ST-segment anomalies detection from compressed sensing based ECG data by means of machine learning

Giovanni Rosa[1], Marco Russodivito[1], Gennaro Laudato[1], Angela Rita Colavita[2], Luca De Vito[4], Francesco Picariello[4], Simone Scalabrino[1,3], Ioan Tudosa[4], and Rocco Oliveto[1,3]

[1] STAKE Lab – University of Molise, Pesche (IS), Italy
{giovanni.rosa, marco.russodivito, gennaro.laudato, simone.scalabrino, rocco.oliveto}@unimol.it
[2] ASREM, Campobasso (CB), Italy
angelarita.colavita@asrem.org
[3] Datasound srl, Pesche (IS), Italy
{rocco, simone}@datasound.it
[4] Dept. of Engineering, University of Sannio, Benevento (BN), Italy
{devito, fpicariello, ioan.tudosa}@unisannio.it

Abstract. Telemedicine allows to constantly monitor patients without the need of hospitalization. Such a practice is enabled by IoMT (Internet of Medical Things) devices, which acquire signals, and by AI (Artificial Intelligence)-based algorithms, able to automatize the analysis carried out on many patients. The large quantity of data produced every minute by IoMT devices, however, makes the use of compression fundamental to reduce the bandwith used to transmit those data and the memory required to acquire them, and lossy compression (specifically, Compressed Sensing) has shown to be the most effective technique to use for the task. Previous work introduced AI-based approaches for automatically detecting hearth-related anomalies based on the electrocardiographic (ECG) signal. However, most of them assume the presence of the complete raw ECG signal. In this paper, we extend our previous work in which we introduced RAST, an approach for detecting ST segment-related anomalies. We present RAST^C, an approach able at identifying the same abnormalities but on a highly compressed ECG signal. The results of our experiment, carried out on the Physionet European ST-T Database, shows that RAST^C is capable of discriminating Normal ECG from ST-depression and ST-elevation with classification metrics around 90%, even with the highest compression ratio experimented, *i.e.,* with an ECG signal compressed by a factor of 16.

Keywords: Machine Learning · Compressed Sensing · Automatic Detection · ECG signal · ST-segment · Anomalies

1 Introduction

Due to the growing popularity of wearable devices applied in the medical field, the Internet of Medical Things (IoMT) paradigm is experiencing rapid growth from the scientific perspective of Research and Development (R&D). Basically, an IoMT network

comprises numerous smart medical devices connected to each other over the internet. Modern telemedicine relies on an IoT-based smart healthcare system [38]. A smart healthcare system built upon the IoMT paradigm is composed of many stages. First, by using smart sensors built into wearable (or implanted) devices that are linked by a Wireless Body Sensor Network (WBSN), medical data will be gathered from the patient's body [29, 42]. Next, the component handling the prediction and analysis step will get this data through internet transmission. After obtaining the medical data, an analysis may be carried out using an Artificial Intelligence (AI)-based data transformation and interpretation approach, [31]. Based on them, the system can decide how to act based on the urgency of the possibly detected issue, *e.g.,* contacting a physician in case of critical conditions [32].

Telemedicine is another name for the process of employing IoMT tools to remotely monitor patients while they're at home. When a patient receives this type of care, they are spared from visiting a hospital or doctor's office every time they have a medical inquiry or a change in their condition. The greatest advantage of telemedicine, however, is having continuous monitoring of the health parameters of the caregivers. Recent telemedicine systems proposed in the literature include ATTICUS [3, 25]. ATTICUS grew out of the development of a prototype hardware and software system based on advanced artificial intelligence techniques, capable of constantly monitoring an individual and reporting anomalies affecting both his or her health status, detected through the automatic measurement and analysis of biological parameters, and in his or her behavior, detected through the monitoring and analysis of the movements the person makes in the normal course of his or her activities. The system consists of one key element: a "smart wearable" device [10]. This consists of a T-shirt made of innovative fabrics, which integrates a data acquisition system (integrated into the T-shirt fabric) capable of measuring the person's vital parameters, such as electrocardiogram tracing, temperature, respiratory behaviors, etc [3, 25].

Among the many signals that can be considered in IoMT devices (and, particularly, in ATTICUS), the electrocardiographic signal acquired through the Electrocardiogram (ECG) is of paramount importance for identifying possible hearth-related issues. Specifically, within the ATTICUS project, several detectors where designed aimed at automatically analysing pathological features derived from the ECG. These detectors allow to automatically identify (i) Atrial Fibrillation (AF) episodes [20, 22, 24], (ii) arrhythmia related conditions [33], and (iii) Congestive Heart Failure [35]. Most of such detectors assume the presence of a raw ECG signal to run their analyses and provide predictions. However, the electronic core components in a wearable device (such as those that capture real-time multi-lead ECG signals) create and transmit large amount of data, which make the adoption of compression techniques necessary to save memory and bandwidth. Specifically, lossy compression techniques based on Compressed Sensing (CS), particularly digital CS techniques, can be adopted to reduce the local power consumption and to reduce memory size [4]. In this context, it is necessary to adapt existing AI-based detection approaches to make them work with the compressed signal [21, 23].

In our previous work [34], we introduced RAST, a detector for ST segment-related anomalies. RAST, however, relies on the uncompressed ECG signal. In this paper,

we extend our previous work, and we introduce RAST^C, an innovative detector of ST segment-related anomalies that directly relies on a compressed version of a single lead digital ECG. The results show that RAST^C is capable of achieving results comparable with the state-of-the-art, by obtaining classification metrics above 90% even with Compression Ratios up to 16.

Specifically, the contributions of this paper can be summarized as follows:

- We introduce RAST^C, which is provided in two versions: BINARY-RAST^C, based on identification of only ST abnormality from Normal ECG and TERNARY-RAST^C, based on discrimination between ST-depression, ST-elevation and Normal ECG. To increase the information component, which is compromised by the compression process, we present 4 newly defined features. These features are measures derived from the ϕ matrix of the Bernoulli compression algorithm application case;
- We present a study aimed at testing whether an approach based on detection of cardiac pathologies derived from ST-segment analysis can be dropped into a context of compressed ECG signal analysis. In order to conduct a comprehensive study, two compression techniques were analyzed: one deterministic (DBBD) and one random (Bernoulli).

This paper is structured as follows: in Section 2 a report on the incidence of ST-segmnent anomalies is offered together with a (i) review of the recent state-of-the-art dedicated to the automatic analysis of ST-segment and (ii) a description of the compressed sensing algorithm. Section 3 describes the workflow of the approach, that includes details on the preprocessing, compression, calculation of the features and the classification. Section 4 proposes details on the design of the study and the results. Finally, Section 5 concludes the paper.

2 Background and related work

In this section an overview on the incidence of ST-segment anomalies on the general population. This enhances the importance of provindig fast and reliable detection of anomalies. After this, an updated review of the state-of-the-art related to recent works dedicated to the detection of ST-segment anomalies.

2.1 ST-related conditions

An electrically neutral region of the complex between ventricular depolarization (QRS complex) and repolarization (T wave) can be represented by the ST segment on an electrocardiogram [37]. It can, however, adopt a variety of waveform morphologies that could point to a benign or clinically severe myocardial infarction or insult. For clinical management since it might affect treatment, it is essential to comprehend the differential diagnosis for alterations in the ST segment [5].

As already described in the introduction, morphological variants for the ST-segment can happen in the form of ST elevation and depression (Figure 1).

ST elevation is frequently caused by three factors. The first situation is when ST elevation is only a common variety. Early repolarization is a common term used to

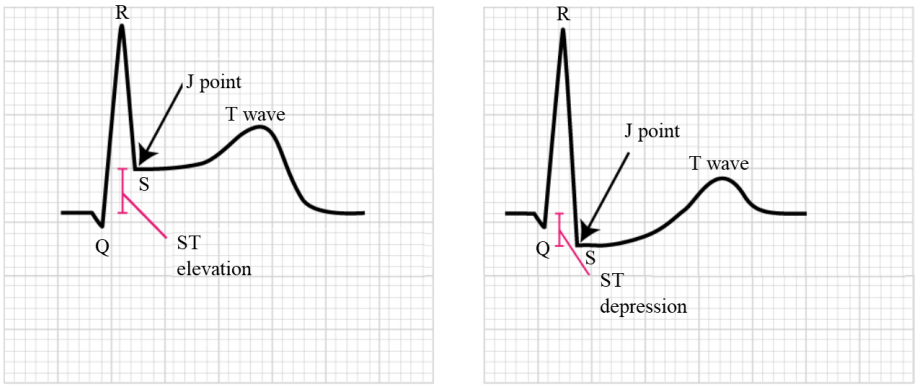


Fig. 1: St depression and elevation phenomena on the left and on the right part of the figure respectively. The amount of depression or elevation can be measured in terms of boxes (or mm).

describe this. In these situations, the J-point is often elevated and the ST section has a normal or quickly ascending slope. Acute ischemia or ventricular dyskinesis damage currents are the second prevalent cause. The third is brought on by pericarditis-related damage currents [17].

When the J point is relocated below baseline, ST depression happens. ST depression is linked to a variety of illnesses. Some of these include heart ischemia, hypokalemia, and drugs like digitalis. ST depression might also result in T wave alterations [17].

Despite not being common, ST-segment elevation problems, often known as STEMI, Even in young people, ST-segment elevation acute myocardial infarction (STEMI) is common. In fact, 6% of STEMI incidents included relatively young persons under the age of 35 [18].

Here, we report some clinical relevance data about the incidence of ST-related pathological conditions [14]:

- Acutely blocked coronary arteries are present in 80% of individuals with typical STEMI;
- acutely occluded coronary arteries were seen in 10% of individuals with ST segment elevation in the electrocardiographic augmented vector right (aVR) lead.

Also, a recent study was conducted to identify any anomalies in the ECG that would indicate cerebral vascular disease [16]. Within 24 to 48 hours, a computed tomography (CT) scan of the brain was obtained, evaluated, and patients were divided into three groups: those with cerebral infarction, intracranial hemorrhage, and subarachnoid hemorrhage (SAH). In patients with cerebral hemorrhage (CH), ST segment alterations were the most typical anomaly seen. More specifically, following cerebral hemorrhage, ST

segment alterations were most frequently observed. ST depression is present in 33% of infarction patients. 50% of individuals with CH had ST elevation [16].

In addition, the COVID-19 outbreak may have increased the incidence of st problems, particularly st depression problems. Indeed, separation and limiting people's movement during the COVID-19 pandemic appeared to impair physical activity and dietary consumption of fresh fruit and vegetables, which may have an impact on long-term cardiovascular results. On the other hand, melancholy, rage, and ongoing stress brought on by separation, mobility restrictions, a lack of supplies and knowledge, financial loss, and a fear of getting sick might have short-term consequences on the cardiovascular system [27]. Numerous studies have revealed connections between vascular health and psychological factors [40]. It has been shown that people with depressive disorders have a higher chance of developing coronary artery disease, increased blood clotting, reduced fibrinolysis, and cardiac death [6, 12, 19, 36]. The COVID-19 pandemic's psychological stress may have contributed to the greater rates of ST-segment depression and aberrant Q waves. These seem the results over a brief observation period [40].

2.2 Automatic detection of ST-related conditions

Several works have been proposed in the literature for the detection of ST-related anomalies.

Meglaveras *et al.* [26] proposed an approach for real-time detection of ischemia episodes using an adaptive Backpropagation Neural Network. In detail, their method is based on the processing of the whole ST pattern, by evaluating the ST segments of the first 10 successive heartbeats. Their results show a sensitivity of 88.62% and 72.22% for detection and duration of ischemia episodes.

Bulusu *et al.* [7] proposed an approach using the Support Vector Machine (SVM) algorithm for the automatic recognition of ST-segment anomalies. In detail, it uses morphological features from ECG, including ST-segment information for the diagnosis of myocardial ischemia and cardiac arrhythmia. Their results report an accuracy of 93.3% for the European ST-T Database. Also, they achieved an accuracy value of 96.4% for the classification of cardiac arrhythmia episodes from the MIT-BIH Arrhythmia Database.

Xiao *et al.* [43] proposed an approach based on image analysis using deep learning techniques for the detection of ischemic ST change for ECGs. In detail, they build a Convolutional Neural Network (CNN) using transfer learning and training it on a dataset, extracted from the Long Term ST database [15], composed of 10 seconds image samples regarding signals having significant ST changes and control samples (no significant changes). Their results show an average ROC AUC score of 89.6%, while selecting an optimum cutoff level they achieve an average sensitivity of 84.4%.

Wang *et al.* [41] proposed a beat-to-beat classification method for ST segment changes using the Random Forest algorithm. They use, for the classification, the morphological features and Poincaré characteristics of the ST segments. The approach was trained and tested on the European ST-D Database, achieving an average sensitivity of 85.2% for normal ST segments, 86.9% for ST depression, and 88.8% for ST elevation.

Moreover, Harun-Ar-Rashid [14] proposed a technique for the automatic classification of ST segments based on cross-correlation for five ST categories (*i.e.,* concave,

Convex, Up slope, Down slope, and Horizontal). In detail, after the first denoising preprocessing phase, it follows a two-step ECG annotation starting with (i) R wave identification, and (ii) the annotation of S and T waves, along with J point. The last step is the identification of ST change categories by performing a cross-correlation with the supervised ST change data, to measure the similarity between the ST segment and the corresponding ST change category. Their approach achieved an accuracy value of 88.2% for the MIT-BIH ST change database, and 96.2% for the European ST-T change database.

All the above-mentioned works process an ECG trace as it is acquired, without any loss of information. To the best of our knowledge, this paper represents the first attempt to define an approach for the detection of ST-segment-related abnormalities from a highly compressed ECG.

2.3 The Compressed Sensing Algorithm

Compressed Sensing (CS) is a technique widely adopted in ECG monitoring with wireless-connected wearable devices [4]. CS allows for reducing the amount of wirelessly transmitted data to the host (e.g., smartphone, tablet, server) and requires a low computational load for the data compression.
In the literature, the CS-based techniques for ECG signals are classified into (i) randomly-based, and (ii) deterministic-based. In the former case, random distributions (e.g., Bernoulli and Gaussian) are adopted for data compression, while, in the latter case, a deterministic number sequence is utilized. According to the CS theory, the data compression is modeled as:

$$\mathbf{y} = \Phi \cdot \mathbf{x} \tag{1}$$

where, $\mathbf{x}$ is a vector of N ECG samples acquired in a certain time window, at Nyquist rate, $\mathbf{y}$ is the M-size vector of the compressed samples, where $M < N$ and Φ is an $M \times N$ matrix called sensing matrix. In randomly-based techniques the matrix Φ is randomly built according to a probability distribution, such as Gaussian and Bernoulli. On the other hand, in the deterministic-based techniques the matrix Φ is built according to a well-defined sequence of numbers. For example, in [11], the Authors implemented Toeplitz, Circulant, and Triangular structured sensing matrices, which do not require the generation of random numbers and therefore are easier to be implemented on wearable devices. In [30], a Deterministic Binary Block Diagonal matrix (DBBD) is adopted as sensing matrix. In this case, each row of Φ consists of a sequence of one and zero according to the imposed compression ratio (i.e., $CR = N/M$).
The compressed samples need to be processed to obtain an estimation of the original vector $\mathbf{x}$. The estimation of $\mathbf{x}$ (i.e., the reconstruction step) requires the definition of a dictionary matrix Ψ selected according to the domain where the signal can be represented by few K non-zero coefficients. According to Ψ and Φ, the coefficients θ, representing $\mathbf{x}$ in the selected domain, are estimated by solving:

$$\hat{\theta} = \arg\min_{\theta} ||\theta||_1, \quad \text{subject to:} \quad \mathbf{y} = \Phi\Psi\theta \tag{2}$$

where, $||\cdot||_1$ indicates the l_1 norm operator. $\hat{\mathbf{x}}$ is estimated from $\hat{\theta}$ as follows:

$$\hat{\mathbf{x}} = \Psi \cdot \hat{\theta} \tag{3}$$

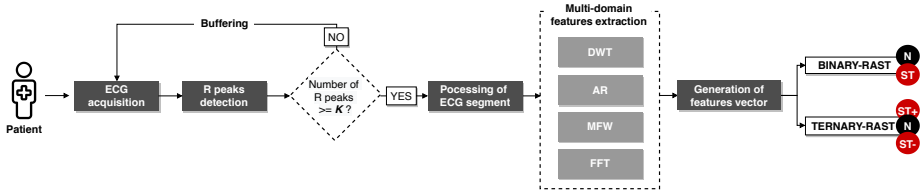


Fig. 2: The workflow of RAST [34]

In terms of reconstruction quality, among the randomly-based techniques, the best results are obtained with the Bernoulli sensing matrix and the two scales Mexican hat-based dictionary matrix, [8]. However, it was demonstrated the deterministic-based technique performances outperform the randomly-based, [1]. In [30], the DBBD with Discrete Cosine Transform (DCT) as a dictionary matrix obtained the best results in terms of reconstruction quality. However, in the literature, more complex deterministic-based techniques are proposed to outperform the performance of DBBD with DCT [9].

All the above-mentioned results require solving of Equation 2 by means of Orthogonal Matching Pursuit (OMP) algorithm that exhibits a computational complexity $O((N+M)S)$, where $S < N$ is the number of iterations. Thus, the reconstruction step limits the use of CS when the application requires the implementation of real-time monitoring and early-warning systems.

3 Automatically Detecting ST-related Conditions from compressed ECG: workflow of the approach

In this section we describe the workflow of the original RAST [34], an approach for the automatic detection of ST-related conditions on digital uncompressed ECG signals. In Figure 2 we report the workflow of RAST. Firstly, RAST takes as input a digital ECG signal containing at least K successive heartbeats (*i.e.,* at least K R peaks detected). After the preprocessing of the ECG segment, it follows the extraction of the feature vector and then the classification of the ECG segment. The parameter K in Figure 2 indicates that in our previous study [34], a study was conducted to determine the best temporal window in terms of successive heartbeats within two cross-validation schemes experimented. The result showed that 4 successive heartbeats are enough to obtain good classification performances in the 80-20 cross validation). The first phase is the processing of the input ECG signal. The Pan-Tomphkins algorithm [28] is applied to detect the number of R peaks in the input ECG signal. The ECG is buffered until it reaches the required size of K successive heartbeats. This to allow the extraction of the features and then the detection of ST-related anomalies. Next, a detrend operation is applied to the ECG segment to improve the extraction of the features. There are two versions of RAST: BINARY-RAST (RAST_{bin}), which classifies the ECG segment as Normal or ST-anomaly, and TERNARY-RAST (RAST_{ter}), which supports the detection of ST depression, ST elevation, and Normal.

After a brief description of the original RAST, we describe in the following subsections the detailed workflow phases of the study proposed in this paper, which led to the definition of RAST^C.

3.1 Introducing RAST^C: Uncompressed vs Compressed Domain

4 The Study

This section provides a full description of the study, including the experimental design, the context of the study and the final results.

4.1 Study Design

The main goal of our study is to evaluate the effectiveness of both RAST_{bin} and RAST_{ter} in the compressed domain.

In particular, we want to answer the following research questions:

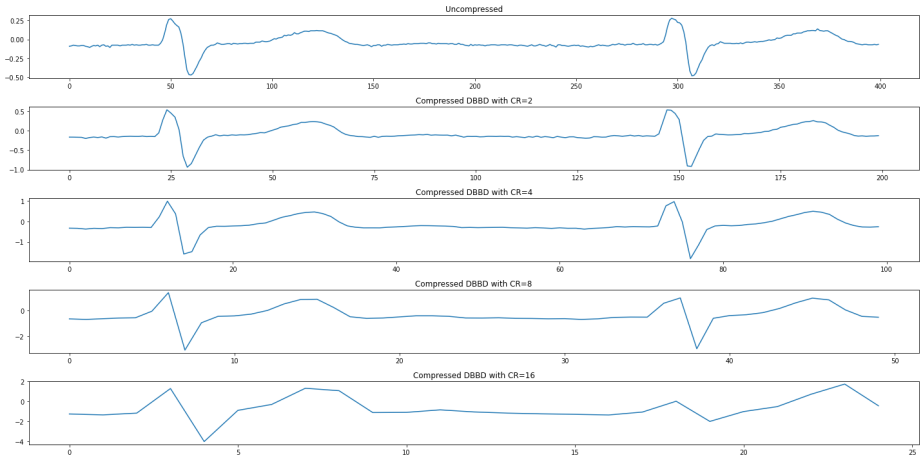
$\text{RQ}_\mathbf{1}$: *Which ECG compression algorithm provides the best classification performances of* RAST*?* With the second research question, we want to compare two different algorithms for ECG compression, namely DBBD (CA_{DBBD}) [30] and Bernoulli compression ($\text{CA}_{Bernoulli}$) [8], using different compression ratios (CRs).

$\text{RQ}_\mathbf{2}$: *What are the classification performances of* RAST *when dealing with* ***compressed*** *ECG signals?* We want to evaluate the applicability of RAST in a context where we have compressed ECG signals, using the same features applied for uncompressed ECGs in the previous study [34].

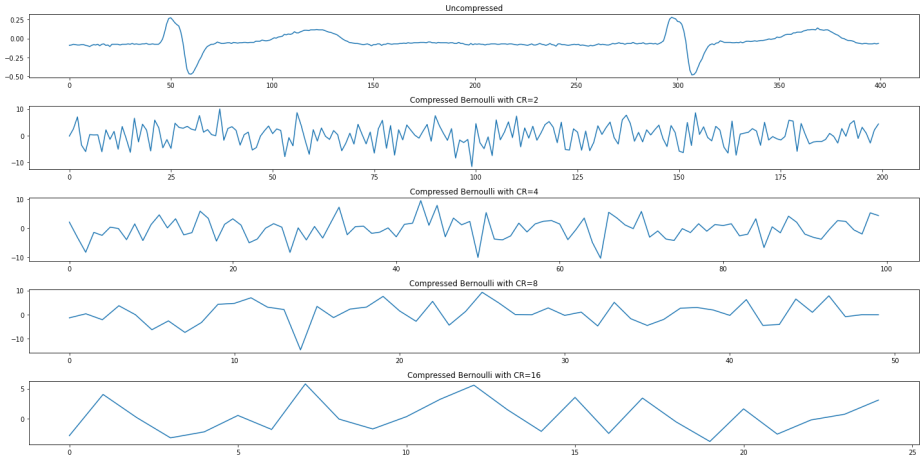
In Figure 3 we report a comparison between uncompressed and compressed ECG using CA_{DBBD} (fig. a) and $\text{CA}_{Bernoulli}$ (fig. b) compression.

Context of the study The context of our study is the Physionet European ST-T Database [13, 39], a state-of-the-art dataset to support the analysis of ST segment and T-wave morphology. This database is made up of 90 ambulatory two-hour ECG signals with annotations from 79 distinct individuals that were collected at a sample rate of 250 Hz and a resolution of 12 bits. Additionally, the information includes 401 events of T-wave transition and 367 instances of ST segment change. ranging from 30 seconds to several minutes in length. Two cardiologists defined such labels after working separately, beat by beat, and discussing any instances of label discrepancies at the conclusion.

In our study, We considered only 3 categories of annotations, related to ST segment changes, namely *NSR*, *ST+* and *ST-*. For the binary variant of RAST, we merged the two *ST* annotations obtaining only two types of heartbeats, *i.e., NSR* and *ST*. The distribution of the heartbeats reporting that labels is described in Figure 4 for both RAST_{bin} (left) and RAST_{ter} (right). It is worth saying that we did not use the MIT-BIH ST Change DB [2] as it does not include ST change annotations.

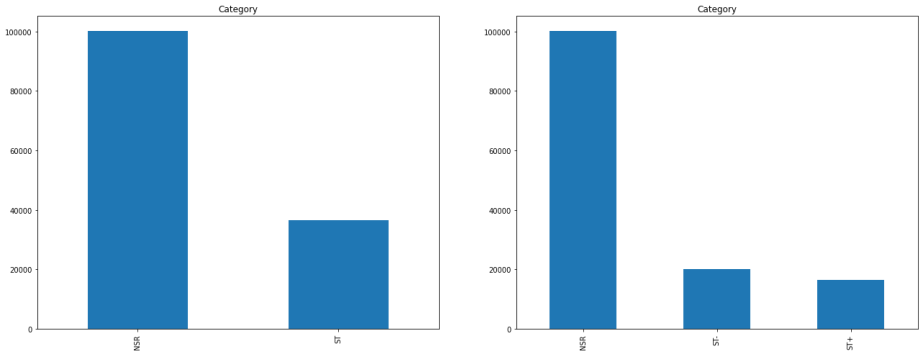


(a)



(b)

Fig. 3: Comparison between uncompressed and compressed ECG signal using CA_{DBBD} (fig. a), and $\mathrm{CA}_{Bernoulli}$ (fig. b) at different values of CR



(a) (b)

Fig. 4: Count of selected heartbeat from the European ST-T Database [39] used for RAST_{bin} (left) and RAST_{ter} (right)

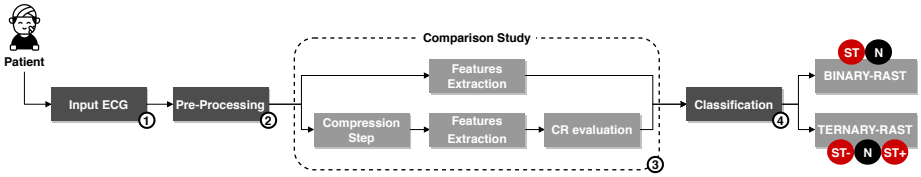


Fig. 5: The experimental workflow conducted in this study.

Experimental Workflow In Figure 5 we represent the experimental procedure to answer our RQs. In summary, the experimental procedure is the following:

1. An input ECG trace is extracted and buffered according to the minimum number of R peaks required for the classification (*i.e.,* 4);
2. Follow a preprocessing step where the signal is filtered and prepared for the feature extraction by dividing the input ECG in 4-beats segments;
3. At this point, we have the features extraction phase. Here, we (i) extract the features used by RAST on both compressed and uncompressed domains, and (ii) compare the two compression algorithm previously described (CA_{DBBD} and $\text{CA}_{Bernoulli}$), selecting the best compression ratio;
4. The last step is the training of the classifier and the detection of ST-related anomalies on both compressed and uncompressed signals for both RAST_{bin} and RAST_{ter}.

For this last step, we use the same validation scheme used in the previous study [34], and the settings which achieved the best results. In detail, we have (i) a TWHO (Temporal Window for the Heartbeat Observation) of 4, (ii) the SMOTE sampling technique applied on the training set, and (iii) the Random Forest classifier as machine learning model. We adopt the same validation scheme, *i.e.,* we perform a 80-20 random split cross validation, repeated 1000 times to avoid any bias due to convenient split. We ap-

ply it, in the two RQs, for the uncompressed and the compressed ECG signal using both CA_{DBBD} and $\mathrm{CA}_{Bernoulli}$ compression algorithms using 4 different CRs.

For convenience, We use RAST^C to identify the modified version of RAST where the compression procedure is applied, with its two variants RAST^C_{bin} and RAST^C_{ter}.

Thus, we apply the described procedure to answer both RQs. To answer the first RQ, we evaluate the classification performance for RAST^C by comparing the two compression algorithms (*i.e.,* CA_{DBBD} and $\mathrm{CA}_{Bernoulli}$) with different CRs. To answer the second, we compare the overall classification performance between RAST and RAST^C having the best performing combination of compression algorithm and CR. In our study, we evaluate both *binary* and *ternary* variants for RAST and RAST^C.

Moreover, we measure the classification performance for both RQs using the following class-level metrics:

- **Accuracy**: the number of correctly categorized instances divided by the total number of instances. $\frac{TP+TN}{TP+TN+FP+FN}$
- **Precision**: the number of correctly categorized positive instances divided by the total number of positive instances. $\frac{TP}{TP+FP}$
- **Specificity**: the number of correctly categorized negative instances divided by the total number of correctly classified negative instances and incorrectly classified positive instances. $\frac{TN}{TN+FP}$
- **Recall**: the number of correctly categorized positive instances divided by the total number of correctly categorized positive instances and incorrectly classified negative instances. $\frac{TP}{TP+FN}$
- **F1 Score**: the harmonic mean of precision and recall. $\frac{2*TP}{(2*TP)+TN+FP}$

4.2 Study Results

In this section we report the results of our empirical study.

RQ_1: Best ECG compression algorithm To answer RQ_1, we compare the overall classification performance between CA_{DBBD} and $\mathrm{CA}_{Bernoulli}$ for both RAST^C_{bin} and RAST^C_{ter}. CA_{DBBD} exhibit an overall better classification performance compared to $\mathrm{CA}_{Bernoulli}$. Also, the achieved results for the first are slightly more stable than the latter, which decreases as CR increases. Those results are the same for both settings of RAST^C, *i.e.,* RAST^C_{bin} , and RAST^C_{ter}. While CA_{DBBD} achieves similar results for both settings, $\mathrm{CA}_{Bernoulli}$ works better overall in RAST^C_{bin} . For RAST^C_{ter}, it achieves good performance only in terms of *specificity* and *accuracy*, which are also higher than the RAST^C_{bin} setting. This means that Bernoulli's features are more effective to identify the absence of ST anomaly, *i.e.,* achieving few false positive results. Thus, it is more useful to distinguish between Normal and ST anomalous ECG segments, than to classify Normal, ST Depression, or ST Elevation.

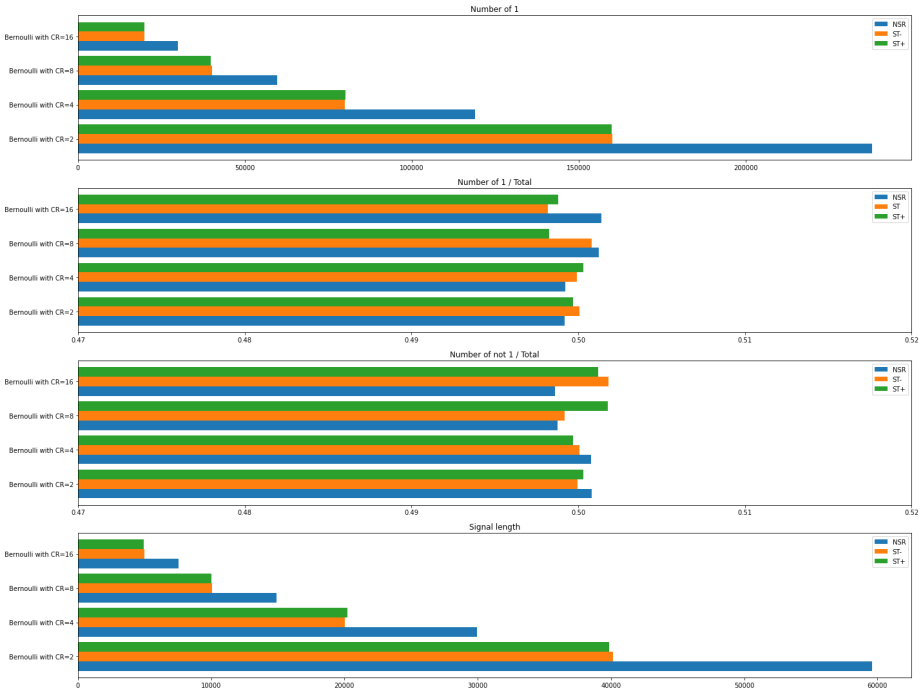
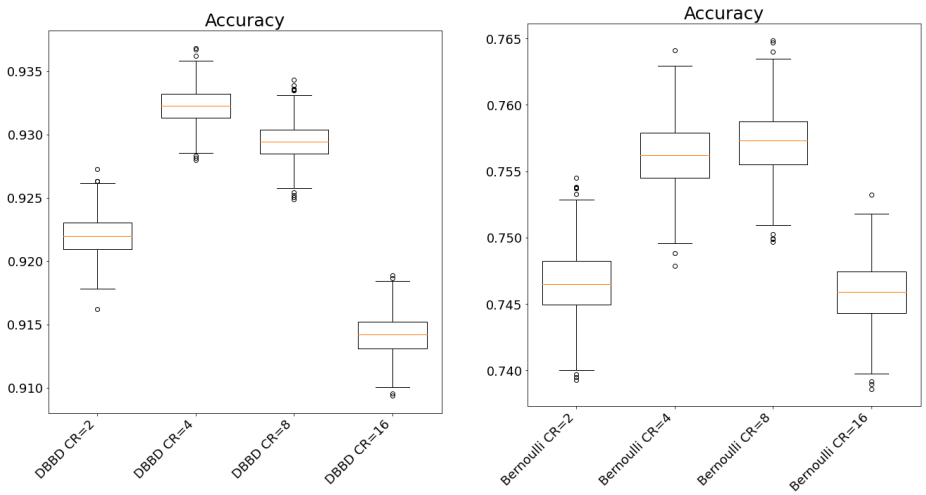
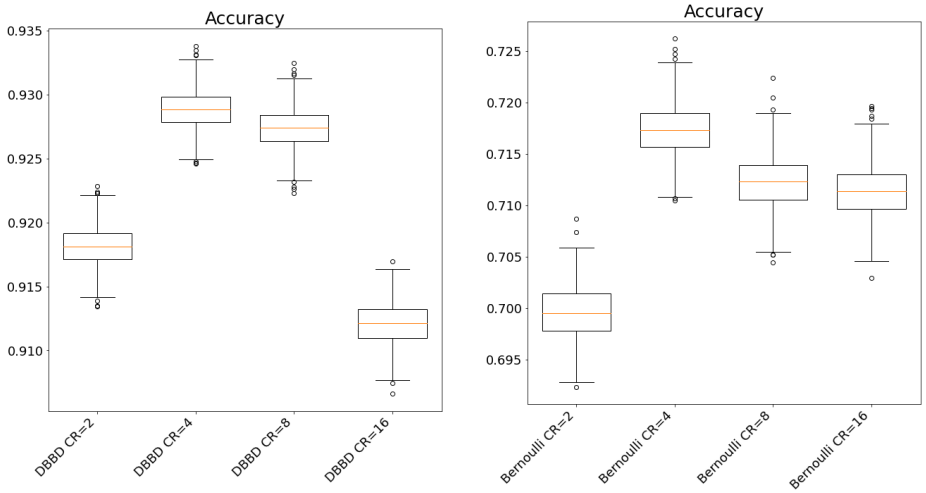


Fig. 6: Bernoulli's features evaluated for an ECG signal compressed using $\mathrm{CA}_{Bernoulli}$ with different CRs.



(a) (b)

Fig. 7: Comparison of the accuracy values for RAST^{C}_{bin} using CA_{DBBD} (fig. a), and $\mathrm{CA}_{Bernoulli}$ (fig. b) with different CRs



(a) (b)

Fig. 8: Comparison of the accuracy values for RAST^{C}_{ter} using CA_{DBBD} (fig. a), and $\mathrm{CA}_{Bernoulli}$ (fig. b) with different CRs

It is observed better in Figure 8, where, for each of the newly introduced Bernoulli's features (*i.e.,* the ones evaluated on the ϕ matrix, we report their value for NSR, ST elevation, and ST depression classes, calculated for different values of compression ratios. Their effectiveness for the NSR class is far higher compared to the others. Indeed, the binary discrimination between normal ECG and ST abnormality can be easily obtained thanks to these new attributes. For this particular case, it is evident that $\mathbf{tot}_1$ and **sequences** are the most suited for the binary scenario. For the ternary scenario, these attributes provide less information. In this case, $\mathbf{perc}_1$ and $\mathbf{tot}_{not1}$ seem the most appropriate.

We computed the boxplots for the values of the accuracy metric collected following our validation scheme (1,000 iterations performing 80-20 random splits). In detail, in Figure 7 we describe the distribution of the *accuracy* values gathered from the experimental procedure of RAST^{C}_{bin} using CA_{DBBD} (Figure 6 a), and $\mathrm{CA}_{Bernoulli}$ (Figure 6 b). The same is reported in Figure 8. We have a larger number of outliers when applying CA_{DBBD} , while for $\mathrm{CA}_{Bernoulli}$ we obtain more stable results among all the CRs in terms of *accuracy* at a price of a lower score overall.

Table 1: Classification metrics for RAST_{bin} in compressed and uncompressed domain. RAST^{C}_{bin} indicates the variant using CA_{DBBD} .

Approach	**Accuracy**	**Precision**	**Specificity**	**Recall**	**F1 Score**
RAST_{bin}	93.05	93.03	88.78	93.05	93.04
RAST^{C}_{bin} (CR=2)	92.20	92.40	**89.81**	92.20	92.27
RAST^{C}_{bin} (CR=4)	**93.23**	**93.23**	89.53	**93.23**	**93.23**
RAST^{C}_{bin} (CR=8)	92.94	92.99	89.76	92.94	92.97
RAST^{C}_{bin} (CR=16)	91.42	91.48	87.29	91.42	91.44

Table 2: Classification metrics for RAST_{ter} in compressed and uncompressed domain. RAST^{C}_{ter} indicates the variant using CA_{DBBD} .

Approach	**Accuracy**	**Precision**	**Specificity**	**Recall**	**F1 Score**
RAST_{ter}	**93.07**	**93.14**	91.20	**93.07**	**93.09**
RAST^{C}_{ter} (CR=2)	91.82	92.22	**91.87**	91.82	91.94
RAST^{C}_{ter} (CR=4)	92.88	92.99	91.23	92.88	92.92
RAST^{C}_{ter} (CR=8)	92.74	92.88	91.26	92.74	92.79
RAST^{C}_{ter} (CR=16)	91.21	91.35	88.92	91.21	91.26

RQ_2: Classification performances of RAST To answer RQ_2, we compare the overall classification performance between uncompressed vs. compressed domains for RAST. We reported the overall results for $RAST_{bin}$ and $RAST^C_{bin}$ in Table 1, and for $RAST_{ter}$ and $RAST^C_{ter}$ in Table 2. For $RAST^C$ we reported the results achieved using the best compression algorithm resulting from the previous RQ, *i.e.,* CA_{DBBD} . In the first table it can be observed that $RAST^C_{bin}$ offers similar or slightly better performance to the uncompressed version ($RAST_{bin}$), particularly with a compression rate of 4. For the second, $RAST_{ter}$ shows higher performance than $RAST^C_{ter}$ for almost all classification metrics, except for *specificity*, which is better for $RAST^C_{ter}$ with a CR of 2. As seen for the *binary* version, $RAST^C_{ter}$ achieves the best performance overall with a CR of 4 for the compressed domain.

In summary, the performance of RAST in both uncompressed and compressed domains are comparable, where the latter works at best using a compression ratio of 4.

5 Conclusion and future works

In this paper, we have proposed a new version of RAST, i.e. the approach proposed in our previous work [34]. The new approach, called $RAST^C$, has the big difference of working with a compressed ECG compared to a full ECG trace, i.e. without any loss of information. Performing detection directly on compressed signals allows IoMT devices to preserve battery life and to offer accurate detection with algorithms at a lower computational cost. To this end, in this paper we experimented with two different signal compression techniques: DBBD and Bernoulli. As the latter is a random compression technique, we introduced 4 new features—calculated on the ϕ matrix—to increase the information potential in the latter case. Furthermore, we evaluated the compression factor in the set [2,4,8,16] to observe the performance of $RAST^C$ even at very high CRs. And finally, we evaluated two classification versions: (i) a binary one, where the objective is simply the discrimination between normal ECG and an ST-abnormality and (ii) a ternary one where the objective of the classification model is to identify normal ECG, ST-depression and ST-elevation. The first major result obtained in this study showed that a deterministic DBBD technique is preferable to a random one. Thus, by applying a compression with a DBBD algorithm, it is possible to obtain—both in the binary and ternary case—average results above 90% even with a CR of 16. Consequently, this paper aims to confidently promote efforts dedicated to the detection of clinical features, such as ST-segment abnormalities, from the analysis of highly compressed ECG signals.

Future works will be dedicated to the study of more compression techniques, in order to achieve comparable results with higher compression ratios.

References

1. A novel compressive sampling method for ECG wearable measurement systems. Measurement **167**, 108259 (2021)
2. Albrecht, P.: St segment characterization for long term automated ecg analysis [dissertation]. Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science: Massachusetts Institute of Technology **378** (1983)

3. Balestrieri, E., Boldi, F., Colavita, A.R., De Vito, L., Laudato, G., Oliveto, R., Picariello, F., Rivaldi, S., Scalabrino, S., Torchitti, P., et al.: The architecture of an innovative smart t-shirt based on the internet of medical things paradigm. In: 2019 IEEE International symposium on medical measurements and applications (MeMeA). pp. 1–6. IEEE (2019)
4. Balestrieri, E., Daponte, P., De Vito, L., Picariello, F., Rapuano, S., Tudosa, I.: A Wi-Fi Internet-of-Things prototype for ECG monitoring by exploiting a novel compressed sensing method. Acta IMEKO **9**(2), 38 – 45 (2020)
5. Bhattarai, S., Chhabra, L., Hashmi, M.F., Willoughby, C.: Anteroseptal myocardial infarction (2019)
6. Brunner, E., Marmot, M., Canner, R., Beksinska, M., Smith, G.D., O'Brien, J.: Childhood social circumstances and psychosocial and behavioural factors as determinants of plasma fibrinogen. The Lancet **347**(9007), 1008–1013 (1996)
7. Bulusu, S.C., Faezipour, M., Ng, V., Nourani, M., Tamil, L.S., Banerjee, S.: Transient st-segment episode detection for ecg beat classification. In: 2011 IEEE/NIH Life Science Systems and Applications Workshop (LiSSA). pp. 121–124. IEEE (2011)
8. Craven, D., McGinley, B., Kilmartin, L., Glavin, M., Jones, E.: Compressed Sensing for Bioelectric Signals: A Review. IEEE Journal of Biomedical and Health Informatics **19**(2), 529–540 (2015)
9. De Vito, L., Picariello, E., Picariello, F., Rapuano, S., Tudosa, I.: A Dictionary Optimization Method for Reconstruction of ECG Signals after Compressed Sensing. Sensors **21**(16) (2021)
10. De Vito, L., Picariello, E., Picariello, F., Tudosa, I., Loprevite, L., Avicolli, D., Laudato, G., Oliveto, R.: An undershirt for monitoring of multi-lead ecg and respiration wave signals. In: 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT). pp. 550–555. IEEE (2021)
11. Dixon, A.M.R., Allstot, E.G., Gangopadhyay, D., Allstot, D.J.: Compressed Sensing System Considerations for ECG and EMG Wireless Biosensors. IEEE Transactions on Biomedical Circuits and Systems **6**(2), 156–166 (2012)
12. Ghiadoni, L., Donald, A.E., Cropley, M., Mullen, M.J., Oakley, G., Taylor, M., O'Connor, G., Betteridge, J., Klein, N., Steptoe, A., et al.: Mental stress induces transient endothelial dysfunction in humans. Circulation **102**(20), 2473–2478 (2000)
13. Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. circulation **101**(23), e215–e220 (2000)
14. Harhash, A.A., Huang, J.J., Reddy, S., Natarajan, B., Balakrishnan, M., Shetty, R., Hutchinson, M.D., Kern, K.B.: avr st segment elevation: acute stemi or not? incidence of an acute coronary occlusion. The American Journal of Medicine **132**(5), 622–630 (2019)
15. Jager, F., Taddei, A., Moody, G.B., Emdin, M., Antolič, G., Dorn, R., Smrdel, A., Marchesi, C., Mark, R.G.: Long-term st database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. Medical and Biological Engineering and Computing **41**(2), 172–182 (2003)
16. Kandala, V.K., Vadaparthi, J.K.: Study of incidence and pattern of ecg changes in cerebrovascular accidents. Radiology **3**(1), 107–109 (2018)
17. Kashou, A.H., Basit, H., Malik, A.: St segment. In: StatPearls [Internet]. StatPearls Publishing (2021)
18. Khoury, S., Soleman, M., Margolis, G., Barashi, R., Rozenbaum, Z., Keren, G., Shacham, Y.: Incidence, characteristics and outcomes in very young patients with st segment elevation myocardial infarction. Coronary artery disease **31**(2), 103–108 (2020)

19. Kop, W.J., Krantz, D.S., Howell, R.H., Ferguson, M.A., Papademetriou, V., Lu, D., Popma, J.J., Quigley, J.F., Vernalis, M., Gottdiener, J.S.: Effects of mental stress on coronary epicardial vasomotion and flow velocity in coronary artery disease: relationship with hemodynamic stress responses. Journal of the American College of Cardiology **37**(5), 1359–1366 (2001)
20. Laudato, G., Boldi, F., Colavita, A., Rosa, G., Scalabrino, S., Torchitti, P., Lazich, A., Oliveto, R.: Combining rhythmic and morphological ecg features for automatic detection of atrial fibrillation. In: 13th International Conference on Health Informatics. pp. 156–165 (2020)
21. Laudato, G., Picariello, F., Scalabrino, S., Tudosa, I., de Vito, L., Oliveto, R.: Morphological classification of heartbeats in compressed ecg. In: 14th International Conference on Health Informatics, HEALTHINF 2021-Part of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2021. pp. 386–393. SciTePress (2021)
22. Laudato, G., Boldi, F., Colavita, A.R., Rosa, G., Scalabrino, S., Lazich, A., Oliveto, R.: Combining rhythmic and morphological ecg features for automatic detection of atrial fibrillation: Local and global prediction models. In: International Joint Conference on Biomedical Engineering Systems and Technologies. pp. 425–441. Springer (2020)
23. Laudato, G., Oliveto, R., Scalabrino, S., Colavita, A.R., De Vito, L., Picariello, F., Tudosa, I.: Identification of r-peak occurrences in compressed ecg signals. In: 2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA). pp. 1–6. IEEE (2020)
24. Laudato, G., Rosa, G., Capobianco, G., Colavita, A.R., Dal Forno, A., Divino, F., Lupi, C., Pareschi, R., Ricciardi, S., Romagnoli, L., et al.: Simulating the doctor's behaviour: A preliminary study on the identification of atrial fibrillation through combined analysis of heart rate and beat morphology (2022)
25. Laudato, G., Scalabrino, S., Colavita, A.R., Chiacchiari, Q., D'Orazio, R., Donadelli, R., De Vito, L., Picariello, F., Tudosa, I., Malatesta, R., et al.: Atticus: Ambient-intelligent telemonitoring and telemetry for incepting and catering over human sustainability. Frontiers in Human Dynamics p. 19 (2021)
26. Maglaveras, N., Stamkopoulos, T., Pappas, C., Strintzis, M.G.: An adaptive backpropagation neural network for real-time ischemia episodes detection: development and performance analysis using the european st-t database. IEEE Transactions on Biomedical Engineering **45**(7), 805–813 (1998)
27. Mattioli, A.V., Nasi, M., Cocchi, C., Farinetti, A.: Covid-19 outbreak: impact of the quarantine-induced stress on cardiovascular disease risk burden (2020)
28. Pan, J., Tompkins, W.J.: A real-time qrs detection algorithm. IEEE transactions on biomedical engineering (3), 230–236 (1985)
29. Quwaider, M., Biswas, S.: On-body packet routing algorithms for body sensor networks. In: 2009 first international conference on networks & communications. pp. 171–177. IEEE (2009)
30. Ravelomanantsoa, A., Rabah, H., Rouane, A.: Compressed Sensing: A Simple Deterministic Measurement Matrix and a Fast Recovery Algorithm. IEEE Transactions on Instrumentation and Measurement **64**(12), 3405–3413 (2015)
31. Rehman, A., Saba, T., Haseeb, K., Larabi Marie-Sainte, S., Lloret, J.: Energy-efficient iot e-health using artificial intelligence model with homomorphic secret sharing. Energies **14**(19), 6414 (2021)
32. Rghioui, A., Lloret, J., Harane, M., Oumnad, A.: A smart glucose monitoring system for diabetic patient. Electronics **9**(4), 678 (2020)
33. Rosa, G., Laudato, G., Colavita, A.R., Scalabrino, S., Oliveto, R.: Automatic real-time beat-to-beat detection of arrhythmia conditions. In: HEALTHINF. pp. 212–222 (2021)
34. Rosa, G., Russodivito, M., Laudato, G., Colavita, A.R., Scalabrino, S., Oliveto, R.: A robust approach for a real-time accurate screening of st segment anomalies. In: HEALTHINF. pp. 69–80 (2022)

35. Rosa, G., Russodivito, M., Laudato, G., Scalabrino, S., Colavita, A.R., Oliveto, R.: A multi-class approach for the automatic detection of congestive heart failure in windowed ecg. Studies in health technology and informatics **290**, 650–654 (2022)
36. Rosengren, A., Hawken, S., Ôunpuu, S., Sliwa, K., Zubaid, M., Almahmeed, W.A., Blackett, K.N., Sitthi-Amorn, C., Sato, H., Yusuf, S., et al.: Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the interheart study): case-control study. The Lancet **364**(9438), 953–962 (2004)
37. Ryu, K.S., Bae, J.W., Jeong, M.H., Cho, M.C., Ryu, K.H., Investigators, K.A.M.I.R., et al.: Risk scoring system for prognosis estimation of multivessel disease among patients with st-segment elevation myocardial infarction. International heart journal **60**(3), 708–714 (2019)
38. Srivastava, J., Routray, S., Ahmad, S., Waris, M.M.: Internet of medical things (iomt)-based smart healthcare system: Trends and progress. Computational Intelligence and Neuroscience **2022** (2022)
39. Taddei, A., Distante, G., Emdin, M., Pisani, P., Moody, G., Zeelenberg, C., Marchesi, C.: The european st-t database: standard for evaluating systems for the analysis of st-t changes in ambulatory electrocardiography. European heart journal **13**(9), 1164–1172 (1992)
40. Tsuji, H., Shiojima, I.: Increased incidence of ecg abnormalities in the general population during the covid-19 pandemic. International Heart Journal **63**(4), 678–682 (2022)
41. Wang, H., Zhao, W., Xu, Y., Hu, J., Yan, C., Jia, D., You, T.: St segment change classification based on multiple feature extraction using ecg. In: 2018 Computing in Cardiology Conference (CinC). vol. 45, pp. 1–4. IEEE (2018)
42. Wei, W., Qi, Y.: Information potential fields navigation in wireless ad-hoc sensor networks. Sensors **11**, 4794–4807 (2011)
43. Xiao, R., Xu, Y., Pelter, M.M., Mortara, D.W., Hu, X.: A deep learning approach to examine ischemic st changes in ambulatory ecg recordings. AMIA Summits on Translational Science Proceedings **2018**, 256 (2018)