. Hence, there are m dissimilar matrices $X = \{X^1, X^2, \dots, X^m\}$ akin to each of the DMs.

Stage2: Get the Matrix of Aggregate Direct-Relation A.

According to Eq. (12), PN-NWBM incorporates the matrix of direct relationships. $X^k = \{X^1, X^2, \dots, X^m\}$ into an incorporated decision matrix $A = \left[a_{ij}\right]_{n\times n}$, where $a_{ij} = (\psi_{ij}, \varsigma_{ij}, \kappa_{ij})$.

Stage3: De-neutrosophication into crisp matrix B.

The matrix of aggregate $A = [a_{ij}]_{m\times n}$ is transformed into a crisp matrix, B, which is de-neutrosophicated in Eq. (13).

$$B = \frac{\psi_A(x) + \varsigma_A(x) + \kappa_A(x)}{3} \tag{13}$$

Stage4: Standardize the matrix to a normalized matrix Z. Eq. (14) is applied to standardize the process of the matrix.

$$Z = \frac{B}{s} \tag{14}$$

Whereas $s = \max\limits_{l\leq i\leq n} \sum_{j=1}^{n} b_{ij}$ and every component in the matrix Z follows $0 \leq z_{ij} < 1$.

Stage5: Construct the Matrix of Total-Influence T

Constructing the matrix of influence using the following expression.

$$T = Z(l - Z)^{-1} \tag{15}$$

During Eq. (15), l depicts the identity matrix.

Stage6: Determine the Sum of the Rows and Columns

The vectors R and C portray the sum of rows, and the entire matrix of influence T is described in Eqs. (16) & (17).

$$R = [\widetilde{r_l}]_{n\times 1} = [\sum_{j=1}^{n} t_{ij}]_{n\times 1} \tag{16}$$

$$C = [\widetilde{c_l}]_{1\times n} = [\sum_{i=1}^{n} t_{ij}]_{1\times n}^{T} \tag{17}$$

Now, t_{ij} denotes the matrix module T. $R + C$ and R-C values fix the relation and significance values, correspondingly.

DOI: https://doi.org/10.54216/IJNS.270208

Stage7: Network Relationship Map (NRM), the Threshold Value

The database model of the graph is $(R+C, R-C)$. $R-C$ and $R+C$ are considered on the vertical and horizontal axes, correspondingly.

D. Parameter Tuning using IDBO Process

At last, the IDBO method is utilized for the parameter-tuning model, resulting in improved classification performance [26]. This model is selected for its enhanced search capability and faster convergence compared to conventional optimization algorithms. This model efficiently balances exploitation and exploration, preventing local optima and guaranteeing a more thorough search of the parameter space. Its bio-inspired mechanisms allow adaptive adjustments, making it appropriate for intrinsic, high-dimensional problems like tuning DL models. This methodology is more effective and requires less evaluation compared to standard models such as grid or random search, thus reducing the computational cost. This results in optimized model performance with improved accuracy and stability in arrhythmia classification tasks. Fig. 4 indicates the phases in the IDBO technique.

Figure 4. Steps involved in the IDBO methodology

The dung beetle's behaviour stimulated the DBO method. DBs feed on creature dung and exhibit an innovative behaviour of rolling dung into balls and directing it, utilizing celestial signs. It is classified into the succeeding 4 categories: small DB, the ball-rolling DB, the thief DB, and the brood balls DB. By prioritizing this natural behaviour over mathematical methods, the DBO enhances an optimizer able to discover optimum solutions effectively. This model relates the natural behaviour principles for solving composite difficulties, connecting the DB's tactics for problem solving in a computational environment.

1) Ball-Rolling DB

In the rolling method, the location upgrade of the ball-rolling DBs can be separated into the subsequent dual settings: with and without difficulties. If the ball-rolling DB does not meet problems, the location upgrade follows the equation below:

$$x_i(t+1) = x_j(t) + \alpha \times k \times x_i(t-1) + b \times \Delta x \qquad (18)$$

$$\Delta x = |x_i(t) - X^w| \qquad (19)$$

Here, t depicts the present iteration count, and $x(t)$ characterizes the location of $i\ th$ DB rolling the ball; k means constant and $k \in (0,0.2]$ signifies the location coefficients deviation; b refers to a continual value characterized by$(0,1)$ α stands for natural coefficients allocated to 1 or -1, X^w represents a globally poor location; Δx represents modifications in light intensities. The location upgrade is as demonstrated:

$$x_i(t+1) = x_j(t) + \tan(\theta)|x_i(t) - x_i(t-1)| \quad (20)$$

Now, θ means deviation angle that comprises the range $[0,\pi]$. During this updated equation, $|x_i(t) - x_i(t-1)|$ characterizes the position dissimilarity of the $i\ th$ DB among the $t\ th$ and $t-1th$ iterations. Hence, the location upgrade of the rolling DB is closely connected with either present or historical information. Note that when $\theta = 0, \pi/2$, or π, the location of the DB shall not be upgraded.

2) Brood Ball DB

In the DB spawn phase, the area chosen by female DBs for laying eggs is represented. The boundary selection approach is shown below:

$$Lowb^* = \max(X^* \times (1-R), Lowb) \quad (21)$$

$$Upb^* = \min(X^* \times (1+R), Upb) \quad (22)$$

Here X^* signifies the present local best location; Upb^* and $Lowb^*$, $Lowb$ and Upb denotes the smaller and higher limits; $R = 1\ - Tmax$; $Tmax$ signifies the maximal iteration counts; correspondingly.

$$B_i(t+1) = X^* + b_1 \times (B_i(t) - Lowb^*) + b_2 \times (B_i(t) - Upb^*) \quad (23)$$

Now $B_i(t)$ refers to the location of the $i\ th$ ball at the iteration i ; b_1 and b_2 are dual self-determining arbitrary vectors, all size $1 \times D$, while D means size.

3) Small DB

Brood balls will turn out to be small DBs, which originate on the ground to forage. The choice of the optimum foraging region is described below:

$$Lowb^b = \max(X^b \times (1-R), Lowb) \quad (24)$$

$$Upb^b = \min(X^b \times (1+R), Upb) \quad (25)$$

whereas X^b signifies the globally optimum location; $Lowb^b$ and Upb^b depicts smaller and higher boundaries, respectively.

Once the best foraging region, the updating of the location for searching by the smaller DBs is demonstrated:

$$x_i(t+1) = x_i(t) + C_1 \times (x_i(t) - Lowb^b) + C_2 \times (x_i(t) - Upb^b) \quad (26)$$

Here, $x_i(t)$ characterizes the place of the ith DB at ith iteration; C_1 refers to randomly generated numbers succeeding normal distributions; C_2 means an stochastic vector within the range of zero and one.

4) Thief DB

From Eq. (25), it is comprehensible that X^b denotes the best source of food. The region around X^b represents the best place for contesting for meal. In the procedure of iteration, the position data of the thieves is upgraded as demonstrated:

$$x_i(t+1) = X^b + S \times g \times (|x_i(t) - X^*| + |x_i(t) - X^b|) \quad (27)$$

Here, $x(t)$ symbolizes the place data of the i^{th} thief at the ith iteration; g means an arbitrary vector resulting in standard distributions, with a dimension of 1xD and S denotes a constant value.

5) Optimal Latin Hypercube Sampling (OLHS)

The first distribution in intelligent optimizer methods has a considerable influence on the subsequent 2 key results: the algorithmic accuracy of convergence and its global searching efficacy. Furthermore, the model might not determine the globally best solution owing to unequal primary distribution, rather than dropping into a suitable local solution. This paper utilizes this model to determine the DB population's original range. In this method, the random, OLHS, and LHS models in a 2D space make the primary distribution of 50 individual beetles $[0,1]$. The initialization results in a more even first population distribution, thus giving more data about the globally best solution. The LHS method is enhanced by the OLHS approach, targeting to improve spatial coverage. The population made with the best LH sampling additionally increases the diversity and quality of the DB population. The succeeding equation signifies the function applied to define the scalar discrimination:

DOI: https://doi.org/10.54216/IJNS.270208

$$\phi_q(X) = \left(\sum_{i=1}^{m} J_i^{d_i^{-q}}\right)^{\frac{1}{q}} \tag{28}$$

Now, $d_i i = 1,2,3, \dots, m$ characterizes the distance of each of the potential groupings of dual points, whereas $J_i, i = 1,2,3, \dots,$ and m denotes the pair counts of points through the distance $d_j; q$ signifies the model, which is an optimistic integer. The IDBO method initiates a fitness function (FF) to accomplish improved classifier outcome. It outlines an advanced value to imply the enriched efficacy of the candidate solution. The minimization of the classifier error rate was measured as FF, as established in Eq. (29).

$$\begin{aligned} fitness(x_i) &= ClassifierErrorRate(x_i) \\ &= \frac{No. of\ misclassified\ instances}{Overall\ instances} \times 100 \end{aligned} \tag{29}$$

4. Validation and Results

The empirical analysis of the DMCA-PCAPNBM system is inspected under the ECG arrhythmia classification database [27]. Table 1 describes the database. The total no. of features is 33, but only 25 features are chosen.

Table 1: Details of the database

Type	Samples
“Normal (N)”	1000
“Supraventricular ectopic beat (SVEB)”	1000
“Ventricular ectopic beat (VEB) ”	1000
“Fusion beat (F) ”	23
“Unknown beat (Q) ”	79
Total Samples	3102

Fig. 5 shows the heartbeat class distribution, which highlights a vital class imbalance amongst dissimilar heartbeat categories. The x-axis embodies five heartbeat classes like N, SVEB, VEB, Q, and F. The y-axis signifies the occurrence count for all classes.

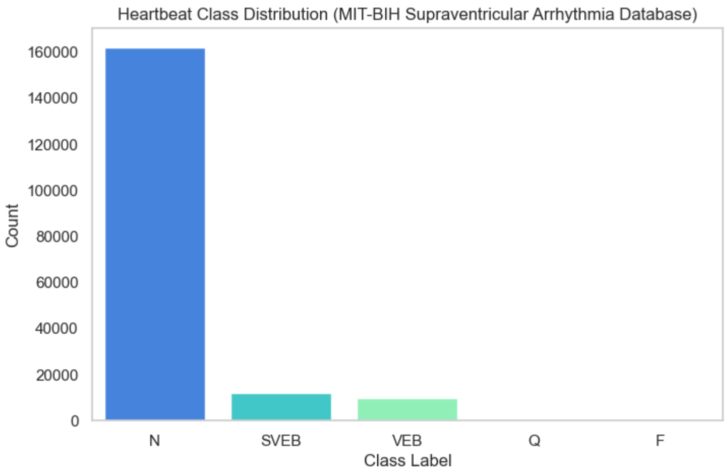


Figure 5. Heartbeat class distribution

DOI: https://doi.org/10.54216/IJNS.270208

Fig. 6 illustrates a comparison of heartbeat class distributions utilizing PCA and t-SNE for the reduction of dimensionality. PCA presents a linear projection with clear class overlapping, limiting separation. On the other hand, t-SNE successfully captures non-linear patterns, resulting in clear class clustering and separation.

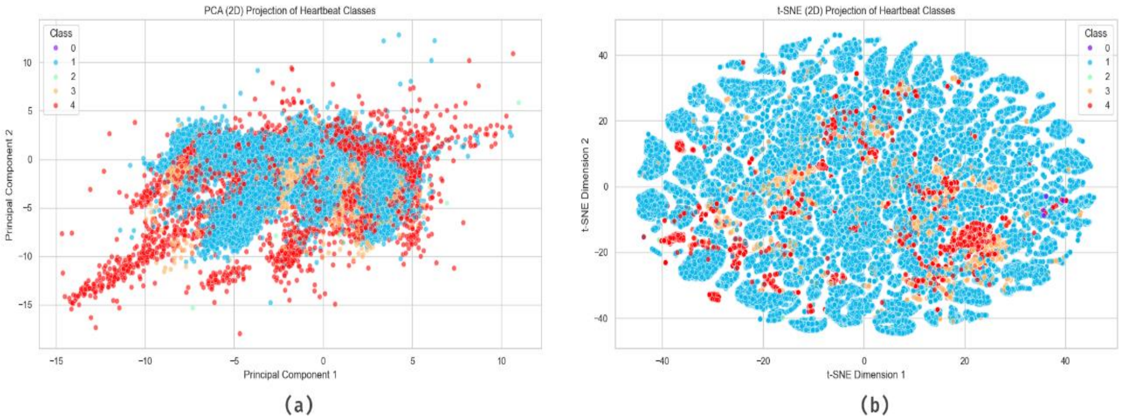
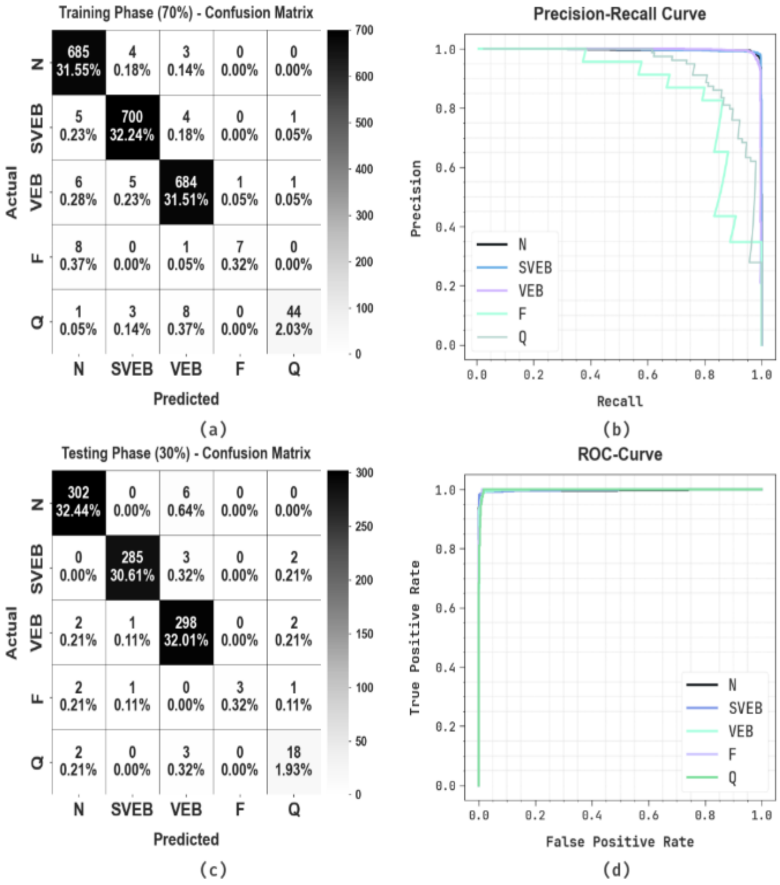


Figure 6. Projection of Heartbeat Classes a) PCA and b) t-SNE

Fig. 7 depicts the classifier outcome of the DMCA-PCAPNBM methodology. Figs. 7a-7c display the confusion matrices at 70:30. Fig. 7b and 7d describes the PR and ROC study, specifying greater performance through each class.


Training Phase (70%) - Confusion Matrix

Actual \ Predicted	N	SVEB	VEB	F	Q
N	685 31.55%	4 0.18%	3 0.14%	0 0.00%	0 0.00%
SVEB	5 0.23%	700 32.24%	4 0.18%	0 0.00%	1 0.05%
VEB	6 0.28%	5 0.23%	684 31.51%	1 0.05%	1 0.05%
F	8 0.37%	0 0.00%	1 0.05%	7 0.32%	0 0.00%
Q	1 0.05%	3 0.14%	8 0.37%	0 0.00%	44 2.03%

Testing Phase (30%) - Confusion Matrix

Actual \ Predicted	N	SVEB	VEB	F	Q
N	302 32.44%	0 0.00%	6 0.64%	0 0.00%	0 0.00%
SVEB	0 0.00%	285 30.61%	3 0.32%	0 0.00%	2 0.21%
VEB	2 0.21%	1 0.11%	298 32.01%	0 0.00%	2 0.21%
F	2 0.21%	1 0.11%	0 0.00%	3 0.32%	1 0.11%
Q	2 0.21%	0 0.00%	3 0.32%	0 0.00%	18 1.93%

Figure 7. Classifier result of (a-c) 70:30 confusion matrices and (b-d) PR and ROC curves

DOI: https://doi.org/10.54216/IJNS.270208

Table 2 depicts the Arrhythmia detection of the DMCA-PCAPNBM approach. On 70% TRPHE, the DMCA-PCAPNBM approach accomplishes an average $accu_y$ of 99.06%, $prec_n$ of 95.27%, $reca_l$ of 83.61%, $F_{Measure}$ of 87.81%, and MCC of 87.97%. Moreover, on 30% TSPHE, the DMCA-PCAPNBM technique gains an average $accu_y$ of 98.93%, $prec_n$ of 94.35%, $reca_l$ of 83.16%, $F_{Measure}$ of 86.47%, and MCC of 86.85%.

Table 2: Arrhythmia detection of DMCA-PCAPNBM model

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F_{Measure}$	MCC
TRPHE (70%)					
N	98.76	97.16	98.99	98.07	97.16
SVEB	98.99	98.31	98.59	98.45	97.70
VEB	98.66	97.71	98.13	97.92	96.94
F	99.54	87.50	43.75	58.33	61.69
Q	99.36	95.65	78.57	86.27	86.38
Average	99.06	95.27	83.61	87.81	87.97
TSPHE (30%)					
N	98.71	98.05	98.05	98.05	97.09
SVEB	99.25	99.30	98.28	98.79	98.24
VEB	98.17	96.13	98.35	97.23	95.88
F	99.57	100.00	42.86	60.00	65.32
Q	98.93	78.26	78.26	78.26	77.71
Average	98.93	94.35	83.16	86.47	86.85

Fig. 8 exemplifies the training (TRAIN) $accu_y$ and validation (VALID) $accu_y$ of a DMCA-PCAPNBM methodology over 50 epochs. At the primary stage, either TRAIN or VALID $accu_y$ rise quickly, demonstrating successful learning of patterns from the data. Around the epoch count, the VALID $accu_y$ somewhat exceeds the training $accu_y$, proposing excellent generalizability without overfitting. As training developments, imitating the maximum and minimum performance gaps between TRAIN and VALID. The neighbouring alignment of both curves in training suggests that the method is generalized and well regularised.

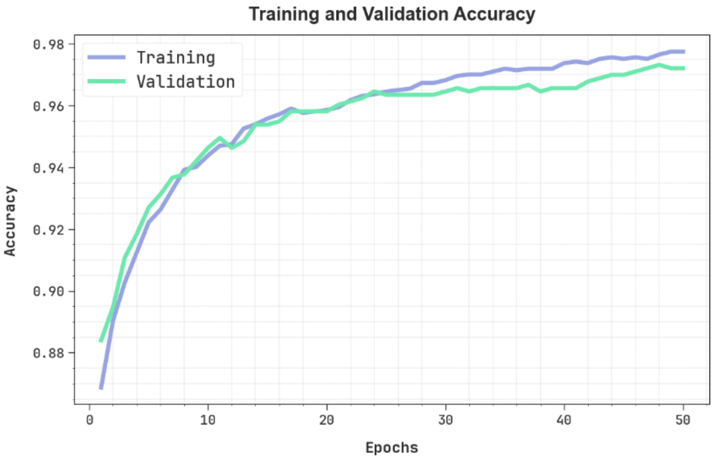


Figure 8. $Accu_y$ Curve of DMCA-PCAPNBM model

DOI: https://doi.org/10.54216/IJNS.270208

Fig. 9 exhibits the TRAIN and VALID losses of the DMCA-PCAPNBM system over 50 epochs. At first, either TRAIN or VALID losses are maximum, specifying that the method begin with a partial understanding of the data. As training advancements, both losses constantly fall, indicating that the approach is efficiently learning and improving its parameters. The closest alignment among the TRAIN and VALID loss curves during training suggests that the technique has not overfitted and preserves good generalization to unseen data.

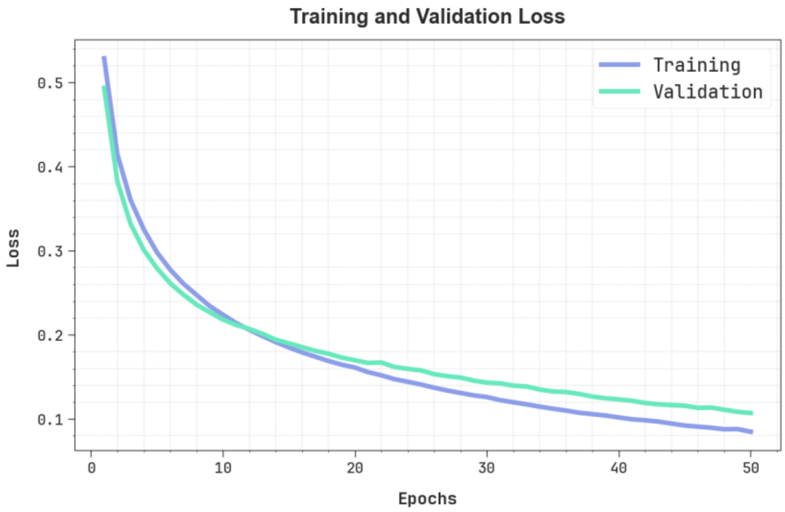


Figure 9. Loss curve of DMCA-PCAPNBM model

Table 3 and Fig. 10 illustrate the comparative analysis of the DMCA-PCAPNBM algorithm with recent methodologies in terms of dissimilar metrics [21-22, 28-29]. The DMCA-PCAPNBM approach has achieved maximum performance with $accu_y$ of 99.06%, $prec_n$ of 95.27%, $reca_l$ of 83.61%, and $F_{Measure}$ of 87.81%. In the meantime, the recent methods such as Bi-LSTM, DTL, XGBoost, 2D-CNN, FFNN, RFT, K-NN, BiLSTM, MLA-CNNBiGRU, and CNN2-SVM have shown poor performance compared to the DMCA-PCAPNBM method.

Table 3: Comparative study of DMCA-PCAPNBM model with recent techniques [21-22, 28-29]

Technique	$Accu_y$	$Prec_n$	$Reca_l$	$F_{Measure}$
Bi-LSTM	97.51	92.94	81.08	80.19
DTL	92.75	92.61	79.41	82.31
XGBoost	97.67	94.87	79.41	83.92
2D-CNN	99.00	91.33	81.85	84.79
FFNN	96.94	92.35	80.36	79.57
RFT	92.16	91.89	78.87	81.79
K-NN	97.00	95.13	78.91	83.36
BiLSTM	97.30	89.72	78.63	82.19
MLA-CNNBiGRU	98.23	93.21	82.16	85.81
CNN2-SVM	97.33	90.62	77.19	80.95
DMCA-PCAPNBM	99.06	95.27	83.61	87.81

DOI: https://doi.org/10.54216/IJNS.270208

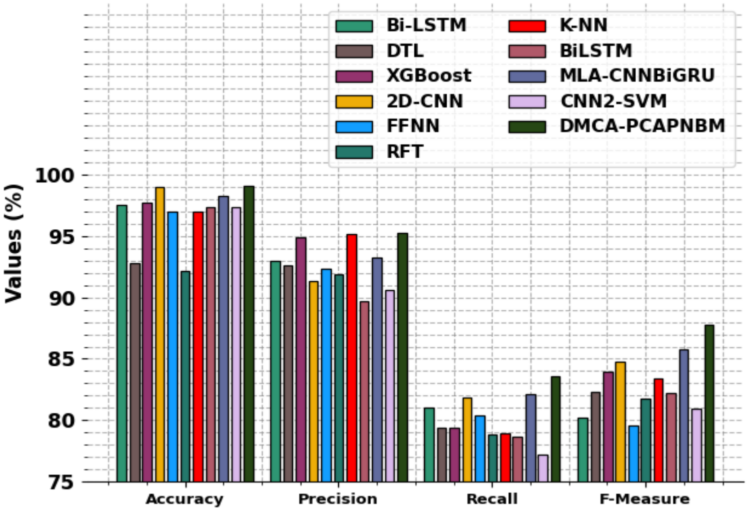


Figure 10. Comparative analysis of the DMCA-PCAPNBM with recent techniques

5. Conclusion

This paper designs an enhanced DMCA-PCAPNBM model in cardiovascular signal processing. The DMCA-PCAPNBM model aims to implement a precise and successful technique for automated arrhythmia detection using advanced algorithms. To obtain this, the DMCA-PCAPNBM model initially uses the min-max scaler-based data pre-processing model for converting the input data into the most appropriate format. Additionally, the PCA method is applied for the feature subset selection model to choose the optimal features. For the arrhythmia detection operation, the PNBM model is employed. Finally, the IDBO method is utilized for the parameter-tuning model, resulting in improved classification performance. A comprehensive experimentation is implemented to verify the superior performance of the DMCA-PCAPNBM on the ECG arrhythmia classification database. The performance assessment of the DMCA-PCAPNBM system represented a better precision value of 99.06% over recent approaches.

5.1. Limitations and Directions for Further Research

The limitations of the DMCA-PCAPNBM methodology encompass its dependence on a single dataset, which may affect the generalizability of the results across various populations. Furthermore, the performance of the model under noisy or incomplete data conditions has not been extensively tested, which may limit its real-world applicability. Computational complexity remains a concern, particularly for distribution in resource-constrained environments. The interpretability of the model's decisions could be further enhanced to support clinical acceptance. Future research should focus on expanding dataset diversity, improving robustness to signal artefacts, and optimizing the model for real-time processing. Exploring explainable AI techniques could also provide greater transparency in diagnostic outputs.

Funding: “The authors gratefully acknowledge with thanks technical support provided by the Department of Computer Science and Engineering at University of Colorado Denver, Denver, Colorado, USA”

Data Availability Statement: The data that support the findings of this study are openly available in the Kaggle repository at https://www.kaggle.com/datasets/sadmansakib7/ecg-arrhythmia-classification-dataset?select=MIT-BIH+Arrhythmia+Database.csv, reference number [25].

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] N. B. Khamis and M. A. Hossain, "Analysis of neutrosophic sets in decision-making problems," *Journal of Fuzzy Logic and Intelligent Systems*, vol. 32, no. 1, pp. 1–12, 2021.

[2] A. K. Gupta, S. R. Kumar, and L. P. Singh, "Applications of neutrosophic logic in medical diagnosis," *International Journal of Health Sciences*, vol. 15, no. 4, pp. 123–130, 2022.

DOI: https://doi.org/10.54216/IJNS.270208

[3] R. S. Patil, A. H. Joshi, and V. K. Pawar, "Neutrosophic decision-making for supply chain management," *Journal of Industrial Engineering and Management*, vol. 14, no. 3, pp. 456–467, 2023.

[4] T. S. Alzahrani and M. F. Alharthi, "Neutrosophic sets in risk assessment: A case study," *Journal of Risk Analysis and Crisis Response*, vol. 13, no. 2, pp. 75–85, 2022.

[5] H. M. Alharbi and F. M. Alhassan, "A novel approach for neutrosophic clustering in data analysis," *Journal of Data Science and Analytics*, vol. 8, no. 1, pp. 45–58, 2023.

[6] A. H. Saheb and R. H. Buti, "A specific category of harmonic functions characterized by a generalized Komatu operator in conjunction with the (RK) integral operator and applications to neutrosophic complex field," *International Journal of Neutrosophic Science*, vol. 23, no. 3, pp. 44–54, 2024.

[7] Y. D. Daydulo, B. L. Thamineni, and A. A. Dawud, "Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, p. 232, 2023.

[8] S. Irfan et al., "Heartbeat classification and arrhythmia detection using a multi-model deep-learning technique," *Sensors*, vol. 22, no. 15, p. 5606, 2022.

[9] R. Anand, S. V. Lakshmi, D. Pandey, and B. K. Pandey, "An enhanced ResNet-50 deep learning model for arrhythmia detection using electrocardiogram biomedical indicators," *Evolving Systems*, vol. 15, no. 1, pp. 83–97, 2024.

[10] N. Katal, S. Gupta, P. Verma, and B. Sharma, "Deep-learning-based arrhythmia detection using ECG signals: A comparative study and performance evaluation," *Diagnostics*, vol. 13, no. 24, p. 3605, 2023.

[11] C. U. Kumari, A. S. D. Murthy, B. L. Prasanna, M. P. P. Reddy, and A. K. Panigrahy, "An automated detection of heart arrhythmias using machine learning technique: SVM," *Materials Today: Proceedings*, vol. 45, pp. 1393–1398, 2021.

[12] R. Hu, J. Chen, and L. Zhou, "A transformer-based deep neural network for arrhythmia detection using continuous ECG signals," *Computers in Biology and Medicine*, vol. 144, p. 105325, 2022.

[13] P. Madan et al., "A hybrid deep learning approach for ECG-based arrhythmia classification," *Bioengineering*, vol. 9, no. 4, p. 152, 2022.

[14] S. Ma, J. Cui, W. Xiao, and L. Liu, "Deep learning-based data augmentation and model fusion for automatic arrhythmia identification and classification algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, p. 1577778, 2022.

[15] Z. Wang et al., "Multi-teacher self-distillation based on adaptive weighting and activation pattern for enhancing lightweight arrhythmia recognition," *Information Fusion*, vol. 122, p. 103178, 2025.

[16] [16] X. Wang et al., "PRA-Net: Polymorphic representation aware network for accurate cardiac arrhythmia detection from electrocardiogram signals," *IEEE Transactions on Instrumentation and Measurement*, early access, 2025.

[17] A. Daduvy et al., "A multiscale convolutional LSTM-dense network for robust cardiac arrhythmia classification from ECG signals," *Computers in Biology and Medicine*, vol. 191, p. 110121, 2025.

[18] R. Tao, L. Wang, Y. Xiong, and Y. R. Zeng, "IM-ECG: An interpretable framework for arrhythmia detection using multi-lead ECG," *Expert Systems with Applications*, vol. 237, p. 121497, 2024.

[19] D. Shekhawat et al., "Binarized spiking neural network optimized with momentum search algorithm for fetal arrhythmia detection and classification from ECG signals," *Biomedical Signal Processing and Control*, vol. 89, p. 105713, 2024.

[20] A. Al Fahoum, "Enhanced cardiac arrhythmia detection utilizing deep learning architectures and multiscale ECG analysis," *Tuijin Jishu/Journal of Propulsion Technology*, vol. 44, no. 6, pp. 5539–5554, 2023.

[21] A. Choudhury, S. Vuppu, S. P. Singh, M. Kumar, and S. N. P. Kumar, "ECG-based heartbeat classification using exponential-political optimizer trained deep learning for arrhythmia detection," *Biomedical Signal Processing and Control*, vol. 84, p. 104816, 2023.

[22] A. J. Prakash and M. Atef, "A lightweight deep learning approach for patient-specific electrocardiogram beat classification using local and long-term dependencies," *Engineering Applications of Artificial Intelligence*, vol. 152, p. 110754, 2025.

[23] A. Panigrahi et al., "Advanced ECG signal analysis for cardiovascular disease diagnosis using AVOA optimized ensembled deep transfer learning approaches," *Computers, Materials & Continua*, vol. 84, no. 1, 2025.

[24] D. Al-Shammary, M. N. Kadhim, A. M. Mahdi, A. Ibaida, and K. Ahmed, "Efficient ECG classification based on Chi-square distance for arrhythmia detection," *Journal of Electronic Science and Technology*, vol. 22, no. 2, p. 100249, 2024.

[25] B. Wu, "Research on the strategy of promoting rural tourism development through IoT technology in rural revitalization," *International Journal of High Speed Electronics and Systems*, p. 2540592, 2025.

[26] A. Ahmed, M. Badawy, and G. F. Gubarah, "Modelling of green human resource management using Pythagorean neutrosophic Bonferroni mean approach," *International Journal of Neutrosophic Science*, vol. 24, no. 2, 2024.

[27] Z. Wang, J. Cai, X. Liu, and Z. Zou, "Optimized BP neural network based on improved dung beetle optimization algorithm to predict high-performance concrete compressive strength," *Buildings*, vol. 14, no. 11, p. 3465, 2024.

[28] Kaggle, "ECG arrhythmia classification dataset," 2025. [Online]. Available: https://www.kaggle.com/datasets/sadmansakib7/ecg-arrhythmia-classification-dataset?select=MIT-BIH+Arrhythmia+Database.csv. [Accessed: Aug. 13, 2025].

[29] J. Huang, B. Chen, B. Yao, and W. He, "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network," *IEEE Access*, vol. 7, pp. 92871–92880, 2019.