Detecting Heart Irregularities Through ECG Sonification

Rahel Joshi

***Abstract*—Cardiovascular disease is a primary cause of death both in the US and worldwide. ECGs are a widespread tool used to monitor heart rhythm and can be used to detect heart irregularities. Applying preprocessing methods to denoise and amplify ECG signals can result in better representation of key ECG features like PQRST waves. Utilizing sonification methods can allow users to listen to a patient's heart rhythm to detect any heart irregularities. MIT BIH Arrhythmia and Normal Sinus Rhythm datasets and self collected ECG data were used to test sonification results. Additionally, convolutional neural networks can be applied to sonified ECG results to detect heart arrhythmia.**

***Index Terms*—Audio Systems, Audio Databases**

I. Introduction

The primary reason for death in the US is heart disease, which is responsible for about 697,000 deaths in 2020 [1] while costing the US about $229 billion [2]. Heart arrhythmia is a commonly diagnosed heart irregularity with atrial fibrillation (AF) being the most common type of heart arrhythmia [3], predicted to affect 12.1 million Americans by 2030 [4,5]. AF occurs when the left and right heart atria beat uncoordinated, resulting in an abnormally slow, fast, or irregular heartbeat, impacting blood flow from the atria to the ventricles. The main symptoms of AF are an irregular heartbeat and heart palpitations [3], and the main risk factors of AF are elderly age, high blood pressure, and obesity [6,7]. Not only does AF increase the risk for stroke but it also increases the severity of strokes [3,8], and the death rate due to AF has been increasing for 2+ decades [9]. However, AF can be treated through medicine that normalizes the heart rate and prevents blood clots [3]. This requires early detection and monitoring to prevent further complications of AF.

Electrocardiography (ECGs) provides a simple noninvasive test to detect AF. Typically a 12 lead ECG is used at hospitals to measure the heart's electrical activity through 10 electrodes attached to the arms, legs, and chest. ECGs detect the potential difference between heart muscle cells in the form of electric amplitudes over time, capturing the heart's rhythm [10]. Heart abnormalities such as AF can be diagnosed by doctors through ECGs [11]. The sinus node (SAN) controls the heart's pace through electric signals detectable by ECGs. The SAN sends a wave of depolarization from the atria to the atrioventricular node (AVN), resulting in atrial contraction. Then, the wave of depolarization travels to the bundle of His and then through the left and right bundle branches along the interventricular septum, resulting in ventricular contraction [10].

The depolarization of the heart can be seen through ECGs. The main key features of ECGs, PQRST waves, correlate to different stages of depolarization. P waves are caused by atrial depolarization. The QRS complex results due to ventricular depolarization. Q waves represent interventricular septum depolarization, R waves depict the depolarization of the main ventricular mass, and T waves show the last depolarization of the ventricles [10].

RR intervals are an important factor in diagnosing heart arrhythmia. Too fast, too slow, or irregular RR intervals can indicate arrhythmia. AF, the most common type of arrhythmia, can be detected by irregular RR intervals [10].

ECG signals can be represented visually but also sonically. Sonification is the process of conveying information within data in audio, for users to hear. The technique of sonification can allow for auditory monitoring of the heart rhythm.

Previous research done on ECG sonification has primarily involved parameter-mapping ECG features to the sonification. One such method utilized the ECG amplitude, mean amplitude, and slope to control audio frequency, volume, and panning. Another method utilized the ST segment amplitude to control audio morphology between a sine and square wave. An additional method utilized the ST segment amplitude to determine the number of water drops present in the sonification [14].

This paper seeks to utilize similar parameter-mapping ECG sonification techniques for monitoring heart rhythm.

II. Methodologies

A. Preprocessing

1) Denoising: One problem regarding ECGs in detecting heart irregularities is the presence of noise within the signal. The main noises found in ECGs are baseline wander resulting from body movements and respiration at frequencies between 0.05 and 1 Hz, power line interference caused by currents in cables and hardware centered at the frequency of 60 Hz, and muscle artifacts produced by electric signals during muscle contractions and body movement between 20 and 1000 Hz. The last noise is especially difficult due to overlapping within the frequency range of the ECG data itself [12].

Through the use of a Fast Fourier Transform, ECG time-domain data can be converted to frequency-domain data. A notch-filter from 59-61 Hz removes any power line noise from the ECG data. Then, a band-pass filter from 1-100 Hz removes any low frequency baseline wander and some high frequency muscle artifacts. Then, a Wavelet Transform filter

utilizes thresholds to futher denoise the ECG. After, a moving average filter of .04 seconds smooths the ECG signal, also removing some high frequency muscle artifacts.

2) PRT Amplification: The main key features located within the now denoised ECG for sonification are the P, R, and T waves. Also, the P and T waves have a much lower amplitude than the R wave and are therefore harder to detect and sonify. Thus, amplifying the P, R, and T waves is important to extract these key features.

The Pan-Tompkins algorithm is utilized to detect the R waves [13]. The next step is to detect the P peaks. Iterating leftwards through sample amplitudes of every detected R peak, the closest relative maxima is recorded as a P peak. The same is done to detect T peaks by iterating rightwards through sample amplitudes of every detected R peak. The first relative maxima greater than the average sample amplitude is then recorded as a T peak.

Then, the average amplitude of all the P, R, and T peaks are calculated. After, the amplification coefficients for the waves are calculated.

$$Amp_P = 4 * Avg_{Amp}/Avg_{Amp_P}$$

$$Amp_R = 7.5 * Avg_{Amp}/Avg_{Amp_R}$$

$$Amp_T = 5 * Avg_{Amp}/Avg_{Amp_T}$$

Then, the R waves are multiplied by Amp_R. The P and T waves are amplified by Amp_P and Amp_T, respectively, at the peaks and up to amplitudes .02 seconds located left of the P peak and right of the T peak.

B. Sonification

Sonification of the denoised and amplified ECG data requires two key features: the amplitude of the ECG sample and the ECG sample distance from the average amplitude of the ECG. The former is mapped to the frequency of a sine wave which is normalized from 50 to 400. The latter is mapped to the amplitude of a sine wave which is normalized from 1 to 200. Then, the resulting sound wave is upsampled to 8000 Hz to match the audio player's sampling rate. This process of ECG sonification was based off of Andrea Lorena Aldana Blanco, Steffen Grautoff, and Thomas Hermann's polarity sonification method [14].

C. Datasets

1) MIT BIH Arrhythmia: The first dataset was created by George Moody and Roger Mark and is titled "MIT-BIH Arrhythmia Dataset" and contains 48 30-minute ECG samples. The dataset's ECGs were sampled at 360 Hz from patients at the Beth Israel Hospital Arrhythmia Laboratory between 1975 and 1979. Samples 100 to 124 capture heart rhythms that a heart arrhythmia detector might be able to detect. Samples 200 to 234 capture heart rhythms, including "complex ventricular, junctional, and supraventricular arrhythmias and conduction abnormalities", that might pose a challenge for heart arrhythmia detectors to detect because of "rhythm, QRS morphology variation, or signal quality". The subjects were 25 men from 32 to 89 years and 22 women from 23 to 89 years.

The utilized ECG channel for sonification was the modified limb lead II (MLII) located on the patient's chest. Also, due to minute differences in orientation of tape heads and wobble of the tape, there is a slight skew between the data of the two ECG channels. Also, tape slippage resulted in ECG flutter.

The dataset was sampled at 360 Hz with noise at 60 Hz due to the battery operated recorders. Samples were recorded from both signals nearly simultaneously at 11 bits with values ranging from 0 to 2047 with 1024 equaling 0 volts.

Two cardiologists annotated this dataset, resolving any conflictions. R peaks are labeled and a digital bandpass was applied to amplify the QRS complexes [15,16].

2) MIT BIH Normal Sinus Rhythm: The second dataset was created by George Moody and is named "MIT-BIH Normal Sinus Rhythm Dataset" and contains 18 long-term 24 hour-plus ECG samples. The subjects included 5 men ages 26 to 45 and 13 women ages 20 to 50.

3) Self Collected ECG: The hardware used in data collection were three electrodes, one placed on the left wrist, one placed on the right wrist, and the last placed on top the right hand. The data collected was single-lead and single-channel through a AD8232 heart rate monitor connected to an Arduino Uno micro controller. The ECG data was sampled at 500 Hz with a baudrate of 230400.

D. RR Interval Outliers

Using the previously calculated R peak threshold to detect R peaks, RR intervals can now be calculated. Then, RR interval outliers can also be detected with a left and right bound:

$LB = Median_{RR\_Interval} - 1.5 * Std\_Dev_{RR\_Interval}$

$RB = Median_{RR\_Interval} + 1.5 * Std\_Dev_{RR\_Interval}$

Any RR intervals outside these bounds are considered to be potential arrhythmic heartbeats. However, some of these RR intervals are actually normal sinus rhythm heartbeats. Thus, to classify heartbeats between normal sinus rhythm and arrhythmic, a one dimensional convolutional neural network (CNN) can be applied here.

E. CNN Irregular Heartbeats

A total of 110 samples of 4 second ECG sonification intervals were taken from patient 100 from the MIT BIH Arrhythmia Dataset. 33 samples were annotated as arrhythmia. 44 samples were detected as potentially arrhythmic but annotated as normal sinus rhythm. 33 samples were the chronologically first annotated normal sinus rhythm heartbeats. The 44 samples and 33 normal sinus rhythm heartbeat samples were combined to create 77 samples of normal sinus rhythm heartbeat. Due to the large class imbalance, the SMOTE algorithm was used to oversample arrhythmic samples by synthetically creating them [17]. Afterwards, there were a total of 154 samples with a 50:50 split between classes.

Table I. Key parameters for 1D model of CNN

Layer (type)	Output Shape	Param #
conv1d (Conv1D)	(None, 31999, 64)	192
max_pooling1d (MaxPooling1D)	(None, 15999, 64)	0
dropout (Dropout)	(None, 15999, 64)	0
batch_normalization (BatchNormalization)	(None, 15999, 64)	256
conv1d_1 (Conv1D)	(None, 15998, 64)	8256
max_pooling1d_1 (MaxPooling1D)	(None, 7999, 64)	0
dropout (Dropout)_1	(None, 7999, 64)	0
batch_normalization_1 (BatchNormalization)	(None, 7999, 64)	256
conv1d_2 (Conv1D)	(None, 7998, 32)	4128
max_pooling1d_2 (MaxPooling1D)	(None, 3999, 32)	0
dropout_2 (Dropout)	(None, 3999, 32)	0
batch_normalization_2 (BatchNormalization)	(None, 3999, 32)	128
conv1d_3 (Conv1D)	(None, 3998, 32)	2080
max_pooling1d_3 (MaxPooling1D)	(None, 1999, 32)	0
dropout_3 (Dropout)	(None, 1999, 32)	0
batch_normalization_3 (BatchNormalization)	(None, 1999, 32)	128
flatten (Flatten)	(None, 63968)	0
dense (Dense)	(None, 1024)	65504256
dropout_4 (Dropout)	(None, 1024)	0
dense_1 (Dense)	(None, 1024)	1049600
dropout_5 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 512)	524800
dropout_6 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 512)	262656
dropout_7 (Dropout)	(None, 512)	0
dense_4 (Dense)	(None, 2)	1026
Total params: 67,357,762 Trainable params: 67,357,378 Non-trainable params: 384		

Above, Table I captures the classification model's architecture. The optimizer used was the SparseCategoricalCrossentropy optimizer, the batch size was kept at default at 32, and 20 total epochs were run. The model consists of four combinations of a one dimensional convolutional layer, followed by a one dimensional max pooling layer, followed by a dropout layer, followed by a normalization layer. Then, after a flattening layer, there are 4 combinations of a densely connected layer followed by a dropout layer. The final layer contains two neurons, one of which represents the probability that the input ECG sonification is arrhythmic and the other which represents the probability that the input ECG sonification is a normal sinus rhythm, and the two probabilities add up to be one. All the activation functions used were relu except for the last one which used sigmoid. All the convolutional layers used a kernal size of 2. The first two convolutional layers had 64 parameters whereas the last two had 32 parameters. All the max pooling layers had pool sizes of 2. The dropout layers following the max pooling layers had values of .2. Then, batch normalization layers were applied to regularize the previous layer's outputs to have a mean close to 0 and a standard deviation close to 1. The first two densely connected layers had 1024 parameters while the next two had 512 parameters. The dropout layers following the densely connected layers had dropout rates of .5. The optimizer used for back-propagation was the adam optimizer. In total, there were 67,357,762 parameters, out of which 67,357,378 are trainable.

The model ran a total of 10 iterations with stratified k-fold cross validation, with k = 5. The purpose of utilizing stratified k-fold cross validation over k-fold cross validation was to ensure as close to a 50:50 ratio between classes within both the training and testing sets of data.

III. Results

As for results, there is denoising of ECG data, amplification of PRT waves in ECGs, sonification of ECGs, and classification of arrhythmia through a CNN.

A. Metrics

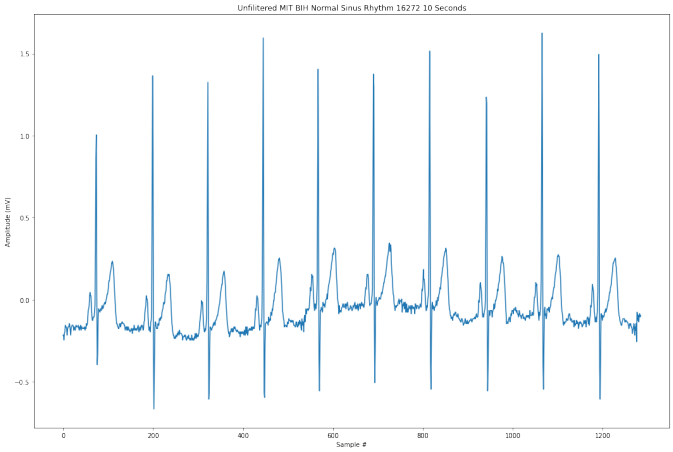
1) Amplification of PRT Waves: To measure the accuracy of the amplification of the PRT waves, the percent of correctly detected R peaks divided by the total number of R peaks annotated is calculated.

$$R_{accuracy} = \frac{Total_{Detected\ R\ Peaks}}{Total_{R\ Peaks}}$$

2) Classification of Arrhythmia: To measure the accuracy of the classification CNN, validation accuracy and loss were used for analysis.

B. Metrics Results and Graphs

1) Denoising of ECG Data: As shown in Figures 1, 2, and 3, the denoising of ECG data results in a smoother and cleaner ECG. The self collected ECG in Fig. 3 had the most noticeable difference after filtering, due to the hardware limitations and preprocessing done within the MIT BIH databases.



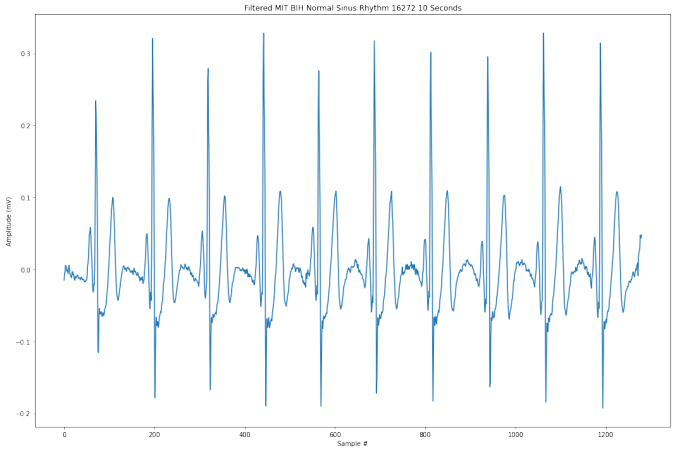


Fig. 1. The unfiltered MIT BIH Normal Sinus Rhythm patient 16272 first 10 seconds ECG (top) and the filtered MIT BIH Normal Sinus Rhythm patient 16272 first 10 seconds ECG (bottom)

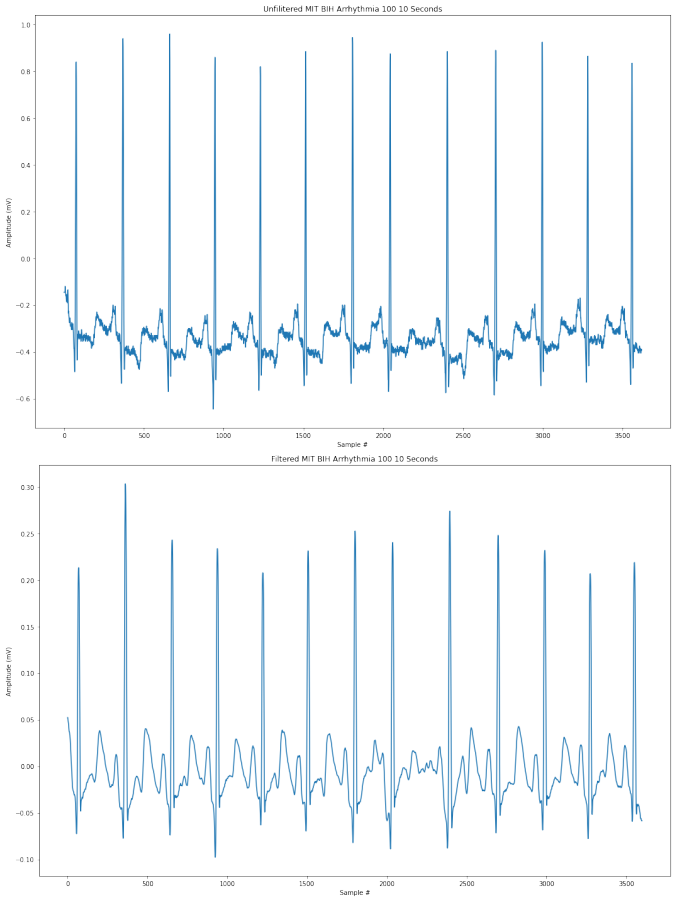


Fig. 2. The unfiltered MIT BIH Arrhythmia patient 100 first 10 seconds ECG (top) and the filtered MIT BIH Arrhythmia patient 100 first 10 seconds ECG (bottom)

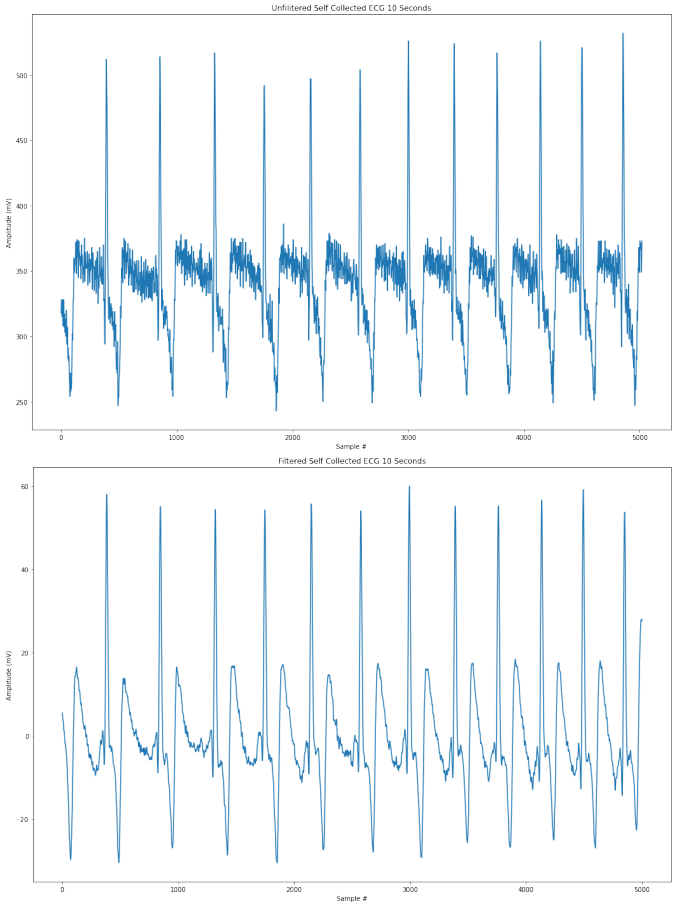


Fig. 3. The unfiltered self collected ECG data first 10 seconds ECG (top) and the filtered self collected ECG data first 10 seconds ECG (bottom)

2) Amplification of PRT Waves: As seen in fig. 4, the PRT Waves have been amplified. This results in easier detection both visually and sonically in ECGs for the PRT Waves.

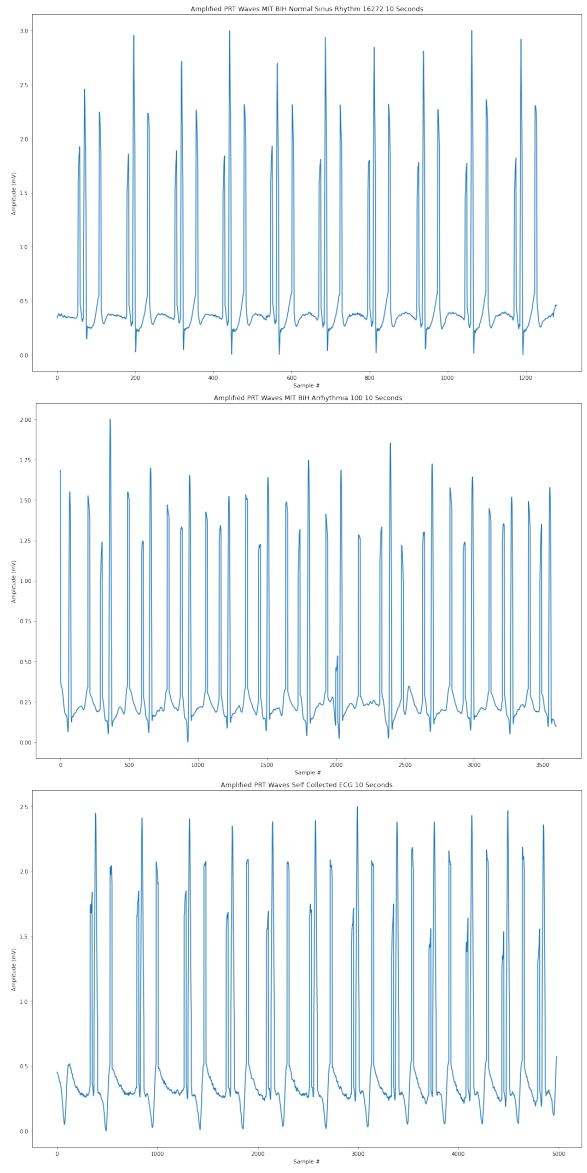


Fig. 4. The PRT amplified MIT BIH Normal Sinus Rhythm Patient 16272 10 Seconds (top), the PRT amplified MIT BIH Arrhythmia Patient 100 10 Seconds (middle), and the PRT amplified self collected ECG 10 seconds (bottom)

As seen in fig. 5, the R Peak detection accuracy for the MIT BIH Normal Sinus Rhythm of 18 Samples had a median of 97.62% and a mean of 92.10%. The R Peak detection accuracy for the MIT BIH Arrhythia of 23 Samples had a median of 96.62% and a mean of 90.17%. Potential reasons for outlier accuracy could be high variability in the R Peak amplitude especially in a long ECG sample as well as an incorrect lead that is not MLII being used.

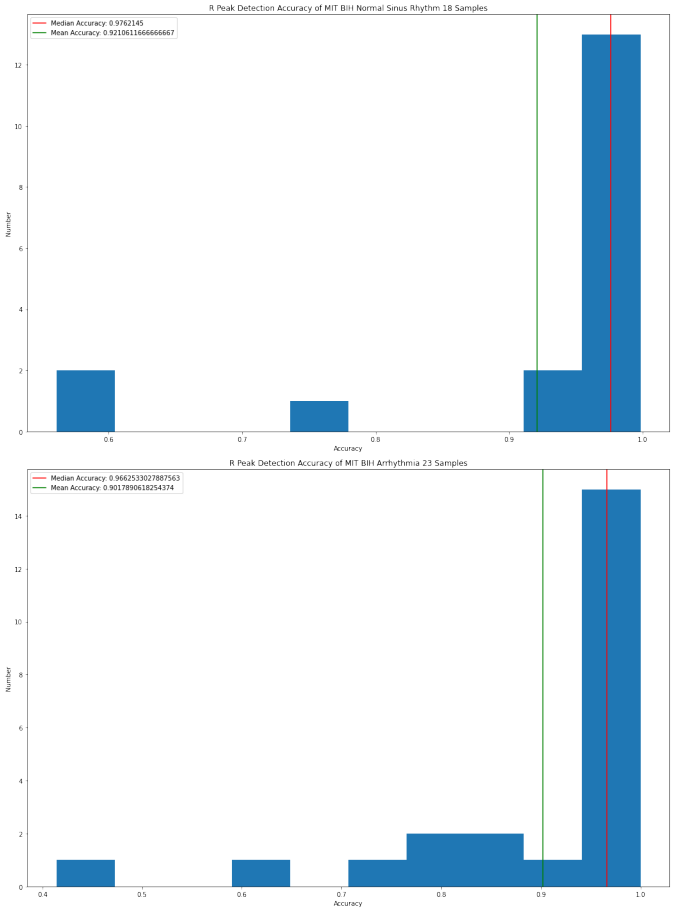


Fig. 5. The R Peak detection accuracy for the MIT BIH Normal Sinus Rhythm of 18 Samples (top) and MIT BIH Arrhythmia of 23 Samples (bottom)

3) ECG Sonification: As seen in fig. 6, the PRT Waves are noticeably visible and audible, allowing for better detection of heart irregularities in ECG sonification.

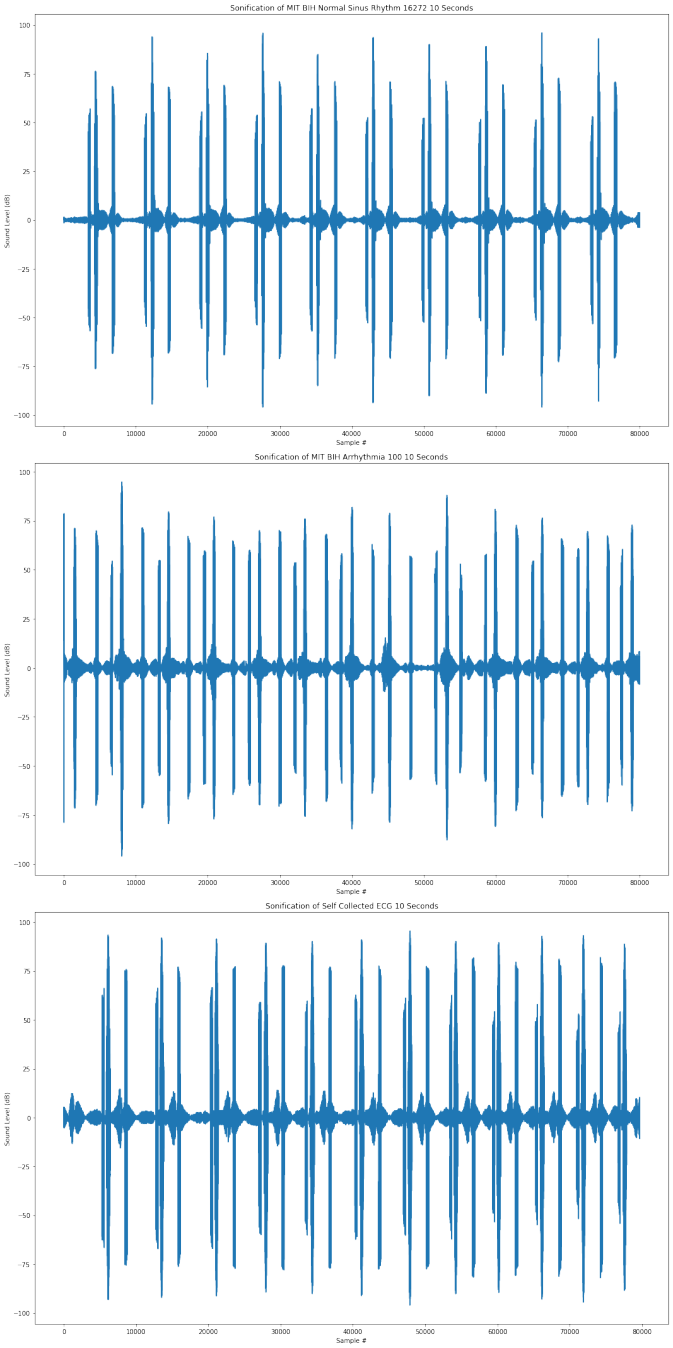


Fig. 6. ECG Sonification for 10 seconds for MIT BIH Normal Sinus Rhythm Patient 16272 (top), MIT BIH Arrhythmia Patient 100 (middle), and Self Collected ECG (bottom)

4) Arrhythmia Classification Model: Across 10 iterations of stratified 5-fold cross validation with 154 samples, the average validation accuracy was 99.28%.

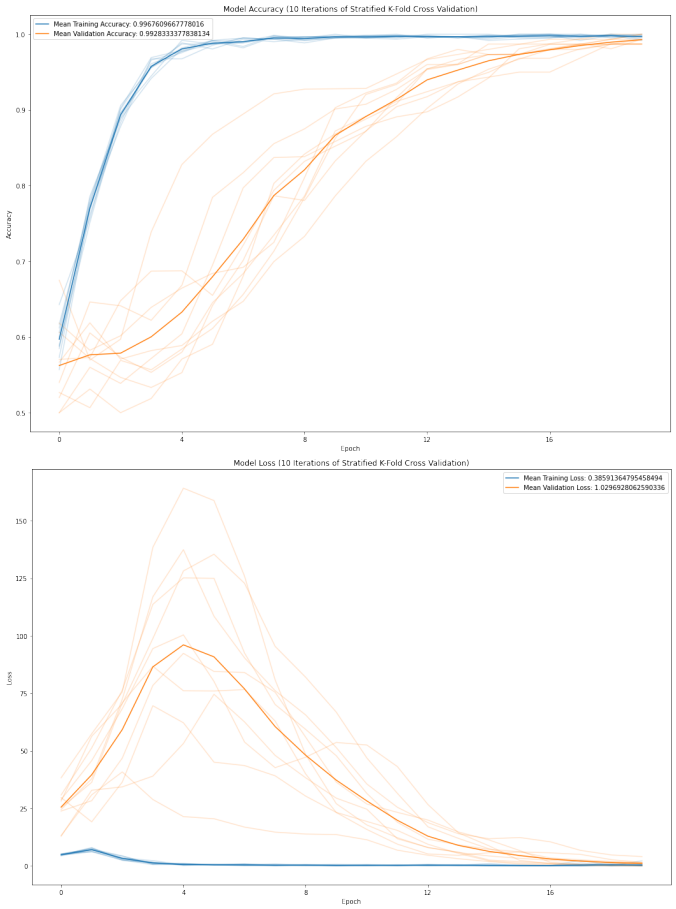


Fig. 7. CNN model accuracy (top) and CNN model loss (bottom) across 20 epochs with 10 iterations of stratified 5-fold cross validation

IV. Discussion

A. Comparison

In addition to mapping ECG features to sound amplitude and frequency, Andrea Lorena Aldana Blanco, Steffen Grautoff, and Thomas Hermann's polarity sonification method was also able to map ECG slope to audio panning between left and right channels as well as sound morphology through harmonics [14]. The addition of more methods of mapping ECG features allows for ECG sonification to contain and play more important features from ECG data.

The expansion of a classification model to include multiple classes instead of just 2 to detect multiple types of arrhythmia can allow for more precise classification of ECG heartbeats and irregularities. Mengze Wu, Yongdi Lu, Wenli Yang, and Shen Yuoung Wong, through a 12 layer CNN, were to achieve a 97.41% accuracy in classifying 5 different types of arrhythmia [18].

B. Heuristics for Irregularity Detection

A problem in detecting an irregular heart rhythm is that for one abnormal heartbeat, there can be much more than one normal sinus rhythm heartbeats. Applying a CNN on each heartbeat would be extremely time consuming. Instead, a

simple heuristic algorithm based on detecting heartbeats that are outliers in RR intervals and then applying a CNN can cut down on total run time and resources required to detect heart irregularities.

V. Conclusion

Through the sonification of ECGs, it is possible to identify heart irregularities. Through both publically available ECG datasets and self collected ECGs, it is possible to denoise ECGs and amplify key features such as the PRT waves. Then, mapping such key features from ECGs to sound is also possible. Machine learning can also be applied to sonified ECG signals to accurately detect heart irregularities. For future work, more attempts can be done on different denoising techniques such as a moving median filter and wavelet transforms and different combinations of such techniques. Detecting PQRST waves can also be improved such as with wavelet transforms, ECG decomposition by frequency, and a variable threshold for R peak detection. Applying a non-linear transformation to the ECG data to amplify key features such as PQRST waves is also possible. However, an exponential transformation amplifies the R waves too much whereas a logarithmic transformation amplifies the baseline signal and noise too much. A sigmoid transformation may allow for lower amplitude key features such as P, Q, S, and T waves to be amplified while also not amplifying the R wave too much. Additionally, utilizing more samples for the classification CNN model or also adding more classes or increasing the model's depth and fine tuning hyper-parameters may allow for better classification of heart irregularities. Such improvements can result in a better ECG sonification and classification model for better detecting heart irregularities.

References

[1] Centers for Disease Control and Prevention, National Center for Health Statistics. About Multiple Cause of Death, 1999–2020. CDC WONDER Online Database website. Atlanta, GA: Centers for Disease Control and Prevention; 2022. Accessed February 21, 2022.

[2] Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey (MEPS): household component summary tables: medical conditions, United States. Accessed April 8, 2021.

[3] "Atrial fibrillation," Centers for Disease Control and Prevention, 12-Jul-2022. [Online]. Available: https://www.cdc.gov/heartdisease/atrial_fibrillation.htm. [Accessed: 02-Aug-2022].

[4] Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. Circulation. 2006;114:199–225.

[5] Colilla S, Crow A, Petkun W, Singer DE, Simon T, Liu X. Estimates of current and future incidence and prevalence of atrial fibrillation in the U.S. adult population. Am J Cardiol. 2013;112:1142–1147. doi: 10.1016/j.amjcard.2013.05.063.

[6] Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics—2019 update: a report from the American Heart Association. Circulation. 2019;139(10):e56–528.

[7] Benjamin E, Levy D, Vaziri SM. Independent risk facors for atrial fibrillation in a population based cohort. JAMA. 1994;271(11):840-844.

[8] Tsao CW, Aday AW, Almarzooq ZI, Beaton AZ, Bittencourt MS, Boehme AK, et al. Heart Disease and Stroke Statistics—2022 Update: A Report From the

[9] American Heart Association. Circulation. 2022;145(8):e153–e639. Heart Rhythm Society. (2019). Complications from Atrial Fibrillation. Accessed May 9, 2019.

[10] A. E. A. N. J; "Cardiology explained," National Center for Biotechnology Information. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/20821845/. [Accessed: 02-Aug-2022].

[11] "Electrocardiogram (ECG)," NHS choices. [Online]. Available: https://www.nhs.uk/conditions/electrocardiogram/. [Accessed: 02-Aug-2022].

[12] Chatterjee, S., Thakur, R.S., Yadav, R.N., Gupta, L. and Raghuvanshi, D.K. (2020), Review of noise removal techniques in ECG signals. IET Signal Process., 14: 569-590. https://doi.org/10.1049/iet-spr.2020.0104

[13] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," in IEEE Transactions on Biomedical Engineering, vol. BME-32, no. 3, pp. 230-236, March 1985, doi: 10.1109/TBME.1985.325532.

[14] Aldana Blanco, A.L., Grautoff, S. & Hermann, T. ECG sonification to support the diagnosis and monitoring of myocardial infarction. J Multimodal User Interfaces 14, 207–218 (2020). https://doi.org/10.1007/s12193-020-00319-x

[15] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Eng in Med and Biol 20(3):45-50 (May-June 2001). (PMID: 11446209)

[16] Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P. C., Mark, R., ... Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation [Online]. 101 (23), pp. e215–e220.

[17] Chawla, N. V., Bowyer, K. W., Hall, L. O., Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research, 16, 321–357.

[18] Wu M, Lu Y, Yang W, Wong SY. A Study on Arrhythmia via ECG Signal Classification Using the Convolutional Neural Network. Front Comput Neurosci. 2021 Jan 5;14:564015. doi: 10.3389/fncom.2020.564015. PMID: 33469423; PMCID: PMC7813686.