A Novel System for Cardiovascular Screening Based on Simultaneous Phonocardiography and Electrocardiography During Auscultation

Sofia Margarida Marques Monteiro
sofia.m.monteiro@tecnico.ulisboa.pt
Instituto Superior Técnico, Lisboa, Portugal

December 2022

Abstract

The combination of the phonocardiogram (PCG) and the electrocardiogram (ECG) allows the evaluation of the electromechanical condition of the heart and could greatly improve the accuracy of an initial cardiovascular disease diagnosis. In this work, the advantages of combining synchronous PCG and ECG were analyzed in depth, resulting in the development of deep learning models for abnormality detection and in a novel prototype of an electronic stethoscope that combines PCG and ECG sensors. The best performing models include a recurrent neural network for murmur detection from heart sounds, which achieved a sensitivity of 82.7% and specificity of 80.1%, and a hybrid neural network based on convolutional and recurrent neural networks for abnormality detection from multi-modal data, which reached a sensitivity of 89.9% and specificity of 78.7%. The developed prototype was evaluated experimentally to assess the sensors, the morphology of the acquired signals, and the effect of the rotation of the stethoscope head, which further extends the state-of-the-art. It was verified that the ECG lead voltages at different angles of rotation follow Kirchhoff's law and, by comparing the ECG waveforms between a reference and the experimental device, the best case reached a Pearson's correlation coefficient of 0.928. The results demonstrate the validity of the device and the potential of combining the two signals, establishing a foundation for further development. The accessibility and low cost of the device, combined with automatic techniques for disease screening, should significantly facilitate its integration in healthcare facilities and telemedicine, as well as in research and education.

Keywords: Electrocardiogram, Phonocardiogram, Auscultation, Electronic Stethoscope, Multimodal Signals

1. Introduction

Cardiovascular diseases (CVDs) are the leading cause of death and morbidity worldwide, lowering the quality of life of the patients and leading to incremental long-term healthcare costs. An early diagnosis of both congenital and acquired heart conditions, together with regular heart monitoring, is essential to reduce the burden of these diseases and prevent premature deaths [1]. Cardiac auscultation remains the most common primary screening method [2], since it is very simple, quick, and inexpensive, and allows doctors to diagnose several cardiac conditions related with the mechanical condition of the heart.

The digital recording of heart sounds is named phonocardiogram (PCG), and it can be obtained with a chest-placed microphone and used for further analysis and processing [3]. The electrocardiogram (ECG), on the other hand, records the electrical activity of the heart and is also one of the most common primary methods of disease screening. It is used to evaluate the electrical condition of the heart and is a very powerful, non-invasive, inexpensive, and easy to use technique. The ECG and the PCG are concurrent phenomena, seeing that the PCG results from the mechanical operation of the heart, which in turn relies on its electrical activation [4]. Thus, their simultaneous acquisition and analysis comes as a natural step. If the two signals can be recorded simultaneously during a routine auscultation exam, it is possible to inspect the electromechanical condition of the heart and obtain a faster diagnosis in a straightforward way. Additionally, there are many cases where the PCG and ECG contain mutually exclusive information about the heart's condition, so a combination of the two signals could also greatly improve the accuracy of an initial screening [5].

The simultaneous acquisition of the PCG and ECG with portable cardiac sensing systems is already a reality and could revolutionize how we address public health policies for cardiology. Current devices are mostly based on integrating ECG electrodes in an electronic stethoscope so the two signals can be acquired at the same time [6]. However, most still present some limitations, which have prevented their widespread implementation in clinical practice, mainly related with inadequate form factors, high costs, or proprietary software and hardware. Advanced signal processing and machine learning techniques for automatic disease detection have also been successfully implemented for both signals, but most research still focuses on exploiting the PCG and the ECG separately [5].

The goal of this work was to perform an in-depth research on the advantages of combining synchronous PCG and ECG for the automatic detection of heart disease during auscultation. Through the integration of deep learning techniques with a hardware architecture for signal acquisition, we hope to contribute to the development of a robust system for the pre-screening of cardiac diseases,

that can improve the accuracy of an initial diagnosis and thus better inform and support the decisions of medical professionals in clinical practice. It includes the design and implementation of deep learning based algorithms for classification of biosignals acquired during auscultation, and the development of a novel open-source and low-cost stethoscope that combines ECG and PCG sensors, overcoming the limitations of current devices.

2. Heart Disease Detection with ECG and PCG

Given the success of deep learning approaches for PCG and ECG classification, the first part of this work focuses on the application of deep neural networks, namely convolutional neural networks (CNNs), recurrent neural networks (RNNs) and hybrid models, to two different tasks: detection of murmurs and clinical outcome from PCG signals (Section 2.1), and abnormality classification based on simultaneous PCG and ECG data (Section 2.2).

2.1. Heart Murmur Detection

The goal of the George B. Moody Physionet Challenge 2022 was to develop an automatic method that could identify the Present, Absent or Unknown cases of murmurs and the Normal or Abnormal clinical outcome from heart sound recordings collected from multiple auscultation locations using a digital stethoscope. It should be noted that the labels correspond to the patient, and not to the recordings, so it is possible that a patient in the Present class has recordings where no murmurs have been detected.

The challenge used the CirCor DigiScope dataset, which consists of 5272 PCG recordings from 1568 patients. This dataset was divided into training, validation, and test sets, with 60%, 10%, and 30% of the data, respectively. The training set was publicly released, while the validation and test sets were hidden and used to evaluate the entries of the challenge [7]. This is the largest publicly available heart sound dataset and, although it doesn't include recordings of simultaneous ECGs, it could help significantly enhance the characterization and classification of the heart sounds, and consequently improve automatic CVD diagnosis based on auscultation.

To evaluate the performance of the algorithms, two metrics were introduced. In the murmur detection task, the evaluation was done with a weighted accuracy metric that assigns more weight to patients that have or potentially have murmurs than to patients that do not have murmurs; in outcome detection, a cost metric that reflects the cost of screening, treatment, and of missed diagnosis was used. For more details we refer the reader to [7].

Preprocessing

Initially, all PCG signals were filtered using a 2nd order Butterworth bandpass filter with cutoff frequencies of 25 and 400 Hz and went through a spike removal process, to remove unwanted noise and friction spikes [8]. Prior to feature extraction, each recording was decomposed into smaller fixed-length segments of 4-seconds. These segments still contain multiple cardiac cycles and enough information for the models to learn, and at the same time are small enough to allow a significant increase in the amount of training data to build a more robust model [9]. To deal with the class imbalance in the murmur detection task, the minority classes (Present and Unknown) were oversampled by extracting their segments with 75% overlap. Recordings from patients in the Present class without an audible murmur were excluded from training.

For each segment, the static, delta, and delta-delta Mel frequency spectral coefficients (MFCCs) were extracted as spectral features, that capture the frequency content of the signals; and the homomorphic, Hilbert, power spectral density, and wavelet envelopes were extracted as temporal features, which capture the variability in morphology. These envelopes have a sampling frequency of 50 Hz and were extracted using the method described in [10].

Neural Network Architectures

Three different neural network architectures were employed for the heart sounds classification: a CNN, which receives the static, delta, and delta-delta MFCCs organized into an array with three channels as input; a RNN consisting of two stacked bidirectional Long Short-Term Memory (BiLSTM) layers with 64 units, that are able to process the long distance dependencies in the data and receive the four envelopes of the PCG signal as input; and a hybrid neural network obtained by combining the CNN and RNN as parallel blocks. With this multi-input hybrid model, the two blocks can complement each other, with the CNN extracting the most relevant spectral features in the signal and the RNN extracting the morphological and temporal information.

In the last layer, the softmax activation function was used for the murmur multiclass classification and the sigmoid activation function for the outcome binary classification. A diagram of the models is presented in Figure 1.

The model was trained using the SGD optimizer and cross-entropy loss, with a learning rate of 10^{-3} and momentum of 0.9. A weighing was applied to the loss function to make the model pay more attention to the underrepresented classes. To avoid overfitting, early stopping was used to monitor the loss of the model on a validation set, obtained by randomly dividing the training data in each fold into 90% for training and 10% for validation.

The evaluation of the performance of the models was done using 5-fold cross validation. In order to obtain the final labels, it was necessary to take into account that these tasks are both cases of multiple-instance classification, given that each patient is represented by a set of instances (i.e. the recordings from the different auscultation locations), but it is only the patient that carries the labels. The models were trained on individual recordings and then the instance-level decisions were combined, assuming that a positive label contains at least one positive instance. The final labels for each patient were therefore generated by selecting the label of the recording with the highest prob-

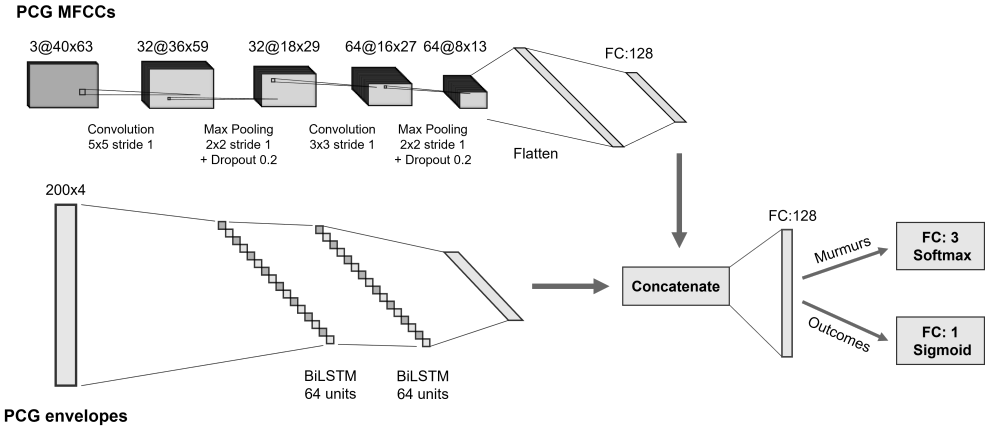


Figure 1: Hybrid CNN and LSTM architecture used for PCG classification.

ability for the positive class.

Results

In Tables 1 and 2 the performance and challenge metrics using the different neural networks are presented for the murmur detection and outcome prediction tasks.

The architecture that reached the best performance in murmur detection was the BiLSTM. It has a high specificity and high sensitivity for the Present class, but a low sensitivity for the Unknown class. A possible explanation for this is the low sample size, since the Unknown class only represents 7.2% of the data, which could make its instances too varied for the model to learn in a way that can be generalized for new examples. Another possibility is the fact that the model could be able to reliably identify the presence or absence of murmurs in these signals, since the Unknown label is only an indicator of the recordings' inferior signal quality [7].

All the tested models had poor results in the outcome prediction task, with a low specificity and sensitivity. This could be because the outcome labels result from an overall assessment of the patient's condition, which is based on multiple examinations (including clinical history, physical examination, and echocardiogram) and not just on auscultation data. It is possible that some of the abnormalities cannot be identified just from the heart sound recordings, meaning that the signals do not contain all the information that is necessary for the neural networks to learn.

Given these results, for scoring and ranking on the challenge the chosen model was the one based on the BiLSTM architecture. The scores and ranks for each task are presented in Table 3. The model showed promising performance in the detection of murmurs, with scores in the official hidden validation and test sets being superior to the scores obtained with cross-validation on the public data.

2.2. Heart Abnormality Detection from the ECG and PCG

The goal of the PhysioNet/CinC challenge 2016 was to develop an algorithm to automatically classify PCG recordings as normal or abnormal. It assembled eight databases from different sources, including the MIT heart sounds database (MITHSDB). The MITHSDB is the only public database with simultaneous PCG and ECG recordings, and it includes 405 samples from 121 subjects, meaning that each patient contributed with more than one recording. The dataset is unbalanced when considering a binary classification: about 32.6% of the data are considered normal, while 71.4% are considered abnormal [11].

Given the low sample size in the MITHSDB, the strategy to develop an algorithm for heart abnormality detection based on the multi-modal signals was to first apply the method described on the previous section to the entire PhysioNet/CinC challenge 2016 database (to determine the best performing neural network for PCG abnormality detection), and then to evaluate whether adding the ECG information improves the classification performance in the MITHSDB.

Since all the recordings in the PhysioNet/CinC challenge 2016 database are treated as independent samples, the final labels were obtained per recording instead of per patient as in the previous task. The results are presented in Table 4, and the top performance metrics are in line with the methods in the state-of-the-art applied to the same dataset. The best performing models are the CNN and hybrid architectures, contrary to the results obtained for murmur detection. It can also be seen that the improvement in performance obtained from adding the parallel LSTM block was not significant, and that the CNN has the best sensitivity. For these reasons, the chosen model for evaluation on the MITHSDB was the CNN.

Multi-modal Signal Classification

RNNs are able to capture the temporal dependencies in the raw ECG signals without any feature extraction steps. The chosen approach was therefore to combine the CNN for the PCG classification with a parallel BiLSTM for ECG classification in a hybrid neural network, in order to verify whether the combination of the two signals enhances the classification performance. Each ECG signal was filtered using a 100th-order band-pass FIR filter with cutoff frequencies of 0.5 Hz and 50 Hz and then normal-

Table 1: Performance metrics from multiclass PCG murmur classification obtained for each neural network architecture. The positive predictive value and the F_1 score refer to the Present class. Best results are in bold.

	Accuracy (%)	Sensitivity (Present) %	Sensitivity (Unknown) %	Specificity (%)	Positive P. Value (%)	F_1 score (%)	Weighted Accuracy (%)
CNN	68.1±6.2	82.1±4.4	14.7±6.7	69.8±8.5	46.0±8.7	59.5±7.6	58.3±3.4
LSTM	**77.1±1.9**	**82.7±3.7**	**31.2±11.9**	**80.1±3.1**	**58.4±4.3**	**68.3±3.0**	**65.25±4.3**
Hybrid	72.2±5.2	81.0±3.1	29.3±9.9	78.1±1.7	57.9±4.3	67.4±2.8	63.45±3.7

Table 2: Performance metrics from PCG clinical outcome classification obtained for each neural network architecture. Best results are in bold.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive P. Value (%)	F_1 score (%)	Cost
CNN	59.9±3.0	**68.2±7.2**	52.1±5.7	57.2±2.7	**62.1±3.8**	**12021±953**
LSTM	**62.6±3.3**	61.4±6.7	**63.8±9.0**	**62.2±4.5**	61.3±2.4	12875±963
Hybrid	60.9±4.7	63.4±5.1	58.7±7.2	59.2±5.1	61.1±4.4	12565±911

Table 3: Challenge metrics for both classification tasks. 5-fold cross validation was used on the public training set, with repeated scoring on the hidden validation set and one-time scoring on the hidden test set.

	Training	Validation	Test	Ranking
Murmurs	0.652±0.043	0.751	0.757	6/40
Outcomes	12875±963	11222	13815	25/39

ized and downsampled to a sampling frequency of 100 Hz.

The used neural network was similar to the one in Figure 1 for the binary classification. It was trained with MFCC features of PCG segments at the input of the CNN block, and with the filtered and normalized raw ECG segments at the input of the BiLSTM block. The length of the segments was equal to 4 seconds, and these were extracted with 75% overlap for the minority class.

Results

In Table 5 the performance metrics for uni- and multimodal signal classification are presented. It is quite clear that the performance using the multi-modal signals is superior to that of the single modalities, and the results are in line with the ones of other deep learning methods in the state-of-the-art [5]. However, unlike the methods applied to uni-modal ECG and PCG classification, the proposed approach and the other deep learning based methods still don't outperform traditional machine learning methods with handcrafted features. This is likely due to the small amount of available data, since deep learning approaches are highly affected by limited datasets. Due to this constraint, the results should be interpreted in a tentative way since, as was mentioned, there is not only a low sample size, but many samples come from the same patient and are treated independently, which could lead to overfitting and an overestimation of the models' performance. Nonetheless, this analysis showed the potential of combining the two signals to detect cardiac abnormalities and provides a good reference for future applications and development.

Larger and more representative datasets are necessary to train and implement deep learning classifiers that can fully take advantage of the complementary information in the two modalities. However, devices that are capable of simultaneous PCG and ECG acquisition are still not widely used due to multiple limiting factors. In order to build more robust classification methods, it is first necessary to develop hardware architectures that can record the synchronous PCG and ECG and overcome the current usability constraints. This motivated the development of a novel hardware prototype based on a low-cost and open-source stethoscope, which is described in Section 3.

3. Novel Electronic Stethoscope for Simultaneous ECG and PCG Acquisition

In this section, a novel electronic stethoscope for simultaneous electrocardiography and phonocardiography during auscultation is proposed.

A core aspect when combining ECG and PCG for disease diagnostic is the temporal synchronization of both signals. Previous work has found restrictions associated with retrofitting external sensors to existing digital stethoscopes [6; 12]. To gain more control over the construction and sensor integration, this work builds upon the Glia stethoscope to devise an end-to-end design of a novel form factor, without significantly changing the standard form factor. The Glia stethoscope[1] is an open-source 3D-printed acoustic stethoscope that is research-validated and operates as well as a Littmann Cardiology III (considered a gold-standard device [13]). It is composed of six different 3D printed parts, the most relevant ones being the stethoscope head and a ring to attach the diaphragm.

In this work, Glia's design was adapted to include phonocardiography and electrocardiography sensing capabilities, while keeping it open-source, with a low-cost, and ensuring that the acoustic properties of the model aren't significantly altered. The form factor of the device was also considered, so it could be used similarly to standard auscultation stethoscopes. As such, the stethoscope head was adapted to include PCG and ECG sensors, with 3D-printed polymer based dry electrodes attached to the ring.

[1] https://github.com/GliaX/Stethoscope

Table 4: Performance metrics from PCG classification obtained for each neural network architecture. Best results are in bold.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive P. Value (%)	F_1 score (%)
CNN	88.5±1.0	**92.8±1.5**	87.3±1.3	65.5±2.3	76.8±1.7
LSTM	79.7±4.3	90.8±3.2	76.8±5.8	50.9±5.0	65.0±4.3
Hybrid	**88.7±0.9**	92.3±2.3	**87.8±0.5**	**66.2±1.5**	**77.1±1.8**

Table 5: Performance metrics from multi-modal signal classification obtained for different signal combinations. Best results are in bold.

	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive P. Value (%)	F_1 score (%)
Only PCG	72.8±2.5	77.4±3.9	61.5±7.5	83.3±2.2	80.2±2.1
Only ECG	84.7±1.5	88.2±2.5	76.2±8.6	90.2±3.3	89.1±0.9
Simultaneous ECG + PCG	**86.7±4.0**	**89.9±3.9**	**78.7±9.3**	**91.3±3.7**	**90.5±2.9**

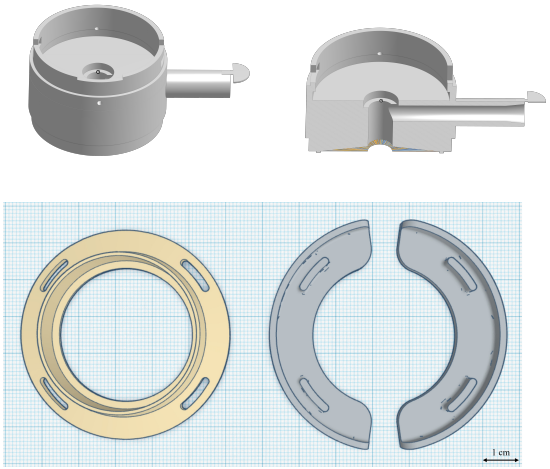


Figure 2: Modified stethoscope head with hole for the electret microphone (top) and modified ring and half-moon ECG electrodes (bottom).

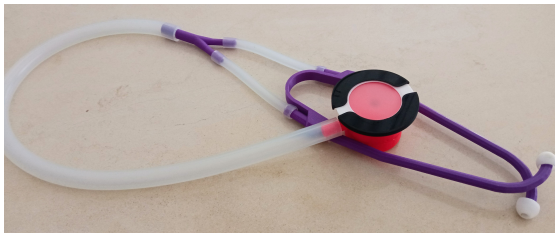


Figure 3: Photo of the developed device.

3.1. Form Factor

The aim was to preserve the standard analog auscultation feature of the stethoscope, to avoid creating barriers in the interaction between the end-user (the physician) and the device [12]. As such, to integrate the PCG sensor, the stethoscope head was modified by adding a small hole in the back (Figure 2), with the same diameter as an electret microphone. With this approach, the sound will still travel through the silicone tubes and ear pieces to be listened to by the doctor during the auscultation procedure, while simultaneously being recorded.

To allow ECG sensing capabilities, two 3D-printed dry electrodes with a half-moon shape were added to the side of the stethoscope head, attached to an altered ring (Figure 2). This design had several advantages, namely: 1) the fact that there was no need to alter the diaphragm of the original stethoscope, which is essential to ensure the quality and intensity of the sound reaching the earpiece and the sensor microphone; 2) a higher inter-electrode distance, which leads to a higher signal magnitude; 3) a higher electrode conductance due to the increased surface area granted by the half-moon shape, which leads to higher SNR and reduced motion artifacts [14]; and 4) better ergonomics, since by fitting the electrodes along the ring the shape of the standard stethoscope isn't being modified.

The modified stethoscope parts (ring and stethoscope head) were printed using polylactic acid (PLA) filaments. For the ECG dry electrodes the electrical conductivity of the material also had to considered, reason for which the Protopasta Conductive PLA (ProtoPlant, Inc) material was chosen, as it combines PLA with carbon black. The electrical interface between the electrodes and the ECG sensor was done by fusing each wire to the material with heat, which has been tested and validated in [15]. The final device is presented in Figure 3.

3.2. Instrumentation

For data acquisition and Bluetooth wireless streaming, a novel acquisition system named ScientISST SENSE[2] was used. It is based on the ESP32 microcontroller architecture and allows multi-channel sampling with high time (up to 16 kHz) and amplitude (12-bit) resolution. It was coupled with two sensors, one for each signal modality:

- A ScientISST ECG sensor [16], which has already been validated in previous studies [17]. It has a gain of 1100 and applies a bandpass filter with cut-off frequencies of 0.5 Hz and 40 Hz. It was used with a virtual ground configuration to measure a single-lead ECG.
- A novel PCG sensor based on an audio amplification circuit that uses an electret microphone and the LM4861 amplifier, which has a gain of 40 and applies a bandpass filter with cut-off frequencies of approximately 21 Hz and 408 Hz (which were chosen based on the frequency content of the heart sounds).

[2] https://www.scientisst.com/sense

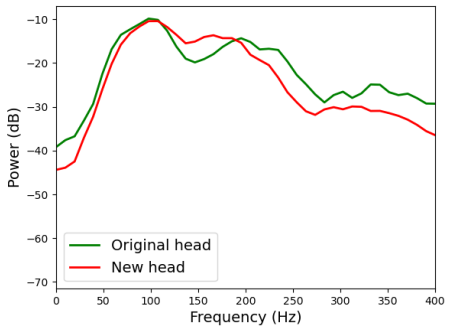


Figure 4: Frequency responses of the Glia stethoscope with the original head and altered heads, for the frequencies of interest in PCG signals.

The SENSE Web App[3] was used to record the data on a base station. Both signal modalities were acquired with a sampling frequency of 2000 Hz.

A total of 19 multi-modal signals were collected from 18 volunteers, 50% of which were female. One person had a previous diagnosis of arrhythmia, and in terms of age distribution 3 participants belonged to the age bracket of 18-24 years, 1 to 25-34 years, 8 to 35-44 years, 1 to 45-54 years, 3 to 55-64 years, and 2 were over 65 years old. Initially the relevant demographic and clinical history data were collected, and after that a simultaneous PCG and ECG were recorded using the proposed device, with the subjects seated and at rest. Written consent was obtained for the participants, and all data was treated anonymously.

3.3. Acoustic Transfer

To verify that the acoustic properties of the stethoscope weren't significantly altered, the frequency response of the proposed device was evaluated with the same experimental setup that was used to compare the frequency response of the Glia stethoscope with the one of the Littmann Cardiology III (described in [13]). A latex balloon was filled with 2 liters of water and used as a phantom, where both the proposed electronic stethoscope with the altered head and the original Glia stethoscope were applied. The phantom excitations were done by an external speaker which was in contact with the balloon, and white noise with frequencies between 0 and 5000 Hz was played for 15 seconds.

The output of the stethoscope was recorded by a microphone placed at the end of the silicone tube that connects the stethoscope head to the earpieces. This experiment was repeated three times, and the average frequency responses are presented in Figure 4. In the low and mid-frequency ranges the devices performed similarly, emphasizing the lower frequency regions between 50 and 200 Hz, with a slight roll-off in the higher frequencies.

3.4. Heart Rate and Auscultation Points

Before further analysis, the acquired signals were filtered using a FIR filter with order equal to half of the sampling frequency, with cutoff frequencies of 0.5 and 40 Hz for the ECG and of 25 and 400 Hz for the PCG.

[3] https://sense.scientisst.com

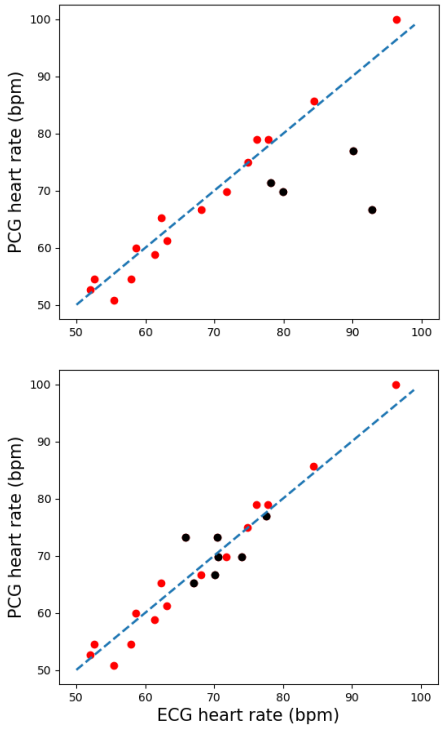


Figure 5: Scatter plots comparing ECG and PCG average heart rates before (top) and after (bottom) artifact removal. The dashed line represents an equal heart rate in both signals. The outlier signals that were corrected are shown in black.

To evaluate the concurrence between the signals, the average heart rates calculated for the PCG and ECG were compared for all subjects. In Figure 5, it can be verified that there is a high agreement between the measurements, although there are a few outliers. They correspond to subjects that had an ECG signal with regions highly corrupted by artifacts, where several R-peaks were erroneously identified. After removing these sections and recalculating the average heart rates for the resulting segments, the ECG-derived heart rates become much closer to the heart rates calculated for the PCG.

As the stethoscope head is placed on different auscultation points, the heart's electrical axis will be viewed from different perspectives and, as such, the morphology of the recorded ECG waves will be different. This is clear in Figure 6, where alterations can also be seen in the PCG signal. S1 and S2 are present in all the signals, as expected, but depending on the area of the chest and its closeness to each heart valve, the stethoscope will amplify the cardiac sounds differently. By rotating the stethoscope head on the same auscultation point the effect on the ECG will be similar, i.e, since the electrode's position is changing, the signal will record the heart's electrical activity from a different perspective. The main fiducial points in the two signals can be clearly identified, and it is possible to verify their alignment: the S1 sound occurs immediately after the R-peak, and the S2 sound immediately after the end of the T-wave.

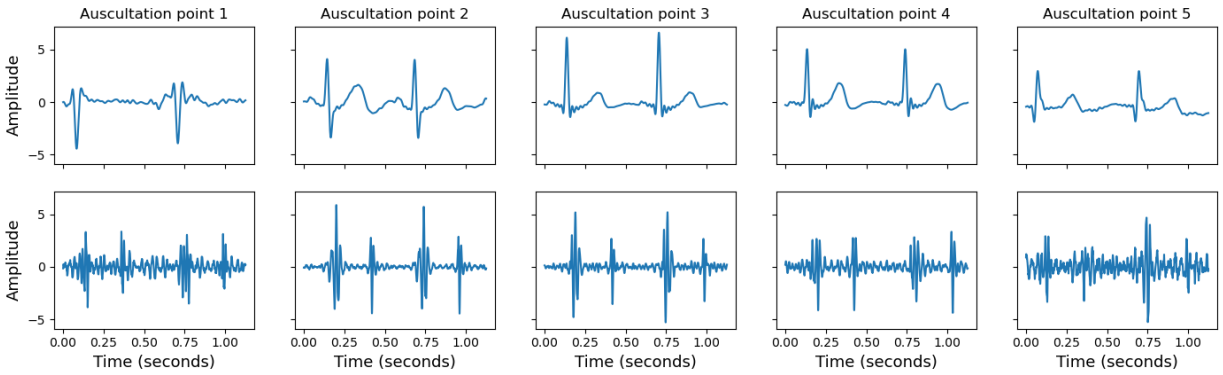


Figure 6: Example of ECG and PCG signals acquired with the stethoscope head at the different auscultation points.

3.5. ECG Leads Characterization

The Einthoven limb leads (Leads I, II, and III) and their corresponding definitions are represented in Figure 7(a). This lead system is based on the assumption that the cardiac sources are represented by a dipole $\bar{p}$ located at the center of the equilateral triangle. Thus, the voltages measured by the three limb leads are proportional to the projections of the electric heart vector on the sides of the lead vector triangle [18]. Considering this, we have:

$$V_I = p\cos(\alpha) = p_y \tag{1}$$

$$V_{II} = \frac{p}{2}\cos(\alpha) - \frac{\sqrt{3}}{2}\sin\alpha = 0.5p_y - 0.87p_z \tag{2}$$

$$V_{III} = -\frac{\sqrt{p}}{2}\cos(\alpha) - \frac{\sqrt{3}}{2}p\sin(\alpha) = -0.5p_y - 0.87p_z \tag{3}$$

and we can verify that Kirchhoff's voltage law is satisfied:

$$V_I + V_{III} = \frac{p}{2}\cos(\alpha) - \frac{\sqrt{3}}{2}\sin\alpha = V_{II} \tag{4}$$

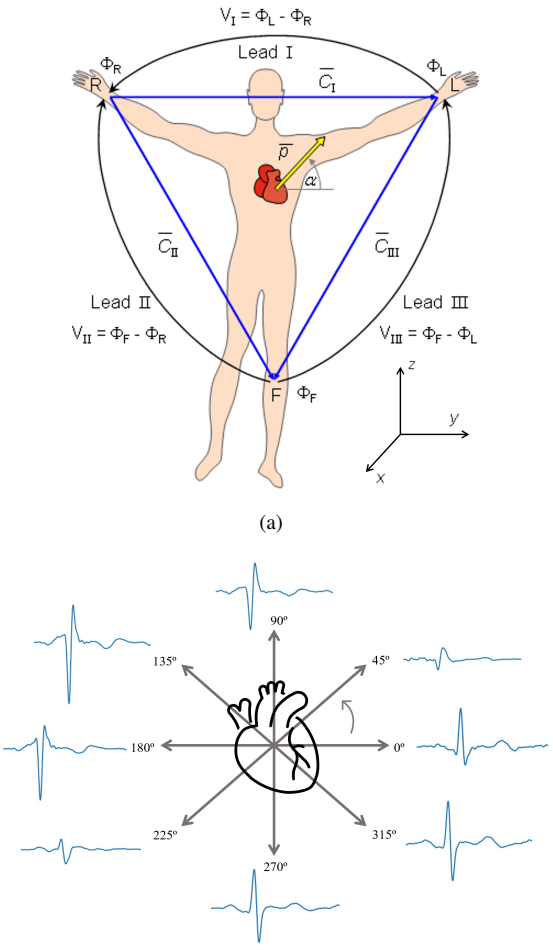
The same assumption can be made for the ECG leads measured with the stethoscope on the chest, since they are in the frontal plane. To verify that the ECG leads obtained at different rotations also follow Kirchhoff's law, eight signals were collected by rotating the stethoscope head counterclockwise in increments of 45º (Figure 7(b)). This data collection was performed always on the same auscultation point, on the left third intercostal space. Using the notation from Figure 8 we have, for example:

$$V_{0^\circ} = p\cos(\alpha) = p_y \tag{5}$$

$$V_{90^\circ} = p\sin(\alpha) = p_z \tag{6}$$

$$V_{45^\circ} = \frac{\sqrt{2}}{2}(p\cos(\alpha) + p\sin(\alpha)) = \frac{\sqrt{2}}{2}(V_{0^\circ} + V_{90^\circ}) \tag{7}$$

To compare the heartbeat waveform morphology, the ECG signals were segmented into individual beats by clipping the signals around the detected R-peaks (at -200 and 400 ms). Then, the DBSCAN algorithm was applied



(a)

(b)

Figure 7: a) Einthoven limb leads and Einthoven triangle. Φ_L and Φ_R are the potentials at the left and right arms, and Φ_F is the potential at the right leg [18]. b) ECG waveforms of leads measured at different angles of rotation of the stethoscope head, relative to the frontal plane.

to detect the outlier beat waveforms, which were excluded from the analysis. The cosine distance metric was used, with a maximum distance between samples equal to 0.06 and a minimum number of samples per cluster equal to 2.

The Mean Squared Error (MSE) and Pearson's cor-

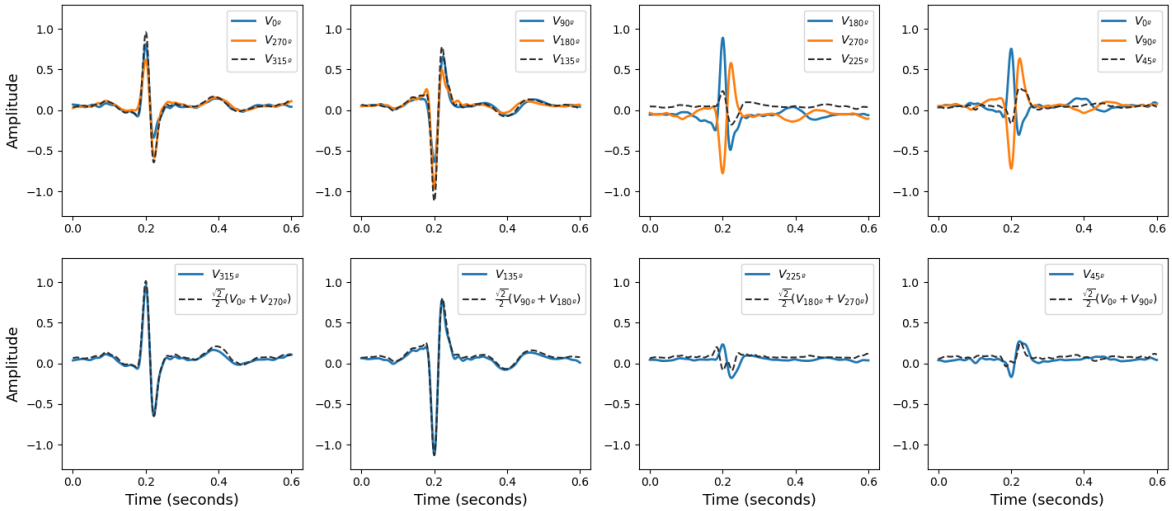


Figure 8: Example of heartbeat waveforms of experimental leads measured at different rotation angles and the equivalent waveforms obtained through Kirchhoff's law (from the adjacent ECG leads, at ± 45º).

Table 6: PCC and MSE between heartbeat waveforms of experimental single-lead ECG at different rotation angles and the equivalent waveforms obtained through Kirchhoff's law (from adjacent leads at ± 45º)

Angle	PCC	MSE
0º	0.948±0.021	2.81E-3±1.19E-3
45º	0.424±0.210	3.99E-3±1.01E-3
90º	0.974±0.012	3.31E-3±1.11E-3
135º	0.991±0.005	1.49E-3±7.99E-4
180º	0.967±0.015	3.80E-3±2.22E-3
225º	0.243±0.264	4.55E-3±1.81E-3
270º	0.985±0.007	1.35E-3±5.71E-4
315º	0.992±0.005	1.28E-3±7.27E-4

Table 7: PCC and MSE between heartbeat waveforms of experimental ECGs at different rotation angles and the equivalent waveforms obtained through Kirchhoff's law (from the adjacent two lead ECG pair).

Angle	PCC	MSE
0º+270º = 315º	0.975±0.014	3.93E-3±1.79E-3
45º+315º = 0º	0.959±0.011	8.32E-3±3.61E-3
180º+90º = 135º	0.994±0.002	4.18E-3±2.91E-4
225º+135º = 180º	0.967±0.004	5.51E-3±4.51E-4

relation coefficient (PCC) were calculated between the heartbeat waveforms of the experimental leads at different rotation angles and the equivalent waveforms obtained through Kirchhoff's law (from the adjacent ECG leads, at ± 45º). The results are presented in Table 6, and are in accordance with what can be observed in Figure 8. All the leads have a very high correlation and a low MSE, except for the ones obtained at rotation angles of 45º and 225º, which have low correlation coefficients. The most likely explanation for this is the fact that the peaks of the adjacent ECG leads at ± 45º have opposite signals and, when combining the waves, the signal amplitudes cancel each other, which makes the resulting morphology more sensitive to small angle changes. As expected, these signals also have a much lower amplitude than the other leads.

This shows that the ECG leads measured with the stethoscope follow Kirchhoff's law, similarly to the Einthoven leads. Therefore, to evaluate all the frontal plane components of the heart's electrical activity it is only necessary to perform two acquisitions in the same auscultation point with the stethoscope head at different angles. As a result, the design was further enhanced by adding two additional dry ECG electrodes perpendicular to the current ones, allowing the recording of two ECG leads at the same time to obtain a full view of the frontal plane components with a single acquisition.

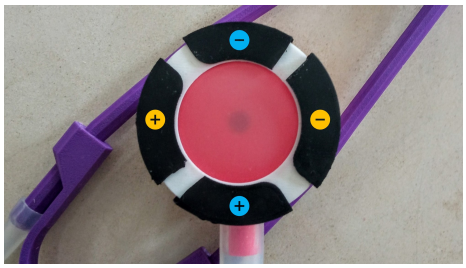


Figure 9: Disposition and polarity of two pairs of ECG dry electrodes on the stethoscope head, for the acquisition of two ECG leads.

A preliminary evaluation to validate this modified design was done by splitting each half-moon electrode part into two separate electrodes, for a total of four electrodes on the stethoscope head. The positive electrodes were placed at the left and bottom sides (Figure 9). Four acquisitions were performed by rotating the stethoscope head counterclockwise at the angles of 0º, 45º, 180º, and 225º. Since there are two perpendicular ECG leads, these measurements correspond to a total of eight signals that span all the angles of rotation represented in Figure 7(b).

Considering the polarity of the ECG electrodes on the stethoscope head, the acquired lead pairs are: 0º and 270º; 45º and 315º; 180º and 90º; 225º and 135º. The results of Table 7 show that it is possible to compute additional leads from the two ECG leads measured perpendicularly.

The calculated waveforms have a high correlation and low MSE when compared with the experimental measurements, although the results were slightly worse than the ones obtained for the single ECG leads (Table 6). This could be because the electrode parts used in this initial study aren't completely symmetrical. By designing new electrodes with a symmetrical shape, the correlation between the acquired and computed signals should increase.

3.6. Comparison with Lead I

The same protocol was used to compare the waveforms of the ECG leads obtained at different angles of rotation with the ones of the standard Lead I, which was acquired simultaneously. All the rotation angles had a high correlation with the Lead I channel, as presented in Table 8. Surprisingly, the leads with the highest correlation were the ones obtained in the vertical axis at 270º and 90º, instead of the ones obtained in the same direction as Lead I, in the horizontal axis. A possible explanation is that the lead vectors created by the electrodes of the device (which are placed at a short distance from each other on the chest) don't span the same volume as the electrodes of Lead I (which are placed on the right and left arms). The projections of the electric heart vector on the lead vector will therefore be different, altering the morphology of the measured ECG signals. Nonetheless, this high correlation suggests that the ECG leads measured with the proposed device are clinically meaningful.

Table 8: PCC and MSE between heartbeat waveforms of reference Lead I channel and experimental leads at different rotation angles.

Angle	PCC	MSE
0º	0.876±0.026	0.166±0.049
45º	-0.854±0.020	2.609±0.409
90º	**-0.922±0.016**	2.464±0.262
135º	-0.887±0.015	2.471±0.209
180º	-0.875±0.015	2.463±0.262
225º	0.812±0.020	0.228±0.055
270º	**0.928±0.007**	0.093±0.014
315º	0.870±0.014	0.186±0.030

3.7. Discussion

These results show that the PCG and ECG are adequately acquired. By using the proposed device on a standard auscultation procedure (through the 5 auscultation points), it would be possible to record 5 ECG leads that view the electrical activity of the heart from different perspectives and 5 PCG signals that will reflect the function of the different heart valves. This provides a high amount of additional information without altering the clinical routine.

An important contribution of this work is the experimental evaluation of the relationships between the leads obtained at different angles, which fills an existing gap in the state-of-the-art. They all have a high correlation with Lead I, suggesting that they all contain meaningful information for CVD screening. It was also found that the ECG leads measured with the stethoscope follow Kirchhoff's law, similarly to the Einthoven leads. A preliminary evaluation was done by recording two ECG leads at the same time with the original electrodes split in two, and it was found that it is possible to calculate additional ECG leads in the frontal plane from just one acquisition. This motivates the creation of a new design for the electrodes, to have four identical parts along the stethoscope's head.

4. Conclusions

Combining ECG and PCG acquisition in the same device allows the simultaneous inspection of the electrical and mechanical condition of the heart, further improving the accuracy of the heart disease diagnosis. The approach developed in this work enables this by incorporating ECG and PCG sensors in a validated 3D-printed acoustic stethoscope without significantly altering its form factor or acoustic transfer properties, which facilitates the integration of the new device in current clinical practice.

The implementation of deep learning architectures for classification of signals acquired during auscultation showed promising results in different tasks, including murmur detection from heart sounds and abnormality detection from multi-modal PCG and ECG data. In the latter, the developed hybrid neural network based on parallel CNN and RNN blocks had a performance at the level of other state-of-the-art methods, with a F_1 score of 90.5%.

Overall, this evaluation demonstrated promising results for further work. The low-cost, portability, and simplicity of the proposed device, allied with computer-aided and deep learning techniques for automatic diagnosis, could significantly facilitate the early screening of CVDs. We believe this system would have the biggest impact in primary care facilities as part of the auscultation routine, in telemedicine, or in areas with difficult access to an integrated health system, where an automatic screening tool that isn't dependent on trained professionals could significantly reduce the mortality and morbidity associated with both congenital and acquired cardiopathies. Furthermore, the device is open-source, which makes it more accessible and easily upgradeable. The ability to collaborate, share the technology, and save the acquired signals for further analysis and model development could also make the device an important tool in both research and education.

5. Future Work

This work has established a basis for further development. Regarding the developed prototype, the next step should be a full, comprehensive evaluation of the device in a real-world setting, with both healthy subjects and patients with diagnosed cardiovascular pathologies. It is especially important to do this evaluation in comparison with standard medical tools for cardiac monitoring (such as the 12-lead ECG), in order to validate the clinical usefulness of the new device. An application has already been made to the health ethics committee of Centro Hospitalar Vila Nova de Gaia/Espinho (CHVNG/E) to evaluate the device and sensors in a hospital environment with cardiac patients.

This can also be articulated with the development of new datasets for the training of deep learning algorithms for automatic CVD screening based on the multi-modal data. With a larger dataset and more detailed annotations, it is expected that the performance of the current deep learning models will significantly increase, and also that they will be able to generalize better for new examples.

More improvements could be done by combining the current models with attention layers, which would not only improve the models' performance (since the neural networks would be able to better extract the most relevant features in the signals and better estimate their contribution for the final classification), but also enhance the interpretability and explainability of the results, which is highly relevant in healthcare applications.

Ackowledgements

This document was written and made publically available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biomedical Engineering of the author at Instituto Superior Técnico. The work described herein was performed at Instituto de Telecomunicações (IT) in Lisbon, during the period from March 2022 to October 2022, under the supervision of professor Hugo Humberto Plácido da Silva. The thesis was co-supervised at Instituto Superior Téecnico by Prof. Ana Luísa Nobre Fred. The work also had the support from doctor Sérgio Laranjo, cardiologist at Centro Hospitalar Universitário de Lisboa Central (CHULC).

References

[1] WHO. Cardiovascular diseases (CVD) fact sheet. Accessed: 22 Aug, 2022.

[2] J. Oliveira, F. Renna, P. D. Costa *et al.*, "The CirCor DigiScope Dataset: From Murmur Detection to Murmur Classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 6, pp. 2524–2535, 2022.

[3] S. Leng, R. S. Tan, K. T. C. Chai, C. Wang, D. Ghista, and L. Zhong, "The electronic stethoscope," *BioMedical Engineering Online*, vol. 14, no. 66, 2015.

[4] W. Phanphaisarn, A. Roeksabutr, P. Wardkein, J. Koseeyaporn, and P. Yupapin, "Heart detection and diagnosis based on ECG and EPCG relationships," *Medical Devices: Evidence and Research*, vol. 4, no. 1, pp. 133–144, 2011.

[5] R. Hettiarachchi, U. Haputhanthri, K. Herath *et al.*, "A Novel Transfer Learning-Based Approach for Screening Pre-Existing Heart Diseases Using Synchronized ECG Signals and Heart Sounds," in *Proc. of the Int'l Symp. on Circuits and Systems (ISCAS)*, 2021, pp. 1–5.

[6] M. Martins, P. Gomes, C. Oliveira, M. Coimbra, and H. P. Silva, "Design and Evaluation of a Diaphragm for Electrocardiography in Electronic Stethoscopes," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 391–398, 2020.

[7] M. A. Reyna, Y. Kiarashi, A. Elola *et al.*, "Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022," *medRxiv*, 2022.

[8] S. E. Schmidt, C. Holst-Hansen, C. Graff, E. Toft, and J. J. Struijk, "Segmentation of heart sound recordings by a duration-dependent hidden Markov model," *Physiological Measurement*, vol. 31, no. 4, pp. 513–529, 2010.

[9] S. Latif, M. Usman, R. Rana, and J. Qadir, "Phonocardiographic Sensing Using Deep Learning for Abnormal Heartbeat Detection," *IEEE Sensors Journal*, vol. 18, no. 22, pp. 9393–9400, 2018.

[10] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-HSMM-based heart sound segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 822–832, 2016.

[11] C. Liu, D. Springer, Q. Li *et al.*, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, pp. 2181–2213, 2016.

[12] R. Baptista, H. Silva, and M. Rocha, "Design and development of a digital stethoscope encapsulation for simultaneous acquisition of phonocardiography and electrocardiography signals: the SmartHeart case study," *Journal of Medical Engineering and Technology*, vol. 44, no. 4, pp. 153–161, 2020.

[13] A. Pavlosky, J. Glauche, S. Chambers, M. Al-Alawi, K. Yanev, and T. Loubani, "Validation of an effective, low cost, Free/open access 3D-printed stethoscope," *PLOS ONE*, vol. 13, no. 3, pp. 1–10, 2018.

[14] E.-R. Symeonidou, A. D. Nordin, W. D. Hairston, and D. P. Ferris, "Effects of Cable Sway, Electrode Surface Area, and Electrode Mass on Electroencephalography Signal Quality during Motion," *Sensors*, vol. 18, no. 4, 2018.

[15] A. dos Santos Silva, H. Almeida, H. P. da Silva, and A. Oliveira, "Design and evaluation of a novel approach to invisible electrocardiography (ECG) in sanitary facilities using polymeric electrodes," *Scientific Reports*, vol. 11, 12 2021.

[16] H. P. da Silva, J. Guerreiro, A. Lourenço, A. Fred, and R. Martins, "BITalino: A Novel Hardware Framework for Physiological Computing," in *Proceedings of the International Conference on Physiological Computing Systems*, Jan 2014.

[17] D. Batista, H. P. da Silva, A. Fred, C. Moreira, M. Reis, and H. A. Ferreira, "Benchmarking of the BITalino biomedical toolkit against an established gold standard," *Healthcare Technology Letters*, vol. 6, no. 2, pp. 32–36, 2019.

[18] J. Malmivuo and R. Plonsey, *Bioelectromagnetism Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, 1995.