COMPUTER VISION LAB

PROJECT REPORT

On

Two class classification of Pascal heart sound signals using the continuous wavelet transform features.

Submitted By:

Boddeda Nehitha Rathan (121CS0215)
Sambit Kumar Rout (121CS0216)
MD Nafis AL Safayet (121CS0217)
Amit Kumar Sah (121CS0218)

Submitted To:

Dr. Puneet Kumar Jain

Department of Computer Science and Engineering
NATIONAL INSTITUTE OF TECHNOLOGY ROURKELA

APRIL, 2024

TABLE OF CONTENTS

ABSTRACT:

Cardiac auscultation plays a crucial role in diagnosing cardiac health by analysing heart sounds. With the advent of electronic stethoscopes capable of recording heart sounds digitally, researchers have gained access to vast amounts of phonocardiogram (PCG) data. In this study, we propose a convolutional neural network (CNN) model to classify paediatric heart sound signals into two categories: normal and abnormal. Leveraging signal processing and machine learning techniques on PCG signals, our model aims to facilitate the diagnosis of various heart disorders. The dataset used in this study comprises heart sound recordings collected from 941 individual. Through our CNN model, we seek to enhance the accuracy and efficiency of cardiac health assessment, thereby aiding clinicians in timely diagnosis and treatment interventions.

LIST OF FIGURES:

LIST OF TABLES:

LIST OF ABBREVATIONS:

1.PCG-Phono cardiogram
2. CNN – Convolutional Neural Network
3. 1D – One Dimensional
4. 2D – Two Dimensional
5. WHO – World Health Organization
6.ECG – Electro cardiography

1.INTRODUCATION

1.1 INTRODUCTION:

From various physiological signals in the human body, important functions and the health status of the body can be indicated through phonocardiogram (PCG), electrocardiography (ECG), electroencephalography (EEG), electromyography (EMG), etc [1].
According to statistics from the World Health Organization (WHO), cardiovascular diseases (CVDs) cause high mortality worldwide, and the current mortality rate is steadily increasing. Among them, PCG refers to the recording of sounds generated by heart valves, atria, and blood flow during the heartbeat, and is a signal that can be identified when there is an abnormality in heart function or condition.
Because the heart plays essential roles and functions in survival, such as temperature control, nutrient delivery, blood pressure maintenance, and oxygen supply, information on the condition and function of the heart can be obtained through this organ, making it possible to diagnose cardiovascular diseases' signals, which are biological signals, can be used to detect and classify heart diseases and abnormalities using machine and deep learning methods [2].

1.2 MOTIVATION:

The motivation behind this project lies in the critical importance of accurate cardiac assessment, especially in paediatric populations where early detection of heart abnormalities can significantly impact patient outcomes.
Traditional methods of cardiac auscultation heavily rely on clinician expertise, which can be prone to subjectivity and inconsistency. By leveraging advanced signal processing and machine learning techniques on digital heart sound recordings, we aim to develop an automated classification system capable of distinguishing between normal and abnormal heart sounds in paediatric patients [3].
Such a system has the potential to augment clinical decision-making, enhance diagnostic accuracy, and ultimately improve the management of paediatric cardiac conditions, leading to better patient care and outcomes.

1.3 PROBLEM STATEMENT:

The project description states that continuous wavelet transform will be used to categorize pascal heart sound signals into normal and abnormal heart sounds. The main goal of this project is to create a Convolutional Neural Network (CNN) model 1D/2D that can distinguish between normal and abnormal cardiac sounds [4].

The dataset comprises paediatric heart sound recordings stored in .wav format, collected from 941 individuals. It includes 176 instances of normal heart sounds and 31 instances of abnormal heart sounds. The recordings are sampled at a rate of 4000Hz, ensuring high fidelity representation of heart sounds. Each recording captures the acoustic characteristics of the heart, providing valuable insights into potential cardiac abnormalities.[5]

2 LITERATURES:

2.1 HEART SOUND/PCG SIGNAL:

Heart sounds refer to the noises generated by the beating heart and the turbulence of blood flow within the cardiovascular system. These sounds are typically heard through a stethoscope placed on the chest, although more advanced diagnostic tools such as echocardiography can also detect and analyse them. Heart sounds are divided into two primary categories:

Normal Heart Sounds: These are the typical sounds produced by the functioning of the heart valves and blood flow. The two main normal heart sounds are:

Example:

lub……….dub……………. lub………. dub……………. lub……….dub……………. lub……….dub…

Abnormal Heart Sounds: These are additional sounds or murmurs that can indicate underlying cardiac conditions, such as valve disorders, structural abnormalities, or abnormal blood flow.

Example:

1.

…lub..****...dub……………. lub..****..dub ……………. lub..****..dub ……………. lub..****..dub

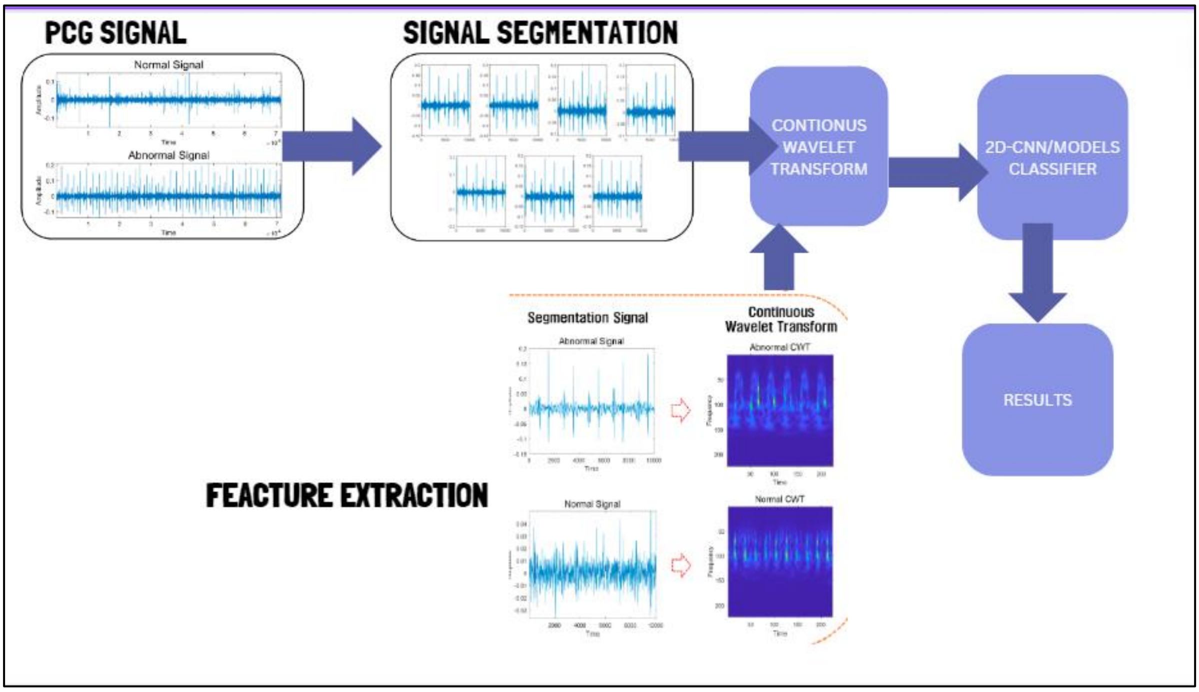
2

….lub.lub……….dub…………………. lub. lub……….dub…………….lub.lub……………dub……

3.PROPOSED MEATHOD :

This study consisted of a preprocessing stage using a PCG dataset for heart sound classification, a feature extraction stage using Continuous Wavelet Transform at regular intervals, a classification 2D-CNN and vgg_16 and a random forest classifier . After extracting features using Continuous Wavelet Transform method, using 2D-CNN. and using vgg_16 and random forest classifier we get required results.

Figure 1. Heart sound classification architecture using wavelet analysis technology and an ensemble of deep learning models



3.1PCG signal:

The Pascal dataset of Phonocardiogram (PCG) signals is a collection of audio recordings capturing heart sounds from various individuals. These recordings are meticulously annotated and curated, providing valuable data for research and development in cardiovascular health and medical diagnostics. The dataset offers a diverse range of PCG signals, enabling researchers to study heart conditions, analyse cardiac abnormalities, and develop algorithms for automated heart sound analysis. Researchers and healthcare professionals utilize this dataset to enhance the accuracy of heart sound analysis systems, improve diagnostic capabilities, and advance the field of cardiovascular medicine.

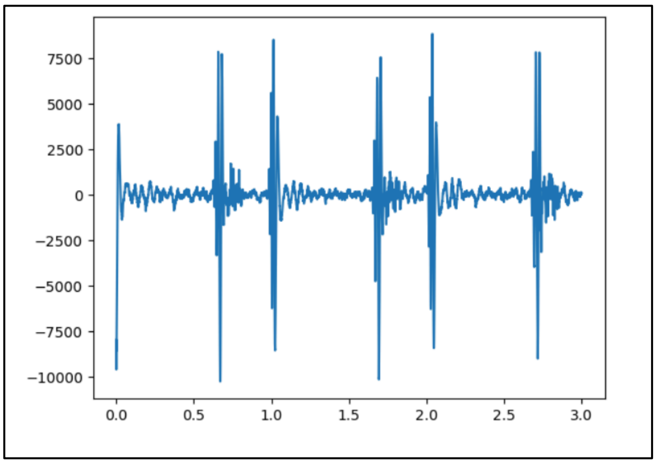


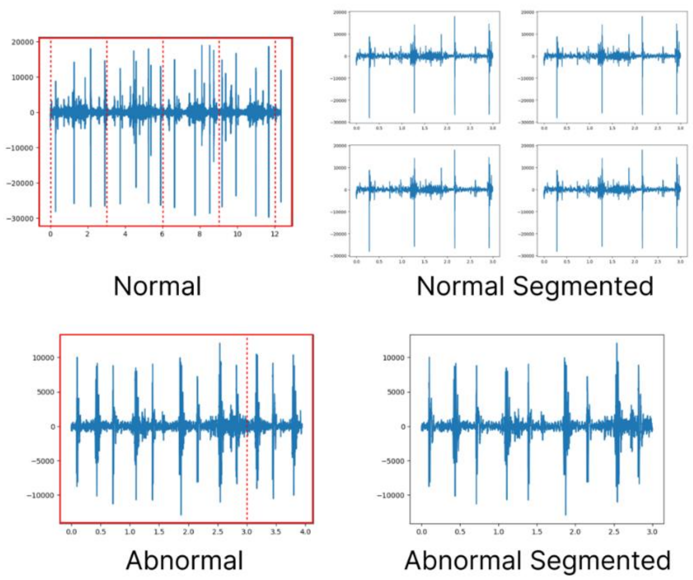
FIGURE 2 : sample picture format of a pcg signals

3.2Down sampling:

Down sampling a signal, such as a heart rate Pascal signal, from a sampling rate (SR) of 4000 to 1000 involves reducing the number of samples taken per second while preserving essential information. This process typically entails taking every fourth sample from the original signal and discarding the rest. Down sampling can be advantageous for various reasons, including reducing computational complexity and memory requirements. However, it's crucial to ensure that the down sampling process does not introduce aliasing, which can distort the signal by folding higher frequencies into the lower frequency range. Proper filtering techniques, such as anti-aliasing filters, are often employed to prevent aliasing effects and maintain signal integrity during down sampling.

3.3 Signal segmentation:

To classify diseases and heart sounds using PCG signals, preprocessing was performed to adjust the original signals containing various signal lengths to the same signal length. Additionally, because PCG signals occur differently for each person depending on the location, shape, size, and health status of the heart and heart valves, features in specific areas can be extracted in more detail through signal segmentation. PCG signals are composed of various signal lengths, and by adjusting the original signal length to a constant length, important features in the signal can be confirmed. Information loss can occur in the process of cropping a long signal to fit a short signal or increasing a short time to a long signal. In addition, if a long signal is used as is without dividing it, important features and patterns that appear minutely in the S3 and S4 sections of the PCG signal, which may indicate heart-related diseases, may be missed; thus, a segmentation process is performed for accurate analysis and diagnosis results. When dividing the PASCAL Classifying Heart Sounds Challenge Dataset A and Dataset B into 1 s, which is the minimum time, a signal that does not include the components S1, S2, S3, and S4 cycles of the PCG signal was included; therefore, it was divided into 3 s intervals.



The segmented size of less than 3sec is drooped and not taken into consideration to our model.

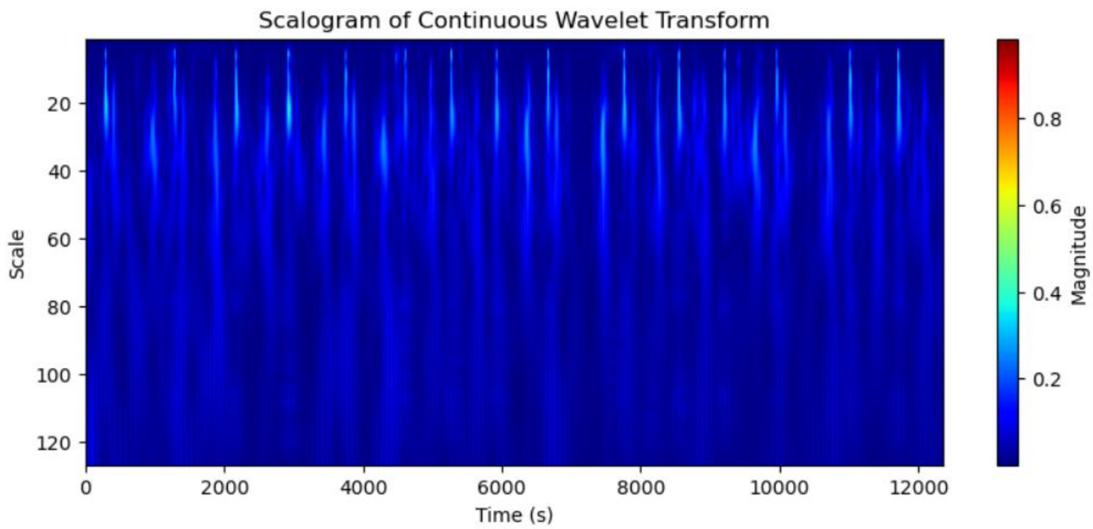
FIGURE 3: Segmented signal each of 3sec.

3.4 FEATURE EXTRACTION:

Continues wavelet transformation:

Continuous Wavelet Transform (CWT) is a signal processing technique utilized to analyse signals that exhibit non-stationary behaviour, such as Phonocardiogram (PCG) signals, which are recordings of heart sounds over time. When implementing CWT on a PCG signal with scales ranging from 1 to 128 and employing the Morlet wavelet, a detailed representation of the signal's time-frequency characteristics is obtained.

The Morlet wavelet, often chosen for its effectiveness in capturing both low and high-frequency components, is a complex exponential modulated by a Gaussian window. This wavelet is particularly well-suited for analysing PCG signals due to its ability to accurately detect both the transient events, such as heart valve closures (S1 and S2 sounds), and the continuous murmurs indicative of abnormal blood flow.
By using scales ranging from 1 to 128, CWT decomposes the PCG signal into different frequency bands and time intervals. The scale parameter determines the width of the wavelet function and thus dictates the frequency resolution. Lower scales provide higher frequency resolution but lower time resolution, while higher scales offer higher time resolution but lower frequency resolution. This range of scales allows for a comprehensive analysis of the PCG signal across various frequency components, enabling the detection of subtle changes and abnormalities.
FIGURE shows scalogram of the wave using continues wavelet transformation.



4.RESULTS AND DISCCUSION:

4.1 EXPERMENTAL SETUP:

SETUP1.1:
MODEL 1:2D/CNN:
Without down sampling
Leaning rate :0.01
1.Conv(2d) (64)—batch normalization--max pooling —conv2d (32 filters) ---batch normalization—max pooling –conv2d(16filters)—batch normalization—max pooling –flatten –dense (8 neurons) – dense (2 neurons)
2.Conv2d (filer size 3x3, stride 1)
3.Max pooling (stride 2,2)

MODEL 1.2: VGG16 WITH RANDOM FOREST: BEST ACCURACY (no down sampling)
VGG16 Model: Pre-trained on ImageNet dataset. Used for feature extraction, excluding the top layers.
1.Compilation:
Adam optimizer, Mean Squared Error loss, and accuracy metric.
2.Feature Extraction:
Extract features from training and test images using VGG16.
3.Random Forest Classifier:
Trained using extracted features and corresponding labels.
Predictions made for both training and test sets.
4.Evaluation:
Accuracy scores computed for training and test predictions.

SETUP 2
DOWNSAMPLED(1000Hz)

MODEL 2.1 :2DCNN
1.CONV(2D) (64 filters) ---dropout (0.5) --- CONV2D(3filters) --- dropout (0.5) -----CONV (16filters) ---dropout (0.5) ----CONV(8filters) ---flatten—dropout (0.5) ------dense (8 neurons) – dense (2 neurons)
2.RELU ACTIVATION FUCTION,
3.KERNEL SIZE 3X3

MODEL2.2: random forest classifier
Model2.3:
This code defines a transfer learning model using the ResNet50 architecture pre-trained on ImageNet. It loads the ResNet50 model with its top (classification) layers excluded, freezes these layers to retain their weights, and then adds custom classification layers on top. The custom layers include a Flatten layer to convert the output of ResNet50 into a 1D tensor, followed by a Dense layer with 256 neurons and ReLU activation function, and finally a Dense layer with 2 neurons (assuming binary classification) and SoftMax activation. The model is

compiled using the Adam optimizer and sparse categorical cross-entropy loss, with accuracy as the evaluation metric.
Model3.4: model1+random forest classifier

4.2 DATASET DESCRIPTION:

The dataset comprises paediatric heart sound recordings stored in .wav format, collected from 941 individuals. It includes 176 instances of normal heart sounds and 31 instances of abnormal heart sounds. The recordings are sampled at a rate of 4000Hz, ensuring high fidelity representation of heart sounds. Each recording captures the acoustic characteristics of the heart, providing valuable insights into potential cardiac abnormalities

4.3 PERFOMANCES METRICS:

Accuracy: Calculated as (TP +TN) / (TP + TN +FN + FP), where TP is true positives and FN is false negatives, TN is true negatives and FP is false positives

Sensitivity: Calculated as TP / (TP + FN), where TP is true positives and FN is false negatives.

Specificity: Calculated as TN / (TN + FP), where TN is true negatives and FP is false positives

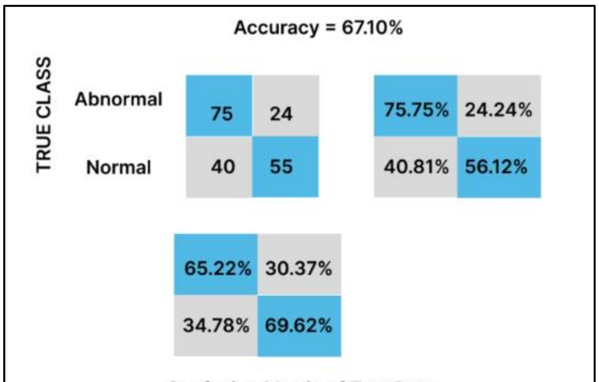
Confusion Matrix: Shows the distribution of true positive, true negative, false positive, and false negative predictions.

F1score: Harmonic mean of precision and recall.

4.4 RESULTS:

The model shows a moderate level of success in correctly classifying the test data, with an experiment accuracy of 64%. It shows some degree of effectiveness in identifying patterns within the image data, even though it is not quite perfect. To improve performance and close the discrepancy between the model's predictions and ground truth labels, more optimization and fine-tuning might be required. The achieved accuracy of 64% suggests that the Continuous Wavelet Transform (CWT) might not be the most optimal choice for feature extraction in this context.

MODEL	ACCURACY
MODEL1.1	55.67
MODEL1.2	67.1
MODEL2.1	65.3
MODEL2.2	57.3
MODEL2.3	60.2
MODEL2.4	57.9

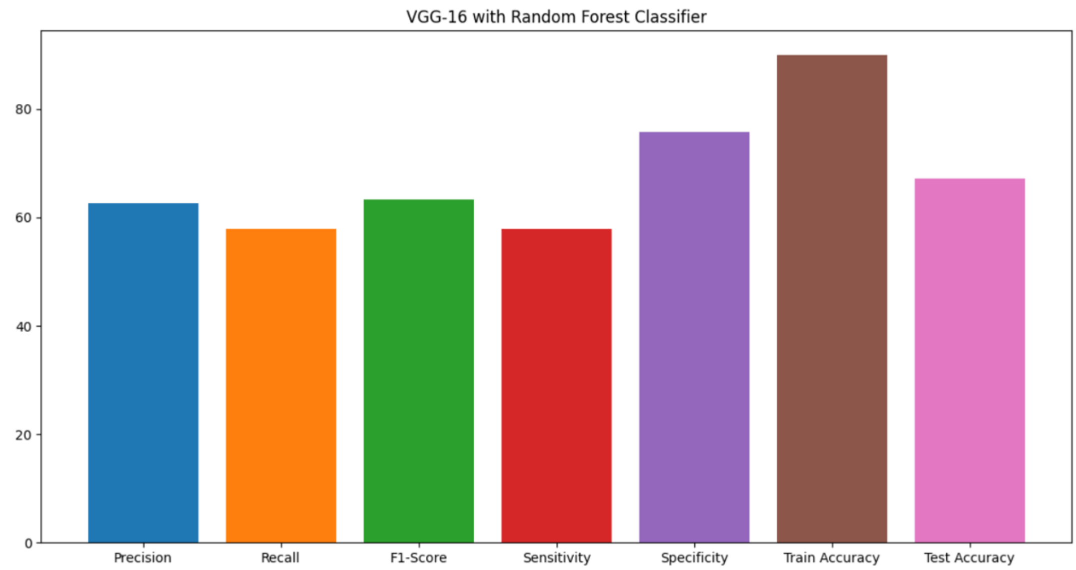


Confusion Matrix of Test Data
VGG-16 with Random Forest Classifier

Dataset Averages

Precision	Recall	F1-Score	Sensitivity	Specificity	Train Accuracy	Test Accuracy
69.62	57.85	63.21	57.90	75.75	89.93	67.10

Table: Performance Results for VGG-16 with random forest classifier



5. CONCLUSION:

The model shows a moderate level of success in correctly classifying the test data, with an experiment accuracy of 64%. It shows some degree of effectiveness in identifying patterns within the image data, even though it is not quite perfect. To improve performance and close the discrepancy between the model's predictions and ground truth labels, more optimization and fine-tuning might be required. The achieved accuracy of 64% suggests that the Continuous Wavelet Transform (CWT) might not be the most optimal choice for feature extraction in this context

6.REFERENCES:

1.Heart Sound Classification Using Wavelet Analysis Approaches and Ensemble of Deep Learning Models

2. https://youtu.be/qoMDpSatG7M?si=uvY4QTDVbVwqvjP4

3. Classifying Heart Sounds Challenge

And many more

7.Task performed to the project and percentage contribution by each project.

TASK	NEHITHA	SAMBIT	NAFIS	AMIT
GATHERING RECOURCES	30%	20%	20%	30%
CODING	30%	20%	20%	30%
PROJECT REPORT	20%	30%	30%	20%
PRESENTATION	20%	30%	30%	20%
OVERALL	25%	25%	25%	25%