Received 20 January 2024, accepted 12 February 2024, date of publication 19 February 2024, date of current version 27 February 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3367597

RESEARCH ARTICLE

Cross Subject Myocardial Infarction Detection From Vectorcardiogram Signals Using Binary Harry Hawks Feature Selection and Ensemble Classifiers

M. KRISHNA CHAITANYA AND LAKHAN DEV SHARMA, (Senior Member, IEEE)
School of Electronics Engineering, VIT-AP University, Amaravati 522237, India
Corresponding author: Lakhan Dev Sharma (devsharmalakhan@gmail.com)

ABSTRACT Myocardial infarction (MI), widely referred to as a heart attack, is a leading reason for deaths worldwide. It is frequently caused by coronary artery occlusion, resulting in inadequate oxygen and blood supply, which damages the myocardial structure and function. Therefore, innovative diagnostic methods are required for reliable and timely identification of MI. The typical 12-lead electrocardiogram (ECG) technology causes patient discomfort and makes cardiac monitoring challenging. The frontal, sagittal, and transverse planes (3 orthogonal planes) are where vectorcardiogram (VCG) renders an edge over 12-lead ECG. This study, proposes a method for detecting MI utilising VCG signals of four seconds. Circulant singular spectrum analysis (CSSA) and four stage savitzky-golay (SG) filter were used in the filtering stage for the removal of power-line interference and base-line wander. The signal was time-invariantly decomposed using the CSSA, then features were extracted. The binary harry hawks-based feature selection method is employed on the extracted features to choose the optimal feature subspace which was followed by supervised machine learning based classification. The 10-fold cross validation, an even more practical leave-one-out (LOO) cross validation approach, and inter dataset cross validation (IDCV) were used to evaluate the reliability of the suggested method. Voting-based ensemble classification was used in LOO, IDCV validation, which improves the accuracy of this method. The proposed technique achieved an accuracy of 99.97%, 91.03%, and 99.41% for 10-fold, LOO cross validation, and IDCV, out-performing the state-of-the-art methods in the cross validation scenarios. The proposed technique results in an accurate detection of MI. Successful accomplishment of the LOO cross validation demonstrates the applicability and dependability of the suggested technique in the health care applications.

INDEX TERMS Vectorcardiography (VCG), myocardial infarction (MI), machine learning, binary Harry Hawks feature selection, ensemble classifier.

I. INTRODUCTION

Myocardial infarction (MI) is the sudden disruption of the heart's normal blood supply. If the normal flow of blood is not restored, then heart's myocardiocytes triggers the development of scar tissues [1]. Hence, Timely detection of MI results in improving patient survival. MI have been diagnosed using the morphological properties of a conventional 12-lead electrocardiogram (ECG) signal [2], [3]. The locations of the ECG electrodes cause significant variations in waveform patterns. Moreover 12-lead recording system could hinder patient's comfort because it makes cardiac monitoring more challenging with 13 electrodes placed on human body [4].

In comparison to 12-lead ECG, vectorcardiogram (VCG) improves the observation of electrical responses of cardio-vascular system from neck alongside spines. VCG is the possible replacement for the recording framework with fewer

The associate editor coordinating the review of this manuscript and approving it for publication was Emanuele Crisostomi.

electrodes and more insight [5]. VCG signals track the electrical activity of the heart along the frontal, transverse, and sagittal planes of the body, which are 3 orthogonal X, Y, and Z planes [6]. VCG is a valuable tool in assessing the heart's electrical axis, identifying conduction abnormalities, and understanding the spatial orientation of the heart's electrical events. Interpreting a VCG requires specialized knowledge and training in cardiology. Hence, automation is required for better understanding detailed electrical processes of the heart and reduce load on the healthcare professionals [7].

VCG demonstrates superior diagnostic abilities in the diagnosis of MI [5]. This is due to the fact that ECGs are unable to precisely record the spatial information contained in cardiac electrical activity [5]. A number of methodologies have been developed recently to identify MI utilising VCG signals [8], [9], [10], [11]. Most earlier MI diagnostic strategies including statistical analysis of the angles, magnitudes, and morphology of the QRS and T waves in the ECG alongside VCG signals was done in [12] and [13]. But these diagnostic techniques don't perform well, leaving scope of research in the field of VCG based MI detection [13].

Dehnavi et al. [8] used independent component analysis (ICA) and principal component analysis (PCA) to project the VCG signal feature vector into a lower-dimensional space. To create the feature vector, they have taken the different morphological characteristics from the VCG signal. The detection of MI through a lesser dimension feature vector from a VCG signal has been accomplished with a neural network-based classifier in [8]. Yang et al. [10], employed a decision tree (DTE) model to identify MI abnormalities and estimated octant as well as vector-based attributes based on VCG signals. The aforementioned techniques needs manual P, Q, R, S, and T-onset point identification in the VCG signal in order to determine the morphological features [14]. Yang et al. [11] used Discrete wavelet transform (DWT) to separate every single channel of the VCG signal into sub-band signals. Recurrent quantification analysis (RQA) dependent non-linear characteristics have been retrieved from each sub-band signal, and a Gaussian discriminant analysis (GDA) classification algorithm is employed to detect MI [11]. Tripathy et al. [14] employed dual-tree complex wavelet transform (DT-CWT) to break down the VCG signal into sub-band signals through each channel, from each sub-band entropy and L1-norm based features were extracted. These VCG signal features have been used by relevance vector machine (RVM) classifier to identify MI [14]. They have obtained sensitivity of 98.4%.

In aforementioned methods for the detection of MI using VCG have used K-fold cross validation. Cross subject or leave-one-out (LOO) method is a technique employs the subject for testing that has not been trained, which occur in real time scenario. Reasat and Shahnaz [15] used raw ECG signals that are fed to convolutional neural network (CNN) for the categorization of MI from ECG signals. They have also performed the LOO approach for the MI classification wih an accuracy of 84.54%. Liu et al. [16] developed multiple-feature-branch convolutional bidirectional recurrent neural network for the classification of MI employing 12-lead ECG using LOO and 10-fold validations. They also used lead random mask optimization technique to reduce overfitting and dropout. They have obtained specificity of 86.29% for the classification of MI. Kapfo et al. [17] decomposed the input ECG signal using variational mode decomposition (VMD) and employed support vector machine (SVM) for the classification of MI through the LOO cross validation. They have obtained an accuracy of 99.88%. Ensemble learning improves the generalisation ability and reliability by exploiting the diversity of classifiers.

Ensemble learning is a method for solving specific computational intelligence problems by strategically generating and combining a number of models, like classifiers. Improved classification is the main objective of ensemble learning. Desai et al. [18] employed bagging, random forest, rotation forest, and adaboost classifiers as an ensemble classifiers for the categorization of MI. They obtained a Matthews correlation coefficient of 0.9886. Khan et al. [19] extracted Mel-frequency Cepstral Coefficients from cardiovascular data for crucial feature extraction and performed classification using the Ensemble Subspace K Nearest Neighbor (ESSKNN) method. They achieved accuracy of 94.9%. Bashir et al. [20] proposed multi-objective weighted voting ensemble classifier for coronary heart disease prediction. It is based on an enhanced bagging approach. Instance based learner (IBL), quadratic discriminant analysis (QDA), naive bayes (NB), linear regression (LR), and support vector machine (SVM) are the five classifiers that make up the proposed ensemble classifier [21]. They have obtained an accuracy of 85.59%. Ravindranath [22] performed a comparative study of various existing clinical decision support systems along with the extended sub-tree approach that was suggested. The obtained accuracy is 80.17% using sub-tree method.

High dimensionality of dataset presents a significant challenge for machine learning [23], [24]. Feature selection techniques (FST) have been shown to improve the prediction of heart disease in the literature. Dun et al.'s [25] investigated the prevalence of heart disease using deep learning methods, random forests, logistic regression, and SVM with hyperparameter tuning and feature selection. The highest accuracy was achieved by neural networks (NN) was 78.3%. Generalized discriminant analysis (GDAS), a binary classifier, and an extreme learning machine (ELME) were used by Singh et al. [26] to reduce cardiovascular features. They boosted the accuracy of detecting coronary heart disease by 100%. Heart rate variability (HRV) was used to categorize arrhythmias by Yaghouby et al. [27]. They used the multilayer perceptron (MLP) NN classifier and the GDAS for feature reduction to achieve 100% accuracy. Initially 15 features were extracted from HRV signal using linear as well as

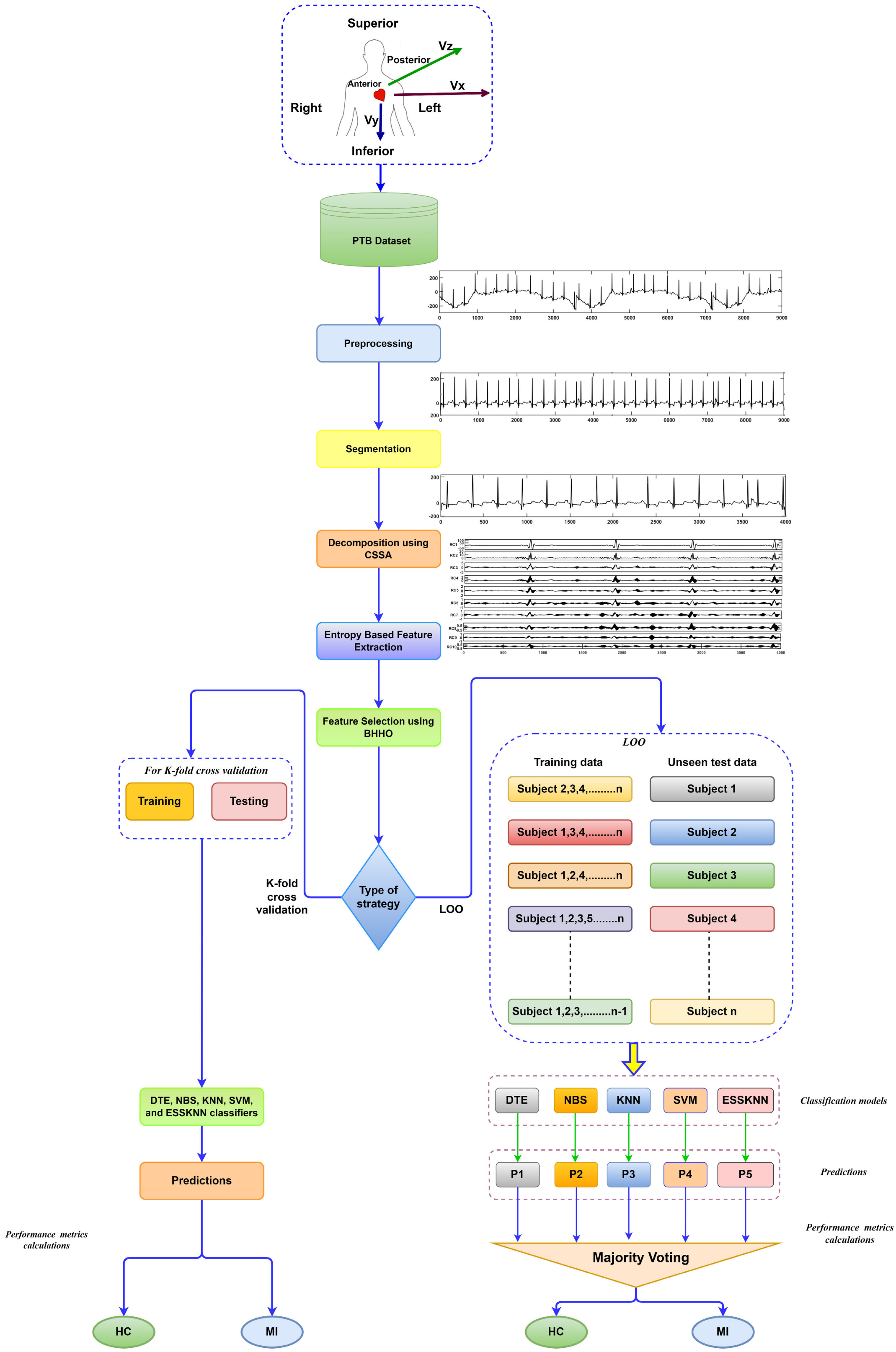


FIGURE 1. Block schematic of the proposed technique for the categorization of MI.

non-linear techniques. GDAS reduced the features to five and computed 100% precision with SVM [28]. The greedy search method takes a long time because it creates and examines every possible feature combination [29]. The random search method, meanwhile, randomly scans the search space in search of the following set of features. These techniques have a number of drawbacks, such as the potential to get stuck at a local optimal location as well as a high level of spatial and temporal complexity. The issues with the previously mentioned FST were addressed using meta-heuristic strategies.

There are several well-known metaheuristic methods with a high capacity for search space optimization, including the whale optimization algorithm (WOO) [30], ant lion optimization (ATLO) [31], salp swarm algorithm (SSWA) [32], grey wolf optimizer (GWFO) and dragonfly algorithm (DFA) [33]. Exploration and exploitation are the two fundamental phases of all metaheuristic methods, irrespective of inspiration algorithms. Heidari et al. developed the distinctive and effective swarm-based method known as the Harris Hawks Optimizer (HHO). To achieve excellent results in resolving a variety of mathematical and engineering problems, it employs two phases of exploration and four phases of exploitation. Therefore, when compared to other optimization methods, binary harris hawks optimizer (BHHO) is an effective metaheuristic feature selection technique for improving accuracy along with other performance metrics.

This manuscript describes a novel method for detecting MI utilising VCG leads using 10-fold cross validation, leave-one-out (LOO) cross validation, and inter dataset cross validation (IDCV) methods. The standard 10-fold cross validation is used for state-of-art comparison whereas the LOO and IDCV are used to test the model on untrained data. The goal of this study is to create a model for categorising MI that does not rely on time-domain fiducial indicators. The proposed approach makes use of shorter VCG segments for MI detection. Entropy based features, BHHO algorithms along with ensemble machine learning models were employed, and they were evaluated in a context that was more realistic approach.

The key contributions of the proposed work are listed below:

- To our knowledge this is the first work employing BHHO feature selection technique on the VCG signals.
- This is the first method to use circulant singular spectrum analysis (CSSA) decomposition on the VCG signals.
- We have performed both the 10-fold and LOO cross validation.
- This is the first work that uses ensemble classifiers for the classification of MI using VCG.
- This is the first work that uses IDCV cross validation for MI detection.

The combination of CSSA, BHHO feature selection, and ensemble classifiers has achieved best classification accuracies. Along with 10-fold cross validation the LOO and IDCV approach reflects the reliability of model in detecting an untrained subject which is a real time requirement in the health care industry. The rest of the article is divided into the following sections. The description of the materials is in Section II. Section III describes the suggested technique, Section IV presents the results, Section V is all about discussions, and Section VI describes a succinct conclusion.

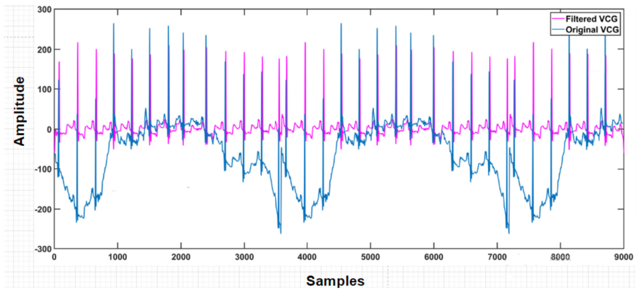


FIGURE 2. Outcomes of the preprocessing employed for MI detection.

II. MATERIALS

A. DATABASE USED

In this study, the Physikalisch-Technische Bundesanstalt Database (PTBDB) [34], [35], a standard to develop MI detection algorithms, is used. It consists of 52 healthy patients who took part in health control (HC) and 148 participants who were hospitalised due to various forms of MI. VCG records of 52 HC and 148 MI subjects were part of the proposed work.

III. METHODOLOGY

A. PROPOSED METHOD

The suggested method involves the following steps: preprocessing, decomposition, feature extraction, feature selection, and classification. The next subsections discuss each of them in extensive detail. Figure 1 illustrates the suggested technique's process flow.

B. PREPROCESSING AND SEGMENTATION

This step enhances the efficiency of the suggested strategy by eliminating the unwanted noise. Baseline drift is removed by decomposing the ECG signal from PTBDB using circulant singular spectrum analysis (CSSA), as described in [36]. Employing a 4 stage Savitzky-Golay (SG) smoothing filter, the signal is further refined in order to eliminate powerline interference. The resulting noise-free ECG is divided into segments with a 4-second duration. As a result 39960 MI segments and 14040 HC segments are produced. The graphical representation of raw and filtered VCG signal are depicted in the Figure 2.

C. DECOMPOSITION

1) SINGULAR SPECTRUM ANALYSIS (SSA)

One dimensional nonlinear time-series data may be studied using singular spectrum analysis (SSA) [37]. SSA builds the trajectory matrix using the time series signal, decomposes it, and reconstructs it to analyse its components. In order to successfully extract the relevant signal from the time series including noise signal, the time series structure is further analysed and rebuilt [38]. The disadvantage of SSA is that it involves identifying the primary trajectory matrix components oscillation frequencies and categorising them to produce the required signals after eliminating them [39].

2) CIRCULANT SINGULAR SPECTRUM ANALYSIS (CSSA)

Older editions of SSA and other decomposition algorithms require determining the oscillating frequencies of the main attributes of trajectory matrix [39]. The second instances of time series are the focus of CSSA, which primarily emphasises on eigenstructure. The circulant matrix has closed-form solutions for both its eigenvalues and vectors [40]. The graphical illustration of the reconstructed components (RCs) obtained by using CSSA is depicted in the Figure 3.

D. FEATURE EXTRACTION

The technique of feature extraction involves reducing down the dimension of the data into useful and manageable groups known as attributes or features. The computational and memory needs are lowered when classification is performed using smaller dimensions. Other advantages of feature extraction include faster training, better data illustration, enhanced accuracy, and more understandable models [41]. Entropy, a nonlinear quantity, can be used to gauge how chaotic a system is [42]. Entropy can be used to learn more about the associated cardiovascular complexity because the human heart is a dynamic, complicated system [41]. In this study we have extracted Shannon's entropy (S_{en}), approximate entropy (A_{en}), differential entropy (D_{en}), renyi entropy (R_{en}), and fuzzy entropy (F_{en}) as mentioned in [43], [44], [45], [46], and [47] respectively.

E. FEATURE SELECTION

The procedure of selecting the most relevant, important, and helpful feature collection is known as feature selection. It requires selecting a small subset of many prominent and significant attributes [48]. By calculating the amount of details, reducing computing requirements and attribute reduction can help with feature selection. In this research, we use binary harris hawks optimisation (BHHO) to improve classification accuracy. To get successful outcomes, it exploits 2 stages of study and four levels of exploitation. The next subsections give thorough explanations of BHHO.

1) BINARY HARRIS HAWKS OPTIMISATION

In 2019, Heidari and his associates unveiled the Harris Hawk Optimisation (HHO), meta-heuristic method [49]. Harris hawks use a variety of attacking strategies in nature, including surprise prey, pounces, and other methods. In HHO, different solutions are represented by hawks, while the most effective (almost optimal) choice is depicted by prey. Prior to making a swift break to capture their prey, Harris hawks use their excellent insightÂ to monitor their prey. HHO is distinctive swarm-based optimizer to solve ongoing problems that has both effective exploratory and exploitative strategies and an evolving framework. The HHO approach may have transitioned from victimisation to investigation, at the point when the investigative behaviour may have changed according to the prey's dwindling energy. The mathematical formula for calculating the prey's escape energy is [49]:

$$E_e = 2E_i(1 - \frac{i}{I}) \tag{1}$$

$$E_i = 2r - 1 \tag{2}$$

where Ei denotes the initial energy collected at random from [0, 1], I is the whole amount of repetitions, and r denotes an integer obtained at random from [0, 1]. Hawks can search for global positions between different locales when their prey's escape energy $|E_e| > 1$. Alternatively, HHO, frequently promotes local population's to look into most effective strategies if the prey's fleeing energy is $|E_e| < 1$.

Two fundamental components of the BHHO should be taken into account: the solution depiction along with the evaluation function. These two objectives are created using the fitness function in equation 3 as we are using a single-objective HHO.

$$\downarrow Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|N|} \tag{3}$$

where $\gamma_R(D)$ is classification error rate, $|R|$ denotes number of selected features, $|N|$ represents number of attributes in the original dataset. The significance of categorization quality and reduction rate is determined by the two parameters $\alpha\epsilon[0, 1]$, and $\beta = (1 - \alpha)$, respectively [50]. Hence, BHHO is a novel algorithm used in the present work for feature selection.

F. MACHINE LEARNING CLASSIFIERS

Machine learning classifiers are used for the classification of various kinds of arrhythmias present in the cardiovascular disorders.

1) DECISION TREE (DTE)

A supervised, non-parametric learning algorithm that is frequently used for classification is called a decision tree (DTE). It makes classifications by inferring decision rules from feature reasoning. A tree begins at the root node and divides into branches and sub-trees until it reaches the leaf or terminal node. A leaf node's parent or root is determined by the feature approximation constant [51]. The approximation constant is the rule, and this node is referred to as the decision node. The approximation constant is calculated by the decision tree using the Gini index. Ranging from zero to one, the Gini index calculates the relative degree of inequality in the distribution of classes. To calculate the Gini Index it uses equation 4.

$$I_G = 1 - \sum_{j=1}^{k} P_j^2 \tag{4}$$

The best features with the most important information about the target are identified using the Gini index. The target

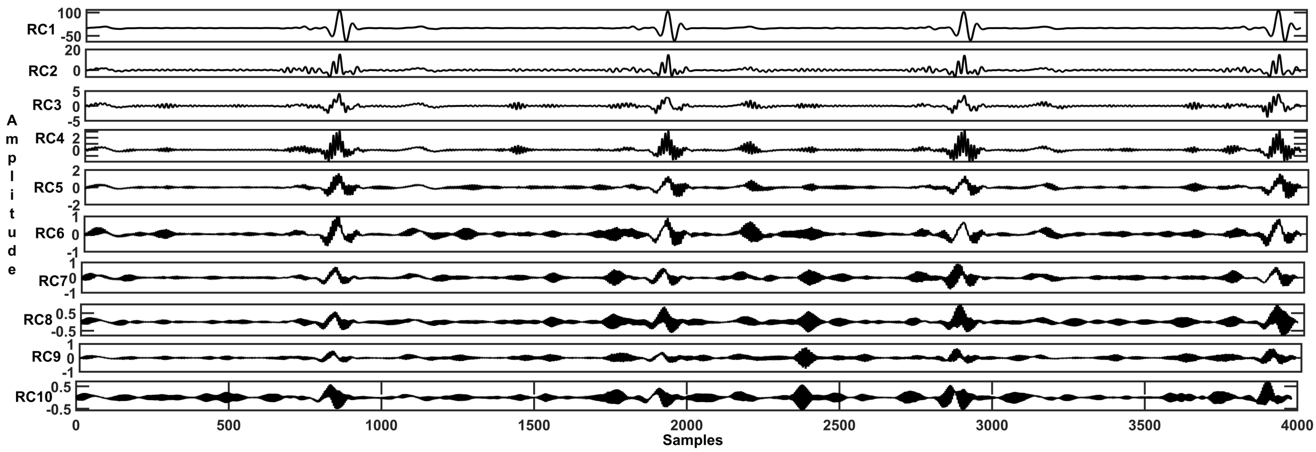


FIGURE 3. Extracted RCs for ECG signal using CSSA.

function values at the leaf nodes are then as precisely as possible determined by using the dataset and the values of these features. This leads to the organisation of a number of decision rules into a distinguishable tree structure. As a result, a collection of classification-capable decision rules are organised in a tree structure [51].

2) NAIVE BAYES (NBS)

Using statistical techniques and probability, the NBS classification model forecasts future prospects by leveraging past experiences [52]. One benefit of using NBS is that it only needs a small amount of training data to calculate the parameter estimates required for classification. For the most part, NBS outperforms expectations in complex real-world scenarios [52].

To calculate the possibility that a variable belongs to a specific class, NBS uses Equation 5, which demonstrates a data classification method based on the Bayes theorem. P(A) stands for the probability calculated before the result is seen, and P(B|A) is the chances that B will happen if A does.

$$P(B|A) = P(A, B)/P(A), P(A, B) = P(B|A)P(A) \tag{5}$$

3) K-NEAREST NEIGHBORS (KNN)

When the target variable is known, the supervised machine learning algorithm K-Nearest Neighbours (KNN) can be used. It doesn't assume that the underlying data follow a non-parametric distribution pattern. In this algorithm, the new data point's similar neighbours are found using a number known as K. To determine which neighbourhood the new data point belongs to, the algorithm considers its K nearest neighbours. The similarity of the features is the basis for this decision. The KNN results that are obtained are significantly impacted by the choice of K [53].

The KNN algorithm operates in four main steps. Initially, an odd number, K, should be selected. The second step is to determine how far each training data point is from the new point. The third step is to determine which K neighbours the new data point is closest to. Counting the number of data points in each category among the k neighbours is the last step in the classification process [53].

Four basic theories—Euclidean, Manhattan, Hamming, and Minkowski—are used to calculate the distance between each training set and the new point.

The square root of the total squared distance between two points is known as the Euclidean distance (ECD).

$$ECD = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{6}$$

Manhattan distance (MTD) is the sum of the absolute values of the differences between two points

$$MTD = \sum_{i=1}^{k} |x_i - y_i| \tag{7}$$

The Minkowski distance (MKD) is a tool used to compare the distances between two points. It becomes the Manhattan distance when p = 1 and the Euclidean distance when p = 2.

$$MKD = \left(\sum_{i=1}^{k} (|x_i - y_i|)^q\right)^{1/q} \tag{8}$$

4) SUPPORT VECTOR MACHINES (SVM)

Support Vector Machines (SVM) employ the hyperplane concept, which involves a hypothetical boundary separating data points of distinct classes. The optimal hyperplane is one that has the lowest generalisation error, the largest margin distance between the closest points, and a good functional margin. In SVM, vectors are classified in multi-dimensional space using the kernel trick method; the hyperplane is not a suitable method [54]. There are several available kernel functions that can be utilised in this situation, including sigmoid, polynomial, linear, and radial basis function. The RBF kernel function yields the lowest computation cost, even in high dimension spaces. The RBF Kernel equation that is utilised is provided below in equation 9.

$$K(y_i, y_j) = exp\left\{-\lambda \left\|y_i - y_j\right\|^2\right\} \tag{9}$$

5) ENSEMBLE SUB-SPACE KNN (ESSKNN)

Using a variety of classification strategies to combine the data into a highly efficient composite model is known as ensemble learning. This strategy aims to obtain a higher accuracy rate from several models than from any one of them alone. Any kind of base classifier algorithm, such as decision trees, k-NN, and other base learner algorithms, can be combined to create an ensemble inducer. In the proposed study, the base learner and ensemble approach have been implemented using the k-NN and random subspace, respectively [55].

The k-NN classifier's perceptive input selection makes it possible for ensemble systems based on random subspaces to improve the performance of individual kNN classifiers [56]. An ensemble technique called random subspace is widely used to produce individual classifiers from data subspaces selected at random [57]. The final result is also generated by eventually integrating the output of each independent classifier.

6) MAJORITY VOTING CLASSIFIER

This classifier functions as a meta-classifier to combine machine learning models that have the same or conceptually different structures in order to make predictions by using a simple majority. The classification accuracy of the system can be increased by voting of different classifiers. As a result, a range of classifiers can choose among available possibilities. The decisions adopted by the majority are taken into account when making the final decision. A better answer can be found if several algorithms are applied to the same issue [58]. DTE, NBS, KNN, SVM, and ESSKNN classifiers have been ensembled in the suggested model. Each model makes a separate prediction using a voting aggregator, and the result is determined by computing the majority vote, which produces the final prediction.

IV. RESULTS

The suggested technique is implemented using MATLAB 2021a on a personal computer with Intel Core i7 processor having 16 GB of RAM. The proposed method's effectiveness is evaluated in terms of accuracy ($A_c\%$), sensitivity ($S_e\%$), and specificity ($S_p\%$).

A metric of MI occurrences is called $S_e\%$, which is defined in equation (10):

$$S_e\% = \frac{t_p}{t_p + f_n} \times 100 \tag{10}$$

where f_n is "false negative" and t_p is "true positive." $S_p\%$ is a measure of normal events and is defined as follows:

$$S_p\% = \frac{t_n}{t_n + f_p} \times 100 \tag{11}$$

"False positive" is denoted here by f_p, and "True negative" is represented by t_n. Equation 12 is used to express accuracy ($A_c\%$):

$$A_c\% = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \times 100 \tag{12}$$

Since it is inappropriate to compare different classifiers using S_e or S_p, the region under the receiver operating characteristic curve (ROC) was also used to draw conclusions. The aforementioned performance measures were derived using two different approaches, namely 10-fold and LOO cross validation. For optimisation strategies, the number of hawks has been set to 10 while the frequency of recurrences to 100. In order to attain the highest classification accuracy feasible, the classifier parameters were updated utilising optimisation techniques. We have run the experiment for different values of the hyper-parameters of the classifiers used. Based on the gird search approach, we have selected the optimum value for the classifiers employed. The Minkowski, Euclidean, and city block distance functions were employed alongside KNN and $K \in [3, 5, 7, 9]$. The KNN-BHHO employing euclidean distance function and K=5 produced best results. Likewise, SVM classifier alongside linear, polynomial, and radial basis function kernel were utilized. The kernel scale σ was varied in steps of 1, i.e from one to five. SVM classifier using radial basis function with a kernel scale of 3 yielded optimum outcomes.

A. K-FOLD CROSS VALIDATION APPROACH

A K-fold cross validation approach is a generalised classification strategy in which a comprehensive dataset made up of all MI-affected and HC subjects is prepared. The proposed approach's testing and training procedures were carried out using the K-fold cross validation method. In the K-fold cross validation we have divided the data into 80% for training and the rest 20% for testing and named it as X-fold cross validation. In K-fold cross validation, apart from X-fold cross validation we have also performed 10-fold cross validation by dividing the entire data into 90% for training and 10% for testing. This experiment is repeated for 10 times, randomly splitting data in training and test sets. To evaluate the performance of the suggested technique, we ran experiments utilising DTE, NBS, KNN, SVM, and ESSKNN classifiers. Table 1 provides a summary of the experimental findings of X-fold cross validation employing SSA and CSSA decomposition, with and without feature selection (BHHO) along with machine learning algorithms. Using SVM, we have obtained $A_c\%$ of 86.93 and 87.63 with and without feature selection employing SSA. Where as using CSSA decomposition and SVM classifier, we have acheived $A_c\%$ of 90.51 and 99.8 with and without feature selection.

As it can be observed from Table 1, the ESSKNN outperformed traditional classifiers like DTE, NBS, KNN, and SVM in terms of $A_c\%$, $S_e\%$, and $S_p\%$. When a BHHO feature selection technique was used, the suggested method produced noteworthy results. We have obtained

TABLE 1. Performance metrics calculations for X-fold cross validation strategy using the proposed methodology.

Classifier	SSA						CSSA					
	With out Feature Selection			Feature Selection using BHHO			With out Feature Selection			Feature Selection using BHHO		
	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$
DTE	82.91	81.41	84.12	83.12	84.11	80.21	96.61	94.70	97.30	96.11	94.21	96.91
NBS	79.72	77.52	78.96	79.81	78.12	79.01	85.30	81.21	87.01	85.31	81.42	86.91
SVM	86.93	85.12	86.19	87.63	88.21	86.24	90.51	89.11	93.10	99.80	99.41	99.90
KNN	84.92	84.10	82.41	86.41	86.04	87.14	90.11	85.14	92.31	99.81	99.32	100.00
ESSKNN	88.91	88.14	86.11	89.01	88.12	89.41	99.40	99.01	99.12	99.90	99.90	100.00

TABLE 2. Performance metrics calculations for 10-fold cross validation strategy using the proposed methodology.

Classifier	SSA						CSSA					
	With out Feature Selection			Feature Selection using BHHO			With out Feature Selection			Feature Selection using BHHO		
	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$	$A_c\%$	$S_e\%$	$S_p\%$
DTE	84.24	83.23	86.43	86.91	85.24	87.41	97.54	96.42	97.21	98.25	98.40	98.56
NBS	82.11	78.41	79.42	84.23	78.24	80.44	90.22	90.10	89.56	93.14	92.42	93.21
SVM	87.13	87.11	87.45	88.47	88.10	88.01	92.42	91.45	92.24	99.90	99.56	99.9
KNN	86.24	86.11	87.01	87.42	87.54	87.42	95.41	95.43	95.12	99.91	99.64	100.00
ESSKNN	90.12	89.14	90.12	93.42	89.45	90.43	99.90	99.81	99.90	99.97	99.95	100.00

$A_c\%$, $S_e\%$, and $S_p\%$ of 99.4, 99.01, and 99.12 using CSSA decomposition and without BHHO. Similarly, we were able to attain 99.9 $A_c\%$, 99.9 $S_e\%$, and a $S_p\%$ of 100 by employing CSSA decomposition and BHHO. Table 2 provides a summary of the experimental outcomes of 10-fold cross validation employing SSA and CSSA decomposition, with and without feature selection (BHHO) along with machine learning algorithms. This experiment is repeated for 10 times, randomly splitting data in training and test sets. Using ESSKNN, we have obtained $A_c\%$ of 90.12 and 93.42 with and without feature selection employing SSA. Where as using CSSA decomposition and ESSKNN classifier, we have acheived $A_c\%$ of 99.9 and 99.97 with and without feature selection. Regardless of the classifier employed, it can be seen that the performance metrics improve significantly when the BHHO feature selection strategy is utilised.

The Figure 4 shows the confusion matrix, ROC, and fitness curve that was derived using a BHHO feature selection technique. From the Figure 4-(a), we infer that we obtained t_p of 99.9%, f_p of 0.1% and t_n as 100%. Similarly from the Figure 4-(b), we have obtained area under the curve value of 1, which indicates that the classifier employed performs well. The Figure 4-(c) also indicates that the optimizer employed performed well.

The data are arranged using a boxplot, which is a standard approach based on a five-number summary ('minimum, median ('Q3'), 'first quartile ('Q1'), and 'maximum'). Boxplots can also demonstrate the degree to which the data are organised, whether the data are skewed, and whether the data are symmetrical. Figure 5 shows box plots for the top 10 features for the HC and MI dimensions using the BHHO-KNN feature selection approach, with the entropy values plotted on the y-axis. From the Figure 5, we infer that the median line of the boxplot exists mostly outside the box of the comparison boxplot, which indicate that the features selected are separable. The most significant 10 features are Shannon's Entropy $S^7_{en_1}$, $S^3_{en_2}$, $S^6_{en_5}$, Approximate Entropy $A^3_{en_2}$, $A^4_{en_2}$, $A^5_{en_9}$, Differential Entropy $D^3_{en_7}$, Renyi Entropy $R^3_{en_{10}}$, and Fuzzy Entropy $F^4_{en_4}$, $F^6_{en_5}$ are chosen using the BHHO feature selection method. The lead as well as decomposition level are indicated, respectively, using superscripts and subscripts.

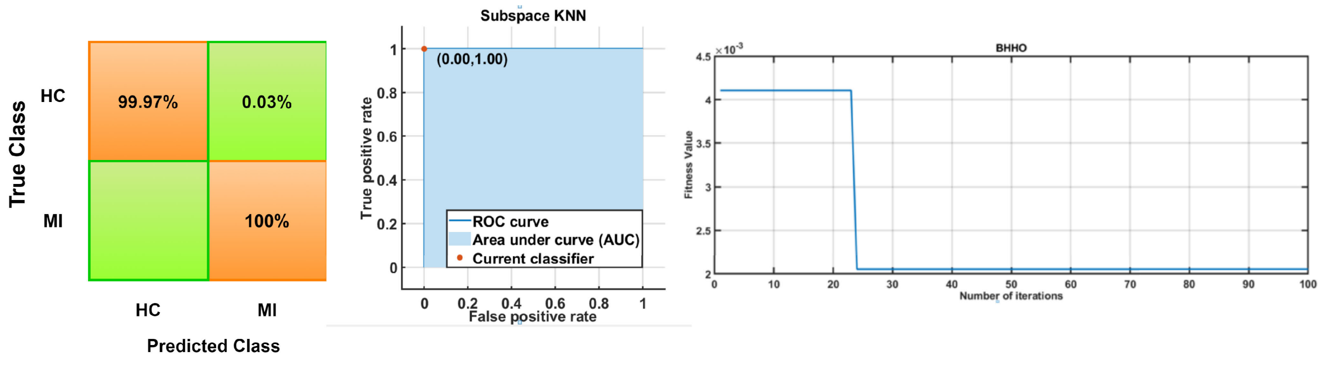
TABLE 3. Performance metrics calculations using LOO cross validation strategy.

Classifier	SSA			CSSA		
	Avg. A_c %	Avg.S_e%	Avg. S_p %	Avg. A_c %	Avg.S_e%	Avg. S_p %
DTE	66.10	62.24	65.10	76.41	72.52	75.10
NBS	68.32	67.10	69.50	72.32	70.10	72.50
SVM	70.14	69.58	71.21	77.11	74.32	76.20
KNN	71.21	71.12	72.00	79.40	75.12	78.31
ESSKNN	75.11	76.51	76.74	85.20	83.40	85.11
Voting	85.24	83.41	86.15	91.03	89.93	90.98

B. LEAVE-ONE-OUT (LOO) CROSS VALIDATION APPROACH

The recommended method can also be used to analyse data from unidentified subjects, for which the classifier was never trained, by employing "LOO cross validation" strategy. The LOO cross validation technique uses the data that was obtained following segmentation and feature selection. Testing is done using data from just one subject, while training is done on the remaining individuals. The procedure is repeated for each participant, and Table 3 shows the average accuracy (Avg. $A_c\%$), average sensitivity (Avg. $S_e\%$), and average specificity (Avg. $S_p\%$).

In order to improve the efficiency of the proposed technique, we have employed voting classifier. The predictions obtained from the DTE, NBS, KNN, SVM, ESSKNN are given as input to voting classifier. Taking into account the



(a) (b) (c)

FIGURE 4. Results obtained for MI classification by 10-fold cross validation with BHHO feature selection (a) Confusion matrix. (b) ROC and (c) Fitness curve.

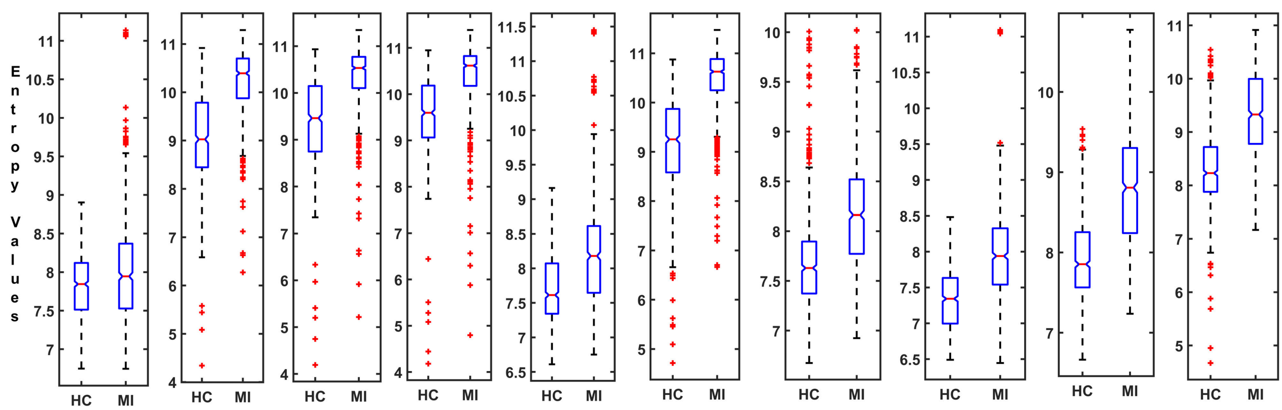


FIGURE 5. Top ten features Boxplot illustration of the BHHO-ESSKNN approach (leftmost feature corresponds to feature 1 and rightmost feature represents feature 10).

majority's selections, a final selection is made. The proposed LOO cross validation approach offers a novel technique to assess whether a technique will be efficient when employed in information pertaining to an unknown subject for which the algorithm for classification has not been trained. The results presented in the Table 3 clearly demonstrate that the technique employed performs remarkably well in the classification of MI.

C. INTER DATASET CROSS VALIDATION (IDCV)

We have also performed the cross validation approach by using IDCV, that is dividing the entire data into the ratio of 70:30 for training and testing. That is 70% of data is utilized for training and the rest unseen 30% subjects (which are not employed for training) are used for testing. The same processes is repeated 10 times and the results obtained using IDCV are demonstrated in the Table 4. Form Table 4, we infer that we have obtained average $A_c\%$ of 82.15 and 95.4 by employing SSA and CSSA decomposition methods. Likewise we have obtained an average $A_c\%$ of 85.31 and 98.54 using SSA and CSSA decomposition techniques, respectively. Similarly, using voting classifier we got 89.26 and 99.41 average $A_c\%$ using SSA as well as CSSA. Hence, the outcomes depicted in the Table 4 clearly show that the technique employed performs remarkably well in categorization of MI. As per literature reviewed by the author's this is the first work to apply IDCV technique to VCG signal.

TABLE 4. Performance metrics calculations from IDCV.

Classifier	SSA			CSSA		
	Avg. A_c %	Avg.S_e%	Avg. S_p %	Avg. A_c %	Avg.S_e%	Avg. S_p %
DTE	71.20	68.4	72	86.14	84.21	85.05
NBS	74.21	75.40	74.31	88.42	86.25	87.65
SVM	79.45	78.81	79.14	91.01	88.92	90.50
KNN	82.15	80.92	83.00	95.40	96.17	94.42
ESSKNN	85.31	84.58	86.41	98.54	98.43	97.43
Voting	89.26	88.61	89.10	99.41	98.91	99.16

V. DISCUSSION

To distinguish MI from the healthy person, we evaluated the suggested classification system with other cutting-edge techniques already in use. The embedding length (L) plays major role in CSSA. Figure 6 shows the outcomes of the

TABLE 5. Comparison of the suggested approach for MI classification with other existing algorithms 10-fold cross validation approach).

Refernce	Year	Signal	NOL	Pre-Processing	Features extracted	Classifiers	$A_c\%$	$S_e\%$	$S_p\%$
Current	2023	VCG	3	CSAA+4 satge SG filter	S_{en}, A_{en},D_en, R_{en},F_{en}	ESSKNN	99.97	99.95	100
[60]	2021	ECG	12	Linear and non-linear filtering	Maximas between T and Q, mean area under QRS and ST.	CNN + Bi-LSTM	99.24	99.25	99.62
[61]	2021	ECG	1	Fourier decomposition	Entropy, kurtosis, and energy	KNN, SVM, ensemble bagged trees and ESSKNN	99.65	99.61	-
[62]	2022	ECG	12	Daubechies 6 wavelet function	R peak detection	CNN and DenseNet	98.9	-	-
[63]	2022	ECG	1	-	-	Single-Lead convolutional generative adversarial network	99.06	-	98.65
[65]	2020	ECG	12	DWT + PCA	Statistical and spectral features	Neural network	98.21	97.5	98.01
[66]	2020	ECG	12	FIR + BPF	-	CNN	99.01	96.75	99.2
[64]	2017	ECG	15	Fuzy information granulation	Multiscale features	CNN	96	95.4	97.37
[67]	2015	ECG	2	DYWT	PolyECG-S	Decision tree	94.4	-	-
[68]	2013	ECG+VCG	15	-	Local octant, Octant residence, Octant transition	Classification and regression tree	91	88	92
[11]	2011	VCG	3	DWT	Trapping time, recurrence rate , linemax, determinism, entropy, and laminarity	RQA	-	96.5	75
[69]	2012	ECG	12	DCTBF	ST segmentation	Latent topic multiple instance learning	-	92.3	88.1
[70]	2012	ECG	12	-	Time-domain features	Pruning algorithm, KNN	-	99.97	99

NOL: Number of leads, DCTBF: Discrete cosine transform based filter, DYWT: Dyadic wavelet transform, FIR: finite impulse response, BPF: band pass filter, DWT: discrete wavelet transform

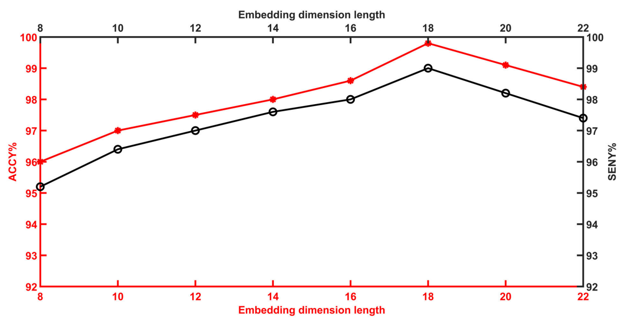


FIGURE 6. Graph illustrating variation in classification $A_c\%$, $S_e\%$ of CSSA tuning parameters.

evaluation of the recommended technique for different values of this parameter. L = 18 displayed superior results, hence it is applied for the course of the experiment. Both the LOO cross validation and 10-fold cross validation categorization results are compared independently. The results from the suggested approach employing 3-leads VCG were compared to those from methods using ECG signal, 3-leads VCG, and 15-leads (ECG+VCG) signal.

It is evident from the Table 5 that the performance of ESSKNN was superior to other machine learning models in the K-fold cross validation classification strategy. Rahul et al. [59] succeeded in obtaining accuracy of 99.1% by employing random forest classifier. Dey et al. [60] utilized 12 lead ECG and obtained $A_c\%$ of 99.24, $S_p\%$ of 99.6, and $S-e\%$ of 99.25 by employing the CNN and Bi-LSTM classifiers. Fatimah et al. [61] used single ECG lead and employed classifiers like KNN, SVM, and ESSKNN and acheived $A_c\%$ of 99.6 and $S_e\%$ of 99.61. Jahmunah et al. [62] used CNN and DenseNet classifiers and obtained $A_c\%$ of 98.9. Li et al. [63] categorized MI with and $A_c\%$ of 99.06 by employing single-Lead convolutional generative adversarial network. Similarly, Liu et al. [64] used fuzzy information granulation in the preprocessing, extracted multi-scale features and were able to obtain the $A_c\%$ of 96, $S_p\%$ of 97.37 by employing CNN classifier utilizing the 15 leads information. Yang [11] used 3 leads of VCG and filtered using DWT; extracted the attributes like trapping time, recurrence rate, linemax, determinism, entropy and laminarity and are fed to the RQA classifier. They have achieved $S_e\%$ of 96.5, 75 $S_p\%$. The proposed method performs impressively well when compared with other state-of-art techniques in terms of the $A_c\%$, $S_e\%$, and $S_p\%$.

Our proposed method outperformed different models in LOO cross validation classification. Rahul et al. [59] used VCG leads and filtered the VCG signal using the SG filter as well as SWT; extracted the features like SE, NSE, LEE,

TABLE 6. Comparison of the suggested approach for MI classification with other existing algorithms (LOO cross validation approach).

Refernce	Year	Signal	NOL	Classifiers	$A_c\%$	$S_e\%$	$S_p\%$	MIs considered
Current	2023	VCG	3	Majority voting	91.03	89.93	90.98	All
[59]	2019	VCG	3	Random forest	89.37	88.02	90.8	IMI
[72]	2019	ECG	8	Fully-CNN	-	87.4	90	IMI+AMI
[71]	2018	ECG	3	SVM	81.7	79.01	79.26	IMI
[15]	2017	ECG+ VCG	15	CNN	84.54	85.33	84.09	IMI

and MS and fed to the random forest classifier and succeeded in achieving an $A_c\%$, $S_p\%$, and $S_e\%$ of 89.37, 90.8, and 88.02 respectively. Rahul et al. [59], Sharma et al. [71] considered only inferior myocardial infarction (IMI) but our proposed technique considers all the 10 cases of MI. Similarly, Strodthoff and Strodthoff [72] utilized limb leads and inferior MI leads as an input to the Fully CNN classifier and obtained a sensitivity of 87.4% and 90% of specificity. Reasat and Shahnaz [15] used all the 15 leads of ECG and preprocessed by two stage median filter, SG filter, and SWT. Thus, the obtained clean signal is given as an input the CNN for classification. They succeeded in achieving $A_c\%$ of 84.54, 84.09 $S_p\%$, and $S_e\%$ of 85.33. Strodthoff and Strodthoff [72] and Reasat and Shahnaz [15] employed CNN, in conventional CNNs, dense layers come after convolutional layers to transform feature maps to the desired output. It is challenging to understand how filters relate to categorical outputs due to the dense layers' role as "black boxes." Dense layers also have a high parameter count, which makes them susceptible to overfitting as well as computationally expensive. When compared to other available techniques, our suggested method outperformed them in terms of $A_c\%$, $S_p\%$, and $S_e\%$.

Comparison of the proposed method with other techniques using LOO method is depicted in Table 6.

VI. CONCLUSION

This article describes a strategy towards MI detection using small duration VCG signal. In this study, we have employed the amalgamation of CSSA and 4 stage SG filter for expunging BW and PLI from the VCG signal. The obtained noise free VCG signal is segmented into four seconds duration. The ensemble classifiers are fed with features extracted as well as with the prominent features selected by the BHHO for categorization. The results clearly demonstrate that by employing feature selection, the classification accuracy is improved. The validation has been done utilising both a 10-fold as well as LOO cross validation approaches. The improved outcomes for both approaches, particularly the LOO cross validation strategy, demonstrate the method's viability and usefulness for automating the identification of MI in unseen subjects on which the proposed model has not been trained. This method may aid in the accurate and reliable detection of MI thereby reducing the load on the health care professionals. The following shortcomings of the suggested technique offer future research directions:

- The accuracy of categorization can be increased by using deep learning models.
- In the preprocessing stage other denoising techniques can be employed.
- Other features, decomposition methods, alongside feature selection techniques can be employed.

REFERENCES

[1] S.-M. Dima, C. Panagiotou, E. B. Mazomenos, J. A. Rosengarten, K. Maharatna, J. V. Gialelis, N. Curzen, and J. Morgan, "On the detection of myocadial scar based on ECG/VCG analysis," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3399–3409, Dec. 2013.

[2] D. Sadhukhan, S. Pal, and M. Mitra, "Automated identification of myocardial infarction using harmonic phase distribution pattern of ECG data," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 10, pp. 2303–2313, Oct. 2018.

[3] J. W. Hurst, R. A. Walsh, V. Fuster, and J. C. Fang, *The Heart Manual of Cardiology*. New York, NY, USA: McGraw-Hill, 2013.

[4] E. Prabhakararao and S. Dandapat, "Automated detection of posterior myocardial infarction from VCG signals using stationary wavelet transform based features," *IEEE Sensors Lett.*, vol. 4, no. 6, pp. 1–4, Jun. 2020.

[5] A. R. P. Riera, A. H. Uchida, C. F. Filho, A. Meneghini, C. Ferreira, E. Schapacknik, S. Dubner, and P. Moffa, "Significance of vectorcardiogram in the cardiological diagnosis of the 21st century," *Clin. Cardiol.*, vol. 30, no. 7, pp. 319–323, Jul. 2007.

[6] G. Liu and H. Yang, "Multiscale adaptive basis function modeling of spatiotemporal vectorcardiogram signals," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 2, pp. 484–492, Mar. 2013.

[7] R. Chen, F. Imani, and H. Yang, "Heterogeneous recurrence analysis of disease-altered spatiotemporal patterns in multi-channel cardiac signals," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 6, pp. 1619–1631, Jun. 2020.

[8] A. R. M. Dehnavi, I. Farahabadi, H. Rabbani, A. Farahabadi, M. P. Mahjoob, and N. R. Dehnavi, "Detection and classification of cardiac ischemia using vectorcardiogram signal via neural network," *J. Res. Med. Sci., Off. J. Isfahan Univ. Med. Sci.*, vol. 16, p. 136, Feb. 2011.

[9] G. Bortolan and I. Christov, "Myocardial infarction and ischemia characterization from T-loop morphology in VCG," in *Proc. Comput. Cardiol.*, vol. 28, Sep. 2001, pp. 633–636.

[10] H. Yang, S. T. S. Bukkapatnam, T. Le, and R. Komanduri, "Identification of myocardial infarction (MI) using spatio-temporal heart dynamics," *Med. Eng. Phys.*, vol. 34, no. 4, pp. 485–497, May 2012.

[11] H. Yang, "Multiscale recurrence quantification analysis of spatial cardiac vectorcardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 2, pp. 339–347, Feb. 2011.

[12] H. Yang, S. T. Bukkapatnam, and R. Komanduri, "Spatiotemporal representation of cardiac vectorcardiogram (VCG) signals," *Biomed. Eng. OnLine*, vol. 11, no. 1, p. 16, 2012.

[13] X. Xi, M. Tang, S. M. Miran, and Z. Luo, "Evaluation of feature extraction and recognition for activity monitoring and fall detection based on wearable sEMG sensors," *Sensors*, vol. 17, no. 6, p. 1229, May 2017.

[14] R. K. Tripathy and S. Dandapat, "Detection of myocardial infarction from vectorcardiogram using relevance vector machine," *Signal, Image Video Process.*, vol. 11, no. 6, pp. 1139–1146, Sep. 2017.

[15] T. Reasat and C. Shahnaz, "Detection of inferior myocardial infarction using shallow convolutional neural networks," in *Proc. IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC)*, Dec. 2017, pp. 718–721.

[16] W. Liu, F. Wang, Q. Huang, S. Chang, H. Wang, and J. He, "MFB-CBRNN: A hybrid network for MI detection using 12-lead ECGs," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 2, pp. 503–514, Feb. 2020.

[17] A. Kapfo, S. Dandapat, and P. K. Bora, "Automated detection of myocardial infarction from ECG signal using variational mode decomposition based analysis," *Healthcare Technol. Lett.*, vol. 7, no. 6, pp. 155–160, Dec. 2020.

[18] U. Desai, C. G. Nayak, and G. Seshikala, "Application of ensemble classifiers in accurate diagnosis of myocardial ischemia conditions," *Prog. Artif. Intell.*, vol. 6, no. 3, pp. 245–253, Sep. 2017.

[19] M. U. Khan, Z. Mushtaq, M. Shakeel, S. Aziz, and S. Z. H. Naqvi, "Classification of myocardial infarction using MFCC and ensemble subspace KNN," in *Proc. Int. Conf. Elect., Commun., Comput. Eng. (ICECCE)*, Jun. 2020, pp. 1–5.
[20] S. Bashir, U. Qamar, and F. H. Khan, "BagMOOV: A novel ensemble for heart disease prediction bootstrap aggregation with multi-objective optimized voting," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 305–323, Jun. 2015.
[21] D. Tomar and S. Agarwal, "Feature selection based least square twin support vector machine for diagnosis of heart disease," *Int. J. Bio-Sci. Bio-Technol.*, vol. 6, pp. 69–82, Apr. 2014.
[22] K. R. Ravindranath, "Clinical decision support system for heart diseases using extended sub tree," in *Proc. Int. Conf. Pervasive Comput. (ICPC)*, Jan. 2015, pp. 1–5.
[23] P. Domingos, "A few useful things to know about machine learning," *Commun. ACM*, vol. 55, no. 10, pp. 78–87, Oct. 2012.
[24] H. Liu and H. Motoda, *Feature Extraction, Construction and Selection: A Data Mining Perspective*, vol. 453. Cham, Switzerland: Springer, 1998.
[25] B. Dun, E. Wang, and S. Majumder, "Heart disease diagnosis on medical data using ensemble learning," *Comput. Sci.*, vol. 1, pp. 1–5, Jan. 2016.
[26] R. S. Singh, B. S. Saini, and R. K. Sunkaria, "Detection of coronary artery disease by reduced features and extreme learning machine," *Med. Pharmacy Rep.*, vol. 91, no. 2, pp. 166–175, Apr. 2018.
[27] F. Yaghouby, A. Ayatollahi, and R. Soleimani, "Classification of cardiac abnormalities using reduced features of heart rate variability signal," *World Appl. Sci. J.*, vol. 6, no. 11, pp. 1547–1554, 2009.
[28] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artif. Intell. Med.*, vol. 44, no. 1, pp. 51–64, Sep. 2008.
[29] P. Langley, "Selection of relevant features in machine learning," in *Proc. AAAI*, Sep. 1994, pp. 245–271.
[30] A. A. Heidari, I. Aljarah, H. Faris, H. Chen, J. Luo, and S. Mirjalili, "An enhanced associative learning-based exploratory whale optimizer for global optimization," *Neural Comput. Appl.*, vol. 32, no. 9, pp. 5185–5211, May 2020.
[31] A. A. Heidari, H. Faris, S. Mirjalili, I. Aljarah, and M. Mafarja, "Ant lion optimizer: Theory, literature review, and application in multi-layer perceptron neural networks," in *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications*, 2020, pp. 23–46.
[32] Q. Zhang, H. Chen, A. A. Heidari, X. Zhao, Y. Xu, P. Wang, Y. Li, and C. Li, "Chaos-induced and mutation-driven schemes boosting salp chains-inspired optimizers," *IEEE Access*, vol. 7, pp. 31243–31261, 2019.
[33] M. Mafarja, A. A. Heidari, H. Faris, S. Mirjalili, and I. Aljarah, "Dragonfly algorithm: Theory, literature review, and application in feature selection," in *Nature-Inspired Optimizers: Theories, Literature Reviews and Applications*, 2020, pp. 47–67.
[34] R. Bousseljot, D. Kreiseler, A. Schnabel, "Nutzung der EKG-signaldatenbank CARDIODAT der PTB Über das internet," Walter de Gruyter, Berlin, Germany, Tech. Rep., 1995, pp. 317–318.
[35] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.
[36] M. Krishna Chaitanya and L. D. Sharma, "Electrocardiogram signal filtering using circulant singular spectrum analysis and cascaded Savitzky–Golay filter," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103583.
[37] V. Oropeza and M. Sacchi, "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis," *Geophysics*, vol. 76, no. 3, pp. V25–V32, May 2011.
[38] S. Sanei, M. Ghodsi, and H. Hassani, "An adaptive singular spectrum analysis approach to murmur detection from heart sounds," *Med. Eng. Phys.*, vol. 33, no. 3, pp. 362–367, Apr. 2011.
[39] J. Bógalo, P. Poncela, E. Senra, "Strong separability in circulant ssa," in: Nonparametric Statistics: 3rd ISNPS, Avignon, France, Jun. 2016 3, Springer, 2018, pp. 295–309.
[40] J. Bógalo, P. Poncela, and E. Senra, "Circulant singular spectrum analysis: A new automated procedure for signal extraction," *Signal Process.*, vol. 179, Feb. 2021, Art. no. 107824.
[41] S. Asgari, A. Mehrnia, and M. Moussavi, "Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine," *Comput. Biol. Med.*, vol. 60, pp. 132–142, May 2015.
[42] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 278, no. 6, pp. H2039–H2049, Jun. 2000.
[43] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 5, pp. 3–55, Jan. 2001.
[44] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Nat. Acad. Sci. USA*, vol. 88, pp. 2297–2301, Mar. 1991.
[45] G. Ariel and Y. Louzoun, "Estimating differential entropy using recursive copula splitting," *Entropy*, vol. 22, no. 2, p. 236, Feb. 2020.
[46] P. Bromiley, N. Thacker, and E. Bouhova-Thacker, "Shannon entropy, Renyi entropy, and information," *Statist. and Inf. Ser.*, vol. 9, pp. 2–8, May 2004.
[47] A. Ishikawa and H. Mieno, "The fuzzy entropy concept and its application," *Fuzzy Sets Syst.*, vol. 2, no. 2, pp. 113–123, Apr. 1979.
[48] E. M. El-Kenawy, M. M. Eid, M. Saber, and A. Ibrahim, "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection," *IEEE Access*, vol. 8, pp. 107635–107649, 2020.
[49] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, pp. 849–872, Aug. 2019.
[50] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54–65, Nov. 2016.
[51] Z. M. Abdullah Al, K. Thapa, and S.-H. Yang, "Improving r peak detection in ECG signal using dynamic mode selected energy and adaptive window sizing algorithm with decision tree algorithm," *Sensors*, vol. 21, no. 19, p. 6682, Oct. 2021.
[52] A. Saleh, "Implementasi metode klasifikasi naive Bayes dalam memprediksi besarnya penggunaan listrik rumah tangga," *Creative Inf. Technol. J.*, vol. 2, no. 3, pp. 207–217, 2015.
[53] I. Wannawijit and T. Yingthawornsuk, "Classifying irregular ECG samples using KNN," in *Proc. 20th Pattaya Int. Conf. Sci., Eng., Technol. Natural Resour.*, Aug. 2020, pp. 17–21.
[54] S. Dehuri, A. K. Jagadev, and S.-B. Cho, "Epileptic seizure identification from electroencephalography signal using DE-RBFNs ensemble," *Proc. Comput. Sci.*, vol. 23, pp. 84–95, Jan. 2013.
[55] M. Rashid, M. Mustafa, N. Sulaiman, N. R. H. Abdullah, and R. Samad, "Random subspace K-NN based ensemble classifier for driver fatigue detection utilizing selected EEG channels," *Traitement du Signal*, vol. 38, no. 5, pp. 1259–1270, Oct. 2021.
[56] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 8, pp. 832–844, Aug. 1998.
[57] M. Rashid, B. S. Bari, M. J. Hasan, M. A. M. Razman, R. M. Musa, A. F. Ab Nasir, and A. P. P. A. Majeed, "The classification of motor imagery response: An accuracy enhancement through the ensemble of random subspace k-NN," *PeerJ Comput. Sci.*, vol. 7, p. e374, Mar. 2021.
[58] Z. Mushtaq, M. F. Ramzan, S. Ali, S. Baseer, A. Samad, and M. Husnain, "Voting classification-based diabetes mellitus prediction using hypertuned machine-learning techniques," *Mobile Inf. Syst.*, vol. 2022, pp. 1–16, Mar. 2022.
[59] J. Rahul, L. D. Sharma, and V. K. Bohat, "Short duration vectorcardiogram based inferior myocardial infarction detection: Class and subject-oriented approach," *Biomed. Eng./Biomedizinische Technik*, vol. 66, no. 5, pp. 489–501, Oct. 2021.
[60] M. Dey, N. Omar, and M. A. Ullah, "Temporal feature-based classification into myocardial infarction and other CVDs merging CNN and bi-LSTM from ECG signal," *IEEE Sensors J.*, vol. 21, no. 19, pp. 21688–21695, Oct. 2021.
[61] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, P. S., and R. B. Pachori, "Efficient detection of myocardial infarction from single lead ECG signal," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102678.
[62] V. Jahmunah, E. Y. K. Ng, R.-S. Tan, S. L. Oh, and U. R. Acharya, "Explainable detection of myocardial infarction using deep learning models with grad-CAM technique on ECG signals," *Comput. Biol. Med.*, vol. 146, Jul. 2022, Art. no. 105550.
[63] W. Li, Y. M. Tang, K. M. Yu, and S. To, "SLC-GAN: An automated myocardial infarction detection model based on generative adversarial networks and convolutional neural networks with single-lead electrocardiogram synthesis," *Inf. Sci.*, vol. 589, pp. 738–750, Apr. 2022.

[64] W. Liu, M. Zhang, Y. Zhang, Y. Liao, Q. Huang, S. Chang, H. Wang, and J. He, "Real-time multilead convolutional neural network for myocardial infarction detection," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1434–1444, Sep. 2018.
[65] K. Jafarian, V. Vahdat, S. Salehi, and M. Mobin, "Automating detection and localization of myocardial infarction using shallow and end-to-end deep neural networks," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106383.
[66] Y. Zhao, J. Xiong, Y. Hou, M. Zhu, Y. Lu, Y. Xu, J. Teliewubai, W. Liu, X. Xu, X. Li, Z. Liu, W. Peng, X. Zhao, Y. Zhang, and Y. Xu, "Early detection of ST-segment elevated myocardial infarction by artificial intelligence with 12-lead electrocardiogram," *Int. J. Cardiol.*, vol. 317, pp. 223–230, Oct. 2020.
[67] B. Liu, J. Liu, G. Wang, K. Huang, F. Li, Y. Zheng, Y. Luo, and F. Zhou, "A novel electrocardiogram parameterization algorithm and its application in myocardial infarction detection," *Comput. Biol. Med.*, vol. 61, pp. 178–184, Jun. 2015.
[68] T. Q. Le, S. T. S. Bukkapatnam, B. A. Benjamin, B. A. Wilkins, and R. Komanduri, "Topology and random-walk network representation of cardiac dynamics for localization of myocardial infarction," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 8, pp. 2325–2331, Aug. 2013.
[69] L. Sun, Y. Lu, K. Yang, and S. Li, "ECG analysis using multiple instance learning for myocardial infarction detection," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 12, pp. 3348–3356, Dec. 2012.
[70] M. Arif, I. A. Malagore, and F. A. Afsar, "Detection and localization of myocardial infarction using K-nearest neighbor classifier," *J. Med. Syst.*, vol. 36, no. 1, pp. 279–289, Feb. 2012.
[71] L. D. Sharma and R. K. Sunkaria, "Inferior myocardial infarction detection using stationary wavelet transform and machine learning approach," *Signal, Image Video Process.*, vol. 12, no. 2, pp. 199–206, Feb. 2018.
[72] N. Strodthoff and C. Strodthoff, "Detecting and interpreting myocardial infarction using fully convolutional neural networks," *Physiolog. Meas.*, vol. 40, no. 1, Jan. 2019, Art. no. 015001.

M. KRISHNA CHAITANYA received the M.Tech. degree from the Kakatiya Institute of Technology and Science, Warangal, Telangana, India, in 2008. He is currently pursuing the Ph.D. degree with the School of Electronic Engineering, VIT-AP University. His research interests include biomedical signal processing, deep learning, machine learning, optimization techniques, and ECG signal filtering.

LAKHAN DEV SHARMA (Senior Member, IEEE) received the Ph.D. degree from the Dr. B. R. Ambedkar National Institute of Technology Jalandhar, India, in 2018, and the M.Tech. degree from the Atal Bihari Vajpayee-Indian Institute of Information Technology and Management, Gwalior, India, in 2012. He has been an Associate Professor with the School of Electronics Engineering, VIT-AP University, Amaravati, India, since 2020. His research interests include biomedical signal and image processing, machine learning, and deep learning. He is a reviewer of many SCI/SCIE journals at IEEE, Elsevier, and Springer.