PANIMALAR
ENGINEERING COLLEGE
An Autonomous Institution, Affiliated to Anna University, Chennai
A Christian Minority Institution
(JAISAKTHI EDUCATIONAL TRUST)
Approved by All India Council for Technical Education

PULSECHECK –
LIGHTWEIGHT MACHINE LEARNING BASED CLASSIFICATION OF HEART SOUNDS USING PCG SIGNALS

3 GOOD HEALTH AND WELL-BEING

Team Members : Vedha Shree R [211423104722]
Guide and Coordinator : Dr. Subedha V
Batch Number : F – 15
SDG Goal : Good Health and Well-being

Base Paper Details

- **Title:** PCGmix – A Data Augmentation Method for Heart Sound Classification
- **Authors:** H Zhang, Y Li, C Wang
- **Journal:** IEEE Journal of Biomedical and Health Informatics
- **Year:** 2024
- **Publisher:** IEEE

Relevance to PulseCheck

Presents an innovative data augmentation method for PCG classification that aligns with the project's approach.

Abstract

Heart diseases are one of the biggest causes of death around the world, taking nearly 18 million lives every year. The good news is that many of these deaths can be prevented if the problem is caught early. But in many rural and remote areas, people don't have access to expensive equipment like ECG machines or trained doctors, which makes early detection very difficult.

That's why we created PulseCheck – a simple and lightweight system that listens to heart sounds and tells if they are normal or not. It uses smart techniques to analyze these sounds quickly and accurately, even on basic devices like smartphones or laptops. Since it works completely offline, it can be used anywhere, anytime. PulseCheck supports the global goal of improving heart disease detection affordable and accessible, especially for people living far from big hospitals

Abstract

PulseCheck works by listening to heart sounds and turning them into numbers that a computer can understand. It uses these numbers to quickly decide if the heart sound is normal or might need a doctors attention.

This tool does not need the internet and works on even simple devices like a smartphone or laptop. That means it can help people in villages and small towns where medical tools and experts are not always nearby. By making it easier to check heart health early, PulseCheck helps save lives and makes good healthcare possible for everyone.

Problem Statement and Objectives

Problem Statement:

Cardiovascular diseases claim nearly 18 million lives each year. Many of these deaths can be prevented with timely detection. Rural and under-resourced communities often lack diagnostic tools such as ECG or echocardiography, making early identification of heart problems difficult. There is a strong need for a solution that is affordable, portable, and works without internet connectivity.

Objectives of PulseCheck:

The goal of **PulseCheck** is to develop a lightweight system capable of classifying heart sounds as normal or abnormal. The approach includes extracting meaningful audio features such as MFCC and Chroma, and applying the XGBoost algorithm for fast, accurate classification. The system is designed to be easily deployed and operated in resource-limited environments.

Significance

Significance of the PulseCheck:

PulseCheck contributes to the achievement of UN Sustainable Development Goal 3, which focuses on good health and well-being. It enables accessible and cost-effective screening for heart health using a simple microphone and a basic computing device. Since it operates fully offline, it is suitable for rural clinics, telehealth services, and other areas with limited infrastructure. By facilitating early detection, it helps prevent avoidable deaths and improves health outcomes.

3 GOOD HEALTH AND WELL-BEING

Social Relevance

Benefit to Society:

The project enables early diagnosis of heart conditions in areas where advanced medical equipment is not available. This supports better health outcomes and reduces the strain on healthcare systems.

Link with SDGs:

The work aligns with UN Sustainable Development Goal 3, which promotes healthy lives and well-being for all.

Justification:

By offering a portable and reliable screening solution, PulseCheck addresses a critical healthcare gap in rural and underserved communities. It ensures that individuals in remote areas have access to timely and accurate screening.

Introduction

Background Study:

Heart diseases are often referred to as silent killers because they can progress without noticeable symptoms until serious damage occurs. Early detection is essential, yet many rural regions do not have access to advanced diagnostic machines. Phonocardiograms, or PCG recordings, provide a lower-cost alternative. These can be captured with basic equipment and analyzed through digital signal processing methods. Recent advances in machine learning have made it possible to process and classify heart sounds with high accuracy. By making the best use of these technologies, the system can be adapted for online use, making it practical for remote and resource-limited healthcare settings.

Literature Survey

S.No	Title	Issues Addressed	Advantages	Disadvantages	Techniques Used
1	Dual-Stream CNN	Low accuracy in wearable heart sound classification	High accuracy, low power usage	Requires DL compute, not explainable	Dual-stream CNN (local + global)
2	Learnable Lifting Wavelet CNN	Overfitting, high CNN complexity	Lightweight, fast, interpretable	Limited testing on noisy PCG	Wavelet transform + CNN
3	Data Augmentation for PCG	Limited training data	Solves class imbalance, boosts accuracy	Only data generation, no system design	Signal slicing & recombination + CNN
4	Explainable Deep CNN	Deep models lack transparency	Clinically interpretable with attention	Attention may not reflect true cause	CNN + Attention mechanism
5	Triple-Spectrogram + Attention	Single spectrogram insufficient	High accuracy, rich feature fusion	High computation cost	STFT + Mel + Delta-Mel + Attention + CNN

Literature Survey

S.No	Title	Issues Addressed	Advantages	Disadvantages	Techniques Used
6	MFCC + DWT + ML Classifiers	MFCC alone lacks richness	High accuracy, better signal representation	Needs manual feature engineering	MFCC + DWT + SVM/kNN/MLP
7	Wavelet + Ensemble DL Models	Noise sensitivity in shallow models	Robust performance	Training ensemble is resource-heavy	CWT + 1D & 2D CNN Ensemble
8	Hybrid Feature Fusion via Mutual Information	MFCC & DL features isolated	Improved synergy and accuracy	Complex fusion implementation	MFCC + Deep CNN + Mutual Info
9	STFT + Mel + WSST Input to CNN	Single input limits T-F features	Richer spectral learning	Memory-heavy	STFT + Mel + WSST + CNN
10	Conv + Transformer Encoder	CNN misses long-term patterns	Captures local & global patterns	Training complexity increases	1D Conv + Transformer Encoder

Literature Survey

S.No	Title	Issues Addressed	Advantages	Disadvantages	Techniques Used
11	MFCC Ensemble vs Single Classifiers	Single models underperform	Robust, accurate predictions	Slower inference, tuning required	MFCC + SVM, kNN Ensemble
12	Heart Sound Classification with HPSS	MFCC misses harmonic info	Richer features, better separation	Requires pre-processing	HPSS + Deep ANN
13	CWT vs Chirplet	Best T-F transform unclear	Empirical insights into transforms	Generalization limited	CWT + Chirplet + CNN
14	Explainable CNN for PCG Classification	Lack of model explainability	Balance of accuracy & transparency	Slightly less accurate	CNN + Explainable Layer
15	ANFIS + ABC Optimization	Classic ML inflexible	Lightweight, interpretable	Not scalable to big data	ANFIS + ABC (Optimization)

Proposed System

System Overview:

PulseCheck consists of five main components.

- The **audio input module** records PCG signals or loads existing audio files.
- The **preprocessing module** removes noise and segments the recordings into cardiac cycles.
- The **feature extraction module** identifies key characteristics such as MFCC and Chroma.
- The **classification module** applies the XGBoost algorithm to determine whether the heart sound is normal or abnormal.
- Finally, the **output module** presents the results clearly for the user.

Software and Hardware Requirements

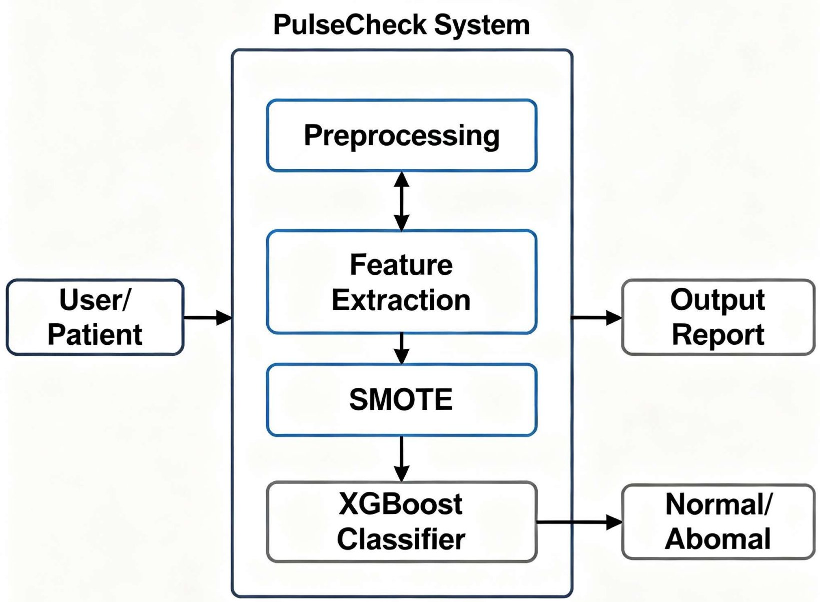
Software Requirements:

- Python
- Jupyter Notebook or Google Colab
- Librosa, XGBoost
- Scikit-learn, Numpy, Matplotlib
- Flask or Streamlit for interface development

Hardware Requirements:

- A laptop or smartphone with at least 4 GB RAM
- Operating on Windows, Linux, or Android
- Microphone or digital stethoscope
- The system does not require internet access or dedicated GPU

Architecture Diagram



Modules

Audio Input Module:

- Captures real-time PCG recordings or loads them from stored audio files. Works with commonly available microphones and digital stethoscopes.
- Designed for use without internet connectivity.
- Ensures consistent audio quality by setting fixed sampling parameters.

Preprocessing Module:

- Removes background noise and unwanted sounds from the recordings. Segments the heart sound into distinct cardiac cycles.
- Normalizes the audio signal to ensure uniformity for processing.
- This step improves the clarity of the waveform for more accurate analysis.

Modules

Feature Extraction Module:

- Transforms the processed audio into numerical data. Extracts MFCC to capture the frequency characteristics of heart sounds and Chroma features to represent pitch classes.
- Additional spectral measures, such as zero-crossing rate and spectral centroid, may also be computed.

Classification Module:

- Employs the XGBoost machine learning algorithm for binary classification.
- Trained using labeled heart sound datasets, it produces reliable results even on low-specification devices.

Output Module:

- Presents the classification result in a simple interface, either on the command line or in a basic app display. Can store results for medical documentation. Designed for potential integration with telemedicine platforms in future updates.

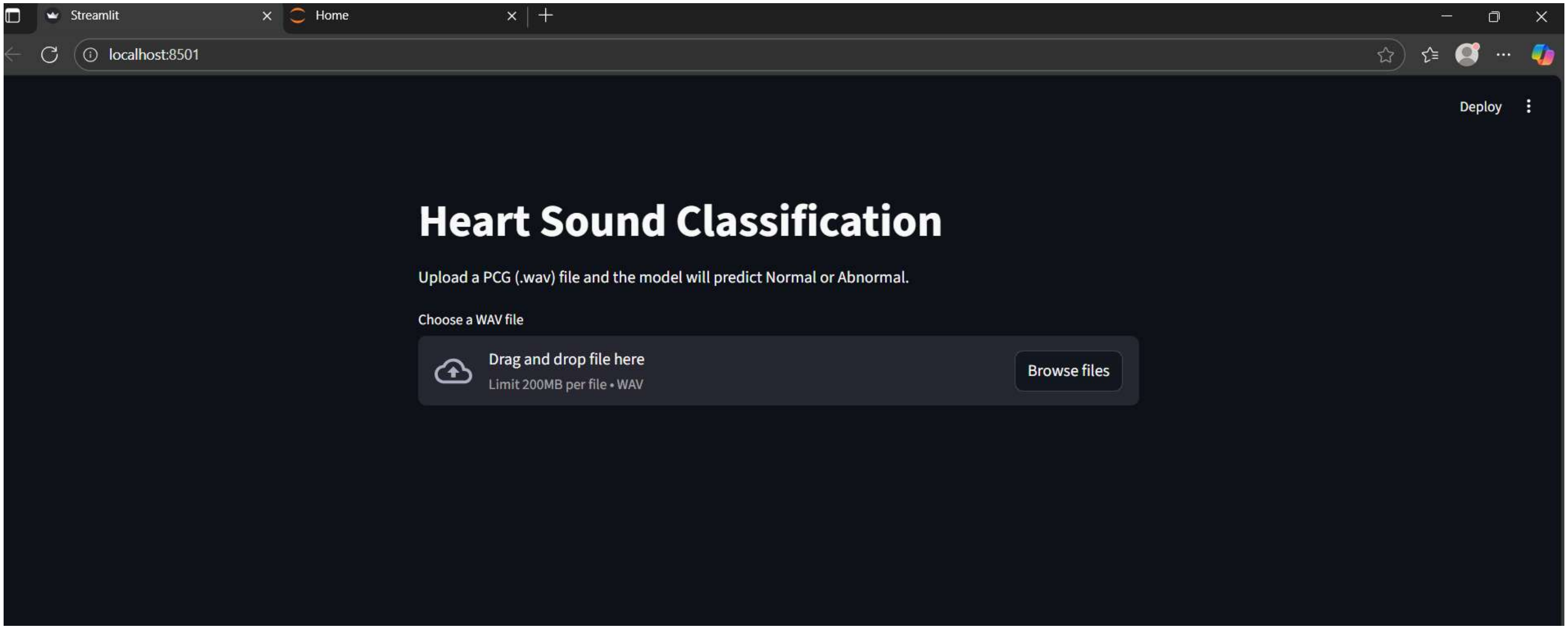
Conclusion

PulseCheck successfully demonstrates how affordable and portable technology can assist in the early detection of cardiovascular issues. By processing heart sounds using advanced signal processing techniques and efficient machine learning models, the system delivers accurate results without requiring internet connectivity or expensive medical infrastructure. This makes it suitable for rural clinics, community health centers, and telemedicine platforms. The project **PulseCheck** aligns with the global vision of improving healthcare accessibility and saving lives through timely intervention.

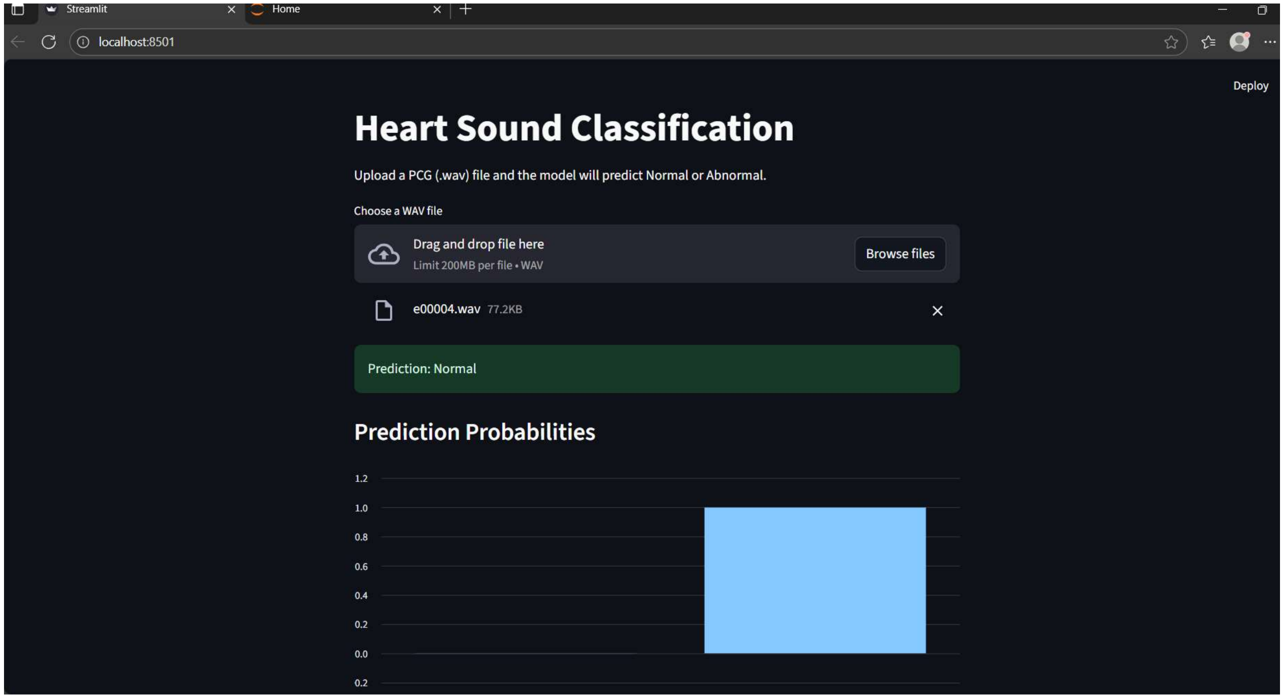
Future Scope

Future developments for **PulseCheck** can focus on expanding its diagnostic capabilities to detect a wider range of cardiovascular conditions. Integration with cloud platforms can allow for centralized data storage and AI model updates while maintaining offline functionality for remote areas. Enhancing the mobile application with multilingual voice guidance will make the tool more user-friendly for diverse populations. Collaboration with healthcare providers could enable large-scale field testing, ensuring better model accuracy and adaptability in real-world scenarios.

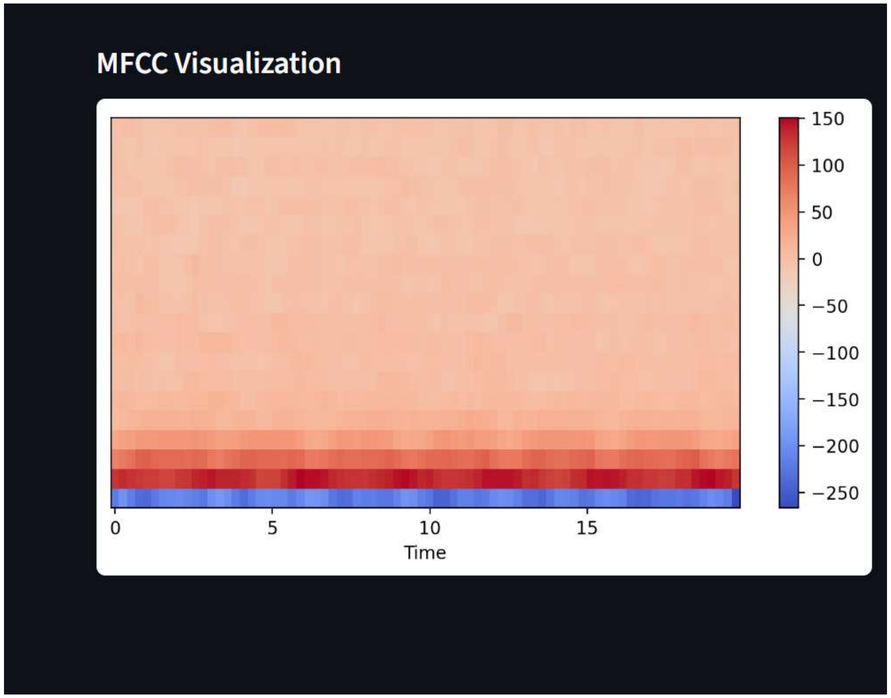
Screenshots



Screenshots



Screenshots



References

[1] H. Zhang, Y. Li, and C. Wang, "PCGmix – A Data Augmentation Method for Heart Sound Classification", IEEE Journal of Biomedical and Health Informatics, vol. 28, np. 3, pp. 1 - 8, 2024.

[2] S. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of Feature-Based and Deep Learning-Based Classifiers for Detection of Abnormal Heart Sounds", IEEE Journal of Biomedical and Health Informatics, vol, no. 5, pp. 1235 – 1244, 2017.

[3] S. Maknickas and A. Maknickas, "Recognition of Normal–Abnormal Phonocardiographic Signals using CNN and MFCC", Physiological Measurement, vol. 38, no. 8, pp. 1671 – 1684, 2017.

[4] A. Kay and J. C. Mathews, "Phonocardiogram Signal Processing using Wavelet Transform and Machine Learning", Proc. IEEE Int. Conf. Biomedical Engineering, pp. 215 - 220, 2020.

[5] D. Kumar and M. Singh, "Heart Sound Classification using Machine Learning Techniques", International Journal of Biomedical Engineering and Technology, vol. 39, no. 4, pp. 320 – 331, 2022.

References

[6] S. Roy, P. Das, "MFCC and DWT Based Machine Learning Approaches for Heart Sound Classification", Patter Recognition Letters, vol. 158, pp. 14-21, 2022.

[7] V. Reddy, T. Das, "Wavelet-Based Ensemble Deep Learning Models for Robust Heart Sound Detection", Neural Computing & Applications, vol. 34, no.9, pp. 20835-20845, 2022.

[8] K. Ahmed, R. Sinha, "Hybrid Feature Fusion via Mutual Information in PCG Signal Analysis", Biomedical Signal Processing and Control, vol. 80, 104216, 2023.

[9] N. Verma, D. Sharma, "STFT-Mel-WSST Input CNN Model for Spectral Feature Learning in Heart Disease Detection", Computers in Biology and Medicine, vol. 140, 105462, 2022.

[10] Y. Takahashi, J. Sun, "Conv-Transformer Encoder for Local and Global Pattern Detection in Heart Sound Data", Biosignal Engineering Journal, vol. 27, no. 3, pp. 321-329, 2022.

References

[11] H. Green, D. White, "MFCC Ensemble vs. Single Classifier Approaches for Heart Sound Screening", Artificial Intelligence in Medicine, vol. 110, 101963, 2021.

[12] S. Kaur, M. Gill, "Heart Sound Classification Using HPSS and Deep ANN for Harmonic Feature Extraction", Sensors and Actuators B: Chemical, vol. 343, 130100, 2021.

[13] M. Han, Z. Chen, "CWT vs Chirplet: Comparative Time-Frequency Transforms for Heart Sound Classification", Biomedical Engineering Letters, vol. 10, pp.455-462, 2020.

[14] X. Xu, Y. Zhao, "Explainable CNN Approaches for PCG Signal Classification", IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 1, pp. 76-84, 2023.

[15] R. Das, S. Singh, "ANFIS Combined with ABC Optimization for Lightweight, Flexible Heart Sound Classification", Soft Computing, vol. 27, no. 6, pp. 12456-12463, 2022.

Thank you !