

AN INNOVATIVE METHOD FOR DIAGNOSING SELECTED CARDIAC PATHOLOGIES BASED ON TIME-FREQUENCY ANALYSIS OF PHONOCARDIOGRAPHIC SIGNALS

Wojciech Lejkowski¹⁾, Andrzej Piotr Dobrowolski²⁾, Robert Olszewski^{3,4)}

1) BITRES Company, T. Chałubińskiego 9/2 Street, 02-004 Warsaw, Poland (✉ wojciech.lejkowski@bitres.pl)

2) Military University of Technology, Faculty of Electronics, Institute of Electronic Systems, ul. Sylwestra Kaliskiego 2, 00-908 Warsaw, Poland

3) Department of Gerontology and Public Health, National Institute of Geriatrics, Rheumatology and Rehabilitation, Spartańska 1 Street, Warsaw, Poland

4) Department of Ultrasound, Institute of Fundamental Technological Research, Polish Academy of Sciences, Pawińskiego 5B Street, Warsaw, Poland

Abstract

The article presents a method for diagnosing selected pathological heart conditions based on the time-frequency analysis of phonocardiographic signals, enabling its application by nursing and medical personnel. The study describes the process of collecting diagnostic data, methods for assessing data quality, and the use of a classifier to distinguish between healthy and diseased cases, where the diseased group includes multiple cardiac pathologies. Phonocardiographic signals were acquired by medical staff using a Littmann 3200 electronic stethoscope and the Littmann *StethAssist* application, and each recording was validated by one of the authors who is a physician. A total of 103 signals acquired in the sitting position at the mitral valve auscultation point were used to determine the performance parameters of the proposed method. Key challenges related to phonocardiographic signal acquisition are discussed, including noise elimination techniques that impact analysis accuracy. The selection of diagnostic features and the optimization of the classifier are presented, with features proposed by the authors and a dedicated feature selection strategy to ensure minimal classification errors while maintaining high signal separation efficiency. The conducted studies confirm the feasibility of developing an effective diagnostic method that allows for the classification of patients as healthy or diseased with sufficient accuracy for medical applications, achieving an AUC of 0.92. The results indicate the potential use of this technology in outpatient settings, opening new perspectives for non-invasive cardiac disease diagnostics.

Keywords: phonocardiogram, spectral analysis, support vector machine, genetic algorithms, distinctive feature selection.

1. Introduction

Cardiovascular diseases remain one of the leading causes of death worldwide. According to the *World Health Organization* (WHO) data, heart diseases are the primary cause of mortality globally, affecting over 13% of the world population [1]. Modern cardiological diagnostics utilize

advanced methods such as *electrocardiography* (ECG), *echocardiography*, *magnetic resonance imaging* (MRI), and *computed tomography* (CT). However, auscultation using a conventional stethoscope still plays a crucial role in primary healthcare. This method, despite its widespread use, is subjective and depends on the physician's experience and individual perceptual abilities. Consequently, even specialists may fail to detect subtle acoustic abnormalities that could indicate the onset of a cardiac pathology.

Phonocardiography (PCG) extends traditional auscultation by enabling graphical recording of heart sounds [2]. By leveraging digital signal processing techniques and spectral analysis, PCG allows for precise detection of pathological changes that might be overlooked in conventional auscultation-based diagnostics. When combined with modern heart sound analysis algorithms, phonocardiography can serve as a valuable tool in cardiological diagnostics, both in clinical and ambulatory settings.

The objective of this article is to present a modern method for diagnosing cardiac pathologies based on analysis of phonocardiographic signals. The presented research results aim to develop an effective heart sound classification system that supports the diagnostic process and enables early detection of abnormalities. This study focuses on key aspects of phonocardiographic diagnostics, including methods for acquiring phonocardiographic signals, assessing the quality of recorded signals, extracting and selecting distinctive features, and applying classification techniques for automatic detection of pathologies. The proposed approach enhances the accuracy of cardiac pathology detection, which could play a significant role in the future of telemedicine and remote patient monitoring.

2. Related work

In recent years, the automated detection of cardiac pathologies from PCG signals has gained significant attention. Existing literature broadly divides into classical machine learning approaches and deep learning techniques [3]. Studies focusing on classical methods often emphasize advanced feature extraction. For instance, Khan *et al.* [4] achieved high classification accuracy by extracting time-frequency features from segmented PCG signals and classifying them using ensemble methods. Similarly, Narváez-Gutiérrez *et al.* [5] successfully utilized modified empirical wavelet transforms combined with power features.

More recently, the focus has aggressively shifted toward deep neural networks and advanced foundation models. Al-Shannaq [6] proposed a real-time abnormal heart sound recognition system combining *support vector machine* (SVM) and *long short-term memory* (LSTM) models. Other recent works have demonstrated the high efficacy of LSTM networks relying on advanced feature extraction [7]. Latifi *et al.* [8] demonstrated the efficacy of multi-branch *convolutional networks* (CNN) and *long short-term memory – convolutional networks* (LSTM-CNN) architectures. Furthermore, the latest advancements in 2025 explore fusing time and frequency domain features using 1D CNNs and Transformer blocks to capture instantaneous dynamics and long-range dependencies [9], as well as adapting state-of-the-art automatic speech recognition systems, such as the Whisper model, to push the boundaries of PCG signal classification [10].

While modern deep learning models offer excellent accuracy, they often operate as “black boxes”, lacking clinical interpretability, and require substantial computational resources. The contribution of our work lies in bridging this gap. We propose a highly interpretable, lightweight diagnostic pipeline based on the extraction of dedicated spectral features from the S1 tone. By employing *Genetic Algorithms* (GA) to strictly select only the most distinctive features, we construct a linear SVM classifier that is computationally efficient, resilient to overfitting, and achieves highly competitive performance, making it uniquely suitable for future implementation in mobile and low-power ambulatory devices.

3. Phonocardiography – significance in diagnostics

The mechanical activity of the heart generates sounds that can be analysed in phonocardiographic diagnostics [11]. The primary heart sounds, designated as S1 and S2, are produced by the closure of the atrioventricular and semilunar valves. Under physiological conditions, additional heart sounds, S3 and S4, may also be present; however, their occurrence does not necessarily indicate pathology. Furthermore, heart murmurs may arise in cases of cardiac abnormalities due to disturbances in blood flow through the heart.

Understanding the mechanics of heart function and the characteristics of the sounds it generates is essential for diagnosing cardiovascular diseases. The analysis of phonocardiographic signals allows for assessment of cardiac performance parameters and detection of abnormalities that may indicate pathological conditions.

The oldest and still widely used diagnostic method is cardiac auscultation using a conventional stethoscope, invented in 1816 by René Laennec [12]. Traditional stethoscopes consist of a chest piece equipped with a diaphragm or bell, a sound-conducting tube, and a binaural headset with earpieces. During cardiac auscultation, the physician places the stethoscope on specific points of the patient's chest where heart sounds and potential murmurs are most audible.

The attenuation of heart sounds by chest wall tissues presents a significant challenge in classical auscultation. Higher frequencies are attenuated more than lower ones, which may result in critical acoustic components becoming inaudible to the physician [13]. Another limitation is the subjectivity of the examination – its effectiveness depends on the physician's experience and perceptual abilities. Age-related hearing loss can further affect the ability to detect subtle acoustic changes, posing a considerable diagnostic challenge.

The advancement of technology has enabled the development of electronic stethoscopes, which allow for the amplification of acoustic signals as well as their recording and digital analysis. An electronic stethoscope is equipped with a transducer that converts acoustic waves into an electrical signal which is then transformed into a digital format.

PCG extends traditional auscultation by enabling the recording and digital analysis of heart sounds. A phonocardiographic recording allows for a detailed examination of components of the cardiac signal, eliminating errors associated with the physician's perceptual limitations.

The characteristics of heart sounds include frequency components ranging from a few to several hundred hertz. Most of the diagnostically relevant energy is concentrated below 100 Hz, which implies that conventional auscultation may be insufficient for detecting critical pathological features [5, 14]. By utilizing digital signal processing methods, it is possible to filter out noise, enhance relevant components, and classify cardiac pathologies using machine learning algorithms. The application of phonocardiography in cardiological diagnostics improves the accuracy of disease detection, particularly in ambulatory settings and telemedicine.

Compared to traditional stethoscopic auscultation, which is a subjective method reliant on the physician's perception, digital PCG signal analysis enables an objective assessment of heart function and the effective detection of abnormalities. The structure of the diagnostic system based on phonocardiographic analysis is presented in Fig. 1.

The core modules of the system include signal acquisition where heart sounds are recorded using a digital stethoscope, with the quality of equipment and environmental conditions being crucial. Signal quality is assessed using a developed classifier that evaluates the reliability of the recorded phonocardiogram. Distinctive features are extracted from the signal in both the time and frequency domains to provide relevant diagnostic information. Finally, signal classification is performed using machine learning, specifically a SVM classifier, to categorize cases as “healthy” or “pathological”.

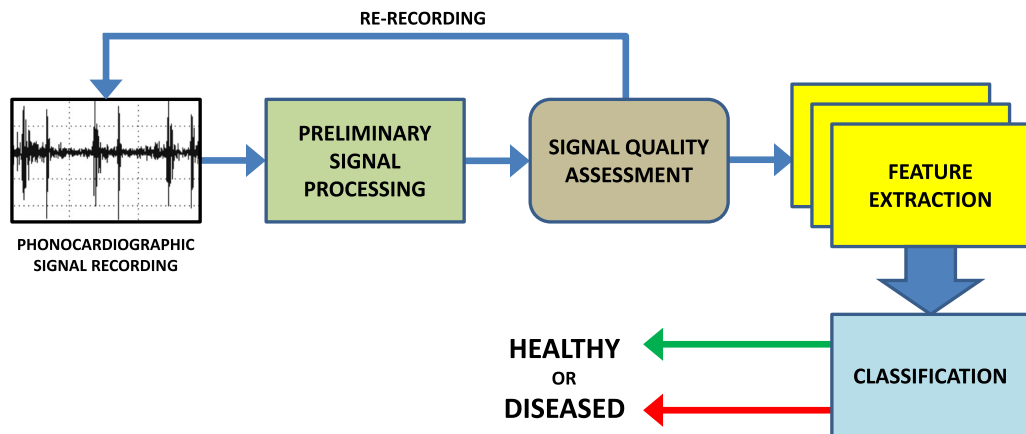


Fig. 1. Structure of the diagnostic system based on phonocardiographic signal analysis.

4. Methods and sources of diagnostic data

Signals were acquired using a *Littmann 3200* electronic stethoscope, leveraging its 20-200 Hz cardiac filtering mode, signal amplification, and active noise reduction capabilities [15].

During PCG signal acquisition, it was essential to maintain appropriate recording conditions to minimize external noise interference. Therefore, the recordings were conducted in low-noise environments, and patients were advised to limit body movements during the recording process. Each recording included several cardiac cycles, allowing for the analysis of both heart sounds and potential murmurs. The duration of a single recording was 15 seconds, providing a sufficient amount of data for further analysis. The signals were recorded using the *StethAssist* application, which is designed by the manufacturer for use with a stethoscope, and subsequently exported in WAV format for processing.

The recorded signals were stored in a specially designed database which includes both healthy and pathological cases, allowing for the verification of PCG signal processing and classification methods. The phonocardiographic recordings were collected from two institutions: the Military Institute of Medicine in Warsaw and the Military University of Technology in Warsaw, ensuring access to data from diverse patient groups. The recordings included both healthy individuals and patients diagnosed with cardiac conditions. Each recording captured heart sounds from two standard auscultation points on the chest (mitral valve auscultation point and Erb's point) and in two patient positions (standing and supine position).

The database structure also included the classification of recordings into primary categories: recordings from healthy individuals *zero-pathology subjects* (ZS) – comprising patients with no detected cardiac pathologies, serving as a reference for comparative analysis and recordings from patients with cardiac pathologies – This group included individuals diagnosed with the following conditions: *coronary heart disease* (CHD), *heart failure* (HF), previous *myocardial infarction* (MI), *atrial fibrillation* (AF), *hypertension* (HT).

All signals were recorded in WAV format with a sampling frequency of 4000 Hz and a 16-bit resolution. Each recording was accompanied by metadata containing: patient demographic data (age, gender), health status and any relevant medical descriptions, recording location on the patient's body and patient position, recording date and the serial number of the stethoscope.

A total of 921 signals were recorded from 231 individuals, including 119 men (52%) and 112 women (48%). The largest group, 41% of patients (94 cases), consisted of healthy individuals. The most common pathological condition was hypertension (21%), followed by patients with a history of myocardial infarction (12%).

All signals were stored in a relational database, ensuring structured organization, easy access, and professional medical annotation. In the next stage, the collected data were used for an initial quality assessment and subsequently analysed for the presence of artifacts associated with cardiovascular diseases.

5. Quality assessment of phonocardiographic signals

In the initial stage of analysis, a quality classifier was implemented to automatically assess the signal's usability and, if necessary, notify the operator to repeat the recording due to the presence of noise or distortions.

Figure 2 presents a relatively high-quality signal where clear periodicity can be observed and the S1 and S2 heart sounds are easily distinguishable. The noise level is minimal, allowing for an accurate determination of heart rhythm. When listening to this signal, heartbeats are clearly audible, and no external interference is present.

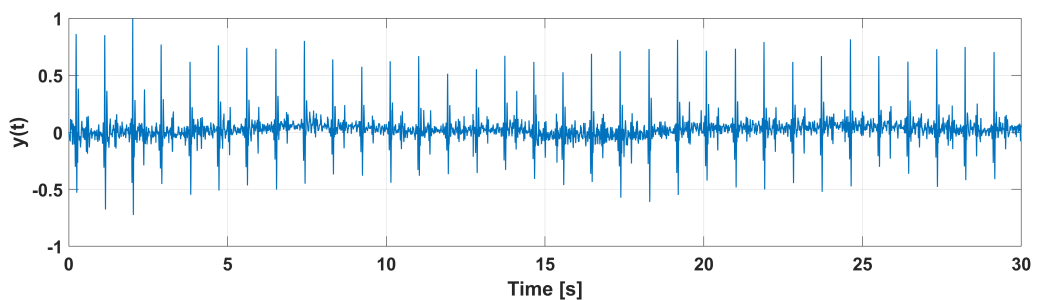


Fig. 2. Example of a relatively high-quality signal.

In contrast, Fig. 3 illustrates a severely corrupted signal, where the periodic structure of heart sounds (visible in Fig. 2) is completely obscured. The signal contamination is characterized by two distinct types of artifacts. First, a high-density, broadband continuous noise is present throughout the entire recording. Second, there are prominent, irregular, non-stationary high-amplitude spikes, most notably at roughly 26s. These characteristics are indicative of severe environmental noise combined with mechanical artifacts, likely from stethoscope head friction or patient movement. Such excessive noise levels render frequency-domain analysis and segmentation ineffective, even before wavelet filtering. The primary purpose of the developed quality classifier is to automatically identify and reject signals resembling Fig. 3, accepting only those resembling Fig. 2, to ensure reliable diagnostic analysis.

In the first stage, the PCG signal undergoes DC component removal and normalization. Next, the signal is filtered using a *discrete wavelet transform* (DWT) [16]. The wavelet decomposition is performed using a Daubechies wavelet of order 10, and the filtered signal is obtained by summing the detail coefficients from levels 5 to 8. This filtering process effectively removes non-stationary low-frequency components, such as those caused by chest movement, as well as high-frequency components beyond the acoustic range of heart sounds [16].

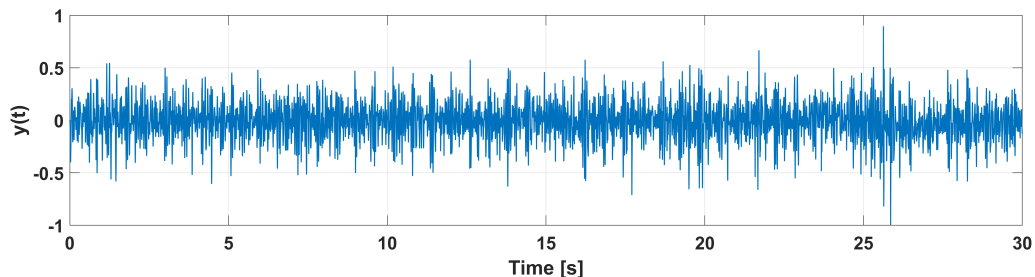


Fig. 3. Example of a relatively low-quality signal.

The parameters used for PCG signal quality assessment are determined based on the wavelet-filtered signal and include: $p1$ – the ratio of the energy of the wavelet-filtered signal to the energy of the original signal; $p2$ – the ratio of the amplitude of the main peak in the autocorrelation function to the amplitude of the subsequent peak; $p3$ – the ratio of the number of detected S1 heart sounds using a threshold method to the number of estimated S1 sounds based on the autocorrelation function; $p4$ – the ratio of the number of detected S1 heart sounds to the number of S1 sounds estimated using the average interval duration; $p5$ – the standard deviation of the intervals relative to their mean value; $p6$ – the *root mean square of successive differences* (RMSSD) between consecutive intervals, normalized to their mean value; $p7$ – the ratio of the number of 2.2-second windows containing 2 to 4 heartbeats to the total number of windows [17]; $p8$ – the ratio of the total duration in which the signal exceeds 40% of its amplitude to the total signal duration; and $p9$ – the ratio of the total duration in which the absolute signal value exceeds the 0.85 quantile to the total signal duration. All nine parameters were determined for a reference group of 52 cases, consisting of 27 signals characterized by relatively low quality and 25 signals of sufficient quality for further analysis. The quality of these signals was subjectively assessed in collaboration with cardiologists involved in the recording process.

Before proceeding with further analysis and after preliminary pilot trials, it was decided to standardize all collected parameters. To evaluate the discriminatory power of individual parameters, Fisher's significance coefficient (γ) [18] was used. This coefficient is defined as the distance between the mean values of a given parameter for both classes, normalized by twice the sum of the corresponding standard deviations.

The larger the value of the γ coefficient, the greater the discriminatory capability of the analysed parameter. The γ coefficient is defined by:

$$\gamma = \frac{|\mu_p - \mu_f|}{2(\sigma_p + \sigma_f)}. \quad (1)$$

The parameter μ represents the mean value of a given parameter for high-quality signals (μ_p) and low-quality signals (μ_f). Meanwhile, σ denotes the standard deviation of the parameter for both high- and low-quality signals.

The values of Fisher's coefficients for all nine quality parameters determined in the reference signal group indicate that the parameters $p5$, $p6$, $p7$, and $p9$ exhibit the highest discriminatory power.

The next step involves selecting a classification method to construct a hyperplane that separates the two data sets. The SVM classifier [4] was chosen, as *linear discriminant analysis* (LDA) produced unsatisfactory results.

From the 921 recorded signals, the reference group of 52 signals used for classifier training was excluded, leaving a total of 869 remaining signals. Nine quality parameters were extracted from each signal, and the trained SVM classifier was then applied to determine the quality coefficient for each case. The “*Quality Coefficient*” is defined as the decision value of the SVM classifier, representing the signed distance of the evaluated sample from the separating hyperplane.

Figure 4 presents the classification results for the remaining database of 869 recorded signals. Blue markers represent signals with high recording quality, while red markers indicate low-quality signals. The horizontal line at level 0 defines the separation threshold. The further a signal is from the 0 point on the vertical axis, the better or worse its quality. It was established that signals within the range $(-0.2, 0.2)$ are of moderate quality. For further analyses, only high-quality signals with a value above 0.2 were considered.

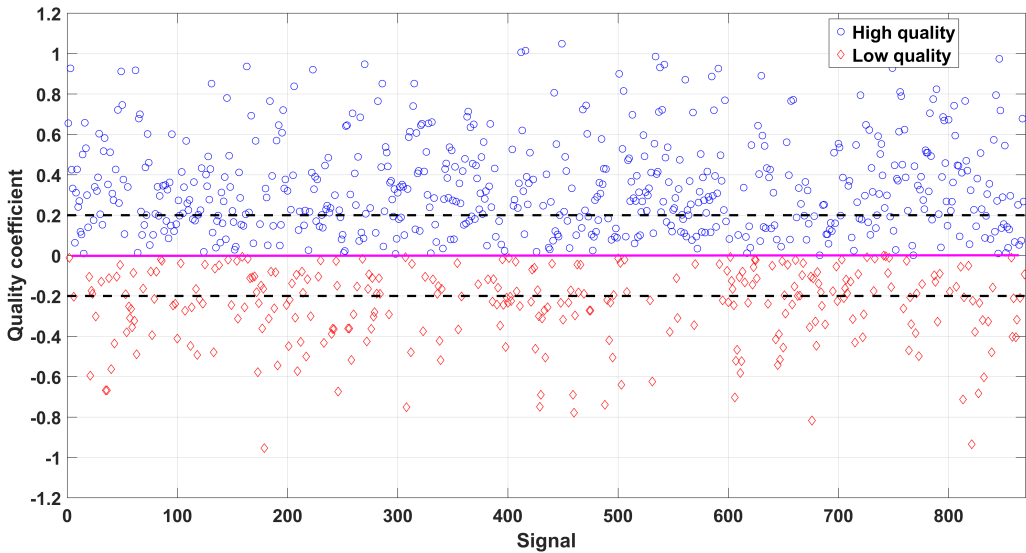


Fig. 4. Quality assessment of the remaining recorded signals.

The quality analysis presented in this section is a crucial step in the process of extracting features indicative of cardiovascular diseases. It ensures that the recorded signal does not contain significant components unrelated to heart activity.

6. Extraction and selection of diagnostic features

To assess the applicability of spectral analysis, a time-frequency analysis of the collected data was performed. The computed spectrogram illustrates the time-frequency characteristics of the analysed signals, providing preliminary confirmation of the suitability of spectral analysis for this type of signal. Figure 5 presents spectrograms of two selected signals, one from a healthy patient and another from a patient with a cardiac pathology [19]. By analysing the spectrograms in Fig. 5, significant differences between them can be observed. The spectrogram of the pathological case exhibits greater variations in the amplitudes of individual heart tones. Additionally, both spectrograms reveal the effect of a low-pass filter with a cutoff frequency of 1 kHz, implemented in the *Littmann 3200* stethoscope.

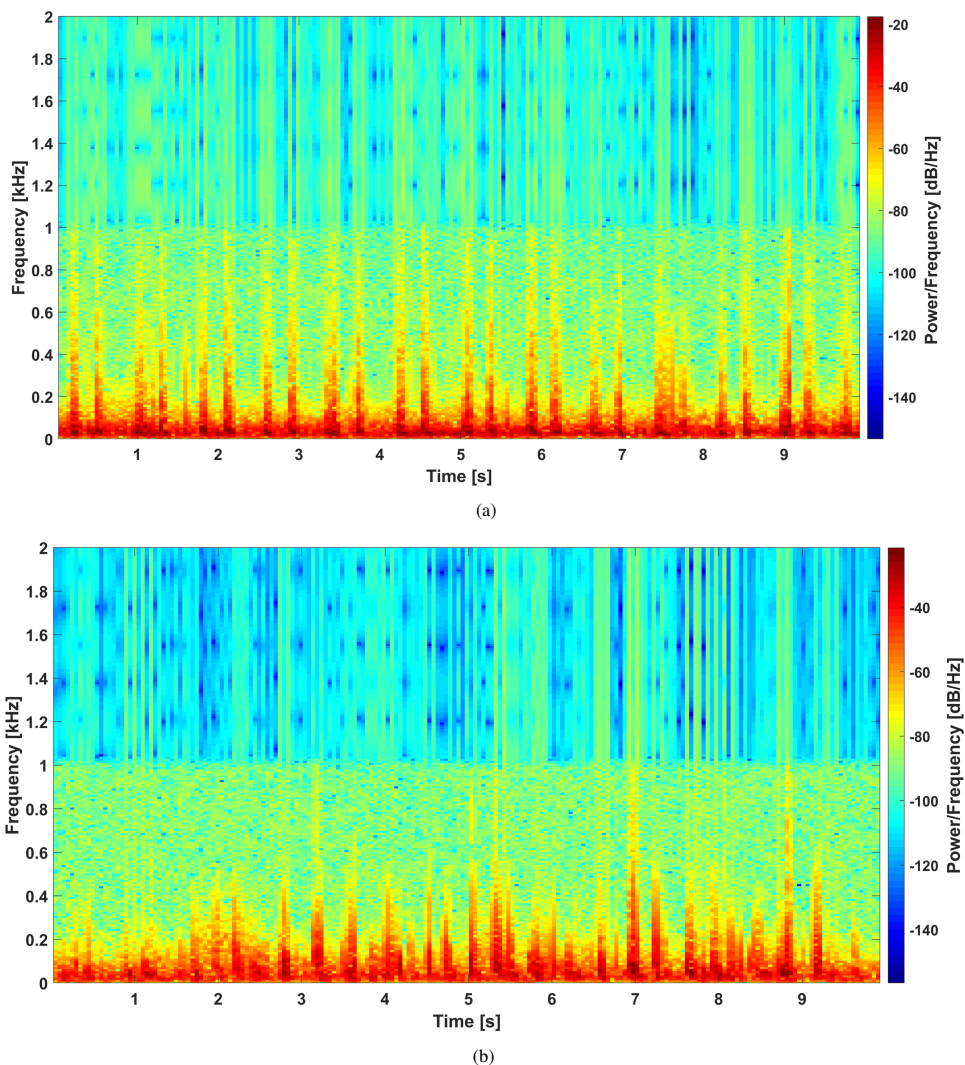


Fig. 5. Example spectrogram of: a) a healthy patient and b) a patient with cardiac pathology.

The recorded PCG signal is a time-domain signal containing several to a dozen heartbeats, consisting of pairs of S1 (systolic tones) and S2 (diastolic tones). Murmurs, which are characteristic of cardiovascular diseases, are present within each repeating tone. PCG signals are often contaminated by internal body noise, so an innovative spectral averaging method was applied, involving several to a dozen consecutive heartbeats. As a result, averaged spectra of S1 and S2 tones were obtained [19]. Figure 6 illustrates the process of averaging individual tones. In the first stage, S1 tones were extracted from each recorded phonocardiographic signal and subsequently subjected to windowing. The application of a *Hamming* window helped minimize spectral leakage.

After extracting and appropriately processing the heart tones, the spectrum of each tone was computed. For this purpose, the *Discrete Fourier Transform* (DFT) algorithm was applied. Subsequently, the power spectrum was calculated.

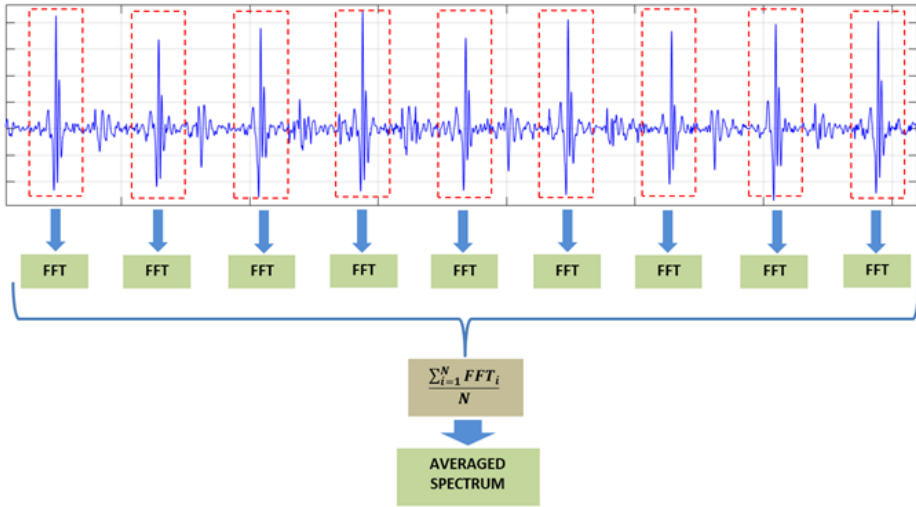


Fig. 6. Innovative method for tone averaging.

The final step involved averaging the obtained power spectra of individual tones and applying normalization. Figure 7 presents the averaged spectra of two selected representative cases in a linear scale: a healthy individual (blue line) and a patient diagnosed with coronary heart disease representing the pathological group (red line) [20]. After medical consultations, the authors ultimately decided to focus on diagnostic feature extraction exclusively from the systolic tone (S1). The S1 tone has higher energy compared to S2, making it easier to isolate from the overall phonocardiographic signal, which is often contaminated by noise [4, 19].

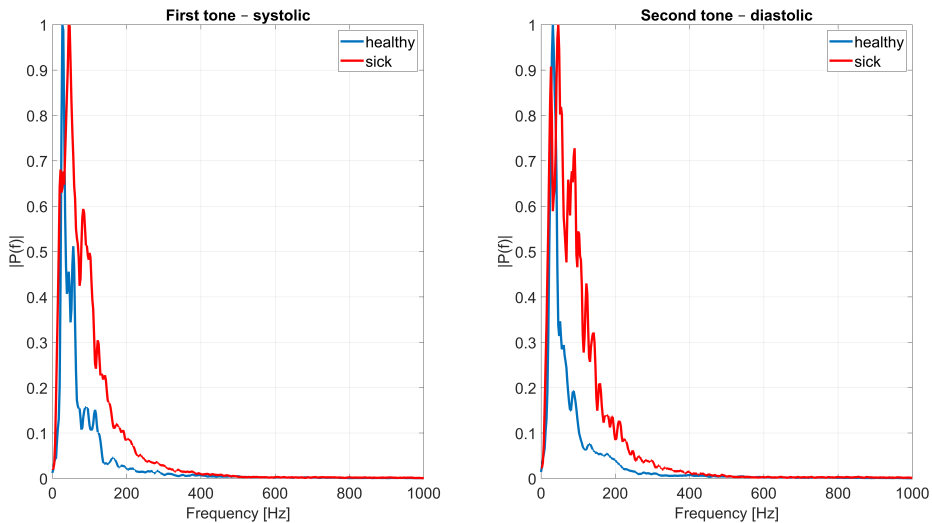


Fig. 7. Averaged amplitude spectra of the first and second tone for a representative healthy individual (blue) and a patient from the pathological group diagnosed with coronary heart disease (red).

To classify the obtained spectra into physiological or pathological cases, 33 spectral features were extracted, including the following:

- c_1 and c_2 – maximum power value and its corresponding frequency;
- c_3 – mean spectral power;
- c_4 – spectral centroid [21], defined by:

$$c_4 = f_c = \sum_{n=1}^N n P'_n \Delta f, \quad (2)$$

where $P'_n = \frac{P_n}{P_c}$ represents the normalized power of the n -th frequency component.

- c_5 – spectral dispersion, defined by:

$$c_5 = f_d = n_d \Delta f, \quad (3)$$

where $n_d^2 = \sum_{n=1}^N (n - n_c)^2 P'_n$;

- c_6 – amplitude spectrum rise rate;
- c_7 – spectral width at 70% of the maximum power;
- c_8 – spectral width at 33% of the maximum power;
- $c_9 - c_{21}$ – maximum power values in defined frequency subbands;
- c_{22} and c_{33} – amplitude and frequency of the first local minimum;
- c_{24} – spectral width at the amplitude of the first local minimum;
- c_{25} – spectral decay rate from the maximum to the first local minimum;
- c_{26} and c_{27} – amplitude and frequency of the first spectral sideband after the first local minimum;
- c_{28} – number of local extrema in the spectrum;
- c_{29} – number of local extrema to the left of the spectral peak;
- c_{30} – number of local extrema to the right of the spectral peak;
- c_{31} – difference between the amplitude of the second spectral sideband and the maximum spectral amplitude;
- c_{32} – difference between the amplitude of the second spectral sideband and the main spectral inflection point;
- c_{33} – frequency difference between the second spectral sideband amplitude and the main spectral inflection point.

The set of 33 features is relatively large and may contain highly correlated features, which is highly undesirable. Therefore, selecting the most relevant features was necessary. Feature selection can be performed in various ways and is one of the most crucial stages in building a classifier, as it directly affects its final structure. There are numerous selection methods, ranging from fast ranking techniques to time-consuming approaches that involve complex classifiers. The most widely used quality measures and selection methods include Fisher's coefficient, t-statistics, mutual correlation, sequential forward selection, genetic algorithms, and LDA [8]. This paper describes the feature selection process using a genetic algorithm as it yielded the best results in selecting relevant features.

In the implemented algorithm, binary encoding was used, meaning that each gene can be represented only by binary values. A logical one at the i -th gene position indicates that the i -th spectral feature will be utilized in generating a new population or evaluated using the fitness function. With this encoding, each chromosome represents a set of features that can be used within a given population.

After generating the pseudo-random initial population, it is evaluated based on the fitness function. The fitness function serves as a measure of how well a given individual in the population performs. It is crucial because it assesses the adaptation level of individual chromosomes (solutions) within the population, allowing the selection of the most well-adapted individuals for generating offspring. The term fitness function originates directly from genetics [22]. The fitness function significantly influences the performance of the genetic algorithm and its proper definition determines the direction of population evolution. In minimization problems, the fitness function (objective function) must be transformed to convert the problem into a maximization task. In this algorithm, the objective function is a weighted average of two coefficients. The first and most decisive coefficient is the classification accuracy, which is determined using LDA applied to the features selected by a given chromosome, as defined by:

$$LDA_{ACC} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (4)$$

where: TP – the number of true positive predictions, TN – the number of true negative predictions, FP – the number of false positive predictions, FN – the number of false negative predictions.

The accuracy of LDA_{ACC} can range from 0 to 1, where 1 represents 100% classification accuracy. The second factor is the number of features used in a given chromosome:

$$Feat_{ACC} = \frac{L - \sum_{i=1}^L g_i}{L}, \quad (5)$$

where: L – chromosome length – the number of genes in the chromosome (the number of features), g_i – the value of the i -th gene in a given chromosome (0 or 1). The accuracy of F will be higher the fewer genes with a value of 1 there are, i.e., the fewer features used for the LDA transformation, the chromosome will contain. The final value of the fitness function is expressed according to the:

$$F(ch_i) = \frac{LDA_{ACC}(ch_i) \cdot w_1 + Feat_{ACC}(ch_i) \cdot w_2}{w_1 + w_2}. \quad (6)$$

In (6) w_1 and w_2 represent the weights of the individual coefficients. The optimization was terminated upon reaching the maximum fitness function value or the predefined maximum number of generations [23, 24]. During this evolutionary process, the algorithm iteratively evaluated the feature set to maximize the classification accuracy – measured via the LDA method – while simultaneously minimizing the total number of features used. By utilizing the genetic algorithm, the original feature set was successfully reduced by half. This significant reduction in dimensionality eliminated components with low diagnostic value while strictly maintaining a high representativeness coefficient for the key spectral properties of the PCG signal, ensuring that essential diagnostic information was preserved. To prevent overfitting and ensure that the classifier is evaluated on data not used during training [3], the dataset must be properly divided. For the classification stage, a pool of 390 high-quality phonocardiographic signals – selected by the quality assessment classifier – was available. These recordings represent two different postures (sitting and supine) and two measurement points (mitral valve and Erb's point), as summarized in Table 1.

After consultations with medical personnel, and due to the largest number of data available for classification, signals recorded in the sitting position and at the apex of the heart were selected, totalling 103 signals from different individuals. Although the quality classifier yielded a total of 390 high-quality signals, it was strictly decided not to mix different measurement points and patient positions for creating a single classifier. Such mixing could lead to severe disturbances in

the distinguishing spectral features. Consequently, the analysis was narrowed down to the largest homogenous group (mitral valve, sitting position; $N = 103$) to preserve the clinical and acoustic integrity of the training and testing sets. The training and test sets were determined according to the 70/30 principle. The training set consisted of 70 signals (35 healthy and 35 sick), while the test set included 33 signals (16 healthy and 17 sick).

Table 1. Quantitative division of signals.

Patient Position	Measurement Point	Quantity	Sick/Healthy
Sitting	Mitral valve	103	52/51
Sitting	Erb's Point	94	39/55
Supine	Mitral valve	101	48/53
Supine	Erb's Point	92	43/49

In this study, a *Linear Support Vector Machine* (LSVM) was implemented. The choice of a linear kernel was motivated by the preceding feature selection stage, which utilized LDA, naturally projecting the data into a linearly separable space. The constant C is a key parameter in SVM that controls the trade-off between maximizing the margin and minimizing classification errors. In practice, it determines how much the model should strive for correct classification of all training examples. A high value of C causes the model to closely fit the training data, which can lead to overfitting and excessive sensitivity to noise. On the other hand, a low value of C increases the tolerance for classification errors, which may improve the model's generalization but simultaneously increase the risk of underfitting. The selection of the optimal value for parameter C is an important part of the optimization process and often requires experimental adjustment to find the best trade-off between accuracy and the model's ability to generalize.

The optimization process of the constant C began with a value of 1, for which the SVM network showed 9 errors in the test group, 9 errors in the training group, and 52 support vectors. During the optimization, the goal was to minimize these parameters. As a result, for the constant $C = 4171$, 4 errors were observed in the test group, 0 errors in the training group, and 13 support vectors.

Figure 8 presents the classification results of the SVM classifier using the optimized value of the regularization parameter C . On the left side, distinctly separated data classes can be observed. The separating hyperplane is located at 0. Only four pathological cases were misclassified as

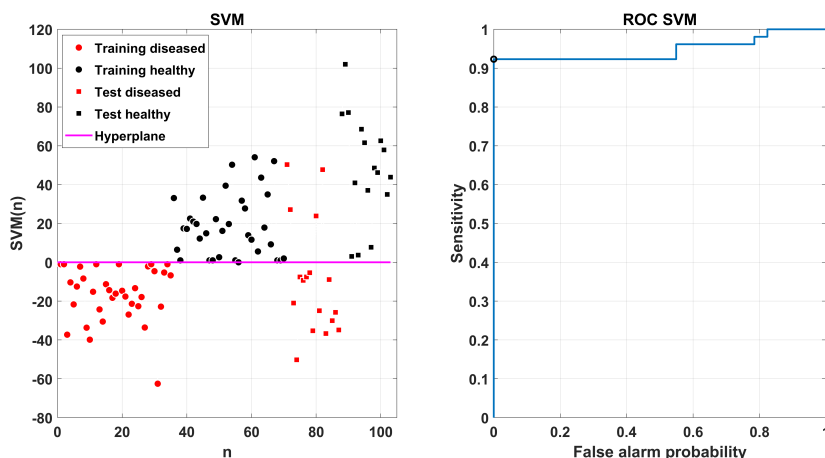


Fig. 8. Classification result using the SVM method.

healthy. The *receiver operating characteristic* (ROC) curve [25] shows the relationship between the true positive rate (sensitivity) and the false positive rate (1– specificity), with pathological recordings treated as the positive class. The *area under the curve* (AUC) was 0.95, indicating very good classifier performance.

Using the SVM network, the maximum classification error in the test set can be estimated according to:

$$P_{\max} = \frac{N_{\text{SVM}}}{M - 1}, \quad (7)$$

where N_{SVM} represents the number of support vectors determined on the full training set.

According to (7), the maximum error is 18.84%. In the test set consisting of 33 cases, 4 errors were obtained, which means the classification error is 12.12%.

While the single 70/30 train-test split presented above effectively illustrates the mechanics of the separating hyperplane (Fig. 9), evaluating machine learning models on relatively small datasets (approximately 100 samples) requires more robust statistical validation to rule out overfitting. Therefore, to definitively confirm the generalization capability of the proposed linear SVM classifier, a stratified 5-fold *cross-validation* (CV) was subsequently performed on the entire subset of 103 signals.

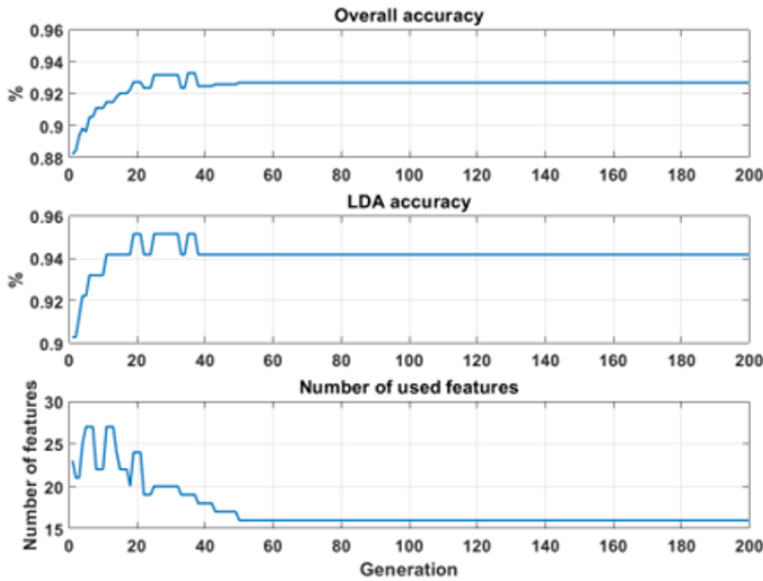


Fig. 9. Classification result using the SVM method.

The average performance metrics across all five folds are presented in Table 2.

Table 2. Average classifier performance parameters (5-fold cross-validation).

Parameter	Value
Sensitivity (Se)	88.36% ($\pm 4.70\%$)
Specificity (Sp)	90.00% ($\pm 7.07\%$)
Accuracy (ACC)	89.29% ($\pm 2.40\%$)
AUC	0.9204 (± 0.0481)

The model achieved a highly stable average accuracy of 89.29% ($\pm 2.40\%$), with a well-balanced average sensitivity of 88.36% and specificity of 90.00%. Most notably, the average AUC reached 0.9204. The extremely low variance across the folds firmly confirms that the selected spectral features provide outstanding and reliable discriminatory power for distinguishing pathological heart conditions from healthy cases, validating the overall effectiveness of the diagnostic system independently of single data partition.

To provide a more detailed and transparent view of the classification distribution, Table 3 presents the aggregated confusion matrix derived from the 5-fold cross-validation across all 103 evaluated cases (52 pathological and 51 healthy signals).

Table 3. Aggregated confusion matrix for the 5-fold cross-validation.

	Predicted: Pathological	Predicted: Healthy
Actual: Pathological	46 (True Positives)	6 (False Negatives)
Actual: Healthy	5 (False Positives)	46 (True Negatives)

As shown in the matrix, the model exhibits a highly balanced error distribution, with only 6 pathological cases misclassified as healthy (false negatives) and 5 healthy cases misclassified as pathological (false positives) out of the entire dataset. This balance further demonstrates the reliability of the system in avoiding significant bias toward either class.

Achieving a 5-fold cross-validated average accuracy of 89.29% and an AUC of 0.9204 places our system competitively within the current state-of-the-art. While modern deep learning algorithms, such as CNN and LSTM architectures, often report high classification accuracies, they typically require massive datasets, lack clinical interpretability, and demand high computational power. Our classical approach using a carefully optimized LSVM achieves comparable reliability (AUC > 0.90) on a smaller dataset while maintaining full feature interpretability, which is a critical factor in medical diagnostics.

However, several limitations of the proposed method must be acknowledged. First, the rigorous quality assessment classifier rejected a significant portion of the initial recordings, resulting in a final evaluated dataset of 103 individuals. While this ensures the model is trained only on reliable data, it restricts the overall sample size. Second, to prevent overfitting on the available data, the current system is designed solely as a binary classifier (healthy vs. diseased) rather than a multi-class system capable of differentiating specific pathologies. Finally, the current diagnostic feature extraction is limited exclusively to the systolic S1 tone.

7. Conclusions

The authors' goal was to design a system for detecting cardiovascular diseases based on a phonocardiographic signal lasting several seconds. As a result of the conducted research, it can be stated that the developed quality assessment system and the final SVM classifier applied in the algorithm for detecting cardiac diseases achieve satisfactory results. The obtained classification error in the test group is smaller than the estimated maximum error derived from the SVM formulation based on the number of support vectors, which supports the correctness of the classifier training. The system was optimized through prolonged optimization using a genetic algorithm. The phonocardiographic recordings were collected by qualified medical personnel from patients with established clinical diagnoses and were subsequently verified by a physician who is a co-author, ensuring the correctness of the assignment of the cases to the disease class and

confirming the effectiveness of the proposed approach. For a controlled subset of 103 recordings acquired in the sitting position at the mitral valve auscultation point, the accuracy of the developed method using 5-fold cross-validation is 89.29%, with a robust average AUC of 0.92. This confirms the stability, generalizability, and high diagnostic potential of the proposed approach. The obtained results form a basis for implementation work, including the development of desktop software for screening tests tailored to the needs of medical personnel. The next step will be to collect a database of phonocardiographic signals recorded with a mobile phone microphone (possibly using a simple acoustic waveguide) and verify whether disease detection can still occur correctly within such an acoustic pathway. Building upon the successful validation of this binary classification model, our future research will focus on two key expansions. First, we plan to incorporate the analysis of the diastolic S2 tone, which, as demonstrated, can be effectively isolated and may provide complementary diagnostic features. Second, through ongoing clinical data collection, we aim to expand the dataset sufficiently to transition from binary classification to robust multi-class algorithms capable of distinguishing specific cardiac conditions, such as hypertension, atrial fibrillation, and heart failure. A positive evaluation of this type of study would allow for the implementation of the system in a mobile application, providing easy access to screening tests for every potential patient.

References

- [1] The top 10 causes of death. (2024). In *World Health Organization*. Retrieved December 19, 2025, from <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- [2] Bialek, W. S. (2012). *Biophysics: Searching for principles*. Princeton, NJ: Princeton Univ. Press, ISBN 978-0-691-13891-6 https://www.researchgate.net/publication/265275031_Biophysics_Searching_for_Principles
- [3] Partovi, E., Babic, A., & Gharehbaghi, A. (2024). A review on deep learning methods for heart sound signal analysis. *Frontiers in Artificial Intelligence*, 7, 1434022. <https://doi.org/10.3389/frai.2024.1434022>
- [4] Khan, F. A., Abid, A., & Khan, M. S. (2020). Automatic heart sound classification from segmented/unsegmented phonocardiogram signals using time and frequency features. *Physiological Measurement*, 41(5), 055006. <https://doi.org/10.1088/1361-6579/ab8770>
- [5] Narváez, P., Gutierrez, S., & Percybrooks, W. S. (2020). Automatic segmentation and classification of heart sounds using modified empirical wavelet transform and power features. *Applied Sciences*, 10(14), 4791. <https://doi.org/10.3390/app10144791>
- [6] Al-Shannaq, M. A., Nasrawi, A., Bsoul, A. a. K., & Saifan, A. A. (2025). Abnormal heart sound recognition using SVM and LSTM models in real-time mode. *Scientific Reports*, 15(1), 9129. <https://doi.org/10.1038/s41598-025-89647-0>
- [7] Sundar, G. N., Subramanian, S., Narmadha, D., Bruntha, P. M., S, I. T. J., & S, S. (2024). Improved Heart Sound Classification Using LSTM Based Deep Learning Technique. In *2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 557–561). <https://doi.org/10.1109/icicv62344.2024.00094>
- [8] Latifi, S. A., Ghassemian, H., & Imani, M. (2025). Multi-branch convolutional network and LSTM-CNN for heart sound classification. *Physical and Engineering Sciences in Medicine*, 49(1), 263–287. <https://doi.org/10.1007/s13246-025-01664-5>

- [9] Liu, N., Chen, X., Yu, Y., Guo, L., Guo, T., Qiu, B., Chen, P., & Wang, Y. (2025). A deep learning model for heart sound classification fusing time-frequency features. *IEEE Transactions on Biomedical Engineering*, PP, 1–11. <https://doi.org/10.1109/tbme.2025.3642718>
- [10] Alotaibi, M., Bazi, Y., Rahhal, M. M. A., Ammour, N., & Zuair, M. (2025). Classification of heart sound signals with Whisper model. *Connection Science*, 37(1). <https://doi.org/10.1080/09540091.2025.2449943>
- [11] McArdle, W. D., Katch, F. I., & Katch, Victor. L. (2014). *Exercise Physiology: Nutrition, Energy, and Human Performance* (8th ed.). Lippincott Williams & Wilkins. ISBN 978-1-4511-9370-2
- [12] Roguin, A. (2006). Rene Theophile Hyacinthe Laennec (1781-1826): The Man Behind the Stethoscope. *Clinical Medicine & Research*, 4(3), 230–235. <https://doi.org/10.3121/cmr.4.3.230>
- [13] Martinek, R., Ladrova, M., Sidikova, M., Jaros, R., Behbehani, K., Kahankova, R., & Kawala-Sterniuk, A. (2021b). Advanced Bioelectrical Signal Processing Methods: Past, Present and Future Approach – Part I: Cardiac Signals. *Sensors*, 21(15), 5186. <https://doi.org/10.3390/s21155186>
- [14] Gradolewski, D., Magenes, G., Johansson, S., & Kulesza, W. J. (2019). A wavelet transform-based neural network denoising algorithm for mobile phonocardiography. *Sensors*, 19(4), 957. <https://doi.org/10.3390/s19040957>
- [15] Leng, S., Tan, R. S., Chai, K. T. C., Wang, C., Ghista, D., & Zhong, L. (2025). Correction: The electronic stethoscope. *BioMedical Engineering OnLine*, 24(1), 117. <https://doi.org/10.1186/s12938-025-01457-7>
- [16] Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6), 674–693. <https://doi.org/10.1109/34.192463>
- [17] De Fazio, R., Spongano, L., De Vittorio, M., Patrono, L., & Visconti, P. (2024). Machine learning algorithms for processing and classifying unsegmented phonocardiographic signals: An efficient edge computing solution suitable for wearable devices. *Sensors*, 24(12), 3853. <https://doi.org/10.3390/s24123853>
- [18] Rajarathinam, A. (2024). Discriminant analysis for heart disease attributes. *International Journal of Statistics and Applied Mathematics*, 9(2), 180–188. <https://doi.org/10.22271/math.2024.v9.i2c.1703>
- [19] Lejkowski, W., Dobrowolski, A. P., Majka, K., & Olszewski, R., (2018). Spectral analysis of major heart tones. In *2017 Radioelectronic Systems Conference* (Vol. 93, p. 9). SPIE. <https://doi.org/10.1117/12.2316785>
- [20] W. Lejkowski, A. P. Dobrowolski, and R. Olszewski, “Automatic analysis of phonocardiogram evaluation in elderly patients with cardiovascular diseases,” presented at EuroPrevent 2019, Lisbon, Portugal
- [21] Ghaheri, A., Shoar, S., Naderan, M., & Hoseini, S. S. (2015). The applications of genetic algorithms in medicine. *Oman Medical Journal*, 30(6), 406–416. <https://doi.org/10.5001/omj.2015.82>
- [22] Taha, Z. Y., Abdullah, A. A., & Rashid, T. A. (2025). Optimizing feature selection with genetic algorithms: a review of methods and applications. *Knowledge and Information Systems*, 67(11), 9739–9778. <https://doi.org/10.1007/s10115-025-02515-1>
- [23] Feng, G. (2024). Feature selection algorithm based on optimized genetic algorithm and the application in high-dimensional data processing. *PLoS ONE*, 19(5), e0303088. <https://doi.org/10.1371/journal.pone.0303088>
- [24] Lejkowski, W. (2017). Multifaceted Spectral Analysis of Phonocardiographic Signals. *Przegląd Elektrotechniczny*, 1(10), 75–78. <https://doi.org/10.15199/48.2017.10.17>
- [25] Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. *PubMed*, 4(2), 627–635. <https://pubmed.ncbi.nlm.nih.gov/24009950>



Wojciech Lejkowski received his PhD in 2020 from the Military University of Technology specializing in digital signal processing. He is an expert in machine learning and artificial intelligence, with a strong focus on the analysis and modelling of signal data. His research and engineering work bridges classical DSP methods with modern ML/AI approaches, including deep learning, feature extraction, and predictive modelling. His interests include designing algorithms robust to noise and variability in measurement conditions as well as implementing efficient solutions for high-performance computing environments. He is a co-owner of BITRES, a company that develops and deploys digital signal processing and AI/ML solutions for industrial and research applications. In his work, he emphasizes rigorous validation, model interpretability, and the practical applicability of proposed.



Robert Olszewski graduated with distinction and first place from the Military Medical Academy (WAM) in 1988. In 1998, he received the First-Degree Award for the Best Doctoral Dissertation at WAM and a distinction from the Polish Ultrasound Society for his Ph.D. thesis in medical sciences. He obtained his D.Sc. degree in 2013. Dr. Olszewski has completed clinical fellowships in Germany, the United States, the United Kingdom, and Canada. He has participated in scientific training programs in the United Kingdom, Italy, France, and the United States, and conducted research in Sweden and the United Kingdom.



Andrzej Piotr Dobrowolski received the D.Sc. degree in electronics in 2010 from the Military University of Technology, Warsaw, Poland. In 2017, the President of Poland awarded him the title of Professor of Technical Sciences. From 2012 to 2016, he served as Deputy Dean for Scientific Issues at the Faculty of Electronics, Military University of Technology, where he also managed doctoral programs. From 2016 to 2020, he served as Dean of the Faculty of Electronics, and in the following four years, he held the position of Vice Rector for Scientific Issues at the Military University of Technology. He is the author or coauthor of four books and more than 200 journal and conference papers. His research interests include digital signal processing, artificial intelligence, medical diagnosis, and biometrics. He has accumulated over 9,000 hours of teaching experience and has supervised eight Ph.D. dissertations.