

# Project Plan for the Design and Implementation of a Heartbeat Segmentation and Classification Model

Boikanyo Radiokana (1386807)  
Elias Sepuru (1427726)  
Group: 19G11

Supervisor: Ellen de Mello Koch

*School of Electrical & Information Engineering, University of the Witwatersrand, Private Bag 3, 2050, Johannesburg, South Africa*

**Abstract:** This paper presents a plan to design and implement a Heartbeat Segmentation and Classification Model. The model is divided into 3 main sections namely: Preprocessing, Segmentation and Classification. The preprocessing stage involves reducing noise from the heartbeat signals and conditioning it for segmentation. The signal will be downsampled with a factor ranging from 10 - 20. Segmentation will be performed by computing the signals average Shannon energy. Classification is to be implemented with machine learning methods such as ANN and RNN. Project management concepts such as work-division and risk management are also taken into consideration.

**Key words:** PCG, Heartbeat-Segmentation, Heartbeat-Classification, ANN, RNN

## 1. INTRODUCTION

Cardiovascular diseases (CVDs) continue to be one of the leading causes of deaths around the world. More people die annually from CVDs than any other cause. In 2016 alone an estimated 31% of the deaths were due to CVDs [1]. Methods to detect early signs of heart diseases could prove to be very helpful in preventing deaths due to CVDs.

This paper presents the plan of Heartbeat Segmentation and Classification project. The project aims to create a first level screening for individuals in the detection of heart diseases. This will aid medical practitioners in their field and possibly home use by individuals.

In the upcoming sections, the background, system overview, design methodology and project management are discussed.

## 2. BACKGROUND

### 2.1 Project Specifications and Requirements

The aim of this project is to create a first level screening to detect signs of heart diseases in individuals. This will be carried out by using audio data sets from two sources, the iStethoscope Pro iPhone app labelled Dataset A and a digital stethoscope labelled Dataset B. The audio data was recorded by the general public and clinical healthcare practitioners respectively [2]. Both data sets A and B each have different categories which all contain various background noises. The audio files vary in length from 1 to 30 seconds as a means of reducing excessive noise [2]. Dataset A is said to consist of the the following categories, namely; Nor-

mal, Murmur, Extra Heart Sound and Artifact whilst Dataset B has Normal, Murmur and Extrasystole categories. The above mentioned categories will be used as classifiers at a later stage.

With the excessive noise present in the audio data, it is required that preprocessing methods, capable of removing noise from the data be implemented before execution of further detection methods. Following the denoising process, a method to locate S1 (lub) and S2 (dub) heart sounds as well as segmentation of Normal audio files from the two data sets is required. A machine learning method to classify heartbeat sounds into normal and diseased categories as mentioned above must be implemented.

### 2.2 Assumptions

The project is to be conducted with the following assumptions:

- The audio data range will be 30 seconds and less.
- Dataset A has only four categories (Normal, Murmur, Extra heart sound and Artifacts) and Dataset B has only three categories (Normal, Murmur and Extrasystole)

### 2.3 Constrains

The following are the project constraints:

- Constrained to only locating S1 (lub) and S2 (dub) heart sounds
- Only the Normal audio data is to be used for the location of S1 and S2.
- Only the data from here <https://www.kaggle.com/kinguistics/heartbeat-sounds/kernels> can

be used, due to ethics clearance

- Heart sounds are only in .wav and .aif format

## 2.4 Success Criteria

For the project to be deemed successful it has to meet all the requirements specified in section 2.1. Since existing solution only have an accuracy of up to 77%, an accuracy of 77% or higher is highly desirable.

## 2.5 Literature Review

Cardiovascular diseases continue to be one of the leading cause of deaths around the world [1]. With the above mentioned there have been various attempts to accurately distinguish between normal and diseased heartbeat sounds using ECG and PCG. Groch uses a microprocessor controlled Heart-Sound Gate (HSG), which automatically identifies heart sounds from PCG alone, using timing relationships [3]. He amplifies the heart sound and passes it through two bandpass filters, folds the negative portions of the waveform into the positive and the envelopes the whole signal. To locate the peaks he uses a Schmitt trigger which then generates a square wave corresponding to the peaks. To identify S1 and S2, he exploits the fact that the diastolic period is longer than the systolic period.

Karraz *et al.* makes use of data from MIT-BIH Arrhythmia database, to classify the heartbeats into five categories. They use an FIR filter set at (0.05 - 40 Hz) cut-off and a notch filter for denoising the signals, for peak location and S1 and S2 identification they use a QRS detection algorithm. To classify the heartbeat sounds into the different categories they picked, they use the Bayesian Artificial Neural Network (BANN) with the following features: i) P-amplitude, ii) P-wide, iii) R-amplitude, iv) Q-amplitude, v) S-amplitude, vi) QRS-wide, vii) T-amplitude, viii) T-wide, ix) PR-period, and x) RR-period [4]. Kampouri also uses the QRS complex for feature extraction but instead of Neural Networks, he uses Support Vector Machine (SVM) [5].

Stunic *et al.* detects and classifies heart murmurs using segmentation techniques and ANN. This study is conducted using simulated and recorded patient heart sounds. The segmentation algorithm identifies individual heart cycles and an average of all cycles is computed to extract components within 195Hz since this band has the most valuable information [6]. The basis of the segmentation algorithm is the fact that the diastolic period is longer than the systolic period. They used an alignment algorithm to ensure that data is always fed in the same order to the ANN input neurons. The ANN algorithm consists of 3 hidden layers with 25 input neurons each and 1 output neuron. The diagnostic system presented an accuracy of  $48,7 \pm 12,7\%$  for real life recorded patient records and an accuracy of  $85 \pm 7,4\%$  for simulated data [6].

To combat the issues of noise in real life recordings of heartbeat sounds Liang *et al.* uses a Chebyshev type I lowpass filter and an algorithm based on the normalised average Shannon energy of a PCG signal. The algorithm is used to find the peak locations and to pick up the locations of S1 and S2. It achieves 93% correct ratio [7]. Debbal *et al.* uses Discrete Wavelet Transform (DWT) to decompose and reconstruct a PCG signal with insignificant loss of information. The error found in reconstructing the signal is considered as an important feature in the classification of diseased categories. It was found that the reconstruction error increases with an increase in murmur intensity [8].

## 3. SYSTEM OVERVIEW

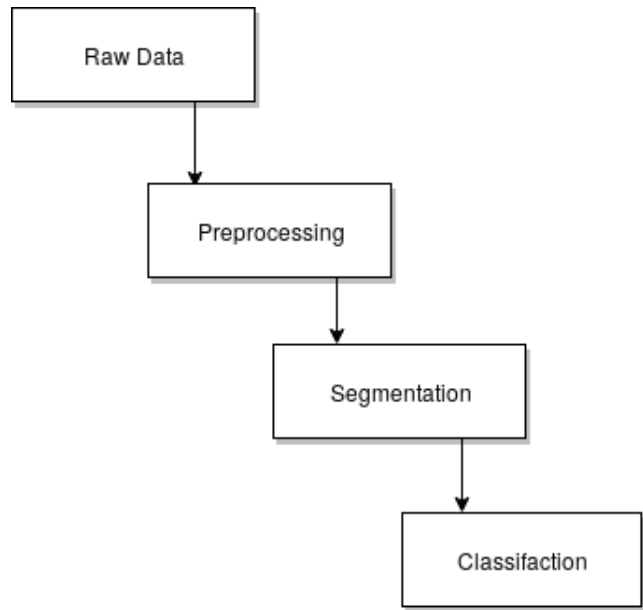


Figure 1 : System Overview

Figure 1 above illustrates an overview of the proposed system. The raw data is read from the two datasets, Dataset A and Dataset B as mentioned in the section 2.1. The audio data can either be in .wav or .aif format. As mentioned previously, the audio data sets contain excessive noise which requires denoising before segmentation can take place. This denoising process is performed in the preprocessing phase to ensure that the data is conditioned for segmentation. Once the data is conditioned, it goes through the segmentation phase where S1 and S2 are located. This is done in order to extract important features useful for classification. The features will be fed into a machine learning algorithm in order to classify the heart sounds into normal or diseased categories.

## 4. DESIGN METHODOLOGY

The project in its entirety mainly consists of two main sections, heartbeat segmentation and heartbeat classification. However, due to the fact that data cleansing

forms the bulk of the project, it was decided that it will be suitable to make preprocessing a standalone section as seen in figure 1. This will also help break down the project into manageable tasks.

#### 4.1 Preprocessing

As mentioned earlier the audio data is from real world situations and contains a lot of background noise, in the processing phase, techniques to remove the background noise are applied to the audio data.

The audio data has a sampling frequency of 44.1 kHz. The first stage of denoising the signal involves downsampling the signal with a factor between 10 - 20. Since it is known that the highest desired frequency is that of the murmur which is 600 Hz [9], downsampling by the highest factor, which is 20, will still give a sampling frequency higher than twice that of a murmur. Downsampling helps with cutting high frequency components, exploiting the Nyquist Theorem. The high frequency components are aliased whilst preserving the desired low frequency components. Downsampling also compresses the size of the audio data, allowing for faster processing time [10].

Most of the information of a PCG signal is contained in the lower frequency components of the signal as mentioned in the above paragraph. With that mentioned a low pass filter of 195 Hz will be applied after downsampling to further curb the noise [2].

The final stage of preprocessing involves normalising the signal to the absolute maximum of the signal using equation 1 [7].

$$x_{norm}(i) = \frac{x_{f_{ds}}(i)}{\max_j(|x_{f_{ds}}(j)|)} \quad (1)$$

where  $x_{f_{ds}}$  is the downsampled signal. Figure 2 summarises the preprocessing phase.

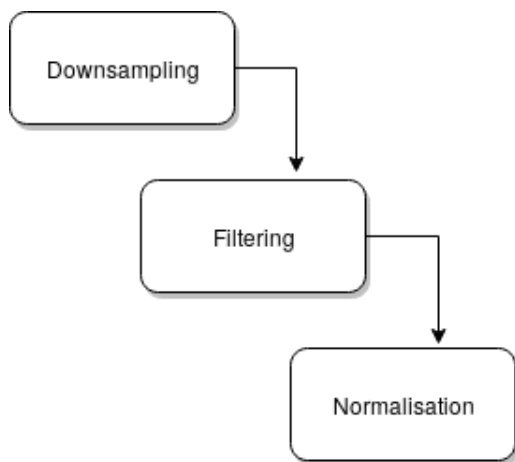


Figure 2 : Flow diagram of the preprocessing phase

#### 4.2 Heartbeat Segmentation

The segmentation phase involves locating S1 and S2 heart sounds, thus determining the diastolic and systolic periods. To segment these heart sounds, the envelope of the normalized signal obtained in the preprocessing phase is computed. The Shannon energy will be used to compute the envelope of the signal with equation [7]:

$$E = -x^2 \log(x^2) \quad (2)$$

The reason for choosing this method is the fact that Shannon energy attenuates small spikes in signals much more than large spikes [7]. Furthermore, a triangular smooth function will be applied to smoothen out the signal since the signal will still contain small spikes. Equation 3 below will then be used to determine the average Shannon energy of the entire signal [7]:

$$E_s = -\frac{1}{N} \sum_{i=1}^N x_{norm}^2(i) \log x_{norm}^2(i) \quad (3)$$

where  $x_{norm}$  is the value of the normalized sample signal and  $N$  is the signal length. This is followed by the computation of the normalized average Shannon energy using equation 4 below [7]:

$$P_a(t) = \frac{E_s(t) - M(E_s(t))}{S(E_s(t))} \quad (4)$$

where  $M(E_s(t))$  is the mean of  $E_s(t)$  and  $S(E_s(t))$  is the standard deviation of  $E_s(t)$ .

##### 4.2.1 Peak Identification

The peaks representing the heart sounds S1 and S2 will be identified with a function in Matlab called peakdet [11]. One advantage of using this function is that it is an open source function and therefore allows one to condition it to best suite their needs. S1 and S2 will be located by exploiting the fact that diastolic period is greater than systolic period.

#### 4.3 Heartbeat Classification

The classification phase involves coming up with a model, given a specific audio file, will successfully identify it as either one of the categories in Dataset A or one of the categories in Dataset B. Before training the model, feature construction and selection has to occur first. The features are to be constructed from the locations of S1 and S2.

Table 1 presents the possible features that could be used to train the machine learning models [11].

Table 1 : Potential feature selection

Feature	Description
N	Number of heartbeats per minute
M <sub>1</sub>	Mean of diastole period
M <sub>2</sub>	Mean of systole period
R <sub>s1</sub>	Standard deviation of S1 over the standard deviation of the whole signal
R <sub>s2</sub>	Standard deviation of S2 over the standard deviation of the whole signal
R <sub>m1</sub>	Mean of S1 over the total mean of the entire signal
R <sub>m2</sub>	Mean of S2 over the total mean of the entire signal
R <sub>med</sub>	Ratio of the median of the three largest segments in the sample over the total mean
R <sub>sq</sub>	Square of the array of the sorted segments of the sample

The features shown in table 1 are to be carefully selected in order to cater and distinguish for the different categories in Dataset A and Dataset B.

For training, two models, ANN and RNN are chosen. ANN is chosen due to its good ability to recognise patterns [12]. The RNN is also chosen for its great pattern recognition and its ability to handle many-to-many predictions problems [13].

## 5. TESTING

In order to allow for testing the machine learning algorithms will be trained according to 7:3 ratio. This means that 70% of the audio data will be used for training whilst the other 30% will be for testing to generate the F-score, Youdens Index and the Discriminant Power.

## 6. PROJECT MANAGEMENT

### 6.1 Work Division

The project will be conducted by a team of two people. Each section in the project as illustrated in figure 1 will be allocated two different techniques. Each team member will then be assigned a technique to follow to solve a particular section. The technique that yields the best accuracy will be the approach chosen for the project. Table 2 below shows a detailed work division amongst the members of the team.

### 6.2 Project Time-line

The project will officially be started on the 15<sup>th</sup> July 2019 following extensive research by both team members. The project deadline is on the 6<sup>th</sup> September

Table 2 : Work Division

Task	Technique	Team Member
Environment Setup	-	Both
Preprocessing	1	Elias
	2	Boikanyo
Segmentation	1	Elias
	2	Boikanyo
Classification, Training and Testing	1	Elias
	2	Boikanyo
Documetation	-	Both

2019. All in all the project is said to run for six weeks. Table 3 shows all the project milestones and completion dates.

Table 3 : Project Milestone Completion

Completion Date	Milestone
15 July 2019	Project Planning
16 July 2019	Environment Setup
31 July 2019	Data Cleansing
09 Aug 2019	Segmentation
30 Aug 2019	Classification(Model Training and Testing)
06 Sep 2019	Documentation (Final Project Report)

### 6.3 Resource Management

The following resources (hardware and software) will be necessary to successfully carry out the project:

- Two compatible computers for data cleansing, training and testing of the models
- Clusters for processing big data
- MATLAB and Jupyter Notebook for signal processing and building models
- KNIME and Anaconda building models
- Internet for research
- Digital Stethoscope.
- Funds for printing of technical report and the poster for open day.

### 6.4 Risk Management

Potential risks that can possibly jeopardise the success of the project have been identified. The risk register in table 4 shows the cause of each risk that is identified, its rating, degree of impact as well as the action that should be taken to mitigate the risk.

Table 4 : Project Risk Register

Risk	Cause	Risk Rating	Impact	Actions	Person Responsible
Machines crashing	Software installation and Processing Big data	H	H	Use clusters provided by Professor Otto for running big data simulations and proper software installations	Both members
Failure to remove noise	Application of incorrect techniques	H	H	Use of alternative secondary open-source data from MIT. Use of predefined methods	Both members
Mislocation of S1 and S2	Using incorrect techniques	M	M	Use of multiple methods of locating S1 and S2 to check for any errors and correspondence of the results.	Both members
Loss of work	Not saving copies of completed work	H	H	Use Github for version control	Both members
Poor milestone estimation and project running behind schedule	Underestimating the complexities of each section	M	M	Work on the project as soon as possible,	Both members

## 7. CONCLUSION

A plan to design and implement a Heartbeat Segmentation and Classification Model is presented. The model is divided into 3 main sections namely; Preprocessing, Segmentation and Classification. The preprocessing stage involves reducing noise from the heartbeat signals and conditioning it for segmentation. The signal will be downsampled with a factor ranging from 10 - 20. Segmentation will be performed by computing the signals average Shannon energy. Lastly, classification will be implemented with machine learning methods such as ANN and RNN.

## REFERENCES

- [1] WHO. “Cardiovascular diseases (CVDs).” [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Last Accessed: 2019-07-14.
- [2] P. Bentley, G. Nordehn, M. Coimbra, and S. Mannor. “The PASCAL Classifying Heart Sounds Challenge 2011 (CHSC2011) Results.” <http://www.peterjbentley.com/heartchallenge/index.html>.
- [3] M. W. Groch, J. R. Domnanovich, and W. D. Erwin. “A new heart-sounds gating device for medical imaging.” *IEEE Transactions on biomedical engineering*, vol. 39, no. 3, pp. 307–310, 1992.
- [4] G. Karraz and G. Magenes. “Automatic classification of heartbeats using neural network classifier based on a Bayesian framework.” In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4016–4019. IEEE, 2006.
- [5] A. Kampouraki, G. Manis, and C. Nikou. “Heartbeat time series classification with support vector machines.” *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 512–518, 2008.
- [6] S. L. Strunic, F. Rios-Gutiérrez, R. Alba-Flores, G. Nordehn, and S. Bums. “Detection and classification of cardiac murmurs using segmentation techniques and artificial neural networks.” In *2007 IEEE Symposium on Computational Intelligence and Data Mining*, pp. 397–404. IEEE, 2007.
- [7] H. Liang, S. Lukkarinen, and I. Hartimo. “Heart sound segmentation algorithm based on heart sound envelopogram.” In *Computers in Cardiology 1997*, pp. 105–108. IEEE, 1997.
- [8] S. Debbal and F. Bereksi-Reguig. “Filtering and classification of phonocardiogram signals using wavelet transform.” *Journal of medical engineering & technology*, vol. 32, no. 1, pp. 53–65, 2008.
- [9] C. Yang. “Music database retrieval based on spectral similarity.” Tech. rep., Stanford, 2001.
- [10] D. Lavry. “Sampling Theory For Digital Audio.” *Lavry Engineering, Inc. Available online: [http://www.lavryengineering.com/documents/Sampling\\_Theory.pdf](http://www.lavryengineering.com/documents/Sampling_Theory.pdf) (checked 24.5. 2010)*, 2004.
- [11] E. F. Gomes and E. Pereira. “Classifying heart sounds using peak location for segmentation and feature construction.” In *Workshop Classifying Heart Sounds, La Palma, Canary Islands*, p. 10. 2012.
- [12] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan. *Neural network design*. Martin Hagan, 2014.
- [13] Y. He and X. Qing. *Automatic Control, Mechatronics and Industrial Engineering: Proceedings of the International Conference on Automatic Control, Mechatronics and Industrial Engineering (ACMIE 2018), October 29-31, 2018, Suzhou, China*. CRC Press, 2019.