11th International Conference on Electrical, Electronics and Computer Engineering (IcETRAN), Nis, 3-6 June 2024

SimEMG database as a tool for testing the preservation of diagnostic ECG-signal features upon the electromyographic noise removal

Vladimir Atanasoski
Vinca Institute of Nuclear Sciences, University of Belgrade, National Institute of the Republic of Serbia
Belgrade, Serbia
vladimir.atanasoski@vin.bg.ac.rs,
https://orcid.org/0000-0003-1695-6341

Lana Popović Maneski
Department of Institute of Technical Sciences of the SASA
Belgrade, Serbia
lanapm13@gmail.com,
https://orcid.org/0000-0002-9498-5546

Marija D. Ivanović
Vinca Institute of Nuclear Sciences, University of Belgrade, National Institute of the republic of Serbia
Belgrade, Serbia
marijap@vin.bg.ac.rs,
https://orcid.org/0000-0002-7218-6575

Marjan Miletić
Vinca Institute of Nuclear Sciences, University of Belgrade, National Institute of the Republic of Serbia
Belgrade, Serbia
marjanmil@yahoo.com,
https://orcid.org/0000-0002-8476-4141

Miloš Babic
Institute for Cardiovascular Diseases Dedinje
Belgrade, Serbia
babicmisa@hotmail.com,
https://orcid.org/0000-0002-3226-003X

Jovana Petrovic
Vinca Institute of Nuclear Sciences, University of Belgrade, National Institute of the Republic of Serbia
Belgrade, Serbia
jovanap@vin.bg.ac.rs,
https://orcid.org/0000-0002-1002-241X

Aleksandra Nikolić
Institute for Cardiovascular Diseases Dedinje
Belgrade, Serbia
nikolicdrsasa@gmail.com,
https://orcid.org/0000-0003-2724-2854

***Abstract*— ECG measurements at rest are susceptible to different kinds of noise which may affect clinical interpretation. Unlike the low-frequency body movements or spectrally well-defined power line interference, the broadband electromyographic (EMG) noise spectrally coincides with the QRS complex, which makes its removal particularly challenging. While a number of algorithms have been developed, their evaluation and comparison have long remained challenging due to the lack of a database with EMG-contaminated and genuine noise-free signals. We have recently published a SimEMG database of the simultaneously recorded EMG-noise-free and -contaminated ECG signals. In this paper, we examine the capability of SimEMG to become a reference test set for different denoising algorithms. We show that the SimEMG is useful for the assessment of the signal morphology distortions leading to potentially misleading changes in the diagnostic markers. We examine the ST segment – a marker of acute myocardial infarction, and R, P_{max}, and T_{max} points– the markers of arrhythmias. The tests are performed with the adaptive wavelet Wiener filter, wavelet transform (WT), finite impulse response (FIR) filter, and iterative regenerative method (IRM). The analysis also reveals the residual noise as a minor limitation of the database and acquisition method. Possible improvements are discussed.**

Keywords*— *ECG acquisition; electromyographic noise; denoising.

I. Introduction

ECG signals measured by mobile devices are frequently contaminated by noises of different origins and spectral characteristics, which makes their removal very challenging [1, 2]. Typical noises are the low-frequency (<1 Hz) baseline wander (BLW), power-line interference (PLI) – a narrowband (50/60 Hz) component, motion artifacts (MA) with a spectrum in the range [1–10] Hz, and electromyographic (EMG) noise with the broad spectrum spreading above 10 Hz. While the BLW and PLI can be relatively easily filtered out, MA and EMG significantly overlap with the frequency content of the diagnostically relevant information stored in the QRS complex. Particularly challenging is the removal of EMG noise as the underlying muscle movements are involuntary and cannot be stabilized by measuring at rest as MAs can. Consequently, several denoising algorithms have been applied in postprocessing, which often results in changes to the signal morphology, diminishing the accuracy of the ECG-feature based diagnostics [1, 3].

However, due to the conundrum of the noiseless signal measurement: "*The fact that the true underlying dynamics of a real ECG can never be known implies that one cannot distinguish between the clean ECG signal and the many sources of noise that can occur during recording*" [4], it is difficult to evaluate the performance of and select an optimum denoising algorithm. The data sets used for algorithm evaluation are typically constructed from the 'noise-free' ECGs selected among the cleanest signals from a given public database or synthesized, and a noise taken from a public database, such as the MIT noise stress test database, or produced by filtering a white Gaussian noise to match the EMG noise [5], [6], [7].

We have recently published SimEMG, a database of genuine noise-free and EMG-noise-contaminated signals simultaneously recorded by a novel measurement scheme [8]. In this paper, we present the possibilities for its use in the assessment of different denoising algorithms, namely adaptive wavelet Wiener filter (AWWF) [7], wavelet transform (WT)

[8], finite impulse response (FIR) [9] and iterative regenerative method (IRM) [10]. The comparison is focused on the investigation of distortion of the diagnostically relevant features: ST elevation – the marker of myocardial infarction (MI), R-point – the indispensable feature in the assessment of arrhythmias [11], P wave – significant in detection of atrial fibrillation [12] and T wave – associated with life-threatening arrhythmias [13] and MI (Fig. 1). Finally, we discuss the limitations of the measurement scheme and propose directions for its further development and use.

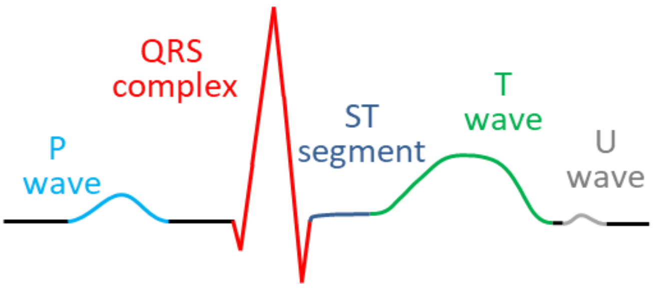

Fig. 1. Components of a single heartbeat.

II. The Method

A. SimEMG database

A detailed description of the SimEMG acquisition method is given in [14]. Here, we briefly review its main assumptions that (i) the arm shortcuts any two potential points along it and, hence, that the potential at every point along the arm is constant when muscles are relaxed [15] and (ii) the EMG generated in hands, in particular fingers, dominates the EMG noise generated at a proximal position on the arm. Hence, the ECG measurement performed with finger electrodes yields noisy signals, while the measurement performed with the electrodes placed near the shoulders yields a signal with a near-zero EMG noise. The corresponding ECG signals are recorded using the limb leads in the standard configuration and the following precordial electrode configuration: V1-V2 electrodes on the shoulders to obtain the reference (noise-free) signal, V3-V4, and V5-V6 pairs on the intermediate and proximal parts of fingers of the left and right arms, respectively. Therefore, we can obtain 4 single-lead signals per recording: one reference and three EMG-noise-contaminated signals. EMG noise was introduced by activating hand muscles either by pressing the fingers of one hand against each other or by pressing the object that causes resistance. To additionally vary the level of EMG noise, we recorded the ECG of healthy subjects in the supine position with arms in three different postures: resting along the body for low noise levels, the forearms leaning on the hips at the approximate 45-degree angle relative to the bed, and the arms pointing upright, with elbows supported on the bed next to the body.

In order to obtain reference signals with good quality, the signals that do not fulfill signal quality criteria are rejected from further analysis [14].

The database is summarized in Table I. The level of noise is quantified as the signal-to-noise ratio (SNR) using the standard definition [10].

This work was approved by the Human Research Ethics Committee of the Institute for Cardiovascular Diseases Dedinje, Serbia.

TABLE I
SIMECG DATABASE SUMMARY

Description	Parameter	Value
Demography	Number of subjects	14
	Female	9
	Male	5
	Age	40 ± 13
Signals	Total number	147
	Noise-free	37
	With EMG noise	110
Signal distribution according to the SNR	SNR [dB]	Number of signals
	<4	33
	4-8	19
	8-12	23
	12-16	28
	16-20	6
	>20	1

B. Denoising methods

The IRM method removed the EMG noise in a number of iterations determined by the initially estimated level of noise. The main idea behind the method is to temporarily remove the main features of the signal, notably the QRS complex, as well as the low-frequency components, upon which the EMG noise is easily extracted and eliminated. The main feature removal is achieved by ensemble averaging. Upon the noise removal, the signal is reassembled, thus restoring its morphology and relevant features [10].

The AWWF method utilizes the dyadic stationary wavelet transform within a Wiener filter framework. It enhances signal quality by incorporating a noise estimation block to track the signal's time-dependent SNR. Moreover, it incorporates an algorithm to optimize parameter values, maximizing the average improvement in SNR [7].

WT, a conventional method for ECG signal filtering, was applied following the procedure in [8]. Furthermore, parameter optimization (wavelet family, lower and higher thresholds) was conducted using the SimEMG dataset, resulting in the utilization of sym4 filter banks, a decomposition level of 5, and hard thresholding for managing cD3 and cD4 coefficients.

For completeness, we applied the conventional FIR filter in the form of a low-pass Butterworth filter with a cutoff frequency set at 40 Hz [9].

The complete numerical analysis was performed in MATLAB, Math Works Inc.

C. Performance metric

The signal quality upon the EMG noise removal was assessed using the SNR improvement factor (SNR_{IMP}) defined as:

$$SNR_{IMP} = SNR_{OUT} - SNR_{IN}\,, \qquad (1)$$

where $SNR_{OUT/IN}$ are the SNR of the output and input signals defined on an interval (here, ST segment) or the whole signal.

The preservation of the amplitudes of the fiducial points, namely the peaks of the QRS complex (R point), P (P_{max}) and T (T_{max}) waves, and J point, was assessed by calculating Pearson coefficients at these points defined by the crosscorrelation:

$$\rho(A,B) = 100 * \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{A_i - \mu_A}{\sigma_A} \right) * \left(\frac{A_i - \mu_B}{\sigma_B} \right) \quad (2)$$

where A_i and B_i are the amplitudes of fiducial points of the denoised and recorded noise-free signal, respectively, while $\mu_{A/B}$ and $\sigma_{A/B}$ are their respective mean and standard deviations on the whole signal. Fiducial points were determined by a human reader and verified by a cardiologist.

Finally, as the general morphology-preservation check, we use the eq. (2) with A and B being the denoised and noise-free signal amplitudes at all samples and $\mu_{A/B}$ and $\sigma_{A/B}$ their respective mean and standard deviations to evaluate the cross-correlation of the whole signal amplitudes (xcorr).

III. Main Results

Figure 2 shows examples of typical signals from the SimEMG database contaminated at different levels of noise. The noise is estimated as the difference between the noisy and noise-free recorded signals. The corresponding SNR_{IN} is used to stratify the data set according to the noise level. The spectral analysis of the data set confirms that the EMG noise above 10 Hz dominates other noises.

While in [10], we use SimEMG to investigate SNR improvement (SNR_{IMP}) over the whole signal, here, we concentrate on the SNR improvement on the ST interval, defined relative to the J point (from J to J+60 ms) [16]. In healthy people, the ECG signal on the ST segment is isoelectrical, i.e., it has near-zero values. Its elevation in either a positive or negative direction is a typical marker of MI and has tremendous significance in detecting acute MI in emergency units. Due to the small amplitudes, the ST segment is susceptible to distortions induced by the filtering operations, which can result in unnecessary referrals to emergency or, worse, in false negative MI.

Table II

Average values of SNR_{IMP} on the whole signal and ST segment obtained by different denoising methods. Analysed are the signals in different input SNR ranges and the whole set

SNR_{in} [dB]	**N**	**IRM [dB]**		**AWWF [dB]**		**WT [dB]**		**FIR [dB]**	
		All	ST	All	ST	All	ST	All	ST
<4	33	**12.4**	**8.3**	11.1	6.3	6.5	0.9	6.1	-0.2
4-8	19	**16.1**	**11.3**	14.6	8.4	9.9	3.9	9.9	3.2
8-12	23	**17.1**	**12.0**	16.3	10.4	12.2	6.6	13.2	6.9
12-16	28	**19.5**	**14.2**	18.8	12.7	12.9	8.4	16.4	10.5
16-20	6	**21.0**	**15.4**	20.7	14.6	15.1	8.1	18.9	13.5
>20	1	22.5	13.9	**23.2**	**14.0**	13.9	2.3	21.4	13.3
total	110	**16.4**	**11.5**	15.4	9.7	10.5	4.9	11.7	5.5

The best results on whole signal or ST level are shown bolded.

In Table II, we compare the denoising algorithms (IRM, AWWF, WT, and FIR) on the SimEMG database to determine the filter performance on this segment. Regardless of the method, the SNR_{IMP} evaluated on ST is significantly lower than on the whole signal, thus confirming the sensitivity of this segment to filtering. The IRM achieves the strongest noise reduction on all signals, especially on the ST segment, except the low-noise (SNR > 20 dB) signals. Standard filtering method FIR can even increase the noise content when applied to signals with SNR < 4 dB). WT is more reliable than FIR, but not at the level of AWWF and IRM. We note that also other metrics, such as RMSE, normalized RMSE, or noise reduction factor, can be used for the assessment of filtering methods [10]. A comparison between IRM and WT methods is shown in Fig. 3.

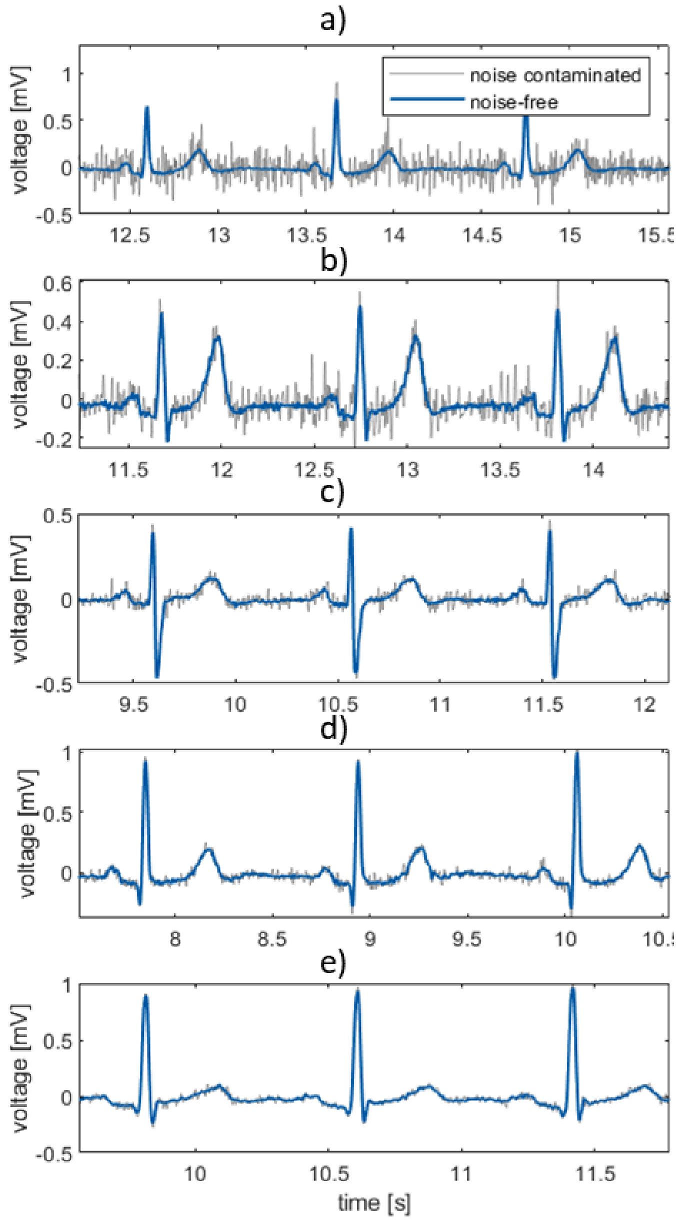

Fig. 2. Examples of EMG-contaminated and noise-free signals with different SNR_{IN}: a) 0.3 dB, b) 6.2 dB, c) 11 dB, d) 14.5 dB, e) 18.2 dB. Displayed signals are from different subjects.

Fiducial points marking the P-wave, QRS-complex, and T-wave maxima and their temporal positions, are customarily used in the detection of arrhythmias. J point marks the end of depolarization and the beginning of repolarization and is used to define the ST segment. Therefore, it is important that the noise removal procedure minimally distorts the signal morphology in the vicinity of these extrema.

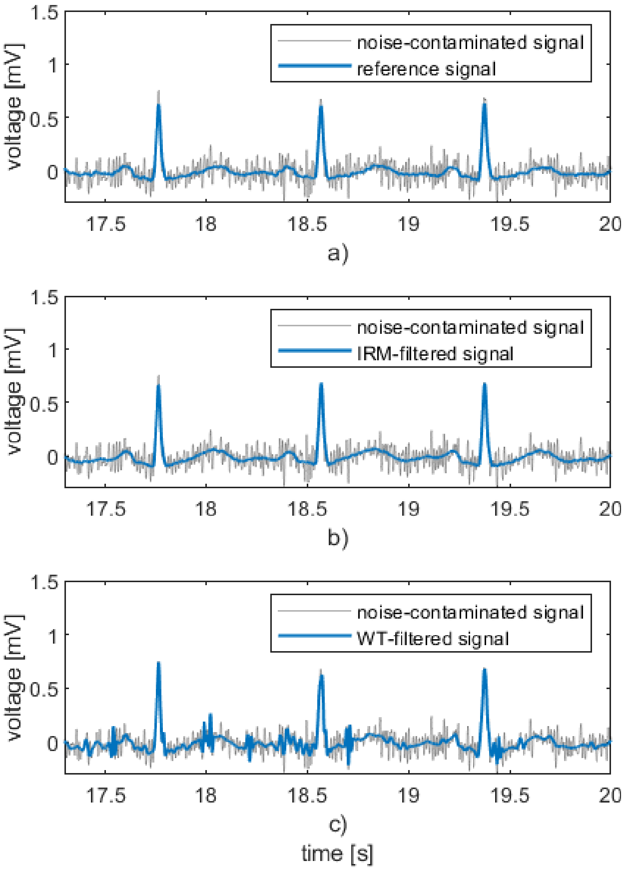

Fig. 3. An example of a) reference (noise-free) signal, b) IRM-filtered signal and c) signal filtered with the WT methods. A signal contaminated with EMG noise is displayed as gray.

The SimEMG database enables evaluation of the denoising algorithms in this aspect by providing noise-free signals with clean ECG segments and clear fiducial points. Fig. 4 illustrates comparison of denoising techniques with the Pearson coefficients of the signal amplitudes at P_{max}, R, J, and T_{max} points as a metric. For this analysis, all fiducial points from all signals were annotated by a biomedical engineer.

The global crosscorrelation evaluated on the whole signal shows excellent morphology preservation by all methods. This is also true for the R point, while at the T_{max} point, the Pearson of WT and FIR filters falls below 90%. Pearson coefficient at P_{max} and J points remains around 80% only for the IRM method, just above 70% for AWWF, while for the WT and FIR methods, it falls below 60%. The assessment shows that the global crosscorrelation is a good indicator of the segments with higher amplitudes (R and T_{max} points), which can be explained by their greater weight in the global crosscorrelation. Importantly, the morphology preservation of the diagnostically relevant segments with lower amplitudes must be assessed separately to ascertain the fair comparison of different methods.

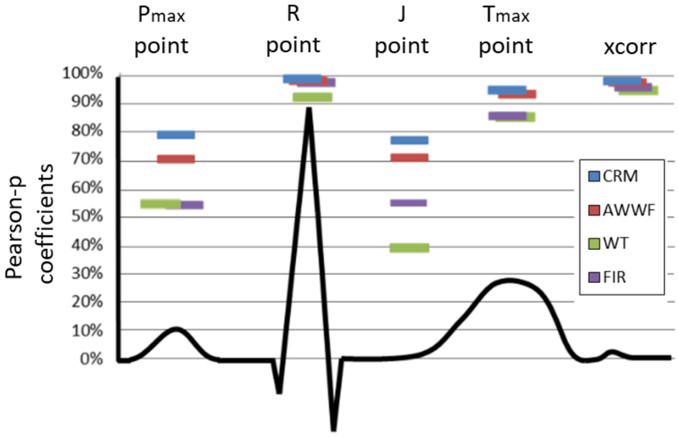

Fig. 4. Pearson coefficients for amplitudes of noise-free and denoised signals at fiducial points evaluated by the IRM (blue line), AWWF (red line), WT (green line), FIR (purple line) methods. Xcorr is the crosscorrelation between these signals calculated on the whole signal. The black line shows an example heartbeat to illustrate the positions of fiducial points.

IV. Conclusion

The open-source SimEMG database recorded using this principle is a unique dataset with genuine noise-free and noisy ECG signals. Here, we have shown that it is suitable for comparison of the denoising techniques in terms of signal morphology preservation. In doing this, particular attention must be paid to a fair assessment of the individual diagnostic features, which do not have large weight in the global metrics, such as P_{max} and J points, and the ST interval. The reference signals in the SimEMG database contain a low level of residual noise. As a further improvement of the acquisition technique, towards the noiseless reference signals, we propose recording a multi-lead ECG signal on the torso in the vicinity of the arms and using it to reconstruct signals without EMG and other noises caused by body movements.

The SimEMG database and acquisition method may be of use to the researchers and practitioners measuring EMG signals, whereby, the noise-free SimEMG ECG signal can be used for elimination of the cross-talk during an EMG assessment.

Acknowledgment

The authors thank volunteers and the Institute for Cardiovascular Diseases Dedinje, Serbia for logistic support of the study.

Vladimir Atanasoski, Jovana Petrovic, Marjan Miletić, and Marija D. Ivanović acknowledge the financial support by HeartBeam Inc. and the Ministry of Science, Technological Development and Innovation Grant No. 451-03-66/2024-03/200017.

Lana Popovic Maneski acknowledges the financial support by the Ministry of Science, Technological Development and Innovation Grant No. 451-03-66/2024-03/ 200175.

References

[1] K. H. C. Li, F. A. White, T. Tipoe, T. Liu M.C.S. Wong, A. Jesuthasan, A. Baranchuk, G. Tse, B. P. Yan, "The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative", *JMIR Mhealth Uhealth*., vol. 7, no. 2, e11606, 2019.

[2] L. Sornmo, and P. Laguna, "Noise and Artifacts," in *Bioelectrical signal processing in cardiac and neurological applications*, 1st ed. Burlington, AP, 2005, pp. 441.

[3] A. D. William, M. Kanbour, T. Callahan et al. "Assessing the accuracy of an automated atrial fibrillation detection algorithm using smartphone technology: The iREAD Study," *Heart Rhythm.* vol. 15, pp. 1561-1565, 2018

[4] Patrick E. McSharry, and Gari D. Clifford, "A comparison of nonlinear noise reduction and independent component analysis using a realistic dynamical model of the electrocardiogram," *Proc. SPIE Fluctuations and Noise in Biological, Biophysical, and Biomedical Systems II*, 5467, 2004.

[5] G.B. Moody, W.E. Muldrow, R.G. Mark, "A noise stress test for arrhythmia detectors," *Comput. Cardiol*. vol. 11, pp. 381-384, 1984.

[6] A.L. Goldberger, A. Amaral, L. Glass et al., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circ.* vol. 101, pp. 215-220, 2000.

[7] L. Smital, M. Vitek, J. Kozumplík, and V. Provazník, "Adaptive Wavelet Wiener Filtering of ECG Signals," *IEEE. Trans. Biomed. Eng.* vol. 60, no. 2, pp. 437-445, 2013

[8] S. Ardhapurkar, R. Manthalkar, and S. Gajre, "ECG Denoising by Modeling Wavelet Sub-Band Coefficients using Kernel Density Estimation," *J Inf Process Syst.* vol. 8, no. 4, pp. 669-684, 2012

[9] V. Atanasoski, J. Petrovic, L. Z. Popovic Maneski et al., "A morphology-preserving algorithm for denoising of EMG-contaminated ECG signals", *IEEE Open J. Eng. Med. Biol*, pp. 1-10, 2024

[10] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, "A Nonlinear Bayesian Filtering Framework for ECG Denoising," *IEEE Trans Biomed Eng.* vol. 54, pp. 2172-2185, 2007

[11] O. Faust and U.R. Acharya, "Automated classification of five arrhythmias and normal sinus rhythm based on RR interval signals," *Expert Sys. Appl.*, vol. 181, p.115031, 2021.

[12] F. Censi, I. Corazza, E. Reggiani, G. Calcagnini, E. Mattei, M. Triventi, and G. Boriani, "P-wave Variability and Atrial Fibrillation," *Sci. Rep.* 6, 26799, 2016.

[13] J. Ramirez, A. Kiviniemi, S. van Duijvenboden, et al., „ECG T-Wave morphologic variations predict ventricular arrhythmic Risk in low- and moderate-risk populations," *J. Am. Heart Assoc*., 11, e025897, 2022.

[14] V. Atanasoski, J. Petrovic, L. Popovic Maneski et al., "A database of simultaneously recorded ECG signals with and without EMG noise", *IEEE Open J. Eng. Med. Biol*, vol. 4, pp. 222-225, 2023.

[15] F.N. Wilson. "The distribution of the potential differences produced by the heart beat within the body and at its surface", *Am. Heart J.* vol. 5, no. 3, pp. 599-616, 1930.

[16] Y. Goletsis, C. Papaloukas, D. I. Fotiadis, A. Likas and L. K. Michalis, "Automated ischemic beat classification using genetic algorithms and multicriteria decision analysis," *IEEE Trans Biomed Eng.* vol. 51, no. 10, pp. 1717-1725, 2004.