Journal of Scientific & Industrial Research
Vol. 84, February 2025, pp. 207-218
DOI: 10.56042/jsir.v84i02.13919

Advancing Cardiac Disease Detection Using Feature Extraction, Feature Selection, and Ensemble Learning Approaches

S R Tripathy[1], A Rath[1], R Sharma[2]*, G Panda[1] & Meenakshi Sharma[3]

[1]Department of CSE, C.V. Raman Global University, Bhubaneswar, 752 054, Odisha, India
[2]Department of ECE, NIT, Durgapur, 713 209, West Bengal, India,
[3]Department of Mathematics, Khalsa College for Women, Ludhiana, 141 001, Punjab, India

Received 24 September 2024; revised 06 November 2024; accepted 25 November 2024

Approximately 26 million individuals globally struggle with cardiac disease, and the incidence is increasing by 2% each year. To reduce the healthcare burden, the researchers propose various CAD models. Feature extraction and Feature selection are essential in reducing the model complexity and memory requirement. In the proposed research, we investigate the performance of different feature extraction and selection methods using two heart sound datasets. The features are extracted using MFCC and DWT methods from heart sounds. Four feature selection methods (Fisher's Score, mRMR, ReliefF, and Gini Index) are analyzed and ranked using the D-CRITIC TOPSIS technique. The two best models based on feature selection are utilized in the weighted average ensemble. The weights in ensemble learning are optimized using the Dwarf Mongoose optimization algorithm. The feature fusion model combining DWT and MFCC with mRMR for feature selection achieved the highest performance on the PhysioNet dataset, with an accuracy of 82.70%, an F1-Score of 0.8369, and an AUC-ROC score of 0.9092. The best accuracy, F1-Score, and AUC-ROC score on the PASCAL CHSC dataset are 79.64%, 0.7826, and 0.8116, respectively. The study compared four feature selection methods. The mRMR-based model achieved the highest TOPSIS score and ranked first in the performance table. The findings demonstrate that the mRMR feature selection performed better than other feature selection for both feature extraction methods evaluated in this study. The ensemble model using mRMR and ReliefF outperformed all base models and achieved the highest performance metrics. This study highlighted the enhanced detection of cardiac disorders through the combined effectiveness of feature extraction, feature selection, classification models, and ensemble strategies.

Keywords: Computer-Aided Diagnosis, Classification, Decision-Making, Metaheuristics, ReliefF

Introduction

Cardiac valve disorder is a serious problem related to Cardiovascular Diseases (CVDs) and a common cause of mortality across the world. In 2020, one in three adults in the USA received treatment for cardiovascular risk factors and related conditions. The annual healthcare costs for these procedures are projected to triple from $400 billion in 2020 to $1,344 billion by 2050. For cardiovascular conditions, costs are expected to nearly quadruple, rising from $393 billion to $1,490 billion. Stroke is anticipated to contribute the most significant increase in costs. Notably, the Asian American and Hispanic American populations are expected to see substantial relative increases in expenses (497% and 489%, respectively) due to population growth.[1] Assistive diagnostic cardiac methodologies like Ultrasound, cardiac CT scan, and other monitoring systems are expensive and in high demand because they are used to investigate various other diseases.[2]

The stethoscope is the main instrument used by physicians to deal with heart sounds. By analyzing the Phonocardiogram (PCG) signals, doctors identify abnormal heart conditions. These signals include vital information about the heart valves.[3] Heart murmurs and other abnormal heart sounds are high-frequency signals, while normal heart sounds are low-frequency, short-term signals produced by the heart valves. Blood flow turbulence via small cardiac valves and backflow through the atrioventricular valves are two primary contributors to heart murmurs. Irregular heart murmurs are often caused by heart valve disorders or congenital heart abnormalities. Aortic stenosis, aortic regurgitation, mitral regurgitation, and mitral stenosis are four common pathological murmur forms. PCG recordings can capture important aspects of heart sounds. There is a growing amount of research carried out on PCG signals using biomedical signal processing techniques.[4]

*Author for Correspondence
E-mail: rohithmr.21791@gmail.com

In recent years, heart sound signals have been used to develop Computer-Aided Diagnosis (CAD) tools for detecting abnormalities. These CAD tools leverage machine learning techniques to identify heart-related issues accurately. The machine learning process involves several key stages: feature extraction, feature selection, and prediction. Research on various machine learning problems has indicated that a good combination of feature selection and feature selection is vital for improving prediction accuracy.[5,6]

Research Contributions

1. This work uses Discrete Wavelet Transform (DWT) and Mel-Frequency Cepstral Coefficient (MFCC) to extract feature vectors from heart sound signals.
2. The performance of four filter-based feature selection methods (Fisher's Score, mRMR, Relief F, GINI Index) is analyzed using extracted heart sound features.
3. The D-CRITICS-TOPSIS method identifies and ranks the most effective models on various evaluation stages.
4. For weighted average ensemble learning, the Dwarf Mongoose Optimization Algorithm (DMOA) is employed to optimize the weights for individual models.
5. In the end, models that are based on feature selection and ensemble learning are assessed for their classification performance.

Background

Machine learning-assisted CAD tools have shown promising results in smart healthcare in detecting different diseases in the early stages.[7–9] Significant research efforts have been devoted to the field of automatic heart sound analysis for the detection of cardiovascular CVDs. Singh *et al.* introduced an innovative pre-processing technique using Variational Mode Decomposition (VMD) to effectively denoise heart sounds.[10] The denoised signals are processed through a gamma-tone filter bank and Short-Time Fourier Transform (STFT) to obtain gammatone and spectrogram images. These images are then categorized using various deep convolutional neural network architectures, including AlexNet, SqueezeNet, GoogLeNet, and VGG19, with data augmentation to address imbalanced datasets, and validated using the publicly accessible PhysioNet 2016 dataset. The researchers obtained the highest accuracy of 99.57 % with a Gammatone filter bank and SqueezeNet classifier. Chen *et al.* introduced an end-to-end technique for heart sound segmentation (first sound, systole, second heart sound, and diastole) using a combination of Convolution Neural Network (CNN) and Contextual Long Short-Term Memory (CLSTM) models.[11] The CLSTM algorithm with two dilated layers achieved an average F1 score of 96.18 ± 0.70%. Zeng *et al.* developed an automatic classification model to detect anomalies in PCG samples without segmenting heart audio signals. For this, they implemented artificial intelligence and hybrid signal processing methods.[12] The suggested dynamical neural network-based classifier obtained sensitivity, accuracy, specificity, and AUC score values of 97.73%, 97.89%, 98.05%, and 97.89%, respectively. Almanifi *et al.*[13] evaluated VGG16, VGG19, and ResNet50 models trained to identify heart murmurs. The researchers employed transfer learning techniques on PCG recordings from the PASCAL CHSC dataset. MFCC and spectrograms are used to clean, process, and transform the data into images. According to the study, spectrograms produce the best classification accuracy, with ResNet50, VGG16, and VGG19 obtaining 87.65%, 83.95%, and 83.95% classification accuracies, respectively. Li and others, in their research, proposed a heart sound classification model using improved MFCC features and deep residual learning in their research. The proposed method achieved 94.43% accuracy on the PhysioNetChallenge 2016 heart sound database.[14] A model leveraging attention mechanisms has been introduced for heart disease detection, employing triple spectrogram representations of PCG signals to improve diagnostic accuracy and interpretability.[15] The proposed method obtained a 97.77% overall accuracy. Oh *et al*.[16] discovered a deep-learning WaveNet model to classify 1000 PCG signals into five categories. The proposed model achieved an overall accuracy of 92.00%. In their study, Arora *et al.*[17] proposed an XGBoost-based framework, a heart sound classifier model applied to unsegmented heart sound signals of the PhysioNet dataset, which includes 3240 recordings varying in length from 5 to 120 seconds. The proposed method is benchmarked against 18 existing techniques and demonstrated high accuracy with a mean score of 0.9290. The model achieved sensitivity and specificity scores of 0.9450 and 0.9130, validating its effectiveness in heart sound classification. The spectral features combined with the Adaptive Neuro-Fuzzy Inference System (ANFIS) have been utilized to identify irregular cardiac valve

sound signals, demonstrating their potential in cardiac disease detection.[18] In this experiment, 1837 heart sound signals—1369 normal and 468 unhealthy signals—from the PhysioNet Challenge 2016 dataset were used. The ANFIS neural network achieved significant results with classification accuracy ranging from 63.00% to 89.00%, sensitivity from 63.00% to 100.00%, and specificity from 62.00% to 100.00%, demonstrating the method's superiority. Krishnan *et al.*[19] introduced a 1D CNN with a Feed-forward Neural Network (FNN) for unsegmented PCG signal classification, aiming to remove separate feature engineering and selection blocks. The five-layered FNN model produced an overall accuracy of 0.8565, sensitivity of 0.8673, specificity of 0.8475, and balanced accuracy of 0.8574. Its performance was evaluated using ROC plots, which produced an AUC value of 0.8570. The model was validated using 1081 PCG records. The 1D and 2D CNN frameworks have been proposed for automatic heart sound classification, utilizing short-segmented heartbeats to enhance detection accuracy.[20] A time frequency CNN ensemble (TF-ECNN) merges outputs from both CNNs through the score-level fusion of class probabilities. The proposed models surpassed traditional classifiers like the support vector machine and hidden Markov models on PhysioNet database, achieving top performance with an ECNN achieving 89.22% accuracy and 89.94% sensitivity.

The generalized prediction capability of a learning algorithm always depends on the presence of relevant features. Feature selection provides many advantages, including improvement in the classifier's performance, improving data understanding, reducing data, limiting storage requirements, reducing costs, and increasing simplicity.[21] Feature selection is typically categorized into three types: 1) Filter methods, 2) Wrapper methods, and 3) Embedded methods. Filter methods are independent of the ML algorithm used. They are computationally efficient and have faster performance. However, they perform poorly compared to the two methods but are employed in the initial stage when the feature space is large. Wrapper-based methods are dependent upon learning algorithms. They are more accurate but computationally more expensive than filter methods. In embedded methods, the feature selection is incorporated into the learning algorithm and hence they are faster, computationally less costly, and more accurate.[22] Different feature selection techniques have shown enhanced performance in the healthcare domain.[23–26]

Cuckoo Optimization performed well with wrapper-based approaches, whereas moth flame optimization worked well with filter-based methods.[27] In the medical field, datasets are typically small, and expanding them is challenging due to limited patient participation. Utilizing an ensemble of several weak learners is one method to increase the accuracy of classification models. Ensemble models combine multiple models to minimize bias, variance, and overfitting. By integrating predictions from diverse models, capture a wider range of patterns in the data, improving accuracy and generalization. The collective decision helps to mitigate the weaknesses of individual models and hence increases the robustness and reliability of the prediction.[28] Although determining the optimal weights in the weighted average ensemble method is a challenging optimization job, the weighted average ensemble technique has yielded encouraging results.[29] In 2022, an optimization method named Dwarf Mongoose Optimization Algorithm (DMOA) optimization has been developed, outperforming other meta-heuristic algorithms on classical benchmarks.[30] The DMO utilized the foraging behavior of dwarf mongoose to explore and exploit the search space. The algorithm is applied successfully in different domains.[31,32] However, neither feature selection nor weighted average ensemble learning has yet to be carried out with this optimization algorithm.

Evaluating and comparing proposed techniques using benchmarking methods is crucial. However, due to conflicting evaluation parameters, it is not easy to select the best models. Multi-Criteria Decision-Making (MCDM) techniques have proven effective in resolving this issue by ranking models based on several performance indicators. A half-quadratic ensemble technique using the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) approach has been introduced for ranking machine learning models in the diagnosis of thyroid cancer.[33] The performance of the ranking model is greatly impacted by the weights assigned to each criterion. Zhang *et al.*[34] analysed and compared the recently proposed Distance Correlation Criteria Importance Through Intercriteria Correlation (D-CRITIC) weighting technique for different engineering problems. However, there is no proof in the literature that distance correlation CRITIC-TOPSIS being used for ranking and selection of the best CAD model in the Heart sound classification task.

Materials, Methods, and Theoretical Overview

Materials

The two independent Heart sound datasets are used to assess the proposed framework: PhysioNet/CinC Challenge 2016 Database and PASCAL CHSC Database.

PhysioNet/CinC Challenge 2016 Database

The PhysioNet/Computing in Cardiology (CinC) Challenge 2016 dataset[35] consolidates nine independent datasets. With a total of 2,435 heart sound recordings from 1,297 patients and PCG signal durations ranging from 8 to 312.5 seconds, the data is split into training and testing sets. We have downsampled each PCG signal to 2000 Hz because the data was gathered with different devices and varied sampling rates. Aortic, pulmonary, tricuspid, and mitral auscultation points are used to collect the heart sounds of both healthy and pathological participants. Heart valve and coronary artery ailments are among the health conditions possessed by the pathological subjects. There are more normal samples in the databases than abnormal ones, causing a class imbalance problem. The recordings are categorized as normal, abnormal, and ambiguous and investigated in clinical and non-clinical scenarios. An automated segmentation approach was used to identify areas of the fundamental heart sounds for all samples except the ambiguous class. The annotations are then manually studied and corrected. It incorporates 4,430 recordings from 1,072 persons, comprising 233,512 heart sounds from individuals in good health and with coronary artery disease and cardiac valve problems. The recordings are made using various equipment in clinical and non-clinical settings (such as patient premise visits), with a time period ranging from a few seconds to several minutes. The proposed research used a total of 22925 normal and 6812 abnormal cardiac heart sound for the classification task.

PASCAL CHSC Database

Two datasets contribute to the PASCAL CHSC database.[36] The heart sounds in Dataset A are recorded at the apex point of individuals using a smartphone application, and the population size is unknown for for this data retrieval. A total of 656 heart sound recordings are obtained from volunteer patients in Dataset B, collected at the Real Hospital Portugues (RHP) in Recife, Brazil, using a digital stethoscope. The PCG signals are captured at 4000 Hz with durations ranging from 1 to 30 seconds. Sounds in Dataset B were gathered from four cardiac auscultation regions on healthy and sick children. Although the details of the specific auscultation areas are missing. Recordings were made in both clinical and non-clinical settings. Dataset A sounds are categorized into normal, murmur, extra heart sound, and artifact categories, while recordings in Dataset B are classified into normal, systole, and murmur categories. In Dataset B, clinicians carefully categorized the fundamental heart sounds. A total of 892 normal and 552 abnormal cardiac heart sound are used in the proposed research for binary classification.

Methodology

To eliminate ripples from the waveform, each standard heart sound signal is first processed using a bandpass filter, whose frequency varies from 50 Hz to 600 Hz, as shown in Fig. 1. A third-order Butterworth filter is used for effective ripple removal. Random noise caused by muscle action, which mixes with heart sound signal, is addressed through zero-phase filtering to cleanse the original signal from contamination. The DWT method is subsequently applied to the waveform, followed by appropriate thresholding of the output. The Daubechies 13 wavelet, with eight decompositions, is then used to convert the data back to the time domain, resulting in the denoised heart sound signal. The proposed algorithms classify heart sounds based on the features, extracted from the pre-processed heart sound input. The non-linear Mel-scale-based short-term power coefficient of the cardiac sound signal is employed as the MFCC feature. The pre-processed heart sound file has a duration of 1325 seconds. The values of 0.5 and 0.125 seconds are selected for window size and window gap, respectively, for sampling purposes. The DWT is used to produce 240 features for each audio file. There are 16 coefficients for each window. A total of 30 windows are applied using the MFCC technique, resulting in the extraction of 480 features from each heart sound file. After DWT and MFCC features fusion, the number of features is 720. In this study, the MFCC and DWT features are extracted and used to benchmark

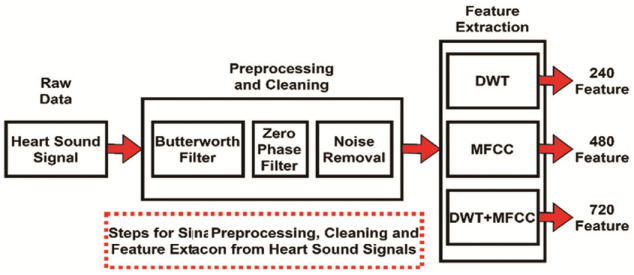

Fig. 1 — Procedure for feature extraction from heart sound signals

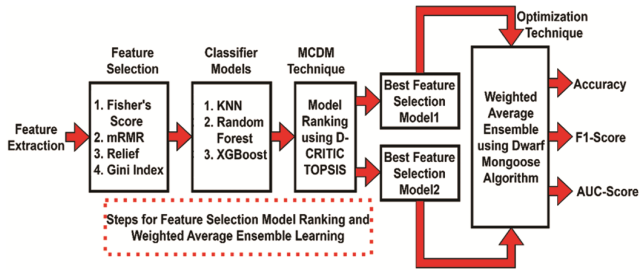

Fig. 2 — A machine learning pipeline for benchmarking the feature extraction, feature selection and ensemble learning models using metaheuristic and MCDM techniques

different models under investigation.[37,38] The procedure followed for the benchmarking of different strategies is illustrated in Fig. 2. The extracted features undergo feature selection strategies, including Fisher's Score, mRMR, ReliefF, and GINI Index. Fisher's Score measures the variance ratio between different classes to the variance within the same class. For each feature, it calculates the difference between the mean values across classes and the variance within each class, selecting features with higher scores as more significant for classification. Minimum Redundancy Maximum Relevance (mRMR) balances the importance of features with their redundancy by enhancing the mutual information between the selected features and the target variable and minimizing the mutual information among the selected features. ReliefF assesses the significance of features by analysing their capacity to differentiate between closely situated instances. ReliefF considers the context of features by accounting for feature interactions and can handle noisy and multiclass data effectively. Features with lower GINI impurity are preferred because they lead to more homogeneous subsets when used in splitting criteria, thus improving the overall predictive power.[39-42] The chosen features are utilized to train three distinct classifiers. The classifiers examined in this study include K-Nearest Neighbor (KNN), Random Forest (RF), and Gradient Boosting (XGBoost). DMOA is used for weighted average ensemble learning. This metaheuristic algorithm was proposed by Jeffrey et al. in 2022. DMOA utilized the foraging behavior of dwarf mongoose to explore and exploit the search space. The proposed algorithm employed three types of mongoose social groups (alpha group, babysitters, and scout group). The alpha females are responsible for initializing foraging, finding the path and distance travelled, and selecting sleeping mounds. The babysitters remain with the young mongoose and are exchanged only during the exploitation phase. The nomadic behavior of mongooses helps them to avoid overexploitation.

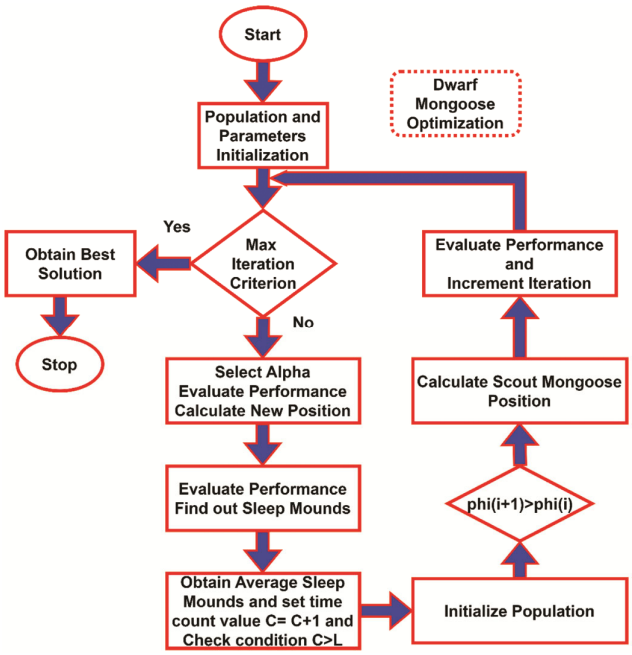

Fig. 3 — Dwarf mongoose algorithm flowchart

Optimization Steps

The flow chart of the optimization algorithm is demonstrated in Fig. 3. The optimization steps are as follows:

- Set up the algorithm parameters: Population Size (*N*), Number of Babysitters (*BS*), Maximum iterations (*MaxT*), Number of Control Variables (*D*), Lower Bound (*LB*), Upper Bound (UB), Babysitter Exchange Parameter (*L*), and Alpha Female Peep Parameter (*peep*). The peep is the vocal sound parameter utilized by the Alpha Female member to guide the mongoose correctly. Also, initialize a time counter (*C*) and increment it with each iteration.
- Evaluate the fitness function of the entire population and select the Alpha Females using Eq. (1).

$$Female\ Alpha\ (\alpha) = \frac{fitness_i}{\sum_{i=1}^{N} fitness_i} \quad \ldots (1)$$

Here, $fitness_i$ is cost value of individual search agent.

- Calculate the new food position using the formula given in Eq. (2).

$$X_{i+1} = X_i + p * peep \quad \ldots (2)$$

X_{i+1} is the new food position whereas X_i is the old food position.

$$p = \frac{peep}{2} * r_1 * w_1 \quad \ldots (3)$$

$$w_1 = \exp\left(-4 * \left(\frac{Current_{iter}}{MaxT}\right)^2\right) \quad \ldots (4)$$

r_1 is a random number, the value lies between [−1, 1], and $Current_{iter}$ is the current iteration

number. Now calculate the new fitness value for the new food positions.

- Obtain the value of the new Sleeping Mound ($SleepingMound_i$) and the average value of the sleeping mound (phi_i) using Eqs (5)- (6).

$$SleepingMound_i = \frac{fitness_{i+1} - fitness_i}{Max[|fitness_{i+1}, fitness_i|]} \quad \ldots (5)$$

$$phi_i = \frac{\sum_i^n SleepingMound_i}{n} \quad \ldots (6)$$

Here, n is the total number of Sleeping Mounds.

- Compare the time counter (C) with the Babysitter Exchange parameter (L) and if $C > L$, then exchange the babysitter with Alpha Females and set the time counter value C = 0. Initialize the population, update the average sleeping mound values, and obtain the Scout Mongoose positions using Eq. (7).

$$X_{i+1} = \begin{Bmatrix} X_i - GF[X_i - N]\ if\ phi_{i+1} > phi_i \\ X_i + GF[X_i - N]\ else \end{Bmatrix} \quad \ldots (7)$$

$$GF = \left(1 - \frac{Current_{iter}}{MaxT}\right)^{2*\frac{Current_{iter}}{MaxT}} \quad \ldots (8)$$

$$N = \sum_{i=1}^{n} \frac{X_i * SleepingMound_i}{X_i} \quad \ldots (9)$$

GF is the dynamic movement and control collective mongoose movement and N is the movement of mongoose to the sleeping mounds.

- Assess the fitness of each search agent and proceed to the next iteration. Identify the best solution and verify if the termination criteria are met.

The distance correlation CRITIC-TOPSIS approach is used to determine the weights for ranking the models examined in this investigation.[43] The necessary steps for D-CRITIC weight calculation are shown in Fig. 4.

Weight Calculation using D-CRITIC

- Normalize the decision matrix using the formula given in Eq. (10).

$$\overline{b}_{ij} = \frac{b_{ij} - b_j^{worst}}{b_j^{best} - b_j^{worst}} \quad \ldots (10)$$

b_j^{best} , b_j^{worst} are the best and worst value of j^{th} criterion, respectively. For non-beneficial criterion

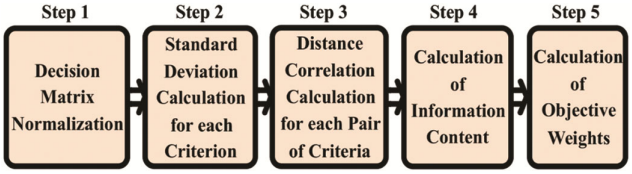

Fig. 4 — D-CRITIC weights calculation process[29]

minimum value is the best value and maximum value is the worst value whereas for beneficial criterion maximum value is the best value and minimum value is the worst value.

- For each criterion standard deviation values are computing using the formula given in Eq. (11).

$$\Sigma_j = \sqrt{\frac{\sum_{i=1}^{n}(b_{ij} - \overline{b_j})^2}{n-1}} \quad \ldots (11)$$

$\overline{b_j}$ is the mean value of the criteria, and n is the total number of alternatives.

- Calculate the distance correlation of every pair of criteria. The main difference between the D-CRITIC and the original CRITIC is that in D-CRITIC distance correlation is employed in place of Pearson correlation. The correlation between the two criteria is calculated using Eq. (12).

$$d\_Cor(\chi_j, \chi_j') = \frac{d\_Cov(\chi_j, \chi_j')}{\sqrt{d\_Var(\chi_j)\left(d\_Var(\chi_j')\right)}} \quad \ldots (12)$$

χ_j, χ_j' are two criteria, $d\_Cov(\chi_j, \chi_j'))$ is the distance covariance. $d\_Var(\chi_j), \left(d\_Var(\chi_j')\right)$ are the distance variance.

- Calculate Information content (Im_j):

$$Im_j = \sigma_j \sum_{j=1'}^{n} \left(1 - d\_Cor\left(\chi_j, \chi_{j'}\right)\right) \quad \ldots (13)$$

- Determination of objective weights (Wt_j):

$$Wt_j = \frac{Im_j}{\sum_{j=1}^{n} Im_j} \quad \ldots (14)$$

TOPSIS Score Calculation and Ranking of Alternatives

The necessary steps involved in TOPSIS method are:

- Calculate the normalized decision matrix.

$$\delta_{jk=} \frac{y_{jk}}{\sqrt{\sum_{j=1}^{m} y_{jk}^2}} \quad \ldots (15)$$

Here, y_{jk} is the term from decision matrix. The $j \in [1,2,3,...,m]$ and k ∈ [1,2,3,...,n], where m is the total number of alternatives arranged along the row and n

is the total number of criteria arranged along the column of decision matrix.

- Determine the weighted normalized decision matrix (Q):-

$$Q = \begin{bmatrix} w_1 t_{11} & w_2 t_{12} & \dots & w_n t_{1n} \\ w_1 t_{21} & w_2 t_{22} & \dots & w_n t_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_1 t_{m1} & w_2 t_{m2} & \dots & w_n t_{mn} \end{bmatrix} \quad \dots (16)$$

The weights are calculated from the D-CRITIC method using Eq. (17) with the constraint

$$\sum_{k=1}^{n} W t_{k=1} \quad \dots (17)$$

The Eq. (16) can be rewritten as:

$$Q = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{bmatrix} \quad \dots (18)$$

- The positive ideal solution (C^*) and negative ideal solution (C^-) for each attributes are determined using Eqs (19) & (21).

$$C^* = \{[max_j g_{jk} | k \in K']\}; j = 1,2, \dots, n \quad \dots (19)$$

$$C^* = \{Q_1^*, Q_2^*, \dots, Q_j^*, \dots, Q_n^*\} \quad \dots (20)$$

$$C^- = \{[max_j g_{jk} | k \in K']\}; j = 1,2, \dots, m \quad \dots (21)$$

$$C^- = \{Q_1^-, Q_2^-, \dots, Q_j^-, \dots, Q_n^-\} \quad \dots (22)$$

- Compute the Euclidean distance from the positive ideal solution (R_j^+) and negative ideal solution (R_j^-) for all alternative.

$$R_j^+ = \sqrt{\sum_{k=1}^{n} (Q_{jk} - Q_j^+)^2}; j = 1,2, \dots, m \quad \dots (23)$$

$$R_j^- = \sqrt{\sum_{k=1}^{n} (Q_j k - Q_j^-)^2}; j = 1,2, \dots, m \quad \dots (24)$$

The TOPSIS Score is given by Eq. (25).

$$C_j^* = \frac{r_j^-}{r_j^* + r_j^-} \quad \dots (25)$$

- Rankings can be established by using the TOPSIS Score. The alternatives that obtain the highest TOPSIS scores will be assigned with higher rankings.

Results and Discussion

The experiments are performed on a Dell Precision 3630 workstation featuring an Intel Xeon E-2136 CPU running at 3.30GHz with 32 GB of RAM.

The performance of various models is assessed using three metrics: Accuracy, F1-score, and AUC-ROC score. Initially, the performance of three classifiers (KNN, RF, and XGBoost) is compared across three feature extraction techniques, resulting in a total of nine models being evaluated. Performance is measured using 5-fold cross-validation with Fisher's score-based feature selection method. For the KNN model, the fine-tuned hyperparameters are configured as follows: 17 neighbors, uniformly distributed weights, and the BallTree method for computing neighbor positions. For RF classifier, the hyperparameters include 120 trees, the Gini index as the evaluation criterion, a maximum depth until all leaves contain fewer than the minimum split samples, a minimum sample split of 2, and balanced class weights. In XGBoost model, the hyperparameters are set to log loss, a learning rate of 0.01, 83 decision trees, Friedman mean squared error for better approximation and a minimum sample split value of 2. In the first stage, 50 features based on Fisher's score were evaluated. The resulting metrics for various feature extraction and classifiers using PhysioNet dataset are presented in Table 1.

Similarly, the resulting metrics for various feature extraction and classifiers using PASCAL CHSC dataset are presented in Table 2. The RF classifier provided the best accuracy, F1-score, and AUC-ROC score. The DWT and MFCC feature extraction models provided accuracies of 80.50% and 77.00%, respectively, while the DWT+MFCC feature fusion model attained an accuracy of 80.57% for the PhysioNet Dataset. The highest F1-score and AUC-ROC values were 0.8114 and 0.8928, respectively. Similarly, for the PASCAL CHSC dataset, the RF classifier achieved the highest accuracies of 77.63%, 72.19%, and 77.81% with DWT, MFCC, and DWT+MFCC-based extracted features, respectively. The best F1-score and AUC-ROC scores obtained were 0.7669 and 0.8523, respectively.

To compare the models, D-CRITIC weights were calculated for both datasets, as presented in Table 3. The different criteria have varying impacts on decision-making, with the F1 score receiving the highest weighting among the three performance evaluation criteria.

In Table 4, the calculated distance positive, distance negative, and TOPSIS score values are

Table 1 — Performance evaluation of different feature extraction and classification techniques using PhysioNet Dataset

Model	Feature Extraction	Classifiers	Accuracy	F1-Score	AUC-ROC
Model1	DWT	KNN	0.6900	0.6943	0.7585
Model2	DWT	XGBoost	0.8050	0.8093	0.8784
Model3	DWT	RF	0.8050	0.8100	0.8897
Model4	MFCC	KNN	0.6992	0.6891	0.7762
Model5	MFCC	XGBoost	0.7650	0.7715	0.8418
Model6	MFCC	RF	0.7700	0.7810	0.8602
Model7	DWT+MFCC	KNN	0.6656	0.6372	0.7289
Model8	DWT+MFCC	XGBoost	0.8010	0.8057	0.8781
Model9	DWT+MFCC	RF	0.8057	0.8114	0.8928

Table 2 — Performance evaluation of different feature extraction and classification techniques using PASCAL CHSC Dataset

Model	Feature Extraction	Classifiers	Accuracy	F1-Score	AUC-ROC
Model1	DWT	KNN	0.7056	0.6590	0.7693
Model2	DWT	XGBoost	0.7763	0.7672	0.8513
Model3	DWT	RF	0.7763	0.7663	0.8446
Model4	MFCC	KNN	0.6612	0.6297	0.7086
Model5	MFCC	XGBoost	0.7183	0.7091	0.7754
Model6	MFCC	RF	0.7219	0.7196	0.7830
Model7	DWT+MFCC	KNN	0.7165	0.6695	0.7829
Model8	DWT+MFCC	XGBoost	0.7609	0.7555	0.8431
Model9	DWT+MFCC	RF	0.7781	0.7669	0.8523

Table 3 — Calculated D-CRITIC weights to obtain the best feature extraction and classification techniques

Criterion	Accuracy	F1-Score	AUC-ROC
D-CRITIC Weights (PhysioNet Dataset)	0.3151	0.3729	0.3120
D-CRITIC Weights (PASCAL CHSC Dataset)	0.2118	0.4957	0.2922

Table 4 — TOPSIS score and Ranks evaluation for feature extraction and classification techniques

Model	Dataset	Distance Positive	Distance Negative	TOPSIS Score	Ranks
Model1	PhysioNet	0.0265	0.0088	0.2496	8
Model2	PhysioNet	0.0023	0.0328	0.9335	3
Model3	PhysioNet	0.0005	0.0343	0.9847	2
Model4	PhysioNet	0.0246	0.0102	0.2925	7
Model5	PhysioNet	0.0097	0.0252	0.7215	6
Model6	PhysioNet	0.0066	0.0282	0.8111	5
Model7	PhysioNet	0.0348	0.0000	0.0000	9
Model8	PhysioNet	0.0025	0.0325	0.9289	4
Model9	PhysioNet	0.0000	0.0348	1.0000	1
Model1	PASCAL CHSC	0.0196	0.0116	0.3724	8
Model2	PASCAL CHSC	0.0002	0.0305	0.9940	2
Model3	PASCAL CHSC	0.0014	0.0294	0.9540	3
Model4	PASCAL CHSC	0.0307	0.0000	0.0000	9
Model5	PASCAL CHSC	0.0156	0.0152	0.4943	6
Model6	PASCAL CHSC	0.0138	0.0171	0.5525	5
Model7	PASCAL CHSC	0.0169	0.0144	0.4593	7
Model8	PASCAL CHSC	0.0022	0.0286	0.9299	4
Model9	PASCAL CHSC	0.0000	0.0307	0.9989	1

presented along with the final ranks. It is shown that for both datasets, Model9 achieved the highest TOPSIS scores and ranks. Model9 consists of DWT+MFCC features combined with the RF classifier.

In the next stage, DWT+MFCC fused features along with RF are kept fixed, and the performance of four feature selection methods (Fisher's Scores, mRMR, ReliefF, GINI Index) is compared on both datasets, as shown in Table 5. For the PhysioNet and PASCAL

Table 5 — Performance evaluation of different Filter-based feature selection methods

Model	Dataset	Feature Selection	Accuracy	F1-Score	AUC-ROC
Model9	PhysioNet	Fisher's Score	0.8057	0.8114	0.8928
Model10	PhysioNet	mRMR	0.8270	0.8369	0.9092
Model11	PhysioNet	ReliefF	0.8235	0.8352	0.9093
Model12	PhysioNet	GINI Index	0.7945	0.8002	0.8816
Model9	PASCAL CHSC	Fisher's Score	0.7781	0.7669	0.8523
Model10	PASCAL CHSC	mRMR	0.7789	0.7670	0.8532
Model11	PASCAL CHSC	ReliefF	0.7782	0.7669	0.8528
Model12	PASCAL CHSC	GINI Index	0.7608	0.7519	0.8421

Table 6 — TOPSIS score and Ranks evaluation for Filter-based feature selection techniques

Model	Dataset	Distance Positive	Distance Negative	TOPSIS Score	Ranks
Model1	PhysioNet	0.0060	0.0031	0.3387	3
Model2	PhysioNet	0.0000	0.0090	0.9980	1
Model3	PhysioNet	0.0003	0.0088	0.9621	2
Model4	PhysioNet	0.0091	0.0000	0.0000	4
Model1	PASCAL CHSC	0.0002	0.0053	0.9627	3
Model2	PASCAL CHSC	0.0000	0.0054	1.0000	1
Model3	PASCAL CHSC	0.0002	0.0053	0.9642	2
Model4	PASCAL CHSC	0.0051	0.0000	0.0000	4

Table 7 — Performance evaluation of random forest classifier using data-spilt, feature selection and feature selection based weighted ensemble technique

Model	Dataset	Data split	Feature Selection/ Feature Selection based Weighted Ensemble	Accuracy	F1-Score	AUC-ROC
Model10	PhysioNet	5 Fold	mRMR	0.8270	0.8369	0.9092
Model11	PhysioNet	5 Fold	ReliefF	0.8235	0.8352	0.9093
Model13	PhysioNet	5 Fold	mRMR+ ReliefF	0.8436	0.8528	0.9243
Model10	PhysioNet	70:30	mRMR	0.8255	0.8312	0.9121
Model11	PhysioNet	70:30	ReliefF	0.8222	0.8313	0.9019
Model13	PhysioNet	70:30	mRMR+ ReliefF	0.8425	0.8514	0.9213
Model10	PhysioNet	80:20	mRMR	0.8275	0.8366	0.9177
Model11	PhysioNet	80:20	ReliefF	0.8209	0.8280	0.9087
Model13	PhysioNet	80:20	mRMR+ ReliefF	0.8466	0.8553	0.9259
Model10	PASCAL CHSC	5 Fold	mRMR	0.7789	0.7670	0.8532
Model11	PASCAL CHSC	5 Fold	ReliefF	0.7782	0.7669	0.8523
Model13	PASCAL CHSC	5 Fold	mRMR+ ReliefF	0.8052	0.7918	0.8631
Model10	PASCAL CHSC	70:30	mRMR	0.7711	0.7581	0.8137
Model11	PASCAL CHSC	70:30	ReliefF	0.7681	0.7586	0.8154
Model13	PASCAL CHSC	80:20	mRMR+ ReliefF	0.7891	0.7742	0.8303
Model10	PASCAL CHSC	80:20	mRMR	0.7964	0.7526	0.8116
Model11	PASCAL CHSC	80:20	ReliefF	0.7420	0.7373	0.8003
Model13	PASCAL CHSC	80:20	mRMR+ ReliefF	0.8054	0.7942	0.8278

CHSC datasets, mRMR feature selection attained the highest accuracy of 82.70% and 77.89%, respectively. The mRMR feature selection provided F1-scores of 0.8369 and 0.7670 for the PhysioNet and PASCAL CHSC datasets, respectively. The performance is compared using TOPSIS scores presented in Table 6. The two top-ranked models for each dataset are selected and evaluated using three data split strategies. Model10 has achieved the best accuracies of 82.70%, 82.55%, and 82.75% for 5-fold, 70:30, and 80:20 data splits on the PhysioNet dataset, respectively. On the PASCAL CHSC dataset, accuracies of 77.89%, 77.11%, and 79.64% are achieved for the 5-fold, 70:30, and 80:20 data splits, respectively. The performance metrics of random forest classifier for various data splits, individual feature selection, as well as feature selection based weighted average ensemble, are presented in Table 7. A weighted average ensemble of the two top-ranked models is performed using the DMOA algorithm. The algorithm is run with a population size of 50, 3

Table 8 — Optimized weights for different split strategies in weighted ensemble model

Dataset	Data Split	W1	W2
PhysioNet	5 Fold	0.5435	0.4565
PhysioNet	70:30	0.4831	0.5169
PhysioNet	80:20	0.6977	0.3023
PASCAL CHSC	5 Fold	0.6035	0.3965
PASCAL CHSC	70:30	0.7383	0.2617
PASCAL CHSC	80:20	0.6250	0.3750

babysitters, and 2 peeps for 50 iterations, achieving the highest accuracies. The Model13 has provided accuracies of 84.36%, 84.25%, and 84.66% for 5-fold, 70:30, and 80:20 splits on the PhysioNet dataset. Similarly, for the PASCAL CHSC dataset, this model has provided accuracies of 80.52%, 78.91%, and 80.54% for the 5-fold, 70:30, and 80:20 splits.

The optimization algorithm runs 20 times and the best weights are provided in Table 8. The results indicate that in the feature-based ensemble model, the mRMR feature selection is provided with more decision weightage than the ReliefF method for the PASCAL CHSC dataset. In contrast, for the PhysioNet dataset, the mRMR-based model demonstrates a greater influence on the decision-making process for the 5-fold and 80:20 splits. This trend is reversed for the 70:30 split, and more decision weightage is provided to the ReliefF feature selection technique. The AUC-ROC curves (5 Fold split) for mRMR, RefiefF, and ensemble model are shown in Fig. 5(a) for PhysioNet, whereas Fig. 5(b) provided the same plotting curves for PASCAL CHSC dataset. The plots show the ensemble model's superior classification capabilities over the individual model for different threshold values. The highest AUC-ROC score obtained for the PhysioNet dataset is 0.9243, which is 1.65% higher than the best individual model. For PASCAL CHSC, the best AUC-ROC score value achieved is 0.8278, which is 2.0% better than the base model. The ensemble model provided the best curve, which demonstrates the model's superior classification capabilities.

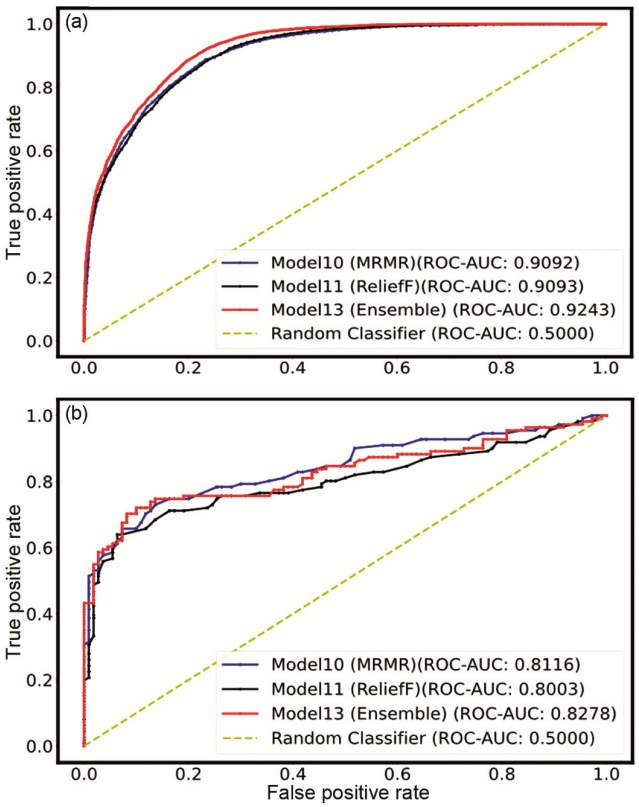

Fig. 5(a) — AUC-ROC plot for PhysioNet dataset, and (b) AUC-ROC plot for PASCAL CHSC Dataset

Conclusions

From the results and discussion, it is evident that feature extraction, feature selection, and classifier choice play crucial roles in enhancing the efficiency of AI-assisted cardiac disorder detection. This study demonstrates the application of the meta-heuristic algorithm for weighted average ensemble learning to improve classification performance. The weighted ensemble model using the Dwarf Mongoose Optimization Algorithm achieved the highest performance. However, a drawback of ensemble models is their large size and the difficulty of deploying them on edge devices due to high computational power requirements. The time and space complexities of the model are not evaluated in this study, indicating a future scope of work. The evaluated feature selection techniques can be utilized in combination with other neuromorphic spiking classification algorithms. Additionally, other metaheuristic algorithms can be explored for weighted average ensembles as well as feature selection for different disorder detection tasks. The suggested framework is versatile and applicable to various machine-learning tasks. The proposed framework can be utilized by the health practitioner during cardiac disorder detection, which may reduce excessive burden and help in decision-making.

References

1 Kazi D S, Elkind M S, Deutsch A, Dowd W N, Heidenreich P, Khavjou O, Mark D, Mussolino M E, Ovbiagele B & Patel S, Forecasting the economic burden of cardiovascular disease and stroke in the united states through 2050: a presidential advisory from the american heart association, *Circulation*, **150(4)** (2024) 1524–4539, https://doi.org/10.1161/ cir.0000000000001258.

2 Abhisheka B, Biswas S K, Purkayastha B, Das D & Escargueil A, Recent trend in medical imaging modalities and their applications in disease diagnosis: a review, *Multimed Tools Appl*, **83** (2023) 43035–43070, http://dx.doi.org/10.1007/ s11042-023-17326-1.

3 Middleton H N, Phonocardiographic studies on heart diseases, *J Natl Med Assoc*, **35** (1943) 119.

4 Ergen B, Tatar Y & Gulcur H O, Time–frequency analysis of phonocardiogram signals using wavelet transform: a comparative study, *Comput Methods Biomech Biomed Eng*, **15(4)** (2012) 371–381, http://dx.doi.org/10.1080/10255842. 2010.538386.

5 Theng D & Bhoyar K K, Feature selection techniques for machine learning: a survey of more than two decades of research, *Knowl Inf Syst*, **66** (2023) 1575–1637, http://dx.doi.org/10.1007/s10115-023-02010-5.

6 Lyu Y, Feng Y & Sakurai K, A survey on feature selection techniques based on filtering methods for cyber attack detection, *Information*, **14(3)** (2023), http://dx.doi.org/10.3390/ info14030191.

7 An Q, Rahman S, Zhou J & Kang J J, A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges, *Sensors*, **23(9)** (2023), http://dx.doi.org/10.3390/s23094178.

8 Sharma R, Mahanti G K, Chakraborty C, Panda G & Rath A, An IoT and deep learning-based smart healthcare framework for thyroid cancer detection, *ACM Trans Internet Technol*, (2023), http://dx.doi.org/10.1145/3637062.

9 Garg A & Mago V, Role of machine learning in medical research: A survey, *Comput Sci Rev*, **40**(2021) 100370, http://dx.doi.org/10.1016/j.cosrev.2021.100370.

10 Singh S A, Singh S A & Singh A D, Enhancing imbalanced heart sound classification through transfer learning and gammatonegram image analysis, (in press), http://dx.doi.org/10.21203/rs.3.rs-3530451/v1.

11 Chen Y, Sun Y, Lv J, Jia B & Huang X, End-to-end heart sound segmentation using deep convolutional recurrent network, *Complex Intell Syst*, **7** (2021) 2103–17, http://dx.doi.org/10.1007/s40747-021-00325-w.

12 Zeng W, Yuan J, Yuan C, Wang Q, Liu F & Wang Y, A new approach for the detection of abnormal heart sound signals using TQWT, VMD and neural networks, *Artif Intell Rev*, **54** (2020) 1613–47, http://dx.doi.org/10.1007/s10462-020-09875-w.

13 Almanifi O R A, Ab Nasir A F, Mohd Razman M A, Musa R M & Majeed Anwar P P Abdul, Heartbeat murmurs detection in phonocardiogram recordings via transfer learning, *Alexandria Eng J*, **61(12)** (2022) 10995–1002,http://dx.doi.org/ 10.1016/j.aej.2022.04.031.

14 Li F, Zhang Z, Wang L & Liu W, Heart sound classification based on improved mel-frequency spectral coefficients and deep residual learning, *Front Physiol*, **13** (2022), http://dx.doi.org/10.3389/fphys.2022.1084420.

15 Chowdhury S, Morshed M & Fattah S A, Spectrocardionet: an attention-based deep learning network using triple-spectrograms of pcg signal for heart valve disease detection, *IEEE Sens J,* **22(23)** (2022) 22799–22807. http://dx.doi.org/ 10.1109/jsen.2022.3196263.

16 Oh S L, Jahmunah V, Ooi C P, Tan R S, Ciaccio E J & Yamakawa T, Classification of heart sound signals using a novel deep WaveNet model, *Comput Methods Programs Biomed,* **196** (2020) 105604, http://dx.doi.org/10.1016/ j.cmpb.2020.105604.

17 Arora V, Leekha R, Singh R & Chana I, Heart sound classification using machine learning and phonocardiogram, *Mod Phys Lett B*, **20** (2019) 1950321, http://dx.doi.org/ 10.1142/s0217984919503214.

18 Al-Naami B, Fraihat H, Gharaibeh NY & Al-Hinnawi ARM, A framework classification of heart sound signals in physionet challenge 2016 using high order statistics and adaptive neuro-fuzzy inference system, *IEEE Access*, **8** (2020) 224852–224859, http://dx.doi.org/10.1109/access. 2020.3043290.

19 Krishnan P T, Balasubramanian P & Umapathy S, Automated heart sound classification system from unsegmented phonocardiogram (PCG) using deep neural network, *Phys Eng Sci Med*, **43** (2020) 505–515, http://dx.doi.org/10.1007/s13246-020-00851-w.

20 Noman F, Ting C M, Salleh S H & Ombao H, Short-segment heart sound classification using an ensemble of deep convolutional neural networks, *IEEE Int Conf Acoust, Speech Signal Process (ICASSP)*, (2019), http://dx.doi.org/10.1109/icassp.2019.8682668.

21 Sánchez-Maroño N, Alonso-Betanzos A & Tombilla-Sanromán M, Filter methods for feature selection – a comparative study, *Intell Data Eng Automat Learn (IDEAL),* (2007) 178–87, http://dx.doi.org/10.1007/978-3-540-77226-2_19.

22 Venkatesh B & Anuradha J, A review of feature selection and its methods, *Cybernet Informat Technol*, **19** (2019) 3–26, http://dx.doi.org/10.2478/cait-2019-0001.

23 Kaur S, Kumar Y, Koul A & Kumar Kamboj S, A systematic review on metaheuristic optimization techniques for feature selections in disease diagnosis: open issues and challenges, *Arch Comput Methods Eng*, **30** (2022) 1863–95, http://dx.doi.org/10.1007/s11831-022-09853-1.

24 Liu H, Lu L, Xiong H, Fan C, Fan L & Lin Z, A novel approach to dual feature selection of atrial fibrillation based on hc-mfs, *Diagnostics*, **14(11)** (2024) 1145, http://dx.doi.org/ 10.3390/diagnostics14111145.

25 Sharma R, Mahanti G K, Panda G, Rath A, Dash S, Mallik S & Ruifeng H, A framework for detecting thyroid cancer from ultrasound and histopathological images using deep learning, meta-heuristics, and MCDM algorithms, *J Imag*, **9(9)** (2023) 173, http://dx.doi.org/10.3390/jimaging9090173.

26 Lee J, Yoon Y, Kim J & Kim Y H, Metaheuristic-based feature selection methods for diagnosing sarcopenia with machine learning algorithms, *Biomimetics*, **9(3)** (2024) 179, http://dx.doi.org/10.3390/biomimetics9030179.2

27 Yab L Y, Wahid N & Hamid R A, A meta-analysis survey on the usage of meta-heuristic algorithms for feature selection on high-dimensional datasets, *IEEE Access*, **10** (2022) 122832–122856, http://dx.doi.org/10.1109/ access.2022.3221194.

28 Dietterich T G, Ensemble methods in machine learning, *Multi Class Syst*, (2000) 1–15, http://dx.doi.org/ 10.1007/3-540-45014-9_1.

29 Sharma R, Kumar Mahanti G, Panda G & Singh A, Thyroid nodules classification using weighted average ensemble and dcritic based topsis methods for ultrasound images, *Curr Med Imag*, **16(20)** (2023), http://dx.doi.org/10.2174/ 1573405620666230405085358.

30 Agushaka J O, Ezugwu A E & Abualigah L, Dwarf mongoose optimization algorithm, *Comput Methods Appl Mech Eng*, **391** (2022) 114570, http://dx.doi.org/10.1016/ j.cma.2022.114570.

31 Moustafa G, El-Rifaie A M, Smaili I H, Ginidi A, Shaheen AM, Youssef AF & T Mohamed, An enhanced dwarf mongoose optimization algorithm for solving engineering problems, *Mathematics*, **11(15)** (2023) 3297, http://dx.doi.org/10.3390/ math11153297.

32 Malathi S R & Kumar P V, Multi-head self-attention-based recurrent neural network with dwarf mongoose optimization algorithm-espoused QRS detector design, *Signal, Image Video Process*, **18** (2024) 4935–44, http://dx.doi.org/10.1007/ s11760-024-03145-w.

33 Sharma R, Mahanti G K, Panda G, Rath A, Dash S, Mallik S & Zhongming Z, Comparative performance analysis of binary variants of FOX optimization algorithm with half-quadratic ensemble ranking method for thyroid cancer detection, *Sci Rep*, **13** (2023), http://dx.doi.org/10.1038/ s41598-023-46865-8.

34 Zhang H & Wei G, Location selection of electric vehicles charging stations by using the spherical fuzzy CPT–CoCoSo and D-CRITIC method, *Comput Appl Math*, **42** (2023), http://dx.doi.org/10.1007/s40314-022-02183-9.

35 Clifford G, Liu C, Springer D, Moody B, Li Q & Mark R, Classification of normal/abnormal heart sound recordings: the physionet/computing in cardiology challenge 2016, *Comput Cardiol Conf (CinC)*, **14** (2016), http://dx.doi.org/10.22489/cinc.2016.179-154.

36 Gomes E F, Bentley P J, Pereira E, Coimbra M T & Deng Y, Classifying heart sounds - approaches to the pascal challenge, *Proc Int Conf Health Informat*, 2013, http://dx.doi.org/10.5220/ 0004234403370340

37 Deng M, Meng T, Cao J, Wang S, Zhang J & Fan H, Heart sound classification based on improved MFCC features and convolutional recurrent neural networks, *Neural Network*, **130** (2020) 22–32, http://dx.doi.org/10.1016 /j.neunet.2020.06.015.

38 Rath A, Mishra D, Panda G & Satapathy S C, Heart disease detection using deep learning methods from imbalanced ECG samples, *Biomed Signal Process Control*, **68** (2021) 102820, http://dx.doi.org/10.1016/ j.bspc.2021.102820.

39 Gu Q, Li Z & Han J, Correlated multi-label feature selection, *Proc 20th ACM Int Conf Info Know Manag*, (2011) 1087–1096, http://dx.doi.org/10.1145/2063576. 2063734.

40 Jo I, Lee S & Oh S, Improved measures of redundancy and relevance for mrmr feature selection, *Computers*, **8(2)** (2019) 42, http://dx.doi.org/10.3390/computers8020042.

41 Urbanowicz R J, Meeker M, La Cava W, Olson R S & Moore JH, Relief-based feature selection: introduction and review, *J Biomed Inf*, **85** (2018) 189–203. http://dx.doi.org/10.1016/j.jbi.2018.07.014.

42 Manek A S, Shenoy P D, Mohan M C & Rao V K, Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier, *World Wide Web*, **20** (2016) 135–154, http://dx.doi.org/10.1007/s11280-015-0381-x.

43 Nguyen H Q, Nguyen V T, Phan D P, Tran Q H & Vu N P, Multi-criteria decision making in the PMEDM process by using MARCOS, TOPSIS, and MAIRCA methods, *Appl Sci*, **12** (2022) 3720, http://dx.doi.org/10.3390/ app12083720.