* Keywords: #📋Projects #[[Science Fair]] Due Date: [[January 31st, 2021]] Goal: Introduce more classes; Realtime recognition + app; noisy environments
* :hiccup [:hr {:style {:borderColor “#bfccd6” :margin “0.75em 0em 0”}}]
* [[TODOs]] #[[📋 Projects]]
  + {{[[DONE]]}} how to do prepcorssing
    - what are the common pcg sampling freq
    - how to do sgementation
    - how to do specrogram, mffc, an mel
  + {{[[TODO]]}} compare pascel dataset a and b metrics
  + {{[[TODO]]}} research audio pre processing teq for reamoval of noise
    - just make a gan lol and feed it noise+pcg and label is pcg lmao
  + {{[[DONE]]}} for speach recognziton (speach to speach)
  + {{[[DONE]]}} learn latex
  + {{[[DONE]]}} look at papers others have refrenced
  + {{[[TODO]]}} look at dangers of ECG authomus recongtiong
  + {{[[DONE]]}} look at exitsing ECG papers
  + {{[[DONE]]}} look at audio processing for speech regonztion
  + {{[[DONE]]}} Study each arrhythmias
  + {{[[DONE]]}} create due date outliner
  + {{[[DONE]]}} motivation for decisions
  + {{[[DONE]]}} GAN vs LSTM vs CNN for time series classification for ECG
    - {{[[TODO]]}} time series vs time frequency(specogram )
  + {{[[DONE]]}} GAN vs LSTM vs CNN for time series classification for PCG
  + {{[[DONE]]}} without and wihout ecg to pcg dataset that trained classifcaito model
  + {{[[DONE]]}} look at papers others have refrenced
  + {{[[DONE]]}} look at exitsing ECG papers
  + {{[[DONE]]}} look at audio processing for speech regonztion
  + {{[[TODO]]}} visualize dense layers like [this](https://playground.tensorflow.org/)
  + {{[[DONE]]}} show effect of each data processing tequinche
  + {{[[DONE]]}} get how long its going to take for printing
  + {{[[DONE]]}} how does audio processing remove unwanted artifacts
    - {{[[TODO]]}} get in contact with audio processing reseracher
  + {{[[DONE]]}} time/space complexity of models (for real-time recognition)
  + {{[[DONE]]}} synthic ecg and pcg signals
  + {{[[TODO]]}} Problem: optimize cross entropy for classifier/discriminator
  + {{[[DONE]]}} frequecy of phome mic and sounds produced by heart

# {{[[DONE]]}} look at ISEF projects

* + - 2nd place
      * why this reaserach is imoprnat?
        + future health care will be automated
  + {{[[DONE]]}} IDEAS
    - LSTM for sequence to sequence for noise redecution
      * adding in intengintal noise Sqe GAN
    - LSTM for locating
    - Testing
      * each test
        + time to train
        + time to test
        + number of tests
        + time and space efficacy
      * time raw
      * time denoised
      * time MFFC
      * local raw
      * local denoise
      * local MFCC
  + {{[[DONE]]}} More IDEAS
  + {{[[DONE]]}} make website for deployment Due: [[February 25th, 2021]]
  + {{[[DONE]]}} Add dataset examples and localization Due: [[February 24th, 2021]]
  + {{[[DONE]]}} Metrics of dataset Due: [[February 24th, 2021]]
  + {{[[DONE]]}} Trifold Finished Reviewed by Le: [[February 24th, 2021]] Submitted: [[February 25th, 2021]] by 12:00am Printed: [[February 26th, 2021]]
  + {{[[TODO]]}} Same domain AE
    - use same encoder/decoder for PCG and ECG
  + {{[[TODO]]}} Bio motivation from human translation
* :hiccup [:hr {:style {:borderColor “#bfccd6” :margin “0.75em 0em 0”}}]
* Meeting Notes
  + [[January 20th, 2021]]
    - add notes from tabs and TODOs
  + [[January 25th, 2021]]
    - Deadline Outline Due: [[February 28th, 2021]]
      * Performance on time series data Due: [[February 3rd, 2021]]
        + LSTM & CNN
        + LSTM & CNN with GAN
        + If time DNN baseline
        + Input type

1d (time series)

2d (spectrogram)

* + - * Segmentation Performance Due: [[February 10th, 2021]]
        + Best model (from LSTM, CNN, GAN) gets segmented data inputted into it
      * ECG to PCG Due: [[February 20th, 2021]]
      * Making Website Due: [[February 24th, 2021]]
      * Making video PJAS Due: [[February 17th, 2021]] CASEF Due: [[February 28th, 2021]]
    - Reasons for structures
      * LSTM
        + Time Series data
      * CNN
        + Feature extraction
      * GAN
        + Emphasize meaningful & deep features
      * Segmentation
        + Positional-based feature extraction

s1 and s2 state are relatively in the same time series position

* + - * ECG database
        + more classes/pathologies
  + [[February 3rd, 2021]]
    - Clarity
    - Proof
      * Compare by implementing techniques from other papers
        + implemented [examples](https://chrisalbon.com/)
        + SVMs
        + XG Boost for feature importance
    - Segmentation Due: [[February 4th, 2021]]
  + [[February 17th, 2021]]
    - Order of implantation
      * Preproessing
        + Denoising using GAN

[Denoising Time Series Data Using Asymmetric Generative Adversarial Networks](https://sci-hub.se/https://link.springer.com/chapter/10.1007/978-3-319-93040-4_23)

* + - * + Resmapleing
        + GAN for augmentation
      * ECG to PCG Due: [[February 24th, 2021]]
        + Stacked BiLSTM

Didn’t work (colab autoencoder)

* + - * + Transformer

raw

specogram (torchaudio?)

* + - * + Transformer + Discriminator
        + Autoencoder + Discriminator
        + (Autoencoder’s Latent Representation + Discriminator) + Transformer
      * Localization Due: [[February 20th, 2021]]
        + implement pervious methods under PCG Segmentation Papers
      * Classification
        + GAN for abnormal and normal

then implement artifact

Conditional GANS

[Colab](https://colab.research.google.com/drive/1VxEGx_G4nuSoeAzNqLFRuRy84yGbX9O3?usp=sharing)

[Paper implamented with ECG](https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8771369)

* + - * + For each dataset test and compare state of the art
        + ^^Visualizations for features extracted^^
  + [[March 3rd, 2021]]
    - Journals - Springer
      * Bioinformatics Due:
      * Journal of Big Data
        + [Manuscript](https://www.editorialmanager.com/bigd/default.aspx)
      * Journal of Signal Processing
        + [Manuscript](https://www.editorialmanager.com/jsig/default.aspx)
    - Conferences (Anonymous)
      * Big Data
    - arXiv
      * used for archiving
    - Find mentor
      * Penn State
      * CMU
* :hiccup [:hr {:style {:borderColor “#bfccd6” :margin “0.75em 0em 0”}}]
* Research Due Date: [[January 16th, 2021]]
  + [[Metrics]] [[source](http://www.damienfrancois.be/blog/files/modelperfcheatsheet.pdf)]
    - **Confusion Matrix** [[source](https://en.wikipedia.org/wiki/Confusion_matrix)]
      * Classification Accuracy: accuracy of classes
        + (TP + TN) / (TP + TN + FP + FN)
    - **Precision**: Predicted positive values which are actually positive
      * TP / (TP + FP)
    - **Recall/Sensitivity/True Positive Rate**: Actual positives which are predicted to be positive
      * TP / (TP + FN)
    - **Specificity/True Negative Rate**: Proportion of actual negatives which are predicted negatives
      * TN / (TN + FP)
    - Positive Likelihood: Probability a predicted positive value is an actual positive
      * sensitivity / (1-specificity)
    - Negative Likelihood: Probability a predicted negative is an ancutal negative
      * (1-sensitivity)/specificity
    - Balanced Classification Rate (BCR): Classification adjusted for classes
      * ½ (TP / (TP + FN) + TN / (TN + FP))
    - Balanced Error Rate (BER):
      * **IDK**
    - Half Total Error Rate: **IDK**
      * 1 - BCR or 2 (precision . recall) / (precision + recall)
    - Fβ: measure weighted harmonic mean between precision and recall
      * (1+β)2 TP / ((1+β)2 TP + β2 FN + FP)
    - Youden’s index: arithmetic mean
      * sensitivity - (1 - specificity)
    - Matthews correlation: correlation between the actual and predicted
      * (TP . TN – FP . FN) / ((TP+FP) (TP+FN) (TP + FP) (TN+FN)) ^ (1/2)
    - **ROC** curve: **IDK**
      * TP / FP
    - **AUC**: area under ROC curve
    - **ROC AUC Scores**
    - DICE
      * Reconstruction task
    - MSE
      * Reconstruction task
    - t-statistic [[source](https://en.wikipedia.org/wiki/T-statistic)]
      * The ratio of the departure of the estimated value of a parameter from its hypothesized value to its standard error. It is used in hypothesis testing via Student’s t-test. The t-statistic is used in a t-test to determine if you should support or reject the null hypothesis.
    - F-score [[source](https://en.wikipedia.org/wiki/F-score)]
      * F-score or F-measure is a measure of a test’s accuracy. It is calculated from the precision and recall of the test, where the precision is the number of correctly identified positive results divided by the number of all positive results, including those not identified correctly, and the recall is the number of correctly identified positive results divided by the number of all samples that should have been identified as positive.
    - F-test [[source](https://en.wikipedia.org/wiki/F-test)]
      * Statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled.
    - z-score
      * used for distribution of models
    - Analysis of variance (ANOVA) [[source](https://en.wikipedia.org/wiki/Analysis_of_variance)]
      * is a collection of statistical models and their associated estimation procedures (such as the “variation” among and between groups) used to analyze the differences among group means in a sample. ANOVA was developed by the statistician Ronald Fisher. The ANOVA is based on the law of total variance, where the observed variance in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form, ANOVA provides a statistical test of whether two or more population means are equal, and therefore generalizes the t-test beyond two means.
    - **Accuracy**
    - **PPV**
      * TP/(TP+FP)
    - **classification rate**
  + [[Visualization]]
    - [[t-SNE]] [[source](https://en.wikipedia.org/wiki/T-distributed_stochastic_neighbor_embedding)]
    - [[Class activation maps]] [[source](http://cnnlocalization.csail.mit.edu/)]
    - [[Z-score]] of all model metrics
    - GAN Dissection: Visualizing and Understanding Generative Adversarial Networks [[source](https://gandissect.csail.mit.edu/)] [[video](https://www.youtube.com/watch?v=yVCgUYe4JTM)]
    - Dense Layer [[source](https://playground.tensorflow.org/)]
  + Training Process
    - [[Bootstrap aggregating]] [[source](https://en.wikipedia.org/wiki/Bootstrap_aggregating)]
    - [[CN2 algorithm]] [[source](https://en.wikipedia.org/wiki/CN2_algorithm)]
      * dealing with imperfect data
    - [[Diffusion map]] [[source](https://en.wikipedia.org/wiki/Diffusion_map)]
    - [[Lasso]] [[source](https://en.wikipedia.org/wiki/Lasso_statistics)]
    - [[Out-of-bag error]] [[source](https://en.wikipedia.org/wiki/Out-of-bag_error)]
  + Papers
    - Autoencoders and Generative Adversarial Networks for Imbalanced Sequence Classification [[source](https://arxiv.org/pdf/1901.02514.pdf)]
    - Autoencoders for Data Augmentation [[source](https://quantdare.com/variational-autoencoder-as-a-method-of-data-augmentation/)]
    - Data Mining Practical Applications (BOOK) [[source](https://tinyurl.com/y49yh2jg)]
    - Learning Deep Features for Discriminative Localization [[source](http://cnnlocalization.csail.mit.edu/Zhou_Learning_Deep_Features_CVPR_2016_paper.pdf)]
    - Visualizing and Understanding Convolutional Networks [[source](https://arxiv.org/pdf/1311.2901.pdf)]
    - Human Attention in Visual Question Answering: Do Humans and Deep Networks Look at the Same Regions? [[source](https://arxiv.org/pdf/1606.03556.pdf)]
    - SmoothGrad: removing noise by adding noise [[source](https://arxiv.org/pdf/1706.03825.pdf)]
    - GAN Dissection: Visualizing and Understanding Generative Adversarial Networks [[source](https://arxiv.org/pdf/1811.10597.pdf)]
    - Localization and classification of heart beats in phonocardiography signals —a comprehensive review [[source](https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0545-9)]
      * [[TODOs]] #[[📋 Projects]]
        + {{[[DONE]]}} PASCAL 2011 dataset (very nosicy and could prove good)
        + {{[[DONE]]}} Real time classification
        + {{[[DONE]]}} Review all datasets used (from tables)
        + {{[[TODO]]}} import all [figures and tables](https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0545-9)
      * [[Annotations]]
        + {{[[TODO]]}} Combine v1 and v2
        + v1

**Extracted Annotations (1/12/2021, 5:13:47 PM)**

“During the pumping action, electrical and mechanical activities are carried out resulting in the flow of blood.” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

^^“According to surveys conducted by the World Health Organization (WHO), 33% of all deaths are the result of CVDs” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))^^

“Different modalities are known to exist to monitor the health of heart. The most popular of these is the electrocardiogram or ECG” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“Another important modality is photoplethysmography (PPG) that employs light-based sensors to estimate the rate of flow of blood by measuring the changes in the reflected/transmitted light” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“phonocardiogram (PCG), the recording of the sounds and murmurs made by heart during a cardiac cycle, can be effectively employed to study and monitor the activities of heart.” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“The mechanical activity of heart due to its physical movement produces four distinct sounds or beats which are named as first normal heart sound (S1), second normal heart sound (S2), third abnormal heart sound (S3), and fourth abnormal heart sound (S4). S1 and S2 are normal sounds while S3, S4, murmurs, and certain other sounds usually refer to some disease or anomaly” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“S1 is usually low frequency and high-amplitude signal while S2 has high frequency and low amplitude.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“ECG, PPG, and PCG, all represent a cyclostationary signal, i.e., a signal in which the statistics of the signal vary but are repetitive with a period” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“ECG and PCG are highly correlated signals [8] and are known to contain more information than the PPG signal” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

^^“PCG, however, enjoys a distinct advantage over ECG and PPG signals as it records the acoustic properties of the signal. These properties are better suited for murmur detection which represents abnormal heart sound [9].” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

“Furthermore, PCG signal also has an excellent starting trigger in the form of S1 wave [10, 11].” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“The mechanical activity of the heart can be heard using a traditional or an electronic stethoscope.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“The time duration from S1 to S2 is known as systole and the one between S2 and S1 is known as diastole.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

^^“The automatic analysis of these signals involves two primary challenges, localization and classification. In localization, the objective is to correctly detect the positioning of beats in the signal while classification deals with the categorizing the beats into S1 and S2 in case of normal heart sounds and into S3, S4, and murmurs etc. in case of abnormal sounds.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

^^“Localization algorithms are generally peak-based algorithms where candidate peaks are localized in a signal and windows are constructed around these potential peaks.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

^^“Features are then calculated from each window, and the peaks are classified.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

“Heart beat is a repetitive or cyclic process (Fig. 2) and a normal PCG signal can be divided into four segments. These include the time window containing S1, the silence window, the segment containing S2, and again the silence window.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))

“However, since the sequence and duration of each state within the cycle can vary, probabilistic modeling is better suited for such scenarios as opposed to deterministic approaches.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))

**Deep Learning probabilisitc modeling (**[**note on p.28**](zotero://open-pdf/library/items/RSE7C8IF?page=3)**)**

^^“As discussed earlier, the present study is organized as function of time, frequency, and time-frequency methods. In time domain, the signal is sampled on time axis (Fig. 4a) and features (like amplitude, mean, and energy) are directly computed from signal. In spectral or frequency domain, signal spectra is divided into various spectral bands (Fig. 4b) and features (for example, spectral flux) are calculated from these bands. Unlike time and frequency methods, the time-frequency methods rely on simultaneous sampling on time and frequency axis. The general representations of time-frequency domain for fixed and variable time/frequency windows are illustrated in Fig. 4c, d, respectively.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))^^

^^“specially, when dealing with noisy signals. Authors highlighted the use of machine learning tools for better classification and suggested that the use of hybrid classifiers serves to enhance the performance. Nabih et al. [22] review the denoising, segmentation, and classification techniques for automated analysis of PCG signals. The key focus of their study is the comparison of different classifiers (artificial neural network (ANN), support vector machine (SVM), self-organizing map (SOM), and hybrid classifiers with the conclusion that while each of the processing steps is important, the choice of classifier is the most critical parameter in enhancing the overall system performance. The initial developments and findings in digital processing of PCG signals have been summarized in works by [18, 19]. In a relatively recent survey, Meziani et al. [20] discussed computerized analysis of PCG signals but limit their discussion to wavelet transforms-based methods only. In another short review [21], authors present a comparative study of EMD and wavelet-based methods for analysis of PCG signals. The study concludes that EMD is better suited for PCG as compared to wavelet, Fig. 3 HMM modeling of human heart cycle” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))^^

“PASCAL HSC dataset was first made available in 2011 for two challenges, segmentation (localization), and classification. The database is divided into two subsets, dataset ‘A’ and dataset ‘B’. Dataset ‘A’ has been collected using the iStethoscope Pro iPhone app while dataset ‘B’ has been gathered in a clinical setting using the digital stethoscope, DigiScope. The two datasets have 176 and 656 total auscultations, respectively. Files in both the datasets contain normal heart beats, murmurs, and extra systoles. In addition to these, dataset ‘B’ also contains files with artifacts and extra sounds and is more challenging.” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“E-General Medical is a well-known vendor of medical equipment and datasets. Among other datasets, a part of the dataset with heart sounds has been freely made available by E-General Medical. The dataset comprises 64 recordings with normal beats, S3, S4, and different pathologies. Majority of pathologies are systolic and diastolic in nature. In many cases, only one sample per signal type is available. Examples of such one sample per signal type include late systolic, early systolic, normal split, open snap, pan systolic, late systolic aortic stenosis, severe systolic aortic stenosis, critical systolic aortic stenosis, and systolic aortic valve replacement. Other sounds like diastolic-fixed S2 split, diastolic wide S2 split, systolic mitral regurgitation, systolic mitral prolapse, systolic split, and diastolic-fixed S2 split have more than one sample. Complete details on the dataset can be found in [23].” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“Healthy individuals as well as those with different pathologies contributed to data collection. The sound classes in the challenge include ‘normal,’ ‘uncertain,’ and ‘abnormal.’ The training set labeled from ‘A’ to ‘E’ has a total of 3126 files while the validation set comprises 300 recordings. The durations of the files vary from 5 to 120 s. The test set has not been made publicly available and was only used to score the participants of challenge. The 2016 PhysioNet CinC database was also developed as a part of a challenge. Heart sounds in the dataset are gathered from clinical as well as non-clinical” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“Heart sound and Murmur Library, University of Michigan Health Systems, is another public dataset that comprises normal heart sounds, normal split S1, systolic click, S4 gallop, S3 gallop, single S2, split S2 persistent, split S2” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

^^“In time domain, analysis is carried out on the signal itself. Statistics like peak value, mean, mode, median, and peak to peak duration of signal are typically employed for analysis.” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))^^

“The general representation of a signal in time domain is presented in Eq. 1; f (t) represents the signal which is modeled by amplitude A, frequency ω, and phase θ. Each of these parameters as well as different statistics derived from these parameters have been used as features for localization and classification. f (t) Asin( t θ ); ω” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

^^“For frequency domain analysis, the signal is first converted from time to frequency domain using fast Fourier transform (FFT) elaborated in Eq. 2. ∞ ∞ 1 jwtdt; ejwtdw F( f (t)e f (t) = F(w) (2) ω) 2π Where f (t) is the signal and F(w) is the fast Fourier transform of f (t).” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))^^

“Bandpass filtering is usually employed in heart beat analysis to remove the low and high frequency noise from stethoscope signals” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“Since the frequency contents of various heart signals are different, the frequency spectra can be exploited for classification of heart beats. Finding the optimal set of filters (along with cutoff frequencies)” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“Time and frequency analysis have been widely investigated to localize and classify normal as well as abnormal heart beats like third and fourth heart sound, clicks and murmurs.” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

^^“time domain usually starts with pre-processing which generally comprises decimation followed by a low pass filter with a cut-off frequency of around 200 Hz. The decimation decreases the computational load on the analysis module. Low pass filtering is used as pre-processing step as heart beat has characteristics of low frequency signal.” ([Ismail et al 2018:31](zotero://open-pdf/library/items/RSE7C8IF?page=6))^^

^^“Normalization removes the amplitude variation for inter and intra signal classification. The signal is then converted to a form that is best suitable for localization and classification. Shannon energy envelop is most widely employed method investigated in most of the time-based techniques [16, 31, 34]. In short, researchers have mostly employed combinations of decimation, filtering, and Shannon energy for pre-processing in the time domain.” ([Ismail et al 2018:31](zotero://open-pdf/library/items/RSE7C8IF?page=6))^^

^^“MFCC is usually used in the frequency-based methods and is found to be effective in speech processing including speech recognition and speaker identification [36]. MFCC has been used in the feature extraction process in [36] and [32] for localization and classification. Likewise, MFCC-based features were used to train HMM classifiers in [16] and reported high classification rates.” ([Ismail et al 2018:32](zotero://open-pdf/library/items/RSE7C8IF?page=7))^^

^^“The general evaluation metrics include accuracy, specificity, sensitivity, and positive predictive value (PPV)” ([Ismail et al 2018:35](zotero://open-pdf/library/items/RSE7C8IF?page=10))^^

“Due to cyclostationary nature of PCG, a signal cannot be localized both in time and frequency due to uncertainty principle limitation. Consequently, time-frequency analysis of PCG signals is investigated in a number of studies. Such methods generally employ short-time frequency transform (STFT), time-frequency representations like Wigner-Ville, wavelet transform, and empirical mode decomposition (EMD).” ([Ismail et al 2018:35](zotero://open-pdf/library/items/RSE7C8IF?page=10))

“It is not uncommon to use both time and frequency scaling. The MP method converts the signal into time-frequency atoms with time and frequency parameters. The summation of atoms gives back the original signal with error. Making use of these parameters, the signal can be transformed. This transformation results in a modified version of the signal which cannot be directly compared with original signal. An inverse scaling is therefore used for comparison. The time-frequency representations (TFR) can be calculated from original, scaled, and inversed versions. The sum of Wigner distribution of all atoms of a signal is termed as MP-based Wigner distribution. MP-based Wigner distribution has the distinct advantage of automatic removal of cross terms over conventional Wigner distribution when applied to signal. Also, MP-based Wigner distribution gives much clear presentation than the spectrogram.” ([Ismail et al 2018:36](zotero://open-pdf/library/items/RSE7C8IF?page=11))

“Debbal and Bereksi-Reguig [47] compared timefrequency methods of STFT (short-time frequency transform), CWT (continuous wavelet transform), and Wigner-Ville (WV) distribution primarily for the analysis of second heart sound. Analysis was also carried out for normal and pathological cases of early stenosis aortic analysis, late aortic stenosis, and pulmonary stenosis, etc. Authors concluded that for second heart sound, WV suffers from cross-term limitation while STFT joins the split. CW outperforms the other two techniques both in terms of localization and split.” ([Ismail et al 2018:36](zotero://open-pdf/library/items/RSE7C8IF?page=11))

^^“Wavelet is a popular method of choice in many fields because of the flexibility it provides in terms of choice of kernel selection which in turn reflects different time and frequency resolutions. Wavelet transform divides the signal into wavelet and scale factor.” ([Ismail et al 2018:37](zotero://open-pdf/library/items/RSE7C8IF?page=12))^^

^^“EMD has been a popular choice for the time-frequency analysis in many fields. Unlike other time-frequency algorithms, EMD operates in time domain and operates directly on the signal. EMD was introduced by Huang et al. [78] along with 2D graphical representation known as Hilbert spectrum for non-linear and non-stationary time series analysis. The algorithm is based on the assumption that each data series is primarily composed of a finite set of simple oscillations which are AM/FM components called intrinsic mode functions (IMF) by sifting process [79, 80].” ([Ismail et al 2018:40](zotero://open-pdf/library/items/RSE7C8IF?page=15))^^

^^“Studying the quantitative performance of time, frequency, and time-frequency methods (presented in the respective tables), it can be seen that a direct comparison of these methods is difficult due to the different types of challenges offered by each dataset. A general observation is that for small datasets and clean signals, time, frequency, and time-frequency-based methods report similar accuracies. For larger datasets and noisy signals, however, the performance of time-frequency methods, especially wavelets, remains relatively stable once compared to other methods. The PASCAL 2011 dataset A, for instance, is considered to be a very challenging set of noisy signals. Only a limited number of studies [30, 52, 53, 62, 67, 77, 95] have been evaluated on this set and among these wavelets-based techniques report the highest accuracies.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))^^

^^“Despite these characteristics, the problem of computerized analysis of PCG signals still remains very challenging for noisy environments. A clear evidence is the performance of different systems on the noisy signals of the PASCAL 2011 datasets which offer a great margin for enhancement” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))^^

^^“It is also worth mentioning that the primary focus of most of the research on automatic analysis of PCG signals has been on enhancing the localization and classification accuracies. From the view point of practical applications, development of computationally efficient solutions which may work in real time is also a challenging problem that needs further exploration.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))^^

^^“In addition, the presently available PCG datasets comprise a limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings. This necessitates the development and labeling of a comprehensive dataset of PCG signals encompassing a variety of signals and covering all major pathologies. Considering the complexity of the problem, modeling the heart beats and various pathologies is likely to offer more robust solutions as opposed to the conventional techniques relying on localization and classification. To the best of authors’ knowledge, very limited efforts have been made in this direction.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))^^

* + - * + v2

**Extracted Annotations (1/19/2021, 7:02:58 PM)**

“During the pumping action, electrical and mechanical activities are carried out resulting in the flow of blood.” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“According to surveys conducted by the World Health Organization (WHO), 33% of all deaths are the result of CVDs” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“Different modalities are known to exist to monitor the health of heart. The most popular of these is the electrocardiogram or ECG” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“Another important modality is photoplethysmography (PPG) that employs light-based sensors to estimate the rate of flow of blood by measuring the changes in the reflected/transmitted light” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

“phonocardiogram (PCG), the recording of the sounds and murmurs made by heart during a cardiac cycle, can be effectively employed to study and monitor the activities of heart.” ([Ismail et al 2018:26](zotero://open-pdf/library/items/RSE7C8IF?page=1))

^^“The mechanical activity of heart due to its physical movement produces four distinct sounds or beats which are named as first normal heart sound (S1), second normal heart sound (S2), third abnormal heart sound (S3), and fourth abnormal heart sound (S4). S1 and S2 are normal sounds while S3, S4, murmurs, and certain other sounds usually refer to some disease or anomaly” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

“S1 is usually low frequency and high-amplitude signal while S2 has high frequency and low amplitude.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“ECG, PPG, and PCG, all represent a cyclostationary signal, i.e., a signal in which the statistics of the signal vary but are repetitive with a period” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“ECG and PCG are highly correlated signals [8] and are known to contain more information than the PPG signal” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“PCG, however, enjoys a distinct advantage over ECG and PPG signals as it records the acoustic properties of the signal. These properties are better suited for murmur detection which represents abnormal heart sound [9].” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

^^“Furthermore, PCG signal also has an excellent starting trigger in the form of S1 wave [10, 11].” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

“The mechanical activity of the heart can be heard using a traditional or an electronic stethoscope.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

^^“The time duration from S1 to S2 is known as systole and the one between S2 and S1 is known as diastole.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

^^“The automatic analysis of these signals involves two primary challenges, localization and classification. In localization, the objective is to correctly detect the positioning of beats in the signal while classification deals with the categorizing the beats into S1 and S2 in case of normal heart sounds and into S3, S4, and murmurs etc. in case of abnormal sounds.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))^^

“Localization algorithms are generally peak-based algorithms where candidate peaks are localized in a signal and windows are constructed around these potential peaks.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“Features are then calculated from each window, and the peaks are classified.” ([Ismail et al 2018:27](zotero://open-pdf/library/items/RSE7C8IF?page=2))

“Heart beat is a repetitive or cyclic process (Fig. 2) and a normal PCG signal can be divided into four segments. These include the time window containing S1, the silence window, the segment containing S2, and again the silence window.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))

“However, since the sequence and duration of each state within the cycle can vary, probabilistic modeling is better suited for such scenarios as opposed to deterministic approaches.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))

**Deep Learning probabilisitc modeling (**[**note on p.28**](zotero://open-pdf/library/items/RSE7C8IF?page=3)**)**

^^“As discussed earlier, the present study is organized as function of time, frequency, and time-frequency methods. In time domain, the signal is sampled on time axis (Fig. 4a) and features (like amplitude, mean, and energy) are directly computed from signal. In spectral or frequency domain, signal spectra is divided into various spectral bands (Fig. 4b) and features (for example, spectral flux) are calculated from these bands. Unlike time and frequency methods, the time-frequency methods rely on simultaneous sampling on time and frequency axis. The general representations of time-frequency domain for fixed and variable time/frequency windows are illustrated in Fig. 4c, d, respectively.” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))^^

“specially, when dealing with noisy signals. Authors highlighted the use of machine learning tools for better classification and suggested that the use of hybrid classifiers serves to enhance the performance. Nabih et al. [22] review the denoising, segmentation, and classification techniques for automated analysis of PCG signals. The key focus of their study is the comparison of different classifiers (artificial neural network (ANN), support vector machine (SVM), self-organizing map (SOM), and hybrid classifiers with the conclusion that while each of the processing steps is important, the choice of classifier is the most critical parameter in enhancing the overall system performance. The initial developments and findings in digital processing of PCG signals have been summarized in works by [18, 19]. In a relatively recent survey, Meziani et al. [20] discussed computerized analysis of PCG signals but limit their discussion to wavelet transforms-based methods only. In another short review [21], authors present a comparative study of EMD and wavelet-based methods for analysis of PCG signals. The study concludes that EMD is better suited for PCG as compared to wavelet, Fig. 3 HMM modeling of human heart cycle” ([Ismail et al 2018:28](zotero://open-pdf/library/items/RSE7C8IF?page=3))

“PASCAL HSC dataset was first made available in 2011 for two challenges, segmentation (localization), and classification. The database is divided into two subsets, dataset ‘A’ and dataset ‘B’. Dataset ‘A’ has been collected using the iStethoscope Pro iPhone app while dataset ‘B’ has been gathered in a clinical setting using the digital stethoscope, DigiScope. The two datasets have 176 and 656 total auscultations, respectively. Files in both the datasets contain normal heart beats, murmurs, and extra systoles. In addition to these, dataset ‘B’ also contains files with artifacts and extra sounds and is more challenging.” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“E-General Medical is a well-known vendor of medical equipment and datasets. Among other datasets, a part of the dataset with heart sounds has been freely made available by E-General Medical. The dataset comprises 64 recordings with normal beats, S3, S4, and different pathologies. Majority of pathologies are systolic and diastolic in nature. In many cases, only one sample per signal type is available. Examples of such one sample per signal type include late systolic, early systolic, normal split, open snap, pan systolic, late systolic aortic stenosis, severe systolic aortic stenosis, critical systolic aortic stenosis, and systolic aortic valve replacement. Other sounds like diastolic-fixed S2 split, diastolic wide S2 split, systolic mitral regurgitation, systolic mitral prolapse, systolic split, and diastolic-fixed S2 split have more than one sample. Complete details on the dataset can be found in [23].” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“sources. Healthy individuals as well as those with different pathologies contributed to data collection. The sound classes in the challenge include ‘normal,’ ‘uncertain,’ and ‘abnormal.’ The training set labeled from ‘A’ to ‘E’ has a total of 3126 files while the validation set comprises 300 recordings. The durations of the files vary from 5 to 120 s. The test set has not been made publicly available and was only used to score the participants of challenge. The 2016 PhysioNet CinC database was also developed as a part of a challenge. Heart sounds in the dataset are gathered from clinical as well as non-clinical” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“Heart sound and Murmur Library, University of Michigan Health Systems, is another public dataset that comprises normal heart sounds, normal split S1, systolic click, S4 gallop, S3 gallop, single S2, split S2 persistent, split S2” ([Ismail et al 2018:29](zotero://open-pdf/library/items/RSE7C8IF?page=4))

“In time domain, analysis is carried out on the signal itself. Statistics like peak value, mean, mode, median, and peak to peak duration of signal are typically employed for analysis.” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“The general representation of a signal in time domain is presented in Eq. 1; f (t) represents the signal which is modeled by amplitude A, frequency ω, and phase θ. Each of these parameters as well as different statistics derived from these parameters have been used as features for localization and classification. f (t) Asin( t θ ); ω” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“For frequency domain analysis, the signal is first converted from time to frequency domain using fast Fourier transform (FFT) elaborated in Eq. 2. ∞ ∞ 1 jwtdt; ejwtdw F( f (t)e f (t) = F(w) (2) ω) 2π Where f (t) is the signal and F(w) is the fast Fourier transform of f (t).” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“Bandpass filtering is usually employed in heart beat analysis to remove the low and high frequency noise from stethoscope signals” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“Since the frequency contents of various heart signals are different, the frequency spectra can be exploited for classification of heart beats. Finding the optimal set of filters (along with cutoff frequencies)” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“Time and frequency analysis have been widely investigated to localize and classify normal as well as abnormal heart beats like third and fourth heart sound, clicks and murmurs.” ([Ismail et al 2018:30](zotero://open-pdf/library/items/RSE7C8IF?page=5))

“time domain usually starts with pre-processing which generally comprises decimation followed by a low pass filter with a cut-off frequency of around 200 Hz. The decimation decreases the computational load on the analysis module. Low pass filtering is used as pre-processing step as heart beat has characteristics of low frequency signal.” ([Ismail et al 2018:31](zotero://open-pdf/library/items/RSE7C8IF?page=6))

“Normalization removes the amplitude variation for inter and intra signal classification. The signal is then converted to a form that is best suitable for localization and classification. Shannon energy envelop is most widely employed method investigated in most of the time-based techniques [16, 31, 34]. In short, researchers have mostly employed combinations of decimation, filtering, and Shannon energy for pre-processing in the time domain.” ([Ismail et al 2018:31](zotero://open-pdf/library/items/RSE7C8IF?page=6))

“Authors in [35], for instance, use support vector machine with a modified cuckoo search (MCS) optimizer using features extracted from linear predictive coding (LPC) for classification. Similarly, Chen et al. [36] investigated deep neural networks for recognition of heart sounds S1 and S2. In another study, Potes et al. [32] applied AdaBoost-abstain (modified version of conventional AdaBoost) and convolutional neural networks (CNN) to detect abnormal heart sounds.” ([Ismail et al 2018:32](zotero://open-pdf/library/items/RSE7C8IF?page=7))

“MFCC is usually used in the frequency-based methods and is found to be effective in speech processing including speech recognition and speaker identification [36]. MFCC has been used in the feature extraction process in [36] and [32] for localization and classification. Likewise, MFCC-based features were used to train HMM classifiers in [16] and reported high classification rates.” ([Ismail et al 2018:32](zotero://open-pdf/library/items/RSE7C8IF?page=7))

“The general evaluation metrics include accuracy, specificity, sensitivity, and positive predictive value (PPV)” ([Ismail et al 2018:35](zotero://open-pdf/library/items/RSE7C8IF?page=10))

“Due to cyclostationary nature of PCG, a signal cannot be localized both in time and frequency due to uncertainty principle limitation. Consequently, time-frequency analysis of PCG signals is investigated in a number of studies. Such methods generally employ short-time frequency transform (STFT), time-frequency representations like Wigner-Ville, wavelet transform, and empirical mode decomposition (EMD).” ([Ismail et al 2018:35](zotero://open-pdf/library/items/RSE7C8IF?page=10))

“It is not uncommon to use both time and frequency scaling. The MP method converts the signal into time-frequency atoms with time and frequency parameters. The summation of atoms gives back the original signal with error. Making use of these parameters, the signal can be transformed. This transformation results in a modified version of the signal which cannot be directly compared with original signal. An inverse scaling is therefore used for comparison. The time-frequency representations (TFR) can be calculated from original, scaled, and inversed versions. The sum of Wigner distribution of all atoms of a signal is termed as MP-based Wigner distribution. MP-based Wigner distribution has the distinct advantage of automatic removal of cross terms over conventional Wigner distribution when applied to signal. Also, MP-based Wigner distribution gives much clear presentation than the spectrogram.” ([Ismail et al 2018:36](zotero://open-pdf/library/items/RSE7C8IF?page=11))

“Debbal and Bereksi-Reguig [47] compared timefrequency methods of STFT (short-time frequency transform), CWT (continuous wavelet transform), and Wigner-Ville (WV) distribution primarily for the analysis of second heart sound. Analysis was also carried out for normal and pathological cases of early stenosis aortic analysis, late aortic stenosis, and pulmonary stenosis, etc. Authors concluded that for second heart sound, WV suffers from cross-term limitation while STFT joins the split. CW outperforms the other two techniques both in terms of localization and split.” ([Ismail et al 2018:36](zotero://open-pdf/library/items/RSE7C8IF?page=11))

“Wavelet is a popular method of choice in many fields because of the flexibility it provides in terms of choice of kernel selection which in turn reflects different time and frequency resolutions. Wavelet transform divides the signal into wavelet and scale factor.” ([Ismail et al 2018:37](zotero://open-pdf/library/items/RSE7C8IF?page=12))

“EMD has been a popular choice for the time-frequency analysis in many fields. Unlike other time-frequency algorithms, EMD operates in time domain and operates directly on the signal. EMD was introduced by Huang et al. [78] along with 2D graphical representation known as Hilbert spectrum for non-linear and non-stationary time series analysis. The algorithm is based on the assumption that each data series is primarily composed of a finite set of simple oscillations which are AM/FM components called intrinsic mode functions (IMF) by sifting process [79, 80].” ([Ismail et al 2018:40](zotero://open-pdf/library/items/RSE7C8IF?page=15))

“These methods need preprocessing steps to make the signal fit for subsequent steps of localization and classification. With the development of machine learning techniques, time domain methods enjoy a renewed interest which is expected to continue in the near future. This trend is very well demonstrated by the recent research worked presented in Table 2. Quantitative comparison of results also shows that time-based methods, in general, are able to localize and classify beats with more than 70% specificity.” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“The frequency-based methods for localization and classification usually start with MFCC which employs preemphasis, windowing and FFT to make the signal appropriate for classification and localization. An analysis of the quantitative results reported in Table 2 shows that on the average, the performance of frequency-based methods is more or less similar to that of time-based methods” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“The time-frequency methods investigated in the literature include wavelets, EMD, and TFR.” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“TFR methods exhibit high variation, mainly as a function of the database employed for evaluation.” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“It can be observed from Table 5 that in general, high accuracies are reported by EMD-based techniques.” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“wavelet-based techniques have been comprehensively studied and evaluated on wide variety of datasets including all three public datasets presented in this paper. These methods consistently report high localization and classification accuracies on multiple datasets.” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“time-frequency based methods like EMD and wavelets was their ability to represent the signal at multiple resolutions unlike time and frequency analysis” ([Ismail et al 2018:48](zotero://open-pdf/library/items/RSE7C8IF?page=23))

“Nevertheless, automatic analysis of noisy heart beats still remains a challenging problem. Time and frequency methods, in most cases, are dependent on machine learning algorithms to enhance the localization and classification performance.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“Features extracted in the time or frequency domain are typically fed to a learning algorithm and the choice of the learning algorithm also influences the overall system performance.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“The timefrequency methods (especially EMD and wavelets), on the other hand, are not too sensitive to the choice of the learning algorithm.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“Furthermore, time and frequency methods have been mostly limited to the two-class problem of classifying the heart beats as normal or abnormal while wavelets and EMD present more sophisticated solutions detecting not only the abnormality but also classifying the pathology.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“A general observation is that for small datasets and clean signals, time, frequency, and time-frequency-based methods report similar accuracies. For larger datasets and noisy signals, however, the performance of time-frequency methods, especially wavelets, remains relatively stable once compared to other methods. The PASCAL 2011 dataset A, for instance, is considered to be a very challenging set of noisy signals. Only a limited number of studies [30, 52, 53, 62, 67, 77, 95] have been evaluated on this set and among these wavelets-based techniques report the highest accuracies.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“EMD and wavelet-based techniques enjoy the advantage over other techniques in the sense that they decompose the signal at multiple resolutions hence removing the highfrequency noise and reducing the energy contribution from low frequencies. These noise handling capabilities make such techniques an attractive choice, especially, when dealing with noisy signals.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“Despite these characteristics, the problem of computerized analysis of PCG signals still remains very challenging for noisy environments. A clear evidence is the performance of different systems on the noisy signals of the PASCAL 2011 datasets which offer a great margin for enhancement” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“It is also worth mentioning that the primary focus of most of the research on automatic analysis of PCG signals has been on enhancing the localization and classification accuracies. From the view point of practical applications, development of computationally efficient solutions which may work in real time is also a challenging problem that needs further exploration.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

“In addition, the presently available PCG datasets comprise a limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings. This necessitates the development and labeling of a comprehensive dataset of PCG signals encompassing a variety of signals and covering all major pathologies. Considering the complexity of the problem, modeling the heart beats and various pathologies is likely to offer more robust solutions as opposed to the conventional techniques relying on localization and classification. To the best of authors’ knowledge, very limited efforts have been made in this direction.” ([Ismail et al 2018:49](zotero://open-pdf/library/items/RSE7C8IF?page=24))

* + - * [[Notable Considerations]]
        + Includes 3 tables of previous studies (split by [time, frequency](https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0545-9/tables/2), [time-frequency](https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0545-9/tables/3) and [wavelet](https://asp-eurasipjournals.springeropen.com/articles/10.1186/s13634-018-0545-9/tables/4)analysis)

Pre-processing techniques

Localization techniques

Classification techniques

Beat types

Data set

Signals statistics

Evaluation metrics

Results

* + - Ensemble of Feature-based and Deep learning-based Classifiers for Detection of Abnormal Heart Sounds [[source](https://www.researchgate.net/publication/311922658_Ensemble_of_Feature-based_and_Deep_learning-based_Classifiers_for_Detection_of_Abnormal_Heart_Sounds)]
    - S1 and S2 Heart Sound Recognition Using Deep Neural Networks [[source](https://pubmed.ncbi.nlm.nih.gov/28113191/)]
    - Phonocardiographic Sensing using Deep Learning for Abnormal Heartbeat Detection [[source](http://arxiv-export-lb.library.cornell.edu/pdf/1801.08322)]
    - An Open Access Database for the Evaluation of Heart Sound Algorithms [[source](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7199391/pdf/nihms-1005138.pdf)]
    - Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool [[source](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4533825/)]
    - Transformer-based Conditional Variational Autoencoder for Controllable Story Generation [[source](https://arxiv.org/pdf/2101.00828.pdf)]
      * can be used for VAE (used in taming transforms) instead of Conv we use transformers
  + Presentation Examples
    - Machine Learning and Microsoft HoloLens by Krithik Ramesh [[source](https://www.youtube.com/watch?v=1aQMbldvd10)]
  + Datasets
    - ECG
      * Cardiologist-labeled test dataset (from [this](https://www.nature.com/articles/s41591-018-0268-3)paper)[source]
      * AF Classification from a Short Single Lead ECG Recording [[source](https://physionet.org/content/challenge-2017/1.0.0/)]
      * PTB-XL, a large publicly available electrocardiography dataset [[source](https://physionet.org/content/ptb-xl/1.0.1/)]
    - ECG & PCG
      * Classification of Heart Sound Recordings [[source](https://physionet.org/content/challenge-2016/1.0.0/)]
    - PCG
      * Classifying Heart Sounds Challenge [[source](http://www.peterjbentley.com/heartchallenge/)]
      * Heart Sound & Murmur Library [[source](http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html)]
      * Cardiac Auscultation of Heart Murmurs (CD) [[source](https://web.archive.org/web/20200120035540/http://www.egeneralmedical.com/listohearmur.html)]
      * Littmann Library [[source](http://www.3m.com/healthcare/littmann/mmm-library.html)]
* Research Paper Outline
  + {{[[TODO]]}} Introduction
    - Topics
      * Statistic of CD
      * Problem
      * Explanation of Arrhythmias studies
      * Explanation of basic structure or ECG and PCG
      * Motivation
      * {{[[TODO]]}} Testing Constraints
      * Engineering Goals
    - 33% of all deaths are the result of Cardiovascular disease ([[Ismail et al]]).
    - Phonocardiograms (PCGs) are sounds that are created by the mechanical movement of the heart. This physical movement produces four distinct sounds: S1, S2, S3, S4, and murmurs. S1 and S2 are sounds created by an health heart; whereas, S3, S4 and murmurs refer to diseases or anomalies [source].
    - The first heat sound, S1, marks the start of Systole. Systole is occurs when heart muscle contracts and pumps blood from the chambers into the arteries. The second heart sound, S2, marks the end of Systole and the start of Diastole. Diastole is phase of the heartbeat when the heart muscle relaxes and allows the chambers to fill with blood [source].
    - Although heart sound databases do exists, these datasets are still limited by the number of pathologies that are collected, often having dividing the dataset into two categories: normal and abnormal. Currently, only three major supervised PCG datasets exist: PhysioNet Classification of Heart Sound Recording Challenge dataset [source], PASCAL Heart Sound Challenge dataset [source], and the Heart Sound and Murmur Library [source]. The presently available PCG datasets have limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings.
    - In diagnosing heart sounds, two major challenges arise: localization and classification. Localization aims to find the position of the aforementioned biomarkers in heart sounds. By doing this, heart sounds can be segmented into signals containing a single heart sound. Furthermore, classification attempts to categorize heart sounds into to normal and abnormal groups by exploiting the information extracted from localization [source].
    - Conventional heart sound localization and classification methods involve time, frequency, or both, and are typically dependent on machine learning algorithms to enhance the results. These algorithms typically include artificial neural networks (ANNs), support vector machines (SVMs), self-organizing maps (SOMs) [source], and are limited to the number of samples and pathologies covered in a given dataset. This leads to surface level analysis of the heart sounds.
    - The main challenge in effective heart sound detection stems from analysis of noisy heart beats, e.g., background noise. For clean datasets, e.g., the PhysioNet Challenge dataset, a varieties time and frequency of methods converged on an localization accuracy of 96.9% (Fernando et al.) and 96.0% classification accuracy (Mostafa et al.). For large datasets with noisy signals, e.g., the PASCAL Challenge dataset, the performance of time and frequency methods remained inconsistent at an localization accuracy of 93.3% (Singh et al.) and 93.3% classification accuracy (Singh et al.).
    - From the view point of practical applications, development of computationally efficient solutions are extremally important to the success of a model’s deployment. Many studies have negated to comment on the practically of their proposed methods. From our research, we have concluded only two studies have noted their time efficiency, (Fernando et al.) and (Messner et al.). The fastest model processed 1000 heart state classifications in 56.88 seconds (Fernando et al.), suggesting the model can processes 18 bps. The average heart rate of a human heart is around 60-100 bps; thus, current models need sever optimization to achieve near to real-time analysis. These results are excluding the classification of heart arrhythmias.
    - Thus the problem of computationally efficient and accurate localization/classification of noisy heart beats and datasets with a variety of pathologies still remains a problem.
      * (LATER) Later on add all points of motivation
        + Nevertheless, automatic analysis of noisy heart beats still remains a challenging problem.
        + From the view point of practical applications, development of computationally efficient solutions which may work in real time is also a challenging problem that needs further exploration.
        + The presently available PCG datasets have limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings.
  + {{[[TODO]]}} Pervious Studies
    - Topics
      * Pre-processing
      * Localization/Segmentation
      * Classification
      * Classes (Beat types)
      * Dataset
      * Metrics
      * Results
  + {{[[TODO]]}} Methods
    - Problem Formulation
      * Localization
        + Categorize beats into: S1/S2 (normal), S3/S4 (abnormal)
        + Techniques

Feature used

peak value

mean

mode

median

peak to peak duration

A = amplitude

= frequencey

= phase

Fast Fourier Transform (FFT)

= signal

= Fast Fourier Transform (FFT)

DNN [[36](https://sci-hub.se/https://ieeexplore.ieee.org/document/7460939)]

Cepstrum [[30](https://www.researchgate.net/publication/260549211_Heart_Sounds_Classification_using_Feature_Extraction_of_Phonocardiography_Signal)]

Wavelet Decomposition [[33](https://www.researchgate.net/publication/220218338_Detection_of_cardiac_abnormality_from_PCG_signal_using_LMS_based_least_square_SVM_classifier)]

Springer Algorithm [[15](https://sci-hub.se/https://ieeexplore.ieee.org/document/7234876)] [38] [40] [41]

= Wigner distribution

= signal

= conjugate of x

= time delay

=frequency.

For a mono component signal, the equation produces a frequency component that is representative of signal frequency. However, for multiple-component signal, this autocorrelation produces extra components known as ‘cross-terms.

As an example, we consider a two-component signal x(t)=x1(t)+x2(t). The autocorrelation of signal x(t) will be composed of terms x(t)∗x(t)=x1(t)∗x1(t)+x2(t)∗x2(t)+x1(t)∗x2(t).

pseudo Wigner-Ville distribution [[43](https://link.springer.com/article/10.1361/154770206X99343)]

= smoothing window.

The peak at 1000th sample represents S1, and the one at 2000th sample represents S2. All other structures represent noise resulting from cross terms due the TFR manipulation. The noise samples result from autocorrelation of multi-spectral signals.

Renyi Entropy Measure (RME) [[46](https://www.researchgate.net/publication/224258806_Segmentation_and_Identification_of_some_Pathological_Phonocardiogram_Signals_using_Time-Frequency_Analysis)]

Continuous Wavelet Transformation [[47](https://sci-hub.se/https://pubmed.ncbi.nlm.nih.gov/16266105/)]

Shannon envelope, ACF Haar Wavelet lifting [[48](https://sci-hub.se/https://pubmed.ncbi.nlm.nih.gov/24418438/)]

signal is first down sampled

coarse detection is performed using PWVD

fine detection using Haar wavelet lifting scheme

normalized average Shannon energy (NASE) algorithm

TSD

Hilbert Vibration Decomposition [[51](https://pubmed.ncbi.nlm.nih.gov/26584485/)]

**Wavelet Transformation** [54]

cj0 d j are coefficients while φj0,k and ψj,k are scaling and wavelet functions

**Segmentation on each phase using Wavelet** [[57](https://link.springer.com/article/10.1007/s10439-006-9187-4)] [[61](https://pubmed.ncbi.nlm.nih.gov/22956159/)]

**EMDand Hilbert-Huang transform** [[78](https://royalsocietypublishing.org/doi/10.1098/rspa.1998.0193)]

Segmentation using variation

Preprocessing

removal of mean and amplitude normalization

VMD Decomposition

Selection of mode

Shannon Entropy Envelope Extraction

Adaptive Thresholding

Segmentation

* + - * Classification
        + Featuers

time duration between S2 and S2

* + - * + Techniques

= signal

= Fast Fourier Transform (FFT)

Frequency Spectra

Reason: freq contects of various heart singal s are different

SVM [[35](https://sci-hub.se/https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0112673)] [[33](https://www.researchgate.net/publication/220218338_Detection_of_cardiac_abnormality_from_PCG_signal_using_LMS_based_least_square_SVM_classifier)]

AdaBoost-abstain [[32](https://archive.physionet.org/challenge/2016/papers/potes.pdf)]

Kurtosis [[40](https://arxiv.org/ftp/arxiv/papers/1810/1810.09253.pdf)]

* + - * + Multiple Arrthimya Can compare?

https://thesai.org/Downloads/Volume7No7/Paper\_40-Improving\_the\_Recognition\_of\_Heart\_Murmur.pdf

* + - Data Management\*\* \*\*
      * Data Collection
        + Topics Location of collection Subject Metadata - Age - Ethnicity Metadata of signals - Average Length - Number of Recordings - Frequency - Commonality of dataset (how many studies used it) Characteristics of Dataset - Noisy Number of classes Number of subjects Sensor
        + Content

PCG

Although PCG signals are analyzed less often than ECG signals, these signals are rather analyzed in real-time by physicians and healthcare workers. Preliminary studies done on PCG segmentation and classification primarily used private datasets. Hence, there existed no publicly available datasets until recently. Since then, many public datasets have been developed aiding researchers in their studies and creating open benchmarks for researchers to use in comparing similar findings. However, these datasets are still limited by the number of classes that are collected, when compared to ECG datasets.

Currently, only three major supervised PCG datasets exist: PhysioNet Classification of Heart Sound Recording Challenge dataset [source], PASCAL Heart Sound Challenge dataset [source], and the Heart Sound and Murmur Library [source]. These datasets are all anonymized and de-identified for the safety of their subjects, and thus includes no personal information such as name, income, age, etc.

The PhysioNet Classification of Heart Sound Recording Challenge dataset was produced as a part of the 2016 PhyisoNet Computing in Cardiology Challenge. The heart sounds were collected from both clinical and non-clinical environments (in-home visits). The challenge focused on creating an accurate dataset of normal and abnormal heart sound recordings, especially in real-world (extremely noisy and low signal quality) scenarios. These recordings were sourced from nine independent databases and in total, contain 4,593 heart sound recordings from 1072 subjects, lasting from 5-120 seconds. Of which, 409 recordings that were collected from 121 patients contain one PCG lead and one simultaneously recorded ECG. Though, all recordings were resampled to 2,000 Hz using an anti-alias filter. Furthermore, the dataset is comprised of 3 classes: normal, abnormal, and unsure (this is due to poor recording quality), and have the following proportion respectively: 77.1%, 12,0%, 10.9% [[source](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7199391/pdf/nihms-1005138.pdf)].

The PASCAL Classifying Heart Sounds Challenge dataset was released to the general public in 2011. The challenge consisted of two sub-challenges: heart sound segmentation, and heart sounds classification; these sub-challenges corresponded with dataset A, and dataset B respectively. Both datasets have recordings of of varying lengths, between 1 second and 30 seconds. Dataset A was collected via the iSethoscope Pro iPhone app, and contained 176 heart sound recording. 124 of which are divided into four classes: Normal (31 recordings), Murmur (34 recordings), Extra heart sound (19 recordings), and Artifact (40 recordings); the rest of the records are unlabeled for testing purposes. Dataset B was collected using a DigiScope (a digital stethoscope), and included 656 heart sounds. All expect 370 were separated into three classes: Normal (320 recordings), Murmur (95 recordings), and Extra-systole (46 recordings). Both datasets A and B vary in sound recordings between lengths of 1 second and 30 seconds. [[source](http://www.peterjbentley.com/heartchallenge/)]

The Heart sound and Murmur Library provided by the University of Michigan Health Systems, is a dataset that contains Normal, Mitral Valve Prolapse, Acute Mitral Regurgitation, Left Ventricular Hypertrophy, Dilated Cardiomyopathy with Mitral Regurgitation, Mitral Stenosis, Aortic Regurgitation, Combined Aortic Stenosis and Regurgitation, Arterial Septal Defect, and Pulmonary Valve Stenosis at the Apex, Aortic, and Pulmonic Area. [[source](http://www.med.umich.edu/lrc/psb_open/html/repo/primer_heartsound/primer_heartsound.html)]

ECG

More than 300 million ECG recordings are analyzed yearly [[source](https://pubmed.ncbi.nlm.nih.gov/10516892/)], and thus create an exceptional tool for arrhythmia classification. Coupled with the recent surge in research interest in 2015, many massive publicly available datasets have been published, notable by PhysioNet - the moniker of the Research Resource for Complex Physiologic Signals [[source](https://physionet.org/about/)]. Numerous, datasets ECG exit, however, many are limited to few classes (Normal and Abnormal). At present, three public datasets exist that have more than 4 classes: AF Classification Challenge 2017 [[source](https://physionet.org/content/challenge-2017/1.0.0/)], PTB Diagnostic ECG [[source](https://www.physionet.org/content/ptbdb/1.0.0/)], and PTB-XL dataset [[source](https://physionet.org/content/ptb-xl/1.0.1/)]. Additionally, iRhythm Technologies have developed a semi-public dataset, that is available upon request, that contains 12 classes [[source](https://www.nature.com/articles/s41591-018-0268-3)].

The PTB-XL is the largest publicly available dataset for ECGs and contains 21,837 clinical 12-lead ECG recordings from 18,885 patients of 10 second length. Theses recordings are separated into 5 super-classes: Normal, Myocardial Infraction, Hypertrophy, ST/T-Change, and Conduction Disturbance. These super-classes are further split into 71 sub-classes that range from AV Block to Posterior Myocardial Infraction. The raw signal data was downsampled to 100 Hz and annotated by up to two cardiologists, who assigned potentially multiple ECG statements to each record. [[source](https://physionet.org/content/ptb-xl/1.0.1/)]

iRhythm Technologies developed a large, 12 classes ECG dataset using raw single-lead ECG inputs. The 12 classes include Atrial fibrillation and flutter, AVB, Bigeminy, EAR, IVR, Junctional rhythm, Noise, Sinus rhythm, SVT, Trigeminy, Ventricular tachycardia, and Wenckebach. The dataset consists of 91,232 ECG recordings from 53,549 patients. This training dataset is available upon request under license from iRhythm Technologies, Inc. The test dataset, which is publicity available, contains 328 records collected from 328 unique patients, split between 6 classes. Both datasets were recorded using a Zio monitor, which monitors the heart through a single-lead sensor at 200 Hz. The annotation was done by a consensus committee of expert cardiologists. [[source](https://www.nature.com/articles/s41591-018-0268-3)]

The PhysioNet AF Classification database, presented in 2017 for the Computing in Cardiology Challenge, contains 8,528 ECG recordings, divided into 4 classes: Normal (5154 recordings), Atrial Fibrillation (771 recordings), Other arrhythmias (2557 recordings), and Noisy (46 recordings). The single-lead recordings last from 9 - 61 seconds, with an a mean of 32.5 seconds and a standard deviation of 10.9 seconds. The ECG recordings were sampled to 300 Hz and provided in MATLAB V4 WFDB-compliant format. [[source](https://physionet.org/content/challenge-2017/1.0.0/)]

* + - * Data Preprocessing
        + Topics

Training Outline (Results for each point) with and w/out segmentation/localization

Results Table

{{[[table]]}}

Unsegmented

Segmented with localization

Raw Signal

Denoising

Data Transformation

Denoising

Adding white gaussian noise

= original signal

= is the noise

= noise strength (standard deviation)

= time

Signal Decomposition

high freq part

passed through high-pass filter

low freq part

passed through low-pass filter

Applying threshold

Acts on single wavelet coefficient at a time

Hard thersholding

absolute value is less than threshold value, then set = 0

Soft thresholding

absolute value is less than threshold value, then set = 0

the remaining nonzero coefficients are shrunk towards zero

reconstruing denoise signal

reconstruction using wavelet or filter

Data Transformation

Freq Domain

Windowing

Algorithms are Peak to Peak

From peaks, windows are calculated

Features are calculated for windows

Spectrum operation

Filtering

Bandpass Filtering

Reason: Remove the low and high freq noise

Mel-Frequency Cepstral Coefficients (MFCCs) [[27](https://ieeexplore.ieee.org/document/1163420)]

Reason: remove noise and frequency domain analysis

Pre-emphasis

Reason: enhance high-freq components of signal

Windowed

FFT to each window

Mel filter bank

Log of all filter bank energies

Mel scale

Maps freq lineraly belwo 1000hz and log above 1000hz

Reason: the acoustic respone of human ear is not lined to all freq

Discrete Cosine Transform (DCT) [[28](https://www.nowpublishers.com/article/Details/SIG-001)]

Time Domain

Decimation (downsampling)

Low-pass filter

Cut-off frequency of ~200Hz

Shannon energy

Localization

Delineate the start of each phase (S1, Systole, S2, Diastole, ect)

Asses signal quality

Filter out baseline changes (low/high band pass filter)

Time/Frequency/Time-Frequency

Normalization

Reason: remove the amplitude variation from inter and intra classification (best sutable for classification)

Standardization

* + - * + Content

PCG recordings often are recording in non-ideal environments that are filled with unwanted background noise and interference. Data preprocessing is the process of altering the data in signal, often by denoising, normalizing, standardizing, and transforming the signal. These steps are crucial for automatic localization and classification tasks. Preprocessing the data allows a model to extract meaning features efficiently and reveals the physiological structure of the heart sounds [Latif et al.]. Furthermore, preprocessing helps ensure that the data that is fed into the model is always in the same domain. This allows the model to generalize to more easily.

We first resample the data to 500 Hz, to decrease spatial resolution of the heart sound recordings, but still retain important features. Thus, helping the model to converge faster. The resampled data is standardized using the standard score:

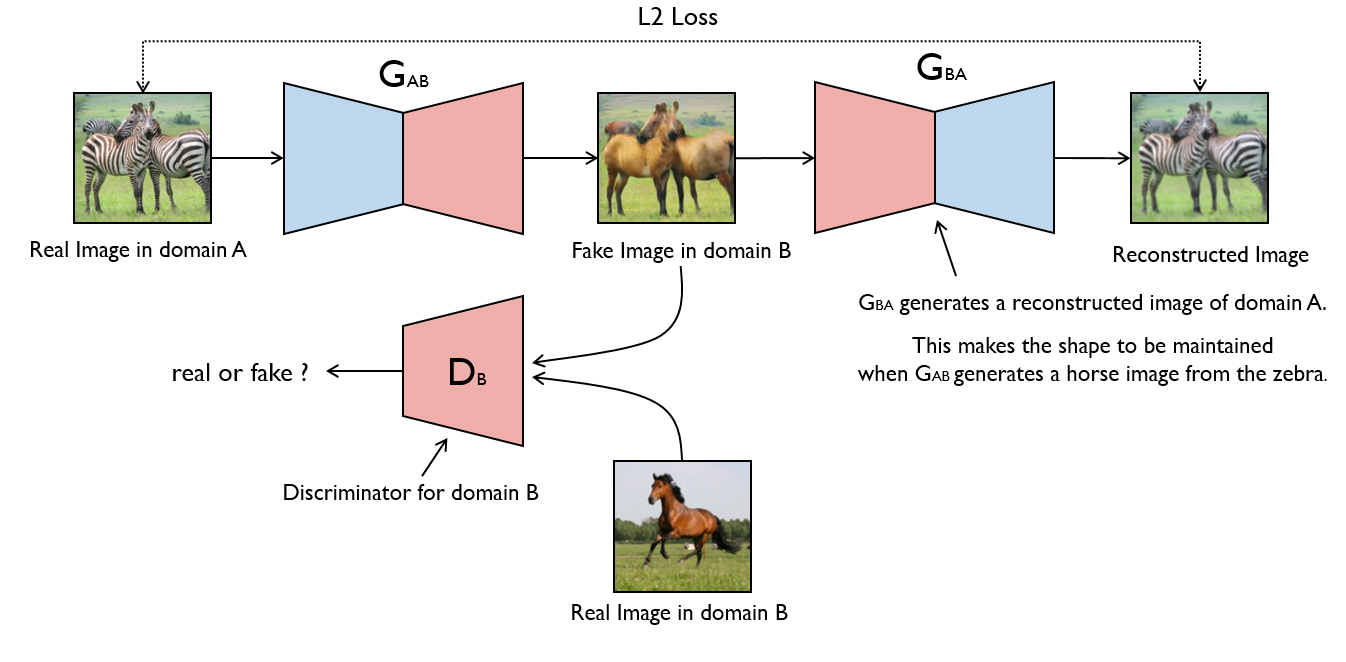
where

represents a single data point,

represents the mean of the signal, and

represents the standard deviation of the data. This scales the mean of the distribution to 0, artificially scaling all data into the similar ranges, thus, helping combat the exploding gradient problem.

The standardized data is then fed into a CycleGAN that the has learned to denoise data. The CycleGAN is fed synthetically noised PCG signal and attempts construct the denoised data from the noisy data. This synthetic noise consists of white noise, pink noise, and real background noise collected from audio recordings. The noise is added to each PCG signal recording and then treated as the input to the CycleGAN. The CycleGAN’s output is compared to the original, non-noise, PCG recording. In this way, the CycleGAN eventually learns to denoise PCG recordings.



{{[[TODO]]}} Add data transformation

{{[[TODO]]}} Add data localization

* + - * Data Augmentation
        + Topics

Resampling

noise injection

* + - * + Content

Data augmentation is a strategy that enables a significant increase in the diversity of data available while training a model, without actually collecting new data. Data augmentation techniques aim to slightly altering existing data to a point where the model cannot recognize the augmented data as one it has trained on before, but still retains the characteristics of the data’s category. This helps in reinforcing important features within the data and is only done during the training portion of the workflow.

A common misconception arises when comparing preprocessing and augmentation. To be clear, preprocessing aims to clean the data of unwanted artifacts that are not meaningful for classification and is done inplace. Augmentation, on the other hand, is solely done for expanding the dataset’s size, often to combat overfitting. Augmenting the data before preprocessing further obscures the data unrealistically, and beyond classification.

We use resample the heart sound recordings to different frequencies to simulate slower and faster beats per minute (bpm). The normal bpm for a human is between 60-100 bpm [[source](https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/all-about-heart-rate-pulse)]. Thus, measuring the sample distance between the first S1 (the start of systole) and the second S1, we calculate the bps:

where

represents the number of beats in the

interval, and

represents the recording frequency of the signal. The constant

coverts from beats per second to beats per minute.

Furthermore, we use noise injection directly to preprocessed PCG recordings [Messner et al.]. This process is identical to the process of synthetically adding noise to PCG recordings described in the preprocessing step. A variety of noises, like white noise, is added to the signal to increase the sample of recordings per class. This method is extremally beneficial for training on small datasets, like the PASCAL dataset.

* + - * Data Segmentation
        + Topics

Training Set

Validation Set

Testing Set

* + - * + Content

Data segmentation refers to the process of creating cross-validation datasets. This process assists in validating if the model is overfitting to the dataset. These datasets include the training set, validation set, and the testing set. Typically, the training set is 70%-80% of the dataset, the reset of the dataset is split among the validation set and testing set. Here, we split the data 80% training, 10% validation, and 10% testing.

* + - Model Development
      * Topics
        + Model Architecture

CNN

Bi-LSTM

DNN

GAN

* + - * + Layers
        + Activation Functions
        + Bach Norm
        + Dropout
      * Content
        + Here we propose using a Generative Adversarial Networks (GANs) for increased success in PCG heart sound detection. GANs poses a unique advantage over tadeonal machine learning and deep learning methods, in that a model learns to mimic a dataset by creating its own data, and tries to fool a discriminator into thinking the generated data is real.
        + In a supervised approach, a GAN consists of two parts, a generator and a discriminator. The generator is responsible for create fake heart sound data, while the discriminator tries to predict where the incoming data is fake or real. In a semi-supervised approach, however, the is fed data from a real dataset and the generator. Here, the discriminator ties to classify the generator’s fake data, as well as predict the classes form the real dataset.
        + Traditionally, generators are dense layers which slowly increase the increase the dimensionality of the generated data to match that of the real dataset. Discriminator, on the other hand, are commonly CNNs because majority of their applications work with images. However, it is possible to use a wide variety of architectures; such as, LSTMs, RNNs, SVMs, DNNs, ANNs, Transformers.
        + As mentioned above, there are many types of model architecture, some are used for classification, and other for feature extraction. Optimizing the combination of feature extraction layers and classification layers is extremally time consuming and computationally taxing. This is because there exits many combinations of hyperparameters, thus making it difficult to optimize each parameter. To optimize hyperparameters, we used hyperparameter sweeps to make the optimization process more efficient. This method involves using one of three methods: grid search, random search, and Bayesian search. Grid search computes each possible combinational of all hyperparameters and test them all. Although this is a very effective, it can be computationally costly. Random search selects a new combination at random, provided a distribution of values. This method is surprisingly effective and scales very well. Bayesian search creates a probabilistic model of metrics and suggests parameters that have a high probability of improving metrics. This works well for small scale projects, but scales poorly as the complexity of parameter relationships increase. Here, we used random search to optimize our hyperparameters.
    - Model Learning
      * Topics
        + Model Training

Cost Function (Loss)

Optimizer

* + - * + Early Stopping

Difference between

* + - * + Model Verification

Tested on testing set

* + - * Content
        + During the training phase, the model is trained using backpropagation in conjunction with a cost function. Backpropagation attempts to calculate the gradient of the cost function with respect to the weight and biases of the model. This process involves an optimizer, which optimize the model’s parameters and a cost function that measure the correctness or incorrectness of the model. The goal of the optimizer is to minimize the cost function’s error by adjusting the parameters to the given label. In this study, we used the Adam optimizer in union with Cross Entropy Loss. The Adam optimizer uses a hyperparameter that dictated the change in the model’s parameters on each backpropagation step, this is called the learning rate. Here we choose a learning rate of
        + .
        + The model is only trained on the training set; thus, backpropagation only occurs on the training set. Additionally, for each step in the training set, the optimizer backpropagates and optimizes the parameters and calculate metrics to further evaluate the model. The amount of steps in the training set is dictated by the batch size, the number of signals the model is trained on, in a single forward pass. Here we use a batch size of 32, meaning that the model is fed 32 signals per input. This significantly speeds up the process of training as more signals are passed through the model every time the model is optimized. A full pass of the training set is called an Epoch, here we train the model on 100 Epochs.
        + To ensure the model is not overfitting, but generalizing to the training set, we use a validation set to track the metrics of the model. In theory, the metrics on the training set equal to that of the validation set. In practicality, after many epochs of training the metrics of the validation set become static, but the metrics of the training set still increase. This suggests that the model is overfitting. Thus, we stop training the model on the training set and test it a testing set.
        + Testing sets or hold-out sets are used to validate the metrics of model, this is because both the validation set and the testing set have been tested by the model; thus, the model has developed a latent bias to both sets. Therefore, a third set is need to assess the model’s ability to generalize on an independent dataset.
    - Model Deployment
      * Topics
        + Integration
        + Monitoring
        + Maintaining
      * Content
        + Model Deployment is one of the last stages of any machine learning project and involves creating releasing the model to the public.
        + Integration

Integration consists of implementing the model in a system, whether it happens on the client-side or the backend. The most popular backend model integration tools involve Flask, Azure, and FastAPI. These tools create APIs that encapsulate the model prediction, given a GET request with the desired input.

* + - * + Monitoring & Maintaining

Following model integration and deployment, we move onto the next phase, monitoring and maintaining the system. As more and more data passes through the model, it increases the opportunity for the model to learn from a more generalized dataset. Though such data would be unsupervised, we could use unsupervised techniques to categories the data. Based on the improvement of the model, the model and be reintegrated and deployed. In essence, looping the whole process from data management to model learning.

* + {{[[TODO]]}} Metrics/Results and Data Analysis
    - Topics
      * Model vs Physicians vs Cardiologist
  + {{[[TODO]]}} Discussion
    - Topics
      * Realtime efficiency
      * Real world implementation
      * Computer Vision
      * Future Applications
  + {{[[TODO]]}} Conclusions
* PJAS and CASEF Outlines
  + {{[[TODO]]}} Refer to PJAS and CASEF Rubrics
    - found onedrive
  + **Title**
    - Generative Adversarial Networks for PCG Arrhythmia Detection
    - Manipulated Variable
      * Data that given to the model generated from Generative Adversarial Networks
    - Responding Variable
      * [[Metrics]] [[source](http://www.damienfrancois.be/blog/files/modelperfcheatsheet.pdf)]
        + Accuracy
        + Precision
    - Purpose of Investigation (motivation)
      * Nevertheless, automatic analysis of noisy heart beats still remains a challenging problem. ([[Ismail et al]])
      * From the view point of practical applications, development of computationally efficient solutions which may work in real time is also a challenging problem that needs further exploration ([[Ismail et al]])
      * The presently available PCG datasets have limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings. ([[Ismail et al]])
  + **Experimental Question / Problem / Rationale / Need**
    - Problem
      * 33% of all deaths are the result of Cardiovascular disease ([[Ismail et al]]). Unfortunately, datasets that contain Cardiovascular disease related PCG signals are noise and difficult to processes.
    - Question
      * Is it possible to use Generative Adversarial Networks to accurately detect Cardiovascular modalities, specifically arrhythmias in PCGs and surpass pervious methods in localization and classification tasks?
  + **Background Research**
  + **Hypothesis**
    - Here, we propose a method to artificially construct false PCG signals to aid our classification model in extracting meaning features from real heart sounds.
      * (LATER) Add information about ECG to PCG
        + In addition, the presently available PCG datasets comprise a limited number of samples and do not cover the complete range of pathologies which are likely to be encountered in clinical settings. This necessitates the development and labeling of a comprehensive dataset of PCG signals encompassing a variety of signals and covering all major pathologies. Considering the complexity of the problem, modeling the heart beats and various pathologies is likely to offer more robust solutions as opposed to the conventional techniques relying on localization and classification. To the best of authors’ knowledge, very limited efforts have been made in this direction ([[Ismail et al]]).
    - If a Generative Adversarial Network is used to create spurious data, then the model will outperform pervious state-of-the-art methods in localization and classification, because the specious data will aid the model in extracting significant features from a ground truth dataset.
    - Null Hypothesis
      * If a model was given falsified data to learn from, then the model will not extract any meaning feature sets from a valid dataset.
  + **Engineering Goals**
  + **Controls and Variables**
  + **Solution and Concept Review**
    - Pros and Cons
  + **Experimental Methods**
    - {{[[TODO]]}} Add stuff discussed in research paper
    - Life cycle of machine learning model
      * Data Management
        + Data Collection
        + Data Preporcessing

Denoising

Method invovlen with data

* + - * + Data Augmentation
        + Data Segmentation

Training Set

Validation Set

Testing Set

* + - * Model Learning
        + Model Hyperparameter Selection

Layers

Activation Functions

Bach Norm

Dropout

* + - * + Model Training

Cost Function (Loss)

Optimizer

Early Stopping

Difference between

* + - * + Model Verification

Tested on testing set

* + - * + Model Deployment

Integration

Monitoring

Marinating

* + - Source
      * Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges (Ashmore et al)
  + **Model A**
    - Outline
      * Methods
        + Data Management

Data Preprocessing

Raw Signal

Resampled to 500hz

Segmented based time (5 sec segments)

Standardization

Denoising

Filter out baseline changes (low/high band pass filter)

Savitzky–Golay flter

Time/Frequency/Time-Frequency

wavelets

Adding white gaussian noise

= original signal

= is the noise

= noise strength (standard deviation)

= time

Signal Decomposition

high freq part

passed through high-pass filter

low freq part

passed through low-pass filter

Applying threshold

Acts on single wavelet coefficient at a time

Hard thersholding

absolute value is less than threshold value, then set = 0

Soft thresholding

absolute value is less than threshold value, then set = 0

the remaining nonzero coefficients are shrunk towards zero

reconstruing denoise signal

reconstruction using wavelet or filter

Data Transformation

Freq Domain

Windowing

Algorithms are Peak to Peak

From peaks, windows are calculated

Features are calculated for windows

Spectrum operation

Filtering

Bandpass Filtering

Reason: Remove the low and high freq noise

Mel-Frequency Cepstral Coefficients (MFCCs) [[27](https://ieeexplore.ieee.org/document/1163420)]

Reason: remove noise and frequency domain analysis

Pre-emphasis

Reason: enhance high-freq components of signal

Windowed

FFT to each window

Mel filter bank

Log of all filter bank energies

Mel scale

Maps freq lineraly belwo 1000hz and log above 1000hz

Reason: the acoustic respone of human ear is not lined to all freq

Discrete Cosine Transform (DCT) [[28](https://www.nowpublishers.com/article/Details/SIG-001)]

Time Domain

Decimation (downsampling)

Low-pass filter

Cut-off frequency of ~200Hz

Shannon energy

Data Segmentation Database = 1.00

Training Set = 0.80

Validation Set = 0.10

Testing Set = 0.10

* + - * + Model Development

CNN

Layers

Input Layer [1x2500]

Embedding Layer [2500x200]

1D Convolutional Layer [32x2496]

Kernel [1x5]

Flattening [1x79872]

Dense [79872x256]

Dense [256x3]

Implementation

Bi-LSTM

DNN

GAN

* + - * + Model Learning

Model Training

Loss: Cross Entropy

Optimizer: Adam

Early Stopping on validation loss

Model Verification

Tested on testing set

* + - * Metrics/Results and Data Analysis
        + Metrics

Accuracy

Loss

FP, FN, TP, TN

Confusion Matrix

Sensitivity

Specificity

ROC (AUC)

* + - * + Visualization

t-SNE

class activation maps (CNN)

LSTM visualization

[Dense layer visualization](https://playground.tensorflow.org/)

* + - * + Statistical Tests (Ask Shaw)

z-score of all models

ANOVA

F-score

t-statistic

* + - **Designing and Concept Sketeches**
    - **Constructing**
    - **Issues & Redesigning**
  + **Model B**
    - **Designing and Concept Sketeches**
    - **Constructing**
    - **Issues & Redesigning**
  + **Model C**
    - **Designing and Concept Sketeches**
    - **Constructing**
    - **Issues & Redesigning**
  + **Results**
    - Outline
    - Confusion Matrix
      * The confusion matrix to the right aids illustrates the performance for each class of the proposed method. Specifically, we evaluated the model’s success on the grounds of average accuracy, specificity, and sensitivity of the classification. We calculated the average true positives, false positives, etc. for all 150 trials of the testing set. Using these floored values, we normalized each label along the y-axis. The matrix reveals that the most common misunderstanding occurs in predicting a normal heart sound, this is not surprising as the model is trained to be biased in detecting abnormal heart sounds. However, this occurs at the cost of specificity, which decreases the average normal misclassification rate to 8.11%. Overall, we conclude that the average accuracy of abnormal heartbeat detection is ~95% with a misclassification rate of just 5%. Thus, the model is extremely accurate in detecting abnormalities in heart sounds and has the capabilities to further classify abnormal heart sounds into labeled arrhythmias.
    - Dataset Visualization (t-sne)
      * Dataset visualization is critical in understanding the dataset’s complexity and model’s effectiveness. Here, we use t-distributed stochastic neighbor embedding (t-SNE), a statistical method for visualizing multi-dimensional data in less computationally expensive dimensions. The method is presented with the raw prediction values for each input of the validation set and maps the corresponding predictions into a 2-dimensional space. Tracked over epochs, the visualization allows us to view the progression of the model’s competence while training. The visualization highlights clear clustering within the dataset, which suggests the model is stable. Though, from epochs 70 and onwards, it is evident that there is overlapping between abnormal and normal signals. Assuming these signals as ground truth, this implies that additional feature engineering is required to adequately classify heart sounds.
    - Time Complexity
      * Model complexity is used to gauge and evaluate the efficacy of a model against an increase in data (n). We mainly focus on time complexity as it is most relevant to the problem at hand (space complexity is O(1)). Depending on model deployment and integration, the complexity can vary. For example, GPUs have parallel processing capabilities, which allow them to process multiple signals at once, efficiently decreasing the model complexity to O(1). For this reason, we use the worst-case scenario (a CPU), for analysis of the proposed method’s time complexity. The graph to the left implies the model’s time complexity is directly and linearly correlated to the input size, suggesting the complexity is O(n).
    - Generalization (Training vs Validation Set)
      * Generalization is important in creating accurate predictions, as it establishes that the model is learning meaningful features that are not just applicable to the training data, but signals overall. The graphs to the left plot the loss and accuracy of the training and validation set over each epoch. The large fluctuations in the training set metrics are caused by logging the metrics after each step in the training set (on every change in the model parameters). Thus, the optimizer is bound to decrease gradients in the wrong direction, thus correcting for such variations cause those fluctuations. Both lines on the loss plot resemble an exponential curve, which suggests the model continues to learn as training progresses. The average deviation between the validation and training set for each epoch is 0.4817; though this deviation is high, it is due to the lack of meaningful surface-level features that would lead to accurate detection. Meaning, the model’s deep feature extraction layers are responsible for the gap. Furthermore, the accuracy plots follow a logarithmic curve. In other words, as the training accuracy increasing, the validation correspondingly increases, though at a slower rate. Both graphs illustrate the model has reached convergence by the end of the training phase. This is confirmed by the static change in metrics in both datasets.
    - Class Activation Map
      * The graphs to the left visualize the channels of the convolutional layers, which are responsible for extracting features from an inputted signal. The color of the vertical line represents the importance of the feature relative to other features. Meaning, the lighter the color, the more important the feature. As the input travels deeper into the model, the extractions get more complex. This is clearly shown by the district increase in coloration between layers one and two. The first convolutional plot depicts the extraction of important biomarkers, such as S1 and S2. In the next two layers, the extractions make it clear that the S1 sound is more important, as those are the brightest throughout all layers. This parallels medical knowledge, as most cardiovascular anomalies occur in Systole, or at the start of S1.
    - Real world testing
      * Testing the model’s viability is crucial for ensuring the model’s success in the real world. Ideally, recording heart sounds are recorded with digital stethoscopes. These tools use transducer technology to convert sound into an electrical signal. Over the past decade, this technology has grown immensely (by cause of speech recognition). Modern phones have the potential to record the sounds at a high resolution, given the microphone is located at the correct position relative to the heart. Such a device will prove to extremely beneficial in providing diagnosis without the need for specialized equipment. The graph to the right shows a heart sound recording from a phone microphone. The plot shows that import biomarkers, like S1 and S2, remain visible. This ensures that the recording doesn’t contain excessive amounts of noise that may hinder the performance of the detection system.
    - Generator
      * The graph to the right shows the input values (seeds) and corresponding generated noise. The variance in amplitudes and periods in the signal mimics that of real heartbeats changing. At a closer examination of the results, the generator appears to resample the seed values to that of the output (2500). This is expected as the generator consists of dense layers, which linearly increase the signal length and complexity.
    - Data Transformation
      * Traditional machine learning reconstruction techniques use images because they contain dense information in a low dimensionality. However, time-series data, such as heart sound signals, are one-dimensional. Thus, we use time and frequency analysis, specifically log spectrograms to represent heart sound recordings in a two-dimensional manner. These spectrograms show the frequency of the signal against time. This process is done on both the PCG and ECG signals and results in a 48 x 48 image.
    - VAE Generator
      * Variational Autoencoders (VAE) is an architecture that attempts to map the initial data into an encoded space (ẑ) that stores latent features, the encoder is responsible for doing this task. The encoded space is regularized to avoid overfitting. In the context of ECG and PCG signals, both signals come from different latent spaces, which makes it difficult to travel from one domain to another. Regularizing this encoded latent space ensures that the latent space has good properties. This space is then reconstructed by a decoder into an encoded-decoded space, identical to the initial data in shape. With respect to ECG and PCG signals, the input, a spectrogram representation of the ECG signals, is encoded into the encoded space. The encoder consists of ResNet and Multi-head attention blocks that introduce a method to negate the vanishing gradient problem and weights for each pixel in the spectrogram respectively. The encoded space represents important structural elements, such as the positions of the PQRST complex. This space is then decoded using a revered structured encoder, into an encoded-decoded state that respectable a spectrogram of the PCG signal. These reconstructed PCG signals are sent to a convolutional discriminator, which attempts to classify the generated PCG spectrogram as real or fake (reconstructed). Of course, the goal of the discriminator is to classify the constructed spectrograms as fake and spectrograms from a PCG dataset as real. Whereas, the generator attempts to fool the discriminator into classifying the generated spectrogram as real. This system ensures that the generated spectrogram looks authentic and genuine.
    - VQGAN Results Viz
      * The plots above show the progression of heart sound construction from ECGs over epochs of the validation set. Ideally, he would want the reconstruction of the PCG spectrogram to identical to that of the ground truth. In practice, this doesn’t occur, some features may be lost in the latent representation of the ECG spectrogram. These missing features will cause a spatial anti-aliasing effect, as the latent space doesn’t have the dimensionality to extract pixel-to-pixel information. The series of spectrograms show the development of features in the latent space through the epochs. For example, the frequency of the first peak (S1) varies significantly until the ~86th epoch. This feature is important because it determines the rate of the S1 or “lub” sound; thus, creating the illusion that the sound is occurring faster relative to the ground truth. Furthermore, the S2 marker is severely softened, this is due to the rapid change in frequencies surrounding the marker and the light vertical bars to the right in each spectrogram. This suggests that much of the information regarding S2 will be lost when converting the spectrogram into a wave signal.
    - Conculsion
      * We proposed a Generative Adversarial Network (GAN), composed of a Dense Generator and a Convolutional Neural Network (CNN) Discriminator to detect abnormal heart sounds in a recording. The model achieved an accuracy of 94.98%, a specificity of 90.30%, and a sensitivity of 99.52% on the testing set. The previous state-of-the-art achieved an accuracy of 86.02%, a specificity of 77.81%, and a sensitivity of 94.24%. This data, along with results from the t-test revealed that the proposed alternative hypothesis was correct and that the null hypothesis should be rejected. This is because the proposed method reached better performance than the previous state-of-the-art methods. Additionally, the model attained a staggering ~2500 classification per second in the worst-case scenario. This is because of the nature of the CNN architecture; unlike other methods, the CNN’s reduce the data dimensionality as it forward propagates through the model. Furthermore, the proposed method showed real-world deployment capabilities for autonomous heart sound abnormality detection with recordings collected from a phone microphone. This test shows extremely promising results for future applications and integrations.
        + We also set out to introduce new pathologies for increased arrhythmia labels in classification. We proposed using a VQGAN for constructing PCG signals from existing ECG datasets that contain a surplus amount of arrhythmia-specific data. The results were promising, in that the VQGAN discriminator was able to construct the general shape of the PCG spectrogram, but missed import details in the fluctuation of important biomarkers (S1 and S2). This caused the PCG waveform representation extracted from the PCG spectrogram to miss rapid oscillations present in the biomarkers.
      * The object of this study was to create a fast and accurate end-to-end heart sound arrhythmia detection system, capable of detecting abnormalities in real-time without specialized equipment. While also increasing the number of cardiovascular pathologies classified. With the data shown, our proposed method accomplishes exemplary statistics in abnormalities detection and shows promising results in increased arrhythmia construction. Hopefully, this study will shed light on PCG construction techniques and give birth to applications with autonomous abnormality detection.
  + **Discussion**
  + **Future Ideas**
  + **Optimization Statistics and Highlights**
  + **References**