Single-Beat Myocardial Infarction Detection and Localization Using PSO-Optimized Extra Trees on 15-Lead ECG

Sourav Kumar Mukhopadhyay, and Sridhar Krishnan, *Senior Member, IEEE*

Abstract—With the advancement of wearable healthcare devices in recent times, there is a growing potential to revolutionize the current strategies of the detection of cardiac anomalies such as myocardial infarction (MI), making it viable to monitor out-of-the-hospital patients using conventional as well as unconventional and pseudo-electrocardiogram (ECG) leads. However, most of the MI detection and localization techniques which are available in the literature to date are based on either all the 12 conventional ECG leads, or subsets of thereof. A few single-lead ECG-based techniques are also there, but their accuracies are not satisfactory. On the contrary, this paper demonstrates a machine learning model that classifies MI from a single ECG beat of *any of the 15 ECG-leads*, including the 12 conventional leads and 3 vectorcardiogram (VCG)-leads, as available in the publicly accessible Physikalisch-Technische Bundesanstalt (PTB) Diagnostic ECG Database. In this proposed technique, first, the R-peaks are detected, and the arrhythmic beats are identified and excluded using a Ramanujan filter bank-based periodicity estimation technique. Next, a few hand-crafted features are extracted from the detected ECG beats and a set of important features are identified. Finally, Extra Trees classifier-based machine learning models with optimized hyperparameters, obtained using the particle swarm optimization technique, are developed utilizing these hand-crafted features for the detection and localization of MI. High F1-score and Matthew's correlation coefficient corroborate the reliability of the proposed algorithm in both intra- and inter-patient paradigms.

Index Terms—15-lead ECG, Extra Trees classifier, hand-crafted features, Matthew's correlation coefficient, multi-class MI, particle swarm optimization, Ramanujan filter-bank, vectorcardiogram leads.

I. INTRODUCTION

IN Ontario, Canada, emergency patients face an average hospital admission wait time of 22 hours, which is 2.75 times the provincial target, and a median wait of 30 weeks to see a specialist [1], [2]. The financial burden of myocardial infarction (MI) in just one Canadian province is significant [3], and while U.S. statistics are slightly better, they remain concerning [4]. At the same time, the U.S. wearable healthcare market is projected to reach USD 18.09B by 2025, reflecting a growing trend of individuals taking proactive health measures. Atherosclerosis, the buildup of fats in coronary arteries, can lead to MI due to restricted blood flow. Early MI detection greatly improves survival and prevention. Wearable devices like Apple, Fitbit, Samsung watches, and Hexoskin smart garments now capture single-lead ECG signals—typically pseudo-lead I or unconventional leads [5], [6]—making cardiac abnormality detection from such signals a growing research priority.

Automated MI detection and localization algorithms can be broadly classified into two categories: those using (i) multi-lead ECG [7]-[21], [43], and (ii) single-lead ECG [22]-[28], [42], based on either classical machine learning or neural networks. Grouping all MI sub-classes against healthy controls (HC) defines detection, while classifying individual MI sub-classes by location against HC defines localization. If training and testing are done on the same subject, it is intra-patient; testing on unseen subjects is inter-patient.

Han et al. [13] proposed a residual network with LSTM for detecting and localizing five types of MI using all 12 conventional ECG leads, evaluating separate models on two databases under both intra- and inter-patient paradigms. Cao et al. [18] introduced an ML-Net-based method using four ECG leads for MI detection and partial localization (three out of six classes), claiming a balance between accuracy and model size. However, [18] offers neither high localization accuracy nor a compact model compared to other state-of-the-art approaches. Overall, the requirement of all the 12 ECG leads or even a group of ECG leads for the detection and localization of MI restricts its applicability only within the hospital or clinical settings.

Compared to multi-lead, the number of published-literature on the detection and localization of MI using single-lead ECG signal to date, is significantly less. In [22], Sahu et al. have proposed two hybrid networks composed of (i) convolutional neural network (CNN) with LSTM and (ii) CNN with gated recurrent unit (GRU) for the detection and localization of MI using single lead ECG. In [22], first, the ECG signal is classified into three categories: HC, MI and non-MI, and then the MI class of ECG is sub-classified into different groups. A number of different combinations of ECG-leads and models have been suggested in [22] for the detection and localization of MI in intra- and inter-patient paradigms. A deep neural network (DNN)-based MI detection and localization technique

from single lead ECG signal utilizing Ramanujan sums and wavelet transform coefficients is proposed in [27].

Vectorcardiogram (VCG) technique records the electrical activity of the heart in a three-dimensional coordinate system: V_x, V_y and V_z. Even though ECG is more frequently used in clinical settings, VCG can also provide additional insights into certain types of cardiac abnormalities such as bundle branch blocks, ventricular hypertrophy and MI. A few MI detection and localization techniques using VCG have also been proposed in the literature [29]-[34]. A depth-wise separable convolution network has been proposed by Xiong et al. in [33] for the detection and localization of 11 types of MI using VCG signals. Another technique employing wavelet transformation, tucker decomposition and TreeBagger classifier has been proposed by Zhang et al. in [32] for the same purpose as in [33]. However, all these VCG-based MI detection and localization techniques mentioned above are trained and tested in intra-patient paradigm only.

A thorough and in-depth literature review suggests that (i) the accuracy of detection and localization of up to 11 classes of MI is very high in intra-patient paradigm, but it drops drastically in inter-patient setting, (ii) the accuracies are high while using 12 or multi-lead ECG signals compared to single lead ECG, but a12 or multi-lead MI detection and localization system is only suitable to use in hospital setting, and (iii) the inter-patient MI detection and localization performance of most of the existing single-lead ECG-based techniques are not satisfactory. In contrast, most of the wearable ECG monitoring devices record single-lead ECG signals, and that too from unconventional locations on the torso. If the correlation between the recorded ECG signals from an unconventional location on the torso and a standard ECG lead position is found to be high, then that unconventional ECG lead is called a pseudo-ECG lead [5].

In contrast to developing a new hardware system, this paper proposes a machine learning algorithm capable of detecting and localizing MI from a *single ECG beat* from *any of the 15 leads*-including the 12 standard ECG leads and 3 VCG leads-available in the publicly accessible PTB Diagnostic ECG Database. Unlike prior algorithms that require either 12 ECG leads [7]-[21], [43] or all three VCG leads [29]-[34], and those algorithms which are trained on a single, fixed ECG lead such as lead V1 [27], the proposed method is lead-agnostic. This approach allows the proposed algorithm to work reliably even when only a few leads, or just one, like V5 or V_x are available. Such flexibility is essential in real-world scenarios where ECG recordings may be incomplete, noisy, or derived from unconventional location on the torso, which is typical for wearable healthcare devices. It also ensures better generalizability across varied clinical setups. The ability to detect and localize MI from any lead rather than a fixed set adds significant flexibility and resilience to the model. By training the model across all 15 leads, this work significantly enhances the generalizability and robustness of MI detection and localization. It allows for diagnostic accuracy to be retained regardless of which lead is available at inference time, making the method suitable for both clinical and

wearable applications. This adaptability also ensures greater resilience in resource-constrained settings, where full multi-lead ECG acquisition may not be feasible. Ultimately, the proposed model's ability to operate effectively with any single lead introduces a valuable level of practicality and reliability for continuous, out-of-hospital cardiac monitoring.

In this research work, first, ECG leads are denoised using Butterworth filters. Then, the QRS-complexes detected using a Ramanujan filter bank (RFB)-based periodicity estimation technique. RFB technique also identifies the arrhythmic beats, which are set apart and are not considered in the subsequent steps. Next, (i) a number of time-domain, frequency-domain, geometrical and statistical features are extracted from each of the detected ECG-beats, (ii) missing feature values are imputed using a *k*-nearest neighbor (*k*-NN)-based technique, (iii) feature values are properly scaled, (iv) the feature-set is balanced, (v) a Random Forest-based feature selection technique is used to identify the dominant feature, (vi) these dominant features are then fed to an Extremely randomized Trees (Extra Trees) classifier with optimized hyperparameters, obtained using the particle swarm optimization technique, for the detection and localization of MI.

Compared to the state-of-the-art techniques, the main contributions and novelties of the proposed research work are: i) A machine learning algorithm is proposed that can detect and localize MI using only one ECG beat from *any of the 15 leads* (12 standard ECG + 3 VCG leads). ii) The model is lead-agnostic, and it can operate effectively regardless of which lead is available. iii) The method is designed to work in both intra and inter-patient paradigms, addressing a key weakness in prior literature. iv) The Ramanujan filter bank-based technique identifies and segregate arrhythmic beats from healthy and MI beats, which are then excluded from further processing. v) The algorithm is designed to be lightweight and fast, making it suitable for real-time implementation on wearable devices.

The paper is organized as follows. The proposed MI detection and localization technique is presented in Section II. Section III demonstrate the performance of the proposed technique, and the MI detection and localization performance of the proposed technique is compared to that of other state-of-the-art algorithms. The ablation study is introduced in Section IV. Finally, the technique is discussed and conclusions drawn in Section V and VI, respectively.

II. METHODOLOGY

The proposed MI detection and localization technique can be divided into four parts: i) ECG dataset description, ii) preprocessing, iii) the machine learning models, iv) QRS-T detection and feature extraction, v) feature conditioning and vi) model development.

A. ECG dataset description

The ECG signals are collected from Physikalisch-Technische Bundesanstalt database (PTBDB) [35]. The database contains 549 ECG records from 290 subjects out of which 148 and 52 subjects are MI and HC, respectively. Each

record contains 15 simultaneously recorded ECG-leads: 3 limb leads (lead I, II and III), 3 augmented limb leads (lead aVR, aVL and aVF), 6 precordial leads (lead V1 to V6) and 3 leads or VCG. The sampling rate of each of the ECG record is 1 kHz. There are 11 sub-classes of MI ECG in PTBDB. The number of ECG records and subjects in each of these HC and MI sub-classes are shown in Table I. It is seen in Table I, that the first six types of MI are the majority both in the number of files as well as in the number of subjects. Moreover, the last five sub-classes of MI files could not even be properly divided into a train-test ratio of 80%:20% in inter-patient paradigm. Therefore, only the HC and first six sub-classes of MI: A, AL, AS, I, IL and IPL are considered in this research. The total number of QRS-complexes in all the 15 leads of each of the MI sub-classes and HC are also shown in Table I.

TABLE I
ECG DATA USED IN THIS RESEARCH WORK

Type		# of files	# of subjects	# of QRS-complexes
HC		110	52	155578
MI	Anterior (A)	71	25	103865
	Antero-lateral (AL)	60	17	94460
	Antero-septal (AS)	108	28	160387
	Inferior (I)	94	35	182809
	Infero latera (IL)	73	25	116959
	Infero postero lateral (IPL)	35	9	39221
	Antero septo lateral (ASL)	2	1	4065
	Infero posterior (IP)	4	4	7425
	Lateral (L)	3	1	6885
	Posterior (P)	4	1	6900
	Postero lateral (PL)	5	2	10395

B. Preprocessing

A second order bidirectional Butterworth lowpass filter of cutoff frequency 100 Hz is used to expel out the high-frequency noise, a second order bidirectional Butterworth highpass filter of cutoff frequency 0.8 Hz is used to expel out the low-frequency noise, and another second order bidirectional Butterworth notch filter of cutoff frequency 60 Hz is used to expel out the powerline interference noise from each of the 15 leads of all the ECG data files.

C. The the machine learning models

The machine learning models which are developed in this research work for the detection and localization of six types of MI are divided into two groups: (i) intra-patient models (*inp*Mdl) and (ii) inter-patient models (*INP*Mdl). The protocols that have been followed in developing these two types of models are described below and also illustrated in Figure 1.

(i) inpMdl

In intra-patient paradigm, first, the features are extracted, as described in Section 2.2, from each of the ECG-beats of all the 15 leads of HC and six MI sub-classes and stored in a matrix named *FeatureMat*. Now, *FeatureMat* is split into training and testing sets in such a way that the training set receives 80% and the test set receives 20% features of all the ECG-beats of all the 15 leads across all the seven ECG classes (HC and six MI sub-classes). The training set is denoted as *Train_FeatureMat* and the test set is denoted as

Test_FeatureMat. In intra-patient paradigm the *Train_FeatureMat* and *Test_FeatureMat* share the same subjects' ECG features.
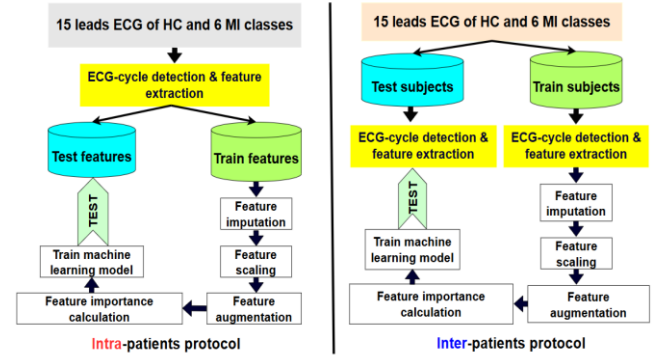

Fig. 1. Protocols for developing intra- (left) and inter-patient (right) MI classification models.

(ii) INPMdl

On the other hand, in inter-patient paradigm, the subjects in HC and each of the six MI sub-classes are first split into training and testing sets with an 80%:20% ratio. Then, features are extracted separately from the train and test subjects, ensuring that *Train_FeatureMat* and *Test_FeatureMat* remain independent of subject.

D. QRS-T detection and feature extraction

Ramanujan filter bank-based periodicity estimation technique, which is proposed in [36] for the detection of ECG R-peaks and identification of two types of arrhythmias: premature ventricular contraction (PVC) and atrial premature contraction (APC), is used in this research work. Ramanujan sums (RS) are defined as the nth power of the qth primitive roots of unity and is expressed as

$$c_q(n) = \sum_{\substack{p=1 \\ (p,q)=1}}^{q} \exp\left(2\pi j \frac{p}{q} n\right) \quad (1)$$

where, $(p,q) = 1$ implies that they are mutually prime and $-\infty \leq n \leq \infty$. The values of $c_q(n)$ can also be derived using Euler's equation; $e^{ix} = \cos x + i \sin x$. If the RS is treated as a digital filter with an impulse response $c_q(n)$, its frequency response over $0 \leq \omega < 2\pi$ can be expressed as

$$c_q = 2\pi \sum_{\substack{1 \leq p \leq q \\ (p,q)=1}} \sigma\left(\omega - \frac{2\pi p}{q}\right) \quad (2)$$

It can be inferred from Equation 2 that the values of c_q are non-zero at only those frequencies where p and q are mutually prime, and zero elsewhere. The center of the passband-frequency of Ramanujan filter is located around its co-prime frequency. A set of such RS filters for periods in the range 1 to P, where P denotes the period of the input signal, with impulse response $\{c_q(n)\}$ is called the Ramanujan filter bank (RFB), and its output is regarded as the time-period representation (TPR) of the input signal. More details about RFB can be found in [37]. Figure 2 shows an example of the detected normal and arrhythmic ECG beats exploiting the TPR of an ECG signal using. From this figure it can be seen that i) every

3

temporal peak in the TPR indicates an R-peaks, and ii) the amplitude of a local trough of the TPR for an arrhythmic QRS-complex is much higher than that of a normal one. Hence, a technique based on thresholding the TPR troughs is used on the TPR in [36] to identify the arrhythmic QRS-complexes. If the arrhythmic QRS-complexes are not pre-segregated, it will hinder the separation of HC and MI in the feature-space, which would lead to an overall degradation in the classification accuracy.
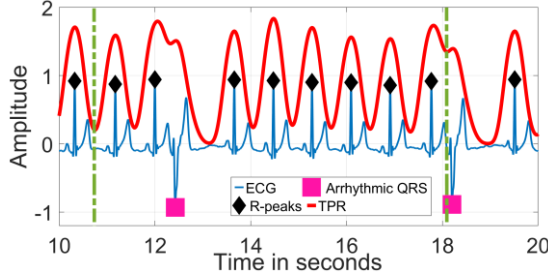

Fig. 2. Detected R peaks (in black-diamond marker) and arrhythmic beats (in magenta-square marker) exploiting the TPR (in red) of an ECG signal (File #s0287, Patient ID 150) taken from PTBDB. Dotted green vertical lines indicate the amplitudes of the local troughs of the TPR of normal and arrhythmic QRS-complexes.

It can be seen in Figure 2 that one complete ECG-beat resides within the one period of the TPR. Hence, the ECG-beats are extracted using the troughs of the TPR as the boundary points. Next, the number of samples in each of the ECG-beats is normalized to 1000 by a fast Fourier transform (FFT)-based interpolation technique [38], and its amplitude is normalized using 2-norm, i.e., Euclidean norm technique ($\|E\|_2 = \sqrt{\sum_{i=1}^{N}|E_i|^2}$, where E is one ECG-beat, N = 1000). Now, traversing to the left of the R-peak of an ECG-beat an instant of slope-reversal is identified a Q-peak, and the same in the right is identified an S-peak. A popular technique of detection of QRS-onset and offset points that is often found in the literature, is to find the next instance of slope reversal immediate to the left and right of the most recently detected Q and S-peaks, respectively. However, in various pathological ECGs, there may not be another instant of slope reversal until the P-offset point to the left and T-onset point to the right. Moreover, the presence of noise makes it even more challenging to accurately detect the QRS-onset and offset points. Therefore, a new technique is used in this research to detect the QRS-onset and offset points accurately. A straight line is drawn using the linear interpolation technique on the ECG-beat from the most recently detected Q-peak to a point which is 45ms left to that Q-peak. Now, the instant where the absolute value of the amplitude difference between the straight line and the corresponding segment of the ECG-beat is found to be the maximum, is identified as the QRS-onset. The QRS-offset point is also detected using the same technique drawing a straight line from the most recently detected S-peak to a point which is 45ms right to that S-peak. The average duration of a QRS-complex in adults is ~90ms. Therefore, selecting half of the duration of a QRS width, i.e., 45ms, allows sufficient boundary coverage for the detection of QRS-onset and offset points [36] and also minimizes false detection. Figure 3A (left) illustrates this proposed technique of

detection of QRS-onset and offset points. For an accurate detection of the T-peak, periodicity of the segment of the ECG-beat from the QRS-offset point to the end of the beat is calculated using RFB technique, and the index of the peak of the TPR is considered as the index of the T peak. The T-onset and offset points are identified using a similar technique which is used for the detection of QRS-onset and offset points, respectively, described above. Figure 3A (right) illustrates this proposed technique of detection of T, T-onset and offset points. The algorithm is also able to detect the ECG characteristic points in the presence of baseline wander and other high frequency noises. Figure 3B illustrates the detected peaks under noisy condition. After the detection of these ECG fiducial points, a total of 34 handcrafted features are extracted from an ECG-beat which are shown in Table II.
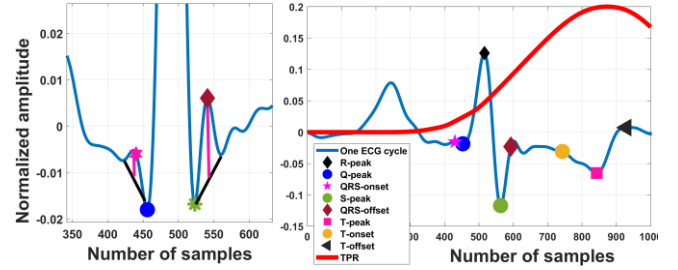

Fig. 3A. (left) The proposed technique of detection of QRS-onset and offset points, (right) detected T, T-onset, T-offset and other ECG fiducial points using the TPR of RFB (File #s0273lrem, patient ID 131).
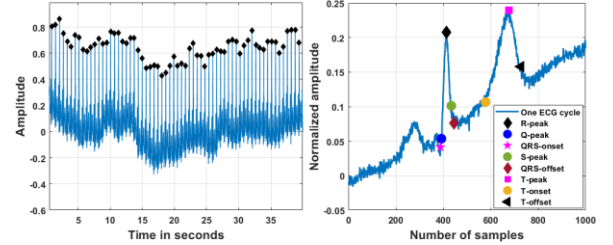

Fig 3B. (left) Detected R-peaks, and (right) other ECG fiducial points (File #s0162lrem, patient ID 38).

TABLE II
LIST OF HANDCRAFTED ECG FEATURES

Feature Type	Features (extracted from one ECG beat)
Time domain: 15 features in total	Amplitude of (i) R-peak (F1), (ii) Q-peak (F2), (iii) S-peak (F3) and (iv) T-peak (F11) Duration of (i) QRS-complex (F4), (ii) QR (F9), (iii) RS (F10), (iv) T-wave (F13), (v) QRS-onset to T-offset (F14), (vi) T-onset to T-peak (F18) and (vii) T-peak to T-offset point (F19). Polarity of the T-wave (F20), slope of the ST-segment (F23), square of the first difference of the ECG beat (F24) and ECG-lead number (F34).
Geometric: 8 features in total	Area of the (i) QRS-complex (F5) and (ii) T-wave (F12). Cosine angle between (i) Q, R and the isoelectric point (F6), (ii) R, S and the isoelectric point (F7), (iii) Q, R and S points (F8), (iv) T-onset, T-peak and the isoelectric point (F15), (v) T-peak, T-offset and the isoelectric point (F16), (vi) T-onset, T-peak and T-offset points (F17).
Frequency domain: 6 features in total	Maximum frequency component of FFT (F25), maximum signal power present in the power spectral density (PSD) (F26), total signal power present in the PSD (F27), median normalized frequency of the power spectrum (F28) and FFT (F29), total spectral entropy (F30).
Statistical: 5 features in total	Kurtosis (F31), skewness (F32), variance (F33), mean of the absolute value (F21) and standard deviation of the absolute (F22) value of one ECG beat.

The amplitudes of the ECG fiducial points are measured with respect to the isoelectric lines. The slope of the ST-segment could either be positive (i.e., upward ST-segment) or negative (i.e., downward ST-segment) or even flat; all three possibilities are considered in this research. The area of the QRS-complex and T-wave is computed using the trapezoidal method of numerical integration and is calculated as follows

$$\int_a^b f(x)dx \approx \frac{b-a}{2(N-1)} \sum_{n=1}^{N-1} \left(f(QRST_n) + f(QRST_{n+1}) \right) \quad (3)$$

where QRST represents either the QRS-complex or T-wave.

The maximum frequency component that is present in an ECG-beat is computed using FFT. The PSD of the ECG-beat is computed using a periodogram-based technique with Hamming window for the suppression of the spectral leakage. The periodogram is a non-parametric estimation of the PSD and is calculated as follows

$$P_e(f) = \frac{1}{Nf_s} \left| \sum_{n=1}^{N} h_n E_n e^{-2j\pi fn/f_s} \right|^2, \quad -\frac{f_s}{2} < f \le \frac{f_s}{2} \quad (4)$$

where f_s is the sampling frequency of the ECG signal and h_n is the window function. The spectral entropy is calculated as follows

$$H = -\sum_{m=1}^{N} P(m)\log_2 P(m) \quad (5)$$

where $P(m) = \frac{S(m)}{\sum_{i=1}^{N} S(i)}$, $S(m) = |X(m)|^2$, and $X(m)$ is the discrete Fourier transform E_n. Out of these 34 handcrafted ECG features shown in Table II, the cosine angle-based features are used in this research work for the first time to the best of our knowledge in classifying cardiac anomalies from the ECG signals. In [39], Mukhopadhyay et al. proposed a QRS-complex detection technique using a cosine angle threshold-based technique, and both in [36] and [39], it has been shown that the various angles between the Q, R and S peaks vary distinctly among the HC and other abnormal classes of ECG. The cosine angle between the Q, R and S peaks is calculated as follows, and these new cosine angle-based features are depicted in Figure 4.

$$\angle QRS = \cos^{-1}\left(\frac{\vec{QR} \cdot \vec{SR}}{\|\vec{QR}\| \cdot \|\vec{SR}\|}\right) \quad (6)$$
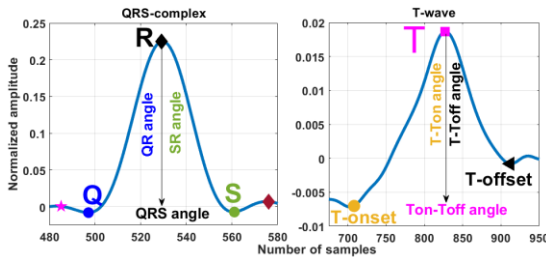

Fig. 4. The cosine angle-based ECG features.

E. Feature conditioning

Before feeding these handcrafted ECG features, i.e., Train_FeatureMat, to a classification module, a few important steps are needed to be followed. These steps are now described below.

(i)   Feature imputation

Deleting rows with missing feature values is a very common practice as many of the machine learning algorithms cannot handle the missing data. However, it might reduce data representativeness and introduce bias as well as. Imputing the missing feature values with the mean or median value of that feature is also very popular, but both these imputation techniques i) reduce the natural variability of the feature, and ii) pull the distribution of the feature towards its mean or median. In this research, the missing feature values in Train_FeatureMat are replaced using a k-nearest neighbors (k-NN)-based imputation technique. Five nearest neighbors have been used, and the Euclidean distance is calculated only over the available feature values as follows

$$d(A, B) = \sqrt{\sum_{i \in v} (A_i - B_i)^2} \quad (7)$$

where d is the Euclidean distance, A_i and B_i are the feature values, v is the total number of feature values in the ith row of Train_FeatureMat, and the missing value is imputed as follows

$$x_{ij} = \frac{1}{k} \sum_{k \in neighbors} x_{kj} \quad (8)$$

where x_{kj} is the imputed feature value, j is the index of the missing feature in the ith row of Train_FeatureMat, and k is the number of selected neighbors. A random forest-based and another multiple imputation by chained equations-based (which uses light gradient-boosting machine) imputation techniques are tested in this research for imputing the missing ECG features but, the k-NN imputer was found to be the optimum one in terms of accuracy and computational efficiency.

(ii)   Feature scaling

Scaling the feature values is of utmost importance to bring them within a common range and also to boost up the overall classification accuracy. Three different types of scaling techniques: (i) Z-score normalization, (ii) min-max scaling, and (iii) robust scaling have been tested in this research, and it has been found that the robust scaling technique works better than the other two methods in terms of getting a better classification accuracy. A plausible reason for the robust scaling technique to perform better in this research work is that its resistiveness to the outliers. The robust scaler technique scales the feature as follows

$$Train\_FeatureMat^i = \frac{Train\_FeatureMat^i - Median^i}{IQR^i} \quad (9)$$

where $Train\_FeatureMat^i$ is the ith feature, $Median^i$ is the median value of the ith feature and IQR^i is the interquartile range of the ith feature of $Train\_FeatureMat$.

(iii)   Feature augmentation

It is seen in Table I that inferior (I) and infero postero lateral (IPL) groups of MIs have the highest and least number of the QRS-complexes, respectively, in PTBDB. The ratio of

5

the number of QRS-complexes of I and IPL is 4.66:1. Hence, an effort has been made to mitigate the data imbalance. Four techniques: (i) synthetic minority oversampling technique (SMOTE), (ii) SMOTE and Tomek links technique, (iii) SMOTE with edited nearest neighbor (ENN), and (iv) adaptive synthetic sampling (ADASYN) [40] have been tested, and ADASYN technique has been chosen based on its performance. A synthetic sample is generated using ADASYN technique as follows

$$x_{syn} = x_i + \delta \cdot (x_{zi} - x_i) \tag{10}$$

where x_{syn} is a synthetic sample, x_{zi} is one of its k-nearest ($k = 5$) minority class neighbors of x_i, and δ ($\delta \in [0,1]$) is a random scaler from a uniform distribution.

(iv)   Feature importance calculation

It's also essential to assess the relevance of each of these 34 handcrafted ECG features in classifying HC and six MI sub-classes. Selecting the important features is crucial as it (i) enhance the performance of the machine learning model by eliminating the irrelevant features, (ii) reduces the computational complexity of the machine learning model, (iii) improves the interpretability of the model, and (iv) helps prevent overfitting. A Random Forest (RF)-based feature selection technique [41] has been used in this research work to evaluate the relevance of these handcrafted features. The RF-based technique of feature importance calculation starts with measuring the Gini importance, which is a measure of how much a feature contributes to reducing the impurities of the nodes across all trees in the forest. The Gini impurity of a node t is defined as

$$G(t) = 1 - \sum_{k=1}^{K} p_k^2 \tag{11}$$

where p_k is the probability of class k in node t, and K is the total number of classes. When a feature x_j is used for splitting a node, the reduction in impurity is expressed as

$$\Delta G_j(t) = G_j(t) - \left(w_L G(t_L) + w_R G(t_R) \right) \tag{12}$$

where w_L and w_R are the proportions of samples in the left and right child nodes. The overall feature importance is the normalized value of the cumulative sum of $\Delta G_j(t)$ across all nodes and trees and is calculated as follows

$$FI_j = \frac{\sum_{t \in nodes\ where\ x_j\ is\ used} \Delta G_j(t)}{\sum_k I_k} \tag{13}$$

where $\sum_k I_k$ represents the sum of the importance values of all the features. Figure 5A shows the importance of all the 34 handcrafted features extracted in this work.
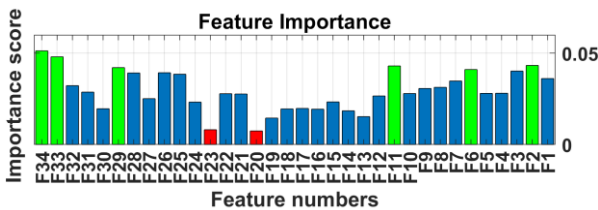

Fig. 5A. Importance of all the 34 handcrafted ECG features.

From this figure, it is seen that the importance of two

features: (i) F20 - polarity of the T-wave, and (ii) F23 - slope of the ST-segment are low compared to that of others. Hence, these two features are dropped from both the train and test sets of features. Among the rest, six features, which are F2 – amplitude of the Q-peak, F6 – the angle between the Q-peak, R-peak and the isoelectric line, F11 – amplitude of the T wave, F29 – median of the normalized frequency of FFT, F33 – variance of one ECG-beat, and F34 – the ECG lead number, carry high importance. To verify whether the importance of the extracted features changes due to imputation, the feature importances are also before imputation, and the results are shown in Figure 5B. From this figure, it can be seen that although the individual feature importance scores vary slightly, the imputation step does not significantly alter their overall importance.
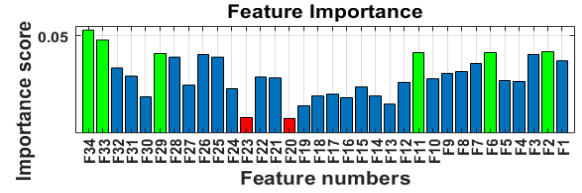

Fig. 5B. Importance of all the 34 handcrafted ECG features before imputation.

Shuffling the data is another important step before finally feeding the features to the machine learning model which help prevent bias in training and improves the model's generalizability. Therefore, the remaining 32 features of *Train_FeatureMat* are shuffled randomly. Now, *Train_FeatureMat*, is used to train machine learning models.

F. Model development

t-distributed stochastic neighbor embedding (tSNE) is a non-linear algorithm for dimensionality reduction and is suitable for visualizing high-dimensional data. In this research, tSNE algorithm is applied to the 32 handcrafted features for the visualization of the likelihood of separability of HC and six types of MI sub-classes in the feature space. First, the features of the six sub-classes of MI, i.e., A, AL, AS, I, IL, IPL, are grouped into a superclass of MI, and the tSNE algorithm is applied to the features of HC and MI superclass. Next, these six MI sub-classes are grouped into two as per their locations or origin. The A, Al and AS classes of MI are grouped into anterior wall MI (AWMI) class, and I, IL and IPL classes of MI are grouped inferior wall MI (IWMI). Then, the tSNE algorithm is applied to the features of HC, AWMI and IWMI. Finally, the tSNE algorithm is applied to the features of HC and the six MI sub-classes. Figure 6 shows the tSNE plots. It can be seen in Figure 6 that the (i) separability of HC and MI superclass is high, (ii) AWMI and IWMI classes are partly overlapped in the feature space, however (iii) the features of the six MI sub-classes are highly overlapped.
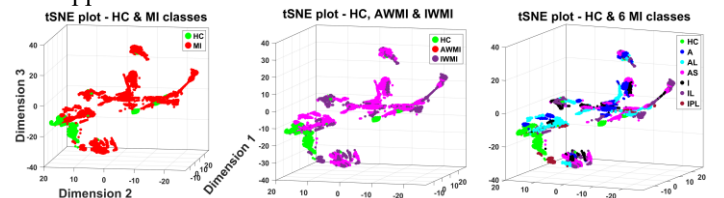
Fig. 6. tSNE plot of - HC and MI superclass features (left), HC, AWMI and IWMI features (middle), and the features of HC and 6 MI classes (right).

In total, 24 classifiers have been tested so as to find the best performing one for this research. Extremely randomized trees (Extra Trees) classifier was found to be the best performing classifier in terms of accuracy. Therefore, Extra Trees classifier-based models are developed for the detection as well as localization of MI in both intra and inter-patient paradigms. Extra Trees classifier is a tree-based ensemble learning method that produces multiple decision trees in a highly randomized fashion. The added randomness helps reduce the variance and also prevent the model from getting overfitted. Let us consider a feature set $X \in \mathbb{R}^{i \times j}$, where i is the number of observations and j is the number of features, and $Y \in \{1,2, \dots, K\}$ is the class labels. At each node, Extra Trees classifier choses m number of features $\{j_1, j_2, \dots, j_m\}$ randomly from X. Now, each feature is split using a random threshold value of $t_j \sim U(j^{max}, j^{min})$. For each feature j and its corresponding threshold value t_j, the impurity of the split is calculated as follows

$$I(j, t_j) = \frac{|D_{\text{left}}|}{|D|} G(D_{\text{left}}) + \frac{|D_{\text{right}}|}{|D|} G(D_{\text{right}}) \quad (14)$$

where D_{left} is the sample where the feature value is less than or equal to t_j, D_{right} is the sample where the feature value is greater than t_j, and $G(D_{left})$ and $G(D_{right})$ are the corresponding Gini impurities. Only the feature with minimum impurity is chosen. When all the decision trees are trained, for a given sample x, the prediction is made as follows

$$y_{\text{pred}} = \arg \max_{c \, \in \, \{1,2, \dots K\}} \sum_{t=1}^{T} I[T_t(x) = c] \quad (15)$$

where y_{pred} is the predicted class, I is an indicator function, whose value is 1 if true, 0 otherwise, K is the total number of classes, and T is the total number of trees.

There are a few hyperparameters of an Extra Trees classifier which are needed to be optimized so as to get the best possible classification performance. Particle swarm optimization (PSO) is a bio-inspired algorithm, which is efficient in finding the maximum or minimum of an objective function defined on a multidimensional space. In PSO, a group of plausible values of the hyperparameters, called the particles, move through the search space by following the highest performing particles. Each particle revises its position based on its own as well as neighbors' experience. The hyperparameters of the Extra Trees classifier, whose values are optimized are the (i) number of trees, (ii) quality of a split, (iii) maximum depth of each decision tree, (iv) minimum number of samples required to split an internal node, (v) minimum number of samples required to be at a leaf node, (vi) maximum number of features to consider, (vii) maximum number of leaf nodes, and (viii) weights to assign to the classes. Matthew's correlation coefficient (MCC) is a performance metric that considers true and false positives and negatives and provides a balanced measure even for an imbalanced dataset. MCC is used as the fitness function during the optimization process, i.e., it has

been tried to maximize value of MCC. In PSO, the velocity and position of a particle is updated as follows

$$v_i(t + 1) = \omega \cdot v_i(t) + c_1 r_1 \big(p_i - x_i(t)\big) + c_2 r_2 \big(g - x_i(t)\big) \quad (16)$$
$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (17)$$

where p_i and g are the local and global best solutions across the swarm. In this research, the following PSO parameter values are used:

Swarm size (i): varied from 1 to 50.

Maximum number of iterations (t): set to 100 for each swarm size.

Inertia weight (ω): set to 0.5 to balance exploration and exploitation.

Acceleration coefficients (c_1 and c_2): both set to 1.5 to equally weigh personal and global best experiences.

Random coefficients (r_1 and r_2): sampled from the uniform distribution $\sim U(0,1)$.

Now, this iterated updating of v_i and x_i refines the values of the hyperparameters so as to minimize the validation loss of the objective function as follows

$$f(x) = \mathcal{L}\big(ExT(x)\big) \quad (18)$$

where $\mathcal{L}$ is the validation loss of the Extra Trees classifier - ExT, with a set of values of hyperparameters defined in x. Three other hyperparameter optimization techniques based on grid search, Genetic and Bayesian algorithms have also been tested. However, the hyperparameter values optimized using PSO were found to yield better classification accuracy in this research. Although grid search and Bayesian optimization are widely used, they often become computationally expensive in high-dimensional hyperparameter spaces. On the other hand, PSO is a gradient-free, population-based metaheuristic that requires fewer objective function evaluations to converge to an optimal solution. PSO's ability to escape local minima and rapidly explore the solution space makes it suitable for tuning a complex model like Extra Trees. Furthermore, unlike Genetic algorithms, PSO does not require crossover and mutation operations, which makes it simpler to implement and computationally efficient.

Classification models that are developed in this research work, are summarized in Table III and Figure 7. Each of these five models are developed following the protocols laid out in Section II, and all the models need only *one* ECG-beat for the task of classification. Models *inp*Mdl#1 and *inp*Mdl#2 (with '*inp*' in italics and lower-case letters) are developed for classifying MI in intra-patient paradigm while *INP*Mdl#1, *INP*Mdl#2 and *INP*Mdl#3 (with '*INP*' in italics and upper-case letters) are developed for classifying MI in inter-patient paradigm. All these five models need only one ECG beat at the input for classifying MI. Now, models (i) *inp*Mdl#1, *inp*Mdl#2 and *INP*Mdl#1 need one ECG beat from any of the 15 ECG leads, and (ii) *INP*Mdl#2 and *INP*Mdl#3 need one ECG beat of a particular ECG lead. While evaluating the performances of the model *INP*Mdl#1 on all the 15 ECG leads, it is found that the model performs best on lead V5, which is denoted as the 'best performing lead'.

TABLE III
DEVELOPED EXTRA TREES CLASSIFIER-BASED MODELS

| Model	Required input	Purpose
*inp*Mdl#1	One ECG-beat of any of the 15 leads	Intra-patient MI detection.
*inp*Mdl#2	-do-	Intra-patient HC and 6 sub-classes of MI localization.
*INP*Mdl#1	-do-	Inter-patient MI detection.
*INP*Mdl#2	One ECG-beat of the best performing lead	Inter-patient HC and 6 sub-classes of MI localization.
*INP*Mdl#3	-do-	Inter-patient HC and 2 classes of MI (AWMI and IWMI) localization.
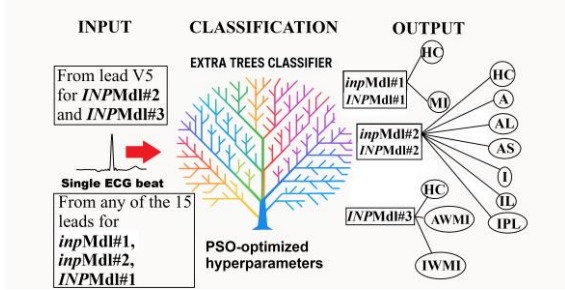
Fig. 7. Input requirements and the outputs of all the five developed classification models.

III. EXPERIMENTAL RESULTS AND COMPARISON

A. Evaluation metrics:

A few reliable mathematical tools are there to evaluate the performance of a classification module precisely. The tools which are considered here are: accuracy (A_{cc}), sensitivity (S_e), specificity (S_p), $F1$ score and Matthew's correlation coefficient (MCC), and are evaluated as follows

$$A_{cc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{19}$$

$$S_e = \frac{TP}{TP + FN} \tag{20}$$

$$S_p = \frac{TN}{TN + FP} \tag{21}$$

$$F1\ score = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{22}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{23}$$

Where TP, TN, FP, and FN are true positive, true negative, false positive and false negative predictions, respectively [41]. The value of MCC varies between ± 1. An MCC value of 1, 0 and -1 represent that all the predictions are correct, random guessing and total disagreement, respectively. On the other hand, $F1\ score$ varies between 0 to 1, where 1 represents that all the predictions are accurate and 0 represents that the model fails to identify any TPs. As can be seen in Equation 22, that the calculation of $F1\ score$ does not include the TN values. Therefore, relying only on the $F1\ score$; particularly when the data is highly imbalance, could lead to an unreliable assessment of the performance of the model. The value of MCC is more realistic in this context as it considers all the elements of the confusion matrix. The number of floating-point operations required per prediction (FLOP/prediction) as well the size of a classification model are important figure of merits, which indicate the complexity as well as the feasibility

of an algorithm to be implemented on resource-constrained hardware platforms such as wearable devices. For an Extra Trees classifier (i) the required numbers of FLOP per prediction during testing are defined as $(T \times D)$, and (ii) the time complexity during testing is $O(T \times D)$ per prediction, where T is the number of trees in the model, and D is the average depth of a tree. This suggests that our models require only $T \times D$ comparisons per prediction. The higher the values of A_{cc}, S_e, S_p, $F1\ score$ and MCC, and the lower the values of FLOP/prediction, O, and size of the model, the better the overall efficiency. Receiver operating characteristic (ROC) curve, area under the curve (AUC) and precision-recall curves are three another figures of merit of a classifier. The ROC curve is a plot of the rate of FP versus the rate of TP that shows the efficiency of a classifier in identifying different classes. The closer the ROC plot gets to the top-left corner of the plot, the better the performance of the classifier. The value of AUC ($0 \leq AUC \leq 1$) summarizes the ROC curve. An AUC value of 1 represents an accurate classification, and an AUC value of 0.5 indicates that the performance of the classifier is no better than that of a random guess. On the other hand, the precision-recall curve plots the recall (also knows as sensitivity) versus precision (the proportion of true positive predictions of a particular class among all positive predictions) for different threshold values of the classification. Unlike the ROC curve, the closer the precision-recall curve gets to the top-right corner of the plot, the better the performance of the classifier. The average values of the performance evaluation metrics of all the developed models are shown in Table IV. The last rwo of Table IV only shows the required time per classification task. However, approximately ~0.01s is also required for the preprocessing and feature extraction from an ECG beat.

B. Performance of the models:

Model *inp*Mdl#2 performs a classification task of seven classes. It not only segregates the MI class of ECG from HC, but also it localizes six MI sub-classes in intra-patient paradigm. The confusion matrices and the performance of the model *inp*Mdl#2 is shown in Table V. In total, 159429 ECG-beats, out of which 30631 are HC, 16093 are A, 17980 are AL, 31416 are AS, 34380 are I, 21205 are IL, and 7724 are IPL classes, taken from all the 15 ECG leads are tested with model *inp*Mdl#2. From Tables V, it can be seen that the localization efficiencies of all the 6 MI sub-classes are high ($A_{cc} > 99\%, MCC > 98\%$) in intra-patient paradigm.

Similarly, 11921 ECG-beats, out of which 2250 are HC and 9671 are MI classes, taken from all the 15 ECG leads are tested with the model *INP*Mdl#1 for the detection of MI. The confusion matrix as well as the performance evaluation metrics of *INP*Mdl#1 for the detection of MI from each of the 15 ECG leads are shown in Table VI in the order of best to worst value of MCC. From this table it can be seen that *INP*Mdl#1 performs: (i) best on the ECG leads V5, V3, V_x, V1 and V2, and (ii) satisfactory on ECG leads II, aVR, I, aVL, V6, V4 and V_y. However, the performance of *INP*Mdl#1 on leads aVF, III, and V_z are not satisfactory. Out of 15

8

standardized ECG leads, MI can be detected from five leads (lead V5, V3, V_x, V1 and V2) with $MCC \geq 90\%$, from six leads (lead II, aVR, I, aVL, V6 and V4) with $80\% \leq MCC < 90\%$, and from one lead (V_y) with MCC of 73.58% in inter-patient paradigm. Moreover, if the first 12 good-performing ECG-leads, as shown in Table VI, are considered, then the average values of A_{cc} and MCC would reach to 96.47% and 89.64%, respectively. These statistics suggest that the proposed INPMdl#1 model can detect MI from most of standardized ECG-lead positions as well as their pseudo-ECG leads with high reliability.

It is evident from Table VI that lead V5 is the best performing ECG lead, i.e., MI can be detected from lead V5 with highest reliability using the proposed technique in inter-patient paradigm. Now, the confusion matrices of the model INPMdl#2, which localizes HC and 6 MI sub-classes from lead V5 is shown in Table VII. This table shows that the model INPMdl#2 could not not localize these 6 MI sub-classes efficiently even though the average value of A_{cc} is 79.07%. Nevertheless, this confusion matrix also reveals that (i) the HC group of ECG-beats are identified precisely, (ii) the measure of A_{cc} does not convey significant information in the event of multiclass classification, whereas the measure of MCC provides better insights into the overall performance, (iii) perhaps the features of A, AL and AS, and I, IL and IPL types of MI are highly overlapped, as shown in Figure 6. This intuition is likely to come true as A, AL, and AS types of myocardial infarction originate in the anterior wall of the myocardium (AWMI), while the I, IL, and IPL types originate in the inferior wall (IWMI). This intuition intrigued us to develop another model; INPMdl#3 - for the localization of AWMI (the group of A, AL, and AS types of MI) and IWMI (the group of I, IL, and IPL types of MI). The confusion matrix of the model INPMdl#3 is shown Table VIII. This table shows that, although the HC and six MI sub-classes could not be localized efficiently, the AWMI and IWMI classes can be localized with high accuracy. The ROC curves, AUC values and the precision-recall curves of INPMdl#1 are shown in Figure 8.

TABLE IV
THE PERFORMANCES OF THE DEVELOPED MODELS

Models	inpMdl#1	inpMdl#2	INPMdl#1	INPMdl#2	INPMdl#3
A_{cc}(%)	99.63	99.67	93.22	79.07	99.83
S_e(%)	99.14	98.85	88.44	26.18	99.66
S_p(%)	99.14	99.80	88.44	73.28	99.81
$F1\ score$(%)	99.41	98.85	88.83	21.99	99.75
MCC(%)	98.82	98.65	77.66	13.37	99.60
Size (MB)	27.9	26.3	27.6	18.9	18.9
$O(T \cdot D)$	$O(1500)$	$O(1500)$	$O(750)$	$O(750)$	$O(750)$
FLOP/ Prediction	1500	1500	750	750	750
Time/ prediction (μs)	1.32	1.63	2.57	4.56	3.21

C. Computer specification:

The proposed MI detection and localization models are developed and tested on Python 3.13 platform with a personal computer configured with Windows 11 operating system, Intel Core i7-14700 CPU and 16GB RAM. Several Python libraries have been utilized in developing the proposed models including scikit-learn (with ADASYN, ExtraTreesClassifier, RepeatedStratifiedKFold), LazyPredict, Optunity and Pyswarm.

TABLE V
CONFUSION MATRIX AND THE PERFORMANCE OF THE MODEL INPMDL#2

		Predicted class						
		HC	A	AS	AL	I	IL	IPL
Actual class	HC	30373	30	23	32	128	37	8
	A	23	15932	27	49	23	33	6
	AL	17	26	17846	43	26	18	4
	AS	89	113	89	30955	106	55	9
	I	137	49	80	103	33882	114	15
	IL	39	30	21	34	54	21019	8
	IPL	16	16	9	14	45	31	7593

	A_{cc} (%)	S_e (%)	S_p (%)	$F1$ (%)	MCC (%)	AUROC
HC	99.64	99.16	99.75	99.06	98.83	1.00
A	99.73	99.00	99.82	98.68	98.54	1.00
AL	99.76	99.25	99.82	98.94	98.80	1.00
AS	99.54	98.53	99.79	98.83	98.54	1.00
I	99.45	98.55	99.69	98.72	98.37	1.00
IL	99.70	99.12	99.79	98.89	98.71	1.00
IPL	99.89	98.30	99.97	98.82	98.76	1.00

TABLE VI
PERFORMANCE EVALUATION OF THE MODEL INPMDL#1

		Predicted class	
		HC	MI
Actual class	HC	1817	433
	MI	375	9296

Performance of INPMdl#1 on 15 ECG leads					
Lead	A_{cc} (%)	S_e (%)	S_p (%)	$F1$ (%)	MCC (%)
V5	100.0	100.0	100.0	100.0	100.0
V3	99.74	99.84	99.84	99.59	99.18
V_x	99.50	99.69	99.69	99.19	98.39
V1	99.25	99.02	99.02	98.77	97.55
V2	96.78	98.00	98.00	95.14	90.71
II	95.75	97.38	97.38	93.57	87.90
aVR	96.24	90.00	90.00	93.31	87.44
I	95.36	97.14	97.14	93.04	86.96
aVL	95.86	94.63	94.63	93.41	86.90
V6	95.63	91.16	91.16	92.57	85.28
V4	93.08	95.74	95.74	90.02	81.81
V_y	90.49	90.81	90.81	86.10	73.58
aVF	82.58	75.19	75.19	73.39	47.18
III	81.03	49.92	49.92	44.76	-1.71
V_z	77.26	48.11	48.11	44.65	-7.20

TABLE VII
CONFUSION MATRIX AND PERFORMANCE OF THE MODEL INPMDL#2 ON LEAD V5

		Predicted class						
		HC	A	AS	AL	I	IL	IPL
Actual class	HC	150	0	0	0	0	0	0
	A	0	0	0	149	0	0	0
	AL	0	32	4	98	0	0	0
	AS	0	135	0	17	0	0	0
	I	0	0	0	0	38	17	0
	IL	0	0	0	2	97	0	0
	IPL	0	0	0	0	0	43	0

Performance of INPMdl#2						
	A_{cc} (%)	S_e (%)	S_p (%)	$F1$ (%)	MCC (%)	AUROC
HC	100	100	100	100	100	1.00
A	59.59	0	73.62	0	-25.28	0.29
AL	83.38	2.99	1	5.80	15.77	0.74
AS	50.90	11.18	60.48	8.13	-23.67	0.59
I	85.42	69.09	86.66	40.00	37.72	0.92

IL	79.67	0	91.22	0	-10.98	0.21
IPL	94.50	0	100	0	0	1.00
Mean	79.07	26.18	73.28	21.99	13.37	0.68

TABLE VIII

CONFUSION MATRIX OF THE MODEL *INP*MDL#3 ON LEAD V5

	Predicted class					
		HC	AWMI	IWMI		
Actual	HC	**150**	0	0		
	AWMI	0	**435**	0		
	IWMI	0	2	**195**		

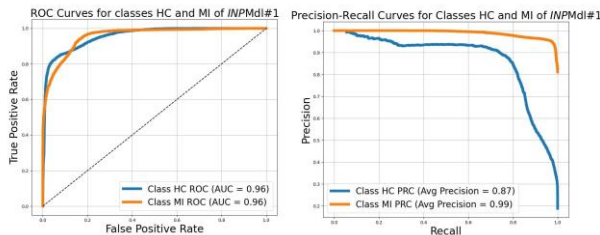| Performance of *INP*Mdl#3						
	A_{cc} (%)	S_e(%)	S_p(%)	$F1$ (%)	MCC (%)	AUROC
HC	100	100	100	100	100	1.00
AWMI	99.74	100	99.42	99.77	99.48	0.78
IWMI	99.74	98.98	100	99.49	99.32	0.40
Mean	99.83	99.66	99.81	99.75	99.60	0.73

Fig. 8. (Left) ROC curves of *INP*Mdl#1, (Right) Precision-recall curves of *INP*Mdl#1.

Due to its superior performance, the Extra Trees classifier has been used to develop models in this research. Extra Trees is a bagging-type ensemble learning method. In contrast, AdaBoost and XGBoost are boosting-type ensemble methods with similar foundations. The performance of these two classifiers has also been evaluated in this study. As an example, model *INP*Mdl#3 was retrained using both AdaBoost and XGBoost, and the results are presented in Table IX. This table demonstrates that the performance of these two classifiers does not match that of the Extra Trees classifier.

TABLE IX

CONFUSION MATRIX OF THE MODEL *INP*MDL#3 ON LEAD V5 USING ADABOOST AND XGBOOST CLASSIFIERS

	AdaBoost Classifier				XGBoost Classifier		
	HC	AWMI	IWMI		HC	AWMI	IWMI
HC	150	0	0		125	0	25
AWMI	0	435	0		435	0	0
IWMI	17	99	81		96	0	101
	A_{cc} (%)	S_e(%)	S_p (%)		$F1$ (%)	MCC (%)	AUROC
AdaBoost	90.11	80.37	89.59		80.90	76.12	0.74
XGBoost	28.89	28.89	28.89		28.89	24.56	0.46

D. Performance comparison:

Performances of cardiac anomaly detection techniques depend on several crucial factors such as the databases used, number of signals or subjects considered in the training or testing phases and the number of target classes. In Table X, the performances of the developed Extra Trees classifier-based intra- and inter-patients MI detection and localization models are compared to fourteen state-of-the-art algorithms that were published in the literature between 2021 and 2025. It can be seen from Figure 6 that the features of the MI sub-classes are highly overlapped. Nevertheless, our proposed models as well as most of the techniques which are proposed in the literature to date achieve high accuracy (> 99%) in localizing up to 11 classes of MI in intra-patient paradigm. But the accuracy

drops drastically in inter-patient paradigm. This strongly suggests that the models, which are designed for intra-patient paradigm are highly biased as they already have the information about the test data. This happens through data leakage during the split of train and test data. Therefore, the true performance of a model can be evaluated using an inter-patient paradigm.

TABLE X

PERFORMANCE COMPARISON OF THE PROPOSED AND EXISTING MI DETECTION-LOCALIZATION MODELS

Method, year	Data and lead	Paradigm	A_{cc} (%)	S_e (%)	S_p (%)
MI detection and localization from multi-lead ECG					
[43] CNN, 2025	12 leads, 2s per lead	Intra, LOC	99.94	99.94	99.99
		Inter, LOC	69.18	66.82	93.84
[19] NN, 2024	1beat, 8 leads	Intra, LOC	99.90	99.49	99.94
[17] CNN, 2024	2s ECG, 4 leads	Intra, DET	97.00	98.10	94.00
[12] DNN, 2023	10s ECG, 12 leads	Intra, DET	82.98	NM	NM
[13] LSTM, 2023	12 leads	Intra, LOC	99.94	99.99	99.80
		Intra, LOC (HC, 5 MI)	99.69	99.58	99.94
		Inter, LOC (HC, 5 MI)	67.89	63.16	93.51
[21] DNN, 2022	1 ECG beat, 12 leads	Intra, DET	95.57	98.01	86.01
		Inter, LOC (HC, 4 MI)	68.90	55.35	67.96
[15] k-NN, 2022	12 ECG lead	Intra, DET	99.83	99.87	99.73
		Intra, LOC (HC, 11 MI)	99.75	99.70	99.97
[18] CNN, 2021	4 ECG lead	Intra, DET	96.65	94.30	97.72
		Inter, LOC (HC, 3 MI)	66.85	62.64	68.70
MI detection and localization from single-lead ECG					
[27] DNN, 2025	1 ECG beat, either lead V1, or V4 or V5	Intra, DET	99.93	99.96	99.86
		Intra, LOC (HC, 5 MI)	99.50	99.38	99.90
		Inter, DET	89.94	91.82	80.08
[42] DNN, 2025	4s ECG, Lead II	Intra, DET	99.82	99.97	99.57
[25] k-NN, 2024	1 ECG beat, Lead III	Intra, DET	99.74	99.81	NM
[24] CNN, 2024	1 ECG beat, Lead II	Intra, DET	99.84	99.70	99.40
MI detection and localization from VCG					
[34] voting, 2024	4s ECG V_x, V_y, V_z	Intra, DET	99.97	99.95	100
[33] CNN, 2023	1 ECG beat V_x, V_y, V_z	Intra, LOC (HC, 11 MI)	99.83	99.96	99.99
[32] TreeBagger, 2023	1 ECG beat V_x, V_y, V_z	Intra, LOC (HC, 11 MI)	99.80	99.97	99.99
Proposed, ExtraTrees	1 ECG beat, from *any* of	*inp*Mdl#1 (Intra,DET)	99.63	99.14	99.14
		*inp*Mdl#2 (Intra,LOC)	99.67	98.85	99.80

10

the 15 leads	*INP*Mdl#1 (Inter,DET)	93.22	88.44	88.44
1 ECG beat, Lead V5	*INP*Mdl#2 (Inter,LOC)	79.07	26.18	73.28
	*INP*Mdl#3 (Inter,LOC)	99.83	99.66	99.81

DET→ detection, LOC→ localization, NM→ not mentioned

In [43], a graph-CNN and multilayer perception-based HC and six MI sub-class localization technique using all the 12 ECG leads achieves accuracies of 99.94% and 69.18% in intra- and inter-patient paradigms, respectively. The technique [43] requires 58.66 MFLOPS/prediction, whereas our intra-patient model *inp*Mdl#2 takes only 1500 FLOPS/prediction. The LSTM-based method proposed in [13] achieves an A_{cc} of 99.94% in detecting MI, whereas our proposed *inp*Mdl#1 achieves an A_{cc} of 99.63% in doing the same. However, in [13], the value of A_{cc} and other performance evaluation metrics are evaluated on a total of 24,157 samples taken from 12 ECG-leads of HC and 5 MI sub-classes. Whereas our proposed *inp*Mdl#1 is evaluated on a total of 159,429 beats, taken from 15 ECG leads of HC and 6 MI sub-classes. Moreover, in inter-patient paradigm, the MI localization accuracy of [13] is 67.89%, whereas our proposed model *INP*Mdl#2 offers an accuracy 79.07% in localizing 6 sub-classes of MI. The CNN-based MI detection and three classes of MI (A, AL and AS) localization technique using leads aVL, V2, V3 and V5, which is proposed in [18] achieves an intra-patients MI detection accuracy of about 96.65%, and inter-patient HC and 3 sub-classes of MI localization accuracy of about 66.85%. The reason that has been mentioned in [18] for choosing only three out of six dominating MI sub-groups is that the total number of subjects that belong to these three MI sub-groups are majority compared to other MI sub-groups. However, in PTBDB, the number of subjects in AWMI and IWMI classes are 70 and 69, respectively. It is interesting to note that none of the confusion matrices is shown in [18]. The size of the model is 45.75MB, and it takes about 28.16ms for the classification of one sample. Our proposed model *inp*Mdl#1, which requires only one ECG beat from any of the 15 ECG leads, achieves an accuracy of 99.63% in detecting MI in intra-patient paradigm. The size of the model *inp*Mdl#1 is 27.9 MB, and it takes about 10.00132ms (0.01s for preprocessing and feature extraction and 1.32µs for classification) for the detection of MI from a single ECG-beat. Also, our proposed model *INP*Mdl#2, with a size of 18.9MB, provides an accuracy of 79.07% in localizing six sub-classes of MI in inter-patient paradigm and takes about 10.00456ms per ECG-beat.

The DNN classifier-based MI detection and 5 sub-classes of MI localization techniques from single ECG-lead proposed in [27], achieves accuracies of 99.93% (from lead V4) and 99.50% (from lead V1) in intra-patient MI detection and localization, respectively. Another model which is developed in [27] for the detection of MI in inter-patient paradigm, achieves an accuracy of 89.94% form lead V5. As reported, the sizes of all the three models that are developed in [27] are indeed very small (around ~1.08MB), but the computational complexities are too high compared to that of the models

proposed in this research work. Nonetheless, our proposed MI detection and localization models outperform the models presented in [27] in terms of accuracy.

The CNN classifier-based model proposed in [33], achieves an accuracy of 99.83% in localizing 11 sub-classes of MI along with HC from the VCG leads in intra-patient paradigm. However, its inter-patient performance is not reported. Moreover, the size of the model is reported to be 80MB, which is 3.04 time more than that of our proposed model *inp*Mdl#2. Another technique, that also localizes HC and 11 types of MI from the VCG signals in intra-patient paradigm is proposed in [32]. The technique [32] uses a TreeBagger classifier and achieves an accuracy of 99.80%. It is reported in [32] that the method is evaluated on a computer having a RAM of *192GB* and yet it takes about 1.77s for localizing one ECG beat, which raises a concern regarding its implementation on resource-limited hardware platforms. However, all these VCG-based MI detection and localization techniques which are included in Table X are evaluated in intra-patient paradigm.

Techniques which are proposed in [12], [13], [15], [17], [18], [19] and [21] use either all the 12 conventional ECG leads or a group of these leads for training as well as testing the multi-ECG-lead models. Similarly, the techniques which are proposed in [32], [33], and [34] uses only the VCG leads, i.e., V_x, V_y, V_z, for developing as well as testing the models. In contrast, the novelty of our proposed models - *inp*Mdl#1, *inp*Mdl#2, and *INP*Mdl#1 - lies in the fact that they are trained using ECG-beats from 12 conventional and 3 VCG leads, which essentially encompasses the working of multi-lead and VCG-lead-based models, and require only one ECG-beat from any of these leads for the detection and localization of MI.

E. Comparison of computational complexity:

The computational complexity of the proposed model is now compared with a few other techniques in Table XI. The number of FLOPs or the time required for classification is available in literature for only five out of the fifteen techniques whose performance is compared with the proposed model in Table X.

TABLE XI
COMPARISON OF COMPUTATIONAL COMPLEXITIES

Model	Computational complexity	# of FLOPS required	Time required/prediction
[27] DNN, 2025	$O(q \times N) + d \times O(n^2)$ $+ d \times O(N) \times h$ $+ h \times O(g^h) + hO(n^2)$ $+ hO(n)$ $+ O(g^o)$	1.01M	984ms (Min) 2s (max)
[42] DNN, 2025	Not available	243M	Not available
[25] *k*-NN, 2024	$O\left(E \times N \times \sum_{l=1}^{L} d_{l-1} \times d_l\right) + O(N)$	Not available	Not available
[43] CNN, 2025	Not available	3754M	530ms
[32] TreeBagger, 2023	Not available	Not available	1.77s
Proposed	$O(750)$ (Min) $O(1500)$ (Max)	750 (Min) 1500 (Max)	10.00132ms (Min) 10.00456ms (Max)

In the equation of computational complexity of [27], q = 4, N is the length of the sample point, d is the number of features, h

11

is the number of neuron, o is the number of output neuron, $O(g^h)$ is the computational complexity of hidden layer activation function, $O(g^o)$ is computational complexity of output layer activation function. Similarly, in [25], E is the number of epochs, N is the number of training samples, L is the number of layers, and d_l is the dimensions of hidden layers. The technique proposed in [32] requires 1.77s per classification, whereas out proposed model *INP*Mdl#2 needs only 4.56μs per classification. It can be seen from the table that our proposed algorithm is significantly faster and lightweight than those in [25], [27], and [42], making it suitable for implementation in wearable devices.

IV. ABLATION STUDY

Detection of QRS complexes and T-waves, feature extraction, imputation, scaling, augmentation, feature importance calculation, and selection of the best-performing classification model are the key components of the proposed MI detection and localization framework. The QRS detection algorithm proposed in [36] is adopted in this study, which not only identifies QRS complexes with high accuracy but also segregates arrhythmic ECG beats—thereby enhancing the classification performance.

Selecting the most important features for a classifier is a crucial task as it offers multiple benefits. Firstly, it reduces the dimensionality of the feature set; secondly, it lowers the algorithm's complexity; and thirdly, by removing features that negatively impact performance, it can enhance the classifier's accuracy. A Random Forest (RF)-based feature selection technique reveals that the contributions of two features, F20 and F23, are minimal in the classification of MI. Although an initial attempt was made to evaluate the model *inp*Mdl#1 using all 34 features, it was found that including F20 and F23 negatively impacted the model's performance. The current version of *inp*Mdl#1, trained without these two features, achieves an accuracy of 99.63%. However, when F20 and F23 are included, the accuracy drops to 99.02%. The maximum, minimum, mean, median, and standard deviation of the remaining 32 features are 0.0529, 0.0142, 0.0308, 0.0291, and 0.0099, respectively. It is also observed that discarding F20 and F23 reduces the algorithm's processing time by approximately 0.62ms per ECG beat. A permutation-based feature importance technique was also evaluated in this work, and the corresponding feature scores are shown in Figure 9. This figure indicates that, similar to the RF-based method, F20 shows low importance, while the importance scores of other features remain largely unchanged. The performance of additional feature importance methods, such as SHapley Additive exPlanations (SHAP), will be explored in future research.
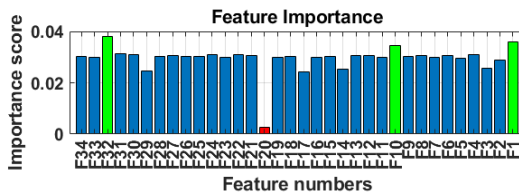

Feature Importance

Fig. 9. Importance of all the 34 handcrafted ECG features using permutation-based technique.

Imputing missing feature values is essential, as many classification algorithms, such as Logistic Regression, Linear Discriminant Analysis, Support Vector Machines, Naive Bayes, XGBoost, and LightGBM, cannot handle missing data. Common imputation techniques using the mean or median can reduce the natural variability of a feature and shift its distribution toward the mean or median. An attempt was also made to impute the missing values using a LightGBM-based MICE (miceforest) technique; however, its processing time was significantly higher compared to the k-NN-based method, which has been used in this study.

Scaling the feature values and bringing them to a similar range is a key preprocessing step in machine learning, as it ensures that all input features contribute equally to the model. In this research, along with the robust scaling technique, Z-score normalization and min-max scaling have also been tested. However, it has been observed that the accuracy of the model *inp*Mdl#1 drops to 95.24% and 94.52% when Z-score normalization and min-max scaling are applied, respectively.

The LazyPredict library in Python has been used in this research to identify the best-performing classifier. Figure 10 illustrates the performance of 24 classifiers in detecting MI under the intra-patient paradigm. Among these, the Extra Trees classifier outperforms the rest. The performances of the AdaBoost and XGBoost classifiers in developing the model *INP*Mdl#3 are also presented in Table IX. Based on the results in Figure 10 and Table IX, it can be concluded that the Extra Trees classifier offers the best overall performance. Thus, it can be inferred that each component of the proposed algorithm has contributed effectively to achieving such high classification accuracy.

From Table VII, it can be observed that the model *INP*Mdl#2 is unable to localize the six MI sub-classes effectively. To investigate the reason behind this poor performance, the histograms of the three most important features, F2, F6, and F11, as identified in Figure 5A, are plotted in Figure 11. From this figure, it is evident that (i) the features corresponding to the healthy control (HC) class (shown in blue) are clearly distinguishable from those of the MI class; however, (ii) the features corresponding to the different MI sub-classes show significant overlap.

	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
ExtraTreesClassifier	0.99109	0.99165	None	0.99108	59.46796
RandomForestClassifier	0.98852	0.98890	None	0.98851	228.27191
BaggingClassifier	0.97487	0.97483	None	0.97486	132.79593
KNeighborsClassifier	0.96157	0.96420	None	0.96148	39.81879
DecisionTreeClassifier	0.92031	0.91932	None	0.92030	19.54321
ExtraTreeClassifier	0.85181	0.85163	None	0.85179	1.10971
LGBMClassifier	0.78809	0.78023	None	0.78751	4.50678
SVC	0.69333	0.67239	None	0.69141	18277.98205
NuSVC	0.58346	0.53844	None	0.57232	39310.77576
LogisticRegression	0.32319	0.28835	None	0.29970	6.80544
CalibratedClassifierCV	0.31950	0.28435	None	0.29532	169.30284
LinearSVC	0.31945	0.28292	None	0.28720	38.62732
LinearDiscriminantAnalysis	0.31678	0.28269	None	0.29822	1.19837
RidgeClassifier	0.31561	0.27893	None	0.28452	0.81982
RidgeClassifierCV	0.31560	0.27891	None	0.28450	1.43005
QuadraticDiscriminantAnalysis	0.28810	0.27412	None	0.23907	0.96780
NearestCentroid	0.29000	0.26634	None	0.26856	0.70499
BernoulliNB	0.29605	0.26411	None	0.26558	0.74830
AdaBoostClassifier	0.29602	0.26095	None	0.27082	70.38083
SGDClassifier	0.22385	0.23611	None	0.22114	5.70528
GaussianNB	0.24907	0.22972	None	0.20115	1.19078
Perceptron	0.17175	0.17104	None	0.17675	4.04609
PassiveAggressiveClassifier	0.17976	0.16984	None	0.16439	4.46930
DummyClassifier	0.17935	0.14286	None	0.05455	0.49150

12

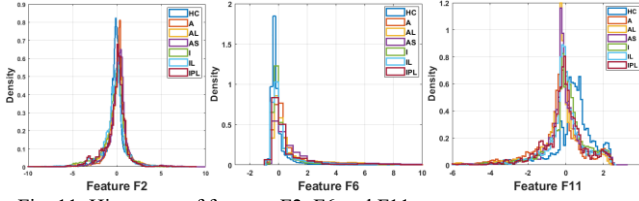Fig. 10. Performance of 24 classifiers in detecting MI in intra-patient paradigm.

Fig. 11. Histogram of features F2, F6 and F11.

V. DISCUSSION

The advantages of using Ramanujan filter bank-based periodicity estimation technique as the R-peak detector are threefold. First, alongside detecting the R-peaks with high accuracy it also identifies the arrhythmic QRS-complexes at the time of its detection, which avoids the need for a separate classifier. Second, it prevents the mixing of arrhythmic QRS-complex features with those of healthy controls (HC) and MI patients. Third, the troughs in the time-period representations serve as the indicators of the start and end points of the ECG-beats, which helps accurately segment the ECG-beats rather than relying on a heuristic number of samples to the left and right of the R-peaks - as done in most previous studies including [15], [28] and [33].

Evaluating the performance of a classifier employing five- or ten-fold cross validation techniques are common in intra-patient paradigm. However, the cross-validation techniques have not been used in this research work. Rather, in this work, in intra-patient paradigm, the train and test sets are separated immediately after the features are extracted from all the ECG-beats. As a result, although the train and test sets share the same subjects' ECG-beats, other information such as the factors of feature normalization and imputation parameters is not shared with the test set. This helps minimize the leakage of information. In contrast, the cross-validation-based evaluation techniques such as [14], [17], [32], [33], [34] and [42] share this information across the train and test sets, which in turn increases model bias. While optimizing the hyperparameters of the Extra Trees classifiers, the number of swarms is set to 50, and each of these swarms are iterated 100 times, which are substantially high values. As a result, the process of finding the optimal hyperparameter values has been time-consuming, given the large size of the training dataset used in this research. However, the PSO-optimized hyperparameters have yielded satisfactory classification results that justify the optimization time.

It is interesting to note that, among the fourteen techniques to which the performances of our proposed models are compared in Table X, four techniques are developed based on classical machine learning models. Techniques which are proposed in [15] and [25] are developed based on k-nearest neighbor (k-NN) classifier, the technique proposed in [34] is developed based on a voting of five classical machine learning models, and the technique proposed in [32] is developed based on a TreeBagger classifier. Moreover, as shown in Table X, these classical machine learning models often outperform other CNN- and DNN-based models.

VI. CONCLUSION

Robust, reliable, and fast machine learning models leveraging the strong regularization ability of the Extra Trees classifier have been developed in this research for the detection and localization of six sub-classes of myocardial infarctions (MI) in both intra- and inter-patient paradigms. Five machine learning models are developed in this research work. Two out of these five models, are trained to detect myocardial infarction (MI), and the other three models are trained to localize MI.

While analyzing the contribution of individual ECG leads in the detection of MI in this research work, it is found that lead V5 outperforms other 14 leads. Although the model *INP*Mdl#2 achieves the highest accuracy of 79.07% in localizing HC and six sub-classes of MI in inter-patient paradigm, it still requires further improvement. Instead of trying to localize the six MI sub-classes at once, as is commonly done in other methods, it would perhaps be more methodical to first classify the MI sub-classes into AWMI and IWMI, and then further subclassify these two groups into their corresponding sub-classes. This is exactly what is tried in this research work, and the model *INP*Mdl#3 attains and average accuracy of 99.83%, which is high indeed, in separating HC, AWMI and IWMI groups. It will be further explored to subclassify the AWMI and IWMI groups in our future research. One of our proposed models - *INP*Mdl#1 can detect MI from a single ECG-beat recorded from any of the 12 different locations on the torso with average A_{cc} and MCC values of 96.47% and 89.64%, respectively. This finding strongly suggests that the model can be adopted in wearable healthcare gadgets to detect and localize MI with conventional, unconventional, and pseudo-ECG leads.

REFERENCES

[1] Canadian Medical Association. [Online]. Available: https://www.cma.ca/latest-stories/insight-why-are-patients-spending-22-hours-emergency-room-waiting-hospital-bed, Accessed: August 2025.

[2] Statista. [Online]. Available: https://www.statista.com/statistics/649600/medical-treatment-wait-times-canada-province/, Accessed: August 2025.

[3] D. T. Tran et al., "The Healthcare Cost Burden of Acute Myocardial Infarction in Alberta, Canada", *PharmacoEconomics*, vol. 2, iss. 4, pp. 433-442, Nov. 2017.

[4] Ebn. [Online]. Available: https://www.benefitnews.com/list/10-states-with-the-longest-er-wait-times, Accessed: August 2025.

[5] Apple. [Online]. Available: https://www.apple.com/healthcare/apple-watch/, Accessed: August 2025.

[6] Hexoskin. [Online]. Available: https://hexoskin.com/pages/health-research, Accessed: August 2024.

[7] X. Ma, X. Fu, Y. Sun, N. Wang and Y. Gao, "Application of Convolutional Dendrite Net for Detection of Myocardial Infarction Using ECG Signals", *IEEE Sensors Journal*, vol. 23, no. 1, 1 Jan. 2023, pp. 460-469.

[8] S. Bhaskarpandit, A. Gade, S. Dash, D. K. Dash, R. K. Tripathy and R. B. Pachori, "Detection of Myocardial Infarction From 12-Lead ECG Trace Images Using Eigen domain Deep Representation Learning", *IEEE Transactions on Instrumentation and Measurement*, vol. 72, ASN 4001812, Feb. 2023, pp. 1-12.

[9] J. Zhang, M. Liu, P. Xiong, H. Du, H. Zhang, G. Sun, Z. Hou and X. Liu, "Automated Localization of Myocardial Infarction of Image-Based Multilead ECG Tensor with Tucker2 Decomposition", *IEEE Transactions on Instrumentation and Measurement*, vol. 71, ASN 2501215, Nov. 2021, pp. 1-15.

[10] N. Sinha and A. Das, "Identification and Localization of Myocardial Infarction Based on Analysis of ECG Signal in Cross Spectral Domain Using Boosted SVM Classifier", *IEEE Transactions on Instrumentation and Measurement*, vol. 70, ASN 4007409, Oct. 2021, pp. 1-9.

[11] E. Prabhakararao and S. Dandapat, "Attentive RNN-Based Network to Fuse 12-Lead ECG and Clinical Features for Improved Myocardial Infarction Diagnosis", *IEEE Signal Processing Letters*, vol. 27, Nov. 2020, pp. 2029-2033.

[12] Q. Yao, L. Zhang, W. Zheng, Y. Zhou and Y. Xiao, "Multi-scale SE-residual network with transformer encoder for myocardial infarction classification", *Applied Soft Computing*, vol. 149, no. 18, Part A, Dec. 2023, pp. 1-12.

[13] C. Han, J. Sun, Y. Bian, W. Que and L. Shi, "Automated Detection and Localization of Myocardial Infarction with Interpretability Analysis Based on Deep Learning", *IEEE Transactions on Instrumentation and Measurement*, vol. 72, ASN 2508512, Mar. 2023, pp. 1-12.

[14] V. S. Parupudi, A. K. Panda and R. K. Tripathy, "A Smartphone-Enabled Deep Learning Approach for Myocardial Infarction Detection Using ECG Traces for IoT-Based Healthcare Applications", *IEEE Sensors Letters*, vol. 7, no. 11, Nov. 2023, pp. 1-4.

[15] G. Sahu and K. C. Ray, "An Efficient Method for Detection and Localization of Myocardial Infarction", *IEEE Transactions on Instrumentation and Measurement*, vol. 71, ASN 4001312, Jan. 2022, pp. 1-12.

[16] J. Yu, J. Gao, N. Wang, P. Feng, B. Zhou and Z. Wang, "Spa-Tem MI: A Spatial-Temporal Network for Detecting and Locating Myocardial Infarction", *IEEE Transactions on Instrumentation and Measurement*, vol. 72, ASN 2522712, Aug. 2023, pp. 1-12.

[17] P. Jain, A. Deshmukh and H. Padole, "Design of an Integrated Myocardial Infarction Detection Model Using ECG Connectivity Features and Multivariate Time Series Classification", *IEEE Access*, vol. 12, Jan. 2024, pp. 9070-9081.

[18] Y. Cao, T. Wei, B. Zhang, N. Lin, J. J. P. C. Rodrigues and J. Li, "ML-Net: Multi-Channel Lightweight Network for Detecting Myocardial Infarction", *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 10, Oct. 2021, pp. 3721-3731.

[19] R. Parmar, K. Yadav, G. Anand and G. Trivedi, "An SNN-Inspired Area- and Power-Efficient VLSI Architecture of Myocardial Infarction Classifier for Wearable Devices", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 71, no. 6, Jun. 2024, pp. 3191-3195.

[20] C. Han and L. Shi, "ML–ResNet: A novel network to detect and locate myocardial infarction using 12 leads ECG", *Computer Methods and Programs in Biomedicine*, vol. 185, iss. 3, Part A, March 2020, pp. 1-10.

[21] W. Pan, Y. An, Y. Guan and J. Wang, "MCA-net: A multi-task channel attention network for Myocardial infarction detection and location using 12-lead ECGs", *Computers in Biology and Medicine*, vol. 150, iss. 11, November 2022, pp. 1-12.

[22] G. Sahu and K. C. Ray, "PSO Optimized Hybrid Deep Learning Model for Detection and Localization of Myocardial Infarction", *IEEE Sensors Journal*, vol. 24, no. 5, Mar. 2024, pp. 6643-6654.

[23] 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", *Information Sciences*, vol. 589, no. 8, Apr. 2022, pp. 738-750.

[24] A. Yousuf, R. Hafiz, S. Riaz, M. Farooq, K. Riaz and M. M. U. Rahman, "Inferior Myocardial Infarction Detection from Lead II of ECG: A Gramian Angular Field-Based 2D-CNN Approach", *IEEE Sensors Letters*, vol. 8, no. 10, ASN 6012104, Oct. 2024, pp. 1-4.

[25] S. M. S. Anwar, D. Pal, S. Mukhopadhyay and R. Gupta, "A Lightweight Method of Myocardial Infarction Detection and Localization from Single Lead ECG Features Using Machine Learning Approach", *IEEE Sensors Letters*, vol. 8, no. 4, ASN 7002204, Apr. 2024, pp. 1-4.

[26] M. Janveja, R. Parmar and G. Trivedi, "MInSC: A VLSI Architecture for Myocardial Infarction Stages Classifier for Wearable Healthcare Applications", *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 70, no. 3, March 2023, pp. 1159-1163.

[27] A. Gupta and K. C. Ray, "An Efficient Method for Detection and Localization of Myocardial Infarction using Ramanujan Sums Wavelet Transform and Machine Learning Framework", *IEEE Transactions on Instrumentation and Measurement*, DOI: 10.1109/TIM.2025.3545189, February 2025, pp. 1-10.

[28] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, S. Pranav, R. B. Pachori, "Efficient detection of myocardial infarction from single lead ECG signal", *Biomedical Signal Processing and Control*, vol. 68, iss. 6, July 2021, pp. 1-9.

[29] 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 Transactions on Biomedical Engineering*, vol. 60, no. 8, Aug. 2013, pp. 2325-2331.

[30] H. Yang, C. Kan, G. Liu and Yun Chen, "Spatiotemporal Differentiation of Myocardial Infarctions", *IEEE Transactions on Automation Science and Engineering*, vol. 10, no. 4, Oct. 2013, pp. 938-947.

[31] H. Yang, "Multiscale Recurrence Quantification Analysis of Spatial Cardiac Vectorcardiogram Signals", *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, Feb. 2011, pp. 339-347.

[32] J. Zhang et al., "Automated Localization of Myocardial Infarction from Vectorcardiographic via Tensor Decomposition", *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 3, Mar. 2023, pp. 812-823.

[33] P. Xiong, K. Li, J. Zhang, C. He, H. Du and J. Yang, "Localization of Myocardial Infarction From 2D-VCG Tensor With DSC-Net", *IEEE Transactions on Instrumentation and Measurement*, vol. 72, ASN 2522110, Jun. 2023, pp. 1-10.

[34] M. K. Chaitanya and L. D. Sharma, "Cross Subject Myocardial Infarction Detection from Vectorcardiogram Signals Using Binary Harry Hawks Feature Selection and Ensemble Classifiers", *IEEE Access*, vol. 12, Feb. 2024, pp. 28247-28259.

[35] PhysioNet. [Online]. Available: https://physionet.org/content/ptbdb/1.0.0/, Accessed: August 2025.

[36] S. K. Mukhopadhyay and S. Krishnan, "Robust Identification of the QRS-Complexes in Electrocardiogram Signals Using Ramanujan Filter Bank-Based Periodicity Estimation Technique", *Frontiers in Signal Processing*, vol. 2, article no. 921973, June 2022, pp. 1-15.

[37] S. V. Tenneti and P. P. Vaidyanathan, "Nested Periodic Matrices and Dictionaries: New Signal Representations for Period Estimation, *IEEE Transactions on Signal Processing*, vol. 63, iss. 14, May 2015, pp. 3736-3750.

[38] S. K. Mukhopadhyay, M. O. Ahmad and M.N.S. Swamy, "Compression of Steganographed PPG Signal with Guaranteed Reconstruction Quality Based on Optimum Truncation of Singular Values and ASCII Character Encoding", *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 7, July 2019, pp. 2081-2090.

[39] S. K. Mukhopadhyay and S. Krishnan, "Robust identification of QRS-complexes in electrocardiogram signals using a combination of interval and trigonometric threshold values" *Biomedical Signal Processing and Control*, vol. 61, iss. 7, August 2020, pp. 1-12.

[40] H. He, Y. Bai, E. A. Garcia and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning", *IEEE International Joint Conference on Neural Networks*, Hong Kong, June 2008, pp. 1322-1328.

[41] R. Iranzad and X. Liu, "A review of random forest-based feature selection methods for data science education and applications", *International Journal of Data Science and Analytics*, February 2024, pp. 1-15.

[42] K. Gupta, V. Bajaj and I. A. Ansari, "CardioNet: A Lightweight Deep Learning Framework for Screening of Myocardial Infarction Using ECG Sensor Data", *IEEE Sensors Journal*, vol. 25, no. 4, February 2025, pp. 6794-6800.

[43] X. Yang, G. Jiang, Z. Zhu, D. Wu, A. He, J. Wang, "MDD2DG-IRA: Multivariate Degree Distribution to Dynamic Graph with Inter-Channel Relevance Attention Mechanism for Multi-Channel Myocardial Infarction ECG Analysis", *IEEE Journal of Biomedical and Health Informatics*, doi: 10.1109/JBHI.2025.3554309, Mar. 2025, pp. 1-12.