Detecting Heart Diseases using a Stethoscope-based Heart Sound Method

Sayeda Farzana Aktar, Stefan Andrei
Department of Computer Science
Lamar University
Email: saktar@lamar.edu, sandrei@lamar.edu

Abstract—Detecting heart diseases has been a research interest for centuries. Many of these techniques are based on stethoscope, but only a few of these are digitally analyzed. In our paper, we propose a new method to detect heart diseases by analyzing heart sounds. Our goal is to help the medical doctor to identify whether a patient has heart disease or not. In general, doctors use acoustic stethoscope to detect abnormalities in the heart sound and predict abnormal conditions of the human heart. One major problem is that the frequency range and intensity of the heart sound is very low. Moreover, there are different types of heart sounds indicating different types of heart diseases. Hence, doctors are facing difficulties while detecting the cardiac sound and its abnormalities. Even the expert doctors may fail sometime to analyze heart sound properly. We developed and applied a novel data analysis to detect heart problems. Our method uses deep architecture features for analyzing heart diseases. We consider the heart sound as our raw data. This approach uses electronic stethoscope also known as e-stethoscope (that is, electronic stethoscope) to collect heart sounds and deep learning approach to identify that a heart has any disease or is healthy. If the heart has a disease, then it is desirable to identify the disease type. It is aimed to design a software known as a heartbeat audio classifier. This software should be able to differentiate normal heartbeats and heart murmurs which would assist the doctors to analyze a heart sound and detect a disease condition of the heart. Though our approach is not perfect, it shows that our approach leads to better results in comparison with others.

Keywords- Cardiovascular Diseases (CVDs), Spectrogram, Machine Learning, Deep Neural Network (DNN), Convolutional Neural Network (CNN).

1 Introduction

According to the World Health Organization, cardiovascular diseases (CVDs) are one of the main causes of death globally. Auscultation, i.e., listening to the mechanical valvular activity is a valuable method for CVDs detection. It is the most primary diagnosis method for the initial detection of heart valves which is also economical and simple screening test, since the only device needed is a stethoscope. Nowadays, with the recent advances in signal processing and electronic stethoscope technologies, the design of automatic classification schemes from phonocardiogram (PCG) recordings appears as a promising diagnosis method for CVDs detection. This paper aims at defining end-to-end pipelines for automatic classification of heart sounds to identify whether a person has a heart disease or not. We use e-stethoscopes to collect heart sounds which have several utilities like an ambient noise reduction, audio recording storage and even remote listening. Moreover, we use deep learning to analyze data and to identify if a person has a heart disease or not. First, we discuss a few concepts about physiology of heart sounds such as the cause of generation, characteristics and later some reviews of important research work that are already done.

1.1 Heart sound signals

1.1.1 Physiology of heart sound

Every heart makes sound. Heart sounds are generated from blood flowing, first the heart chambers, cardiac valves opening and closing during the cardiac cycle. Vibrations of these structures create heartbeat sounds — the more valves contract and relax allowing blood flow to and from the heart, the more vibrations will get generated. The variables- fluid viscosity, density, velocity, and the diameter of the column through which the fluid is traveling, determine the turbulence of blood flow. A healthy heart beats in a repeated rhythm. But if the heart sound is abnormal then- some rhythms are harmless, but some rhythms can indicate serious heart problems.

1.1.2 Different types of heart sounds and their generations

Heart sounds are noises generated by the beating heart and the flow of blood through it. The sounds are mainly the reflection of turbulence created when the heart valves snap shut. Heart sounds provide important auditory data which indicate the condition of the heart. In healthy adults, there are two heart sounds often described as: lub and dub (or dup) [3].

In each heartbeat, lub and dub are sequential. The closing of the AV valves and semilunar valves produces the first heart sound (S1) and second heart sound (S2), respectively. For an abnormal heart, there are some additional sounds present like-Cardiac murmurs: Heart murmurs are produced by the turbulent flow of blood, and they may occur either inside or outside of the heart [3].

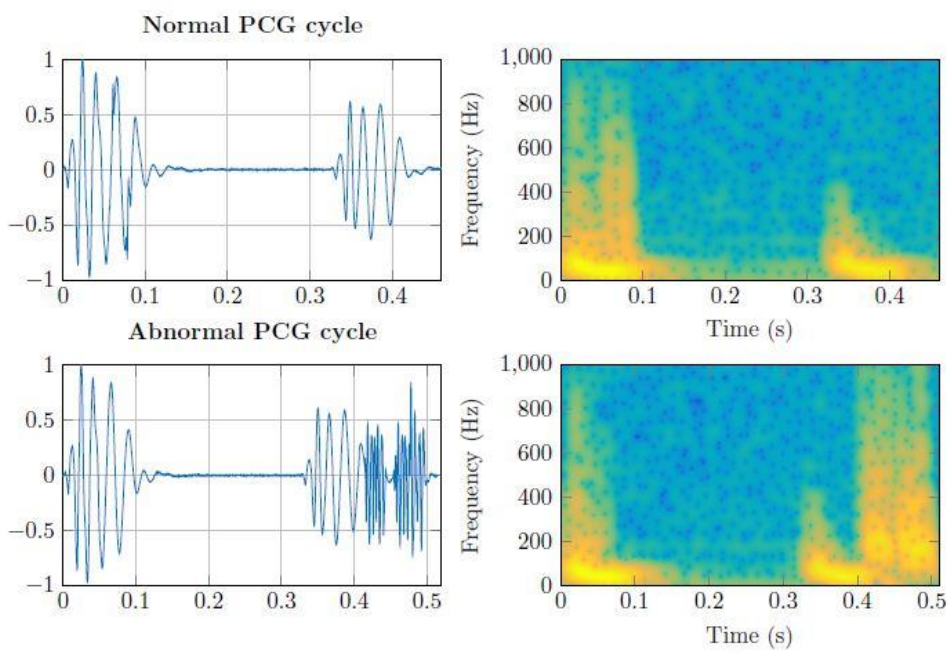

Figure 1. The time waveform and spectrogram representative of a PCG cardiac cycle in normal (top) and abnormal (bottom) conditions [6].

There are two main types of murmurs - physiological (benign) and pathological (abnormal). There are several reasons of an abnormal murmur. First, the stenosis restricting the opening of a heart valve, as a result turbulence as blood flows through it. Second, valvicular insufficiency (as known as regurgitation) allows backflow of blood when an incompetent valve closes with partial effectiveness. Several murmurs are audible in different parts of cardiac cycle, which depend on the reason of murmur. For young people, sometimes there is another heart sound known as, the third heart sound (S3) and it occurs from 0.1sec to 0.2sec, after the second heart sound. The third heart sound occurs because of the rush of blood from the atria into the ventricles, and it produces the turbulence and some vibrations of the ventricular walls. Finally, the fourth heart sound known as (S4) is an advanced diastolic sound, which corresponds to the late ventricular filling through an active atrial contraction. This heart sound has a low-intensity with the bell of the stethoscope.

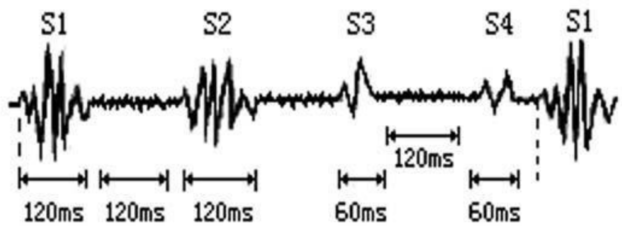

Figure 2: S1 30 to 45 hertz, S2 has frequency range of 50 to 70 hertz. The third heart sound is an extremely weak vibration ranges below 30 Hz.

1.2 Prior research efforts

There are a few studies that have been done about classifying various heartbeat sounds using machine learning methods. Chao and his team proposed a feature extraction and classification methods for heart sounds [1]. Their paper presented a process for classifying audio heart recordings to five most commonly occurring classes- artifact, extra heart sound, extrasystole, murmur and normal heartbeat [1]. The paper also compared the precision and F-scores of six machine learning models, which included naive bayes, support vector machines and decision trees [1]. Instead of F-scores our proposed approach using CNN and AUROC score in order to improve the classifier performance.

Milad and his team proposed a new deep feature selection method based on deep architecture [2]. Their method uses stacked auto-encoders for feature representation in higher-level abstraction [2]. They focused on assessing and prioritizing risk factors for hypertension (HTN) in a vulnerable demographic subgroup (African American) [2]. Their approach is to use deep learning to identify significant risk factors affecting left ventricular mass indexed to body surface area (LVMI) as an indicator of heart damage risk [2]. Although our proposed approach also based on deep feature selection method depending on deep architecture, but we focused on detection of heart diseases by analyzing the heart sounds.

In 2018, Roilhi et al. proposed a pipeline and benchmark for binary heart sounds classification [4]. Their proposed architecture was focused on the use of matching pursuit time-frequency decomposition using Gabor dictionaries and the linear predictive coding method of a residual [4]. They compared seven classifiers with two different approaches: feature averaging and cycle averaging [4]. They used the PhysioNet/CinC challenge 2016 database, which comprises a wide variety of heart sounds recorded from patients with normal and different pathological heart conditions [4]. Roilhi and his team conducted a 10-fold stratified cross-validation method to evaluate the performance of different classification algorithms [4]. Although Roilhi et al. was about the heartbeat classification, their work is different from ours. We propose a heart sound classification based on a CNN algorithm and for improving the classifier performance we use differentiable approximation of the AUROC score.

Yaseen et al. created the database of five categories of heart sound signal (PCG signals) from various sources, where one category was normal and four other abnormal categories [5]. Their study proposed an automatic classification algorithm for cardiac disorder by the heart sound signal [5]. They extracted features from phonocardiogram signal and then processed those features using machine learning techniques for classification [5]. In features extraction, they used Mel Frequency Cepstral Coefficient (MFCCs) and Discrete Wavelets Transform (DWT) features from the heart sound signal [5]. As learning and classification, they used a support vector machine (SVM), a deep neural network (DNN) and a centroid displacement-based k nearest neighbor [5]. For accuracy, they combined MFCCs and DWT features for training and classification using SVM and DWT [5]. Compared with Yaseen et al. work our approach for classification of a heart sound is different. We take the heart sound data as WAV format and convert it to a spectrogram PNG. We use CNN instead of Deep Neural Network (DNN) and performs a series of 2D convolutions and max-pooling operations prior to a series of fully connected layers in order to get a good result.

2 Our Methodology

Our method has mainly two parts. First, we process the audio file which is a heart sound and in the second part we build our model to analyze the heart data.

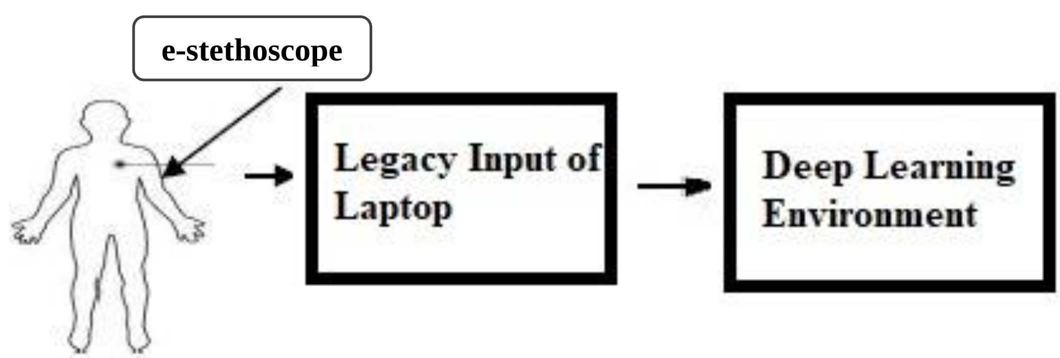

Figure 3: System architecture.

At the beginning, we take heartbeat data from a patient using e- stethoscope. The e-stethoscope gives us data that will be generated the WAV format. To read the information from this encoded audio WAV data, we convert these heartbeat audio recordings (WAV files) into images. For each WAV file, one PNG of a 2D spectrogram is generated.

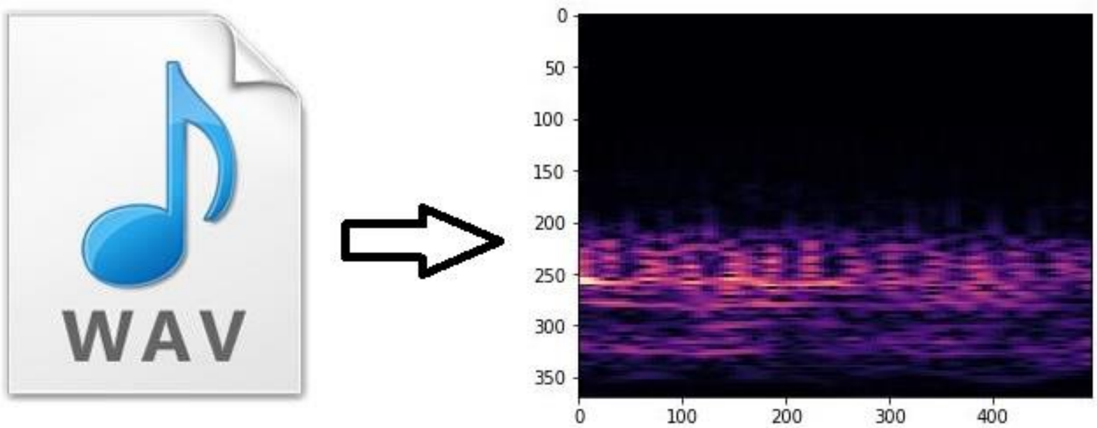

Figure 4: Conversion of the heart sound of WAV file format to PNG OF A 2D spectrogram.

Spectrograms are convenient for representing these heartbeat recordings (WAV files) because they take the intensity of the frequencies from a given sound byte. Recent works have shown that one can recreate the original audio closely from respective spectrograms. As a result, it can be assumed that spectrograms represent effective representations of an audio recording. After converting WAV files to PNG, these images are then further trimmed to reduce their size, before being fed to the model.

Now, we prepare data for our model using- training validation data sets, testing validation data sets and validation data sets. The training set is used by the algorithm to learn. Data validation ensure us that the data is cleaned, corrected and useful. The test set is a set of observations that we used to evaluate the performance of the model.

After preparing the data, we apply the model checkpoint from `keras` library to make sure there are no faults. Then apply the model checkpoint, which is a tolerance technique for long running processes. The API allows us to specify which metric to monitor, such as loss or accuracy on the training or validation dataset. We use this API to define where to checkpoint the model weights, under what circumstances to make a checkpoint of the model. Then we use `EarlyStopping`, this callback allows you to specify the performance measure to monitor, the trigger. Once this triggers it will stop the training process. Our model is a traditional CNN and our model performs a series of 2D convolutions and max-pooling operations prior to a series of fully connected layers. Due to the class imbalance, dropout and kernel regularizers are employed selectively to prevent overfitting.

While binary cross-entropy and KL-divergence are effective loss functions for optimizing the accuracy of our model, they fail to properly optimize AUC-ROC, or the AUROC score. To improve the classifier performance, we use differentiable approximation of the AUROC score instead of the loss function. For using AUROC score there is a minor decrease in accuracy. However, there is a significant improvement in the AUROC score of the trained classifier after this change. All these considerations are embedded in Figure 5.

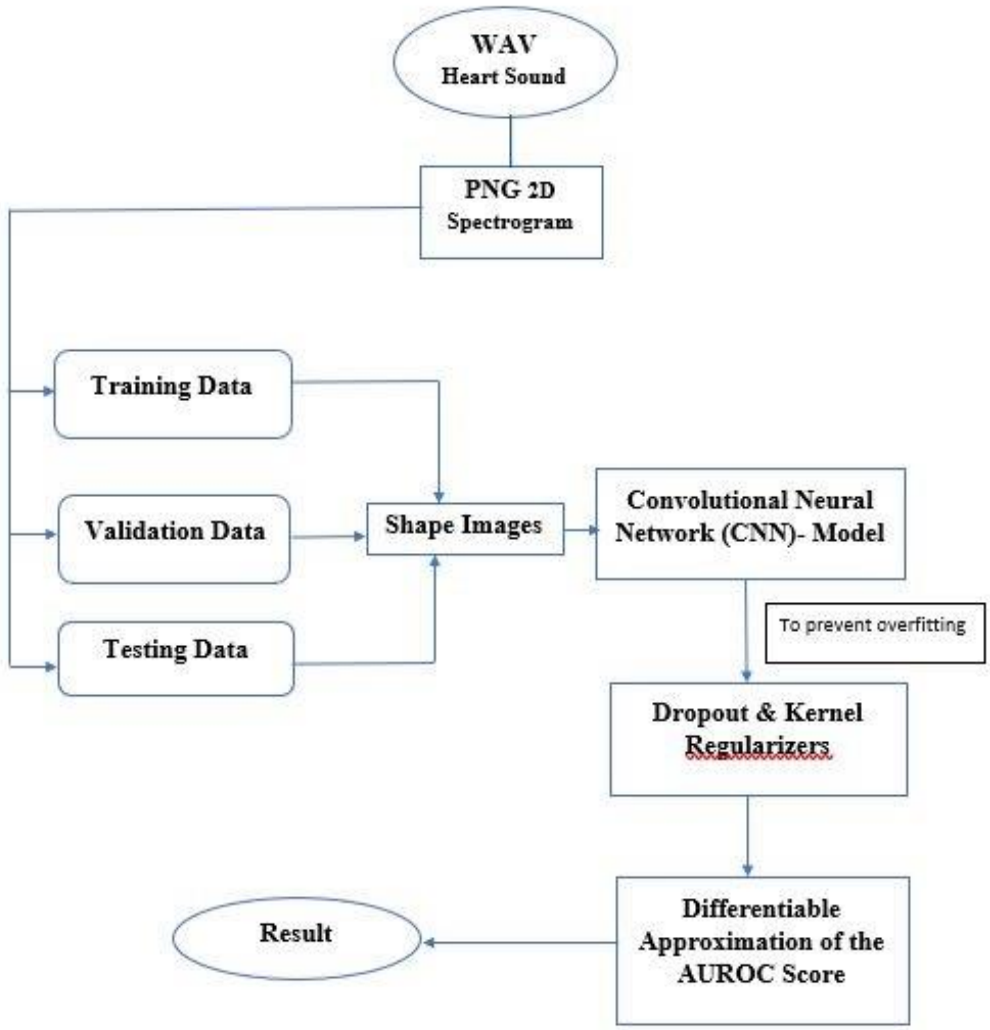

Figure 5: The flow chart about determining AUROC score

For this experiment we use three files - lhbutils.py, lhbmodels.py, HeartbeatClassification.ipynb. The file [`lhbutils.py`] (`lhbutils.py`) contains utility methods for processing the audio and image data. The file [`lhbmodels.py`] (`lhbmodels.py`) holds the Keras model architecture. The file [`HeartbeatClassifier.ipynb`](`HeartbeatClassifier.ipynb`) contains of the data preprocessing and model development/training process.

Followings are the codes of `HeartbeatClassifier`.ipynb.

Step 1: Take WAV file (heart sound) as input and convert it to PNG

Get the heartbeat data as audio WAV data. After that, convert these heartbeat audio recordings (WAV files) into images. For each WAV file, one PNG of a 2D spectrogram is generated.

```
import lhbutils as lhb

lhb.unzip_data('data/set_b.zip', 'data/wavfiles')

lhb.convert_wavfiles_to_spectrograms('data/wavfiles', 'data/pngfiles',
trim=2)

df = lhb.load_df('data/set_b.csv', 'data/wavfiles', 'data/pngfiles')

df.head()
```

Step 2: Separate training data, validation data and test data

Split all data as training data, validation data and test data for using them in our model.

```
X_train, X_test, X_val, y_train, y_test, y_val =
hb.get_train_test_validation_split(df, random_state=17)
```

Step 3: CNN model implementation

`ModelCheckpoint` is the weights of the model. We used to monitor the model such as the loss on training data.

`EarlyStopping` callback allows us to specify the performance measure to monitor, the trigger and once triggered, it will stop the training process.

`HeartbeatClassifier` is the main model class. Heartbeat Classification is a walkthrough of the data preprocessing and model development/training process.

```
fom lhbmodels import HeartbeatClassifier

from keras.callbacks import ModelCheckpoint, EarlyStopping

checkpoint = ModelCheckpoint('best_weights.h5', monitor='val_loss',
mode='min', save_best_only=True, period=2, save_weights_only=True)

earlystop = EarlyStopping(monitor='val_loss', min_delta=0, patience=20,
verbose=0, mode='min')

model = HeartbeatClassifier(loss=lhb.pair_loss,
input_shape=X_train.shape[1:], num_classes=2)

history = model.fit(X_train, y_train, validation_data=(X_val, y_val),
batch_size=4, callbacks=[checkpoint, earlystop], epochs=300, verbose=0)
```

Step 4: Optimizing the accuracy of our model

The code below represents the plot accuracy and the loss information.

```
import matplotlib.pyplot as plt

%matplotlib inline

plt.figure(1)

# summarize history for accuracy
```

```
plt.subplot(211)
plt.plot(history.history['acc'])
plt.plot(history.history['val_acc'])
plt.title('model accuracy')
plt.ylabel('accuracy')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
# summarize history for loss
plt.subplot(212)
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'test'], loc='upper left')
plt.ylim(ymax=2, ymin=0)
plt.show()
```

First, we do audio preprocessing – we convert heartbeat recordings into an image, a PNG of a 2D spectrogram. All the data is split into stratified training, validation, and testing sets.

Model Development -The model is a traditional convolutional neural network (CNN) and performs a series of 2D convolutions and max-pooling operations prior to a series of fully connected layers. At last, we plot our model output data to visualize model accuracy and loss.

3 Experimental Results

The CNN achieves an accuracy of approximately 79% on unseen test data, with an approximate AUROC score of 0.78. For our model performance improvement, we use differentiable approximation of the AUROC score because differentiable approximation of the AUROC significantly improve AUROC score of the trained classifier.

Conclusion

In our approach, we converted (WAV files) into images, namely 2D spectrograms (PNG files). These images are then further trimmed to reduce their sizes, before being fed to the mode (Shape image). After that, we split the data into stratified training, validation and testing sets. Our model is a traditional CNN that performs a series of 2D convolutions and max-pooling operations prior to a series of fully connected layers. We implement a differentiable approximation of the AUROC score in order to improve the classifier performance in this regard. Although these results are not convincing enough to use the model in a diagnostic or a commercial setting, they are still strong and offer an impressive baseline for future competing methods. A future work will include our efforts to classify heartbeat recordings.

References

1. Angela Chao, Shirley Ng and Linda Wang. 2016. Listen to Your Heart: *Feature Extraction and Classification Methods for Heart Sounds*, [accessed online on November 12, 2019 at https://lindawangg.github.io/projects/classifying-heartbeats.pdf]
2. Milad Zafar Nezhada, Dongxiao Zhub, Xiangrui Lib, Kai Yanga, and Phillip Levyc. 2016. *SAFS: A Deep Feature Selection Approach for Precision Medicine.* IEEE, [accessed online on November 12, 2019 at https://arxiv.org/pdf/1704.05960.pdf]
3. Joyanta Kumar Roy and Tanmay Sinha Roy. 2017. *A Simple technique for heart sound detection and real time analysis.* IEEE [accessed online on October 30, 2019 at https://ieeexplore.ieee.org/document/8538255]
4. Roilhi Ibarra-Hernández, Nancy Bertin, Miguel Alonso-Arévalo and Hugo Guillén-Ramírez. 2018. *A benchmark of heart sound classification systems based*. [accessed online on October 30, 2019 at https://hal.inria.fr/hal-01935058/document]
5. Yaseen, Gui-Young Son and Soonil Kwon. 2018*. Classification of Heart Sound Signal Using Multiple Features.* [accessed online on October 30, 2019 at https://www.researchgate.net/publication/329139130_Classification_of_Heart_Sound_Signal_Using_Multiple_Features]
6. Jonathan Rubin, Rui Abreu and Anurag Ganguli. 2017. *Recognizing Abnormal Heart Sounds Using Deep Learning*. [accessed online on October 30, 2019 at https://arxiv.org/abs/1707.04642]