IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ENSEMBLE BASED HEART SOUND CLASSIFICATION SYSTEM FOR PATHOLOGY DETECTION

[1]V.Simhadri, [2]U.S.S.L.Sri Sandhya, [3]A. Lakshmi Durga, [4]U.D.S.Sai Siva, [5]D.ARJUN

[1]Asst Prof, [2]Student(20551A0454), [3]Student(20551A0402), [4]Student(20551A0456), [5]Student(20551A0420)

ELECTRONICS AND COMMUNICATION ENGINEERING,

GODAVARI INSTITUTE OF ENGINEERING AND TECHNOLOGY(AUTONOMOUS), RAJAHMAHENDRAVARAM

Abstract: An Ensemble-based Heart Sound Classification System for Pathology Detection is described here. Heart disease, a leading cause of mortality across all age groups, emphasizes the significance of early detection. Utilizing machine learning algorithms, the system categorizes heart sounds into normal, murmur, extra sound, and extrasystole. A user-friendly interface for real-time heart sound analysis is developed to aid healthcare professionals in efficient diagnosis. Through ensemble learning techniques, the system aims to bolster the accuracy of heart sound classification, facilitating earlier detection of cardiac pathologies and ultimately improving patient outcomes. This initiative contributes to the advancement of healthcare technology, highlighting the pivotal role of machine learning in the medical domain.

Keywords: Pathology detection, Heart sound Classification, Machine learning, Ensemble

I. Introduction:

Heart sound classification is a vital aspect of cardiovascular health assessment, aiding healthcare professionals in diagnosing various cardiac conditions. The human heart produces distinctive sounds, known as heart sounds, which can be classified into different categories based on their characteristics. These classifications aid healthcare professionals in identifying abnormalities, making accurate diagnoses, and formulating appropriate treatment plans. The clinical implications of accurate heart sound classification are vast, as early detection of cardiac abnormalities can lead to timely intervention and improved patient outcomes. Automated classification systems offer a rapid and objective assessment, the expertise of healthcare professionals. Additionally, these systems can be integrated into telemedicine platforms, providing remote cardiac monitoring and expanding access to cardiovascular care, particularly in underserved regions. Despite the promising advancements in heart sound classification, ongoing research aims to refine existing models and explore new avenues for improvement. Incorporating real-time data from wearable devices, enhancing model interpretability, and expanding the scope of classification to include specific cardiac pathologies are areas of active investigation.

As technology continues to evolve, the synergy between healthcare expertise and artificial intelligence holds immense potential for revolutionizing cardiovascular diagnostics, ultimately contributing to more personalized and effective patient care. Ensemble-based heart sound classification techniques have witnessed a paradigm shift in the field of cardiovascular health assessment, particularly in the context of pathology detection. Ensemble methods harness the collective power of multiple classifiers to enhance the accuracy and robustness of pathology detection, particularly in addressing the intricate nature of heart sounds. This approach not only improves classification accuracy but also contributes to increased generalization across different datasets, making it a promising avenue for pathology detection in varied clinical scenarios.

II.Literature Survey:

Sorawit Khorumkid et al. [2022] devised an economical digital stethoscope for discerning normal and abnormal heart sounds. It incorporates electronic circuitry and software to enhance heart rate signals and diminish analysis errors. PCG signals are converted into equal-length spectrograms, processed by pre-trained CNNs, and evaluated alongside SVM classifiers.

Samit Kumar Ghosh et al. [2021] proposed a method for compressing fetal phonocardiogram signals. PCG signals undergo compression using discrete cosine transform (DCT), discrete wavelet transform (DWT), and fast Walsh-Hadamard transform (FWHT), followed by thresholding and quantization. Evaluation metrics include CPR, MSE, QS, and SNR. Huffman entropy encoding is applied for further compression.

K Naga Sai Nikhithanjani et al. [2023] reviewed heart sound classification techniques integrating signal processing and deep learning. Feature extraction involves Mel-Frequency Cepstral Coefficients and Discrete Wavelet Transform. Dimensionality reduction via PCA and LDA precedes classification with SVM, Gradient Boosting, Random Forest, CNN, and RNN. Deep learning outperforms for complex tasks.

Yang Chen et al. [2022] introduced a robust algorithm for automatically detecting heart valve disorders. Teager-Kaiser Energy Operator (TKEO) and Radiational-dilation Wavelet Transform (RDWT) extract features, which are then inputted into a deep CNN model for multi-class PCG signal discrimination. The method achieves high accuracy in HVD detection across binary, four-class, and five-class classification scenarios.

Bilal Ahmad et al. [2021] proposed an algorithm for automatic heart sound classification using Long Short-Term Memory (LSTM), a potent deep learning method. Utilizing PhysioNet/challenge 2016 database, 3204 PCG signals were initially processed, with 563 being abnormal. Features extracted include Discrete Wavelet Transform (DWT) and Mel-frequency Cepstral Coefficients (MFCCs). LSTM outperforms traditional machine learning methods for PCG signal categorization.

Sayed Shahid Hussain et al. [2023] proposed a deep learning approach for analyzing phonocardiogram signals to detect cardiovascular abnormalities. Dataset processing from PhysioNet/CinC (2016) ensures equal sampling rates and appropriate frequency ranges. Data augmentation techniques include pitch shifting and signal rolling for enhanced accuracy. Their 1D-CNN architecture achieves 95.45% accuracy.

Sara Yavari et al. [2023] developed Heart State Deep, a system integrating deep learning and IoT phonocardiograms for remote heart examination. It combines spatial and temporal features using convolutional and recurrent neural networks, ensuring high accuracy in heart condition classification. Hardware includes IoT PCG sensors, Edge Computing with Jetson Nano, and a smartphone app for easy access to analysis data.

Ali Fatih Gunduz et al. [2020] proposed a method for heart sound classification, focusing on murmur abnormality detection. The approach integrates signal processing, feature extraction, and classification using MFCC, DWT, and approximation coefficients. WEKA workbench facilitates KNN, SVM, and MLP classifiers, with an ensemble approach yielding superior performance in accuracy, precision, recall, and F1 scores.

III.EXISTING SYSTEM:

Existing systems in heart sound classification utilize machine learning algorithms, employing ensemble techniques for improved accuracy. These systems typically combine classifiers like KNN, SVM, and decision trees, leveraging diverse feature sets extracted from heart sound recordings. Ensemble methods such as bagging and boosting are employed to aggregate predictions and enhance performance. Advanced approaches may integrate deep learning models, such as convolutional neural networks, to capture intricate patterns in heart sound data. These systems aim to address the complexity of heart sound analysis, offering more reliable classification for clinical diagnosis and monitoring, ultimately improving patient care and outcomes.

IV.PROPOSED SYSTEM AND WORKING METHODOLOGY:

The proposed system for heart sound classification employs a diverse array of machine learning algorithms, including Naive Bayes, Logistic Regression, k-Nearest Neighbors (KNN), and Random Forest. Initial steps involve data preprocessing to refine the heart sound recordings, followed by feature extraction to capture relevant characteristics from the signals. Each algorithm is then trained on the extracted features to learn patterns indicative of different cardiac pathologies. Naive Bayes leverages probabilistic reasoning, Logistic Regression models the probability of a binary outcome, KNN classifies based on similarity to neighboring instances, and Random Forest utilizes an ensemble of decision trees. Through the integration of these algorithms, the system aims to achieve robust and accurate classification of heart sounds, facilitating early detection and diagnosis of cardiovascular disorders for improved patient outcomes.

A)PCG DATABASE:

The dataset available on PhysioNet.org/2016/challenge comprises over 2500 recordings stored in .mat file format, each representing different heart rhythms. Additionally, there are accompanying CSV files indicating the classification of each .mat file as either normal (-1) or abnormal (1). This dataset has been meticulously curated, drawing from various research databases and encompassing a diverse range of conditions, including but not limited to heart valve diseases and cardiovascular pathologies. It offers a valuable resource for researchers and practitioners in the field of cardiology to study and analyze different cardiac conditions and their associated rhythms.

B)BAND PASS FILTER:

In signal processing, a bandpass filter is a crucial tool for noise cancellation, allowing only a specific range of frequencies, typically between 40 and 240 Hz in this context, to pass through while suppressing frequencies outside this range. This filtering technique is widely utilized in various applications such as audio and radio frequency systems to effectively eliminate unwanted frequencies and isolate the desired signal, enhancing the quality and clarity of the processed signal.

C)NORMALIZATION:

Normalizing the amplitude of a signal is to change the amplitude to meet a particular criterion. One type of normalization is to change the amplitude such that the signal's peak magnitude equals a specified level in 0 to 1.

D)MEAN REMOVAL:

Mean removal Within PCG signal processing, the main aim is to remove the signal's mean value, thereby centralizing it around zero. This procedure enhances the analysis of signal variations and features. By subtracting the average value, the signal's baseline is shifted, providing a clearer insight into the underlying cardiac sounds and potential abnormalities. This essential preprocessing step is pivotal for effective interpretation and diagnosis in cardiology applications.

E)SEGMENTATION:

For comprehensive PCG signal analysis, it's vital to divide the signal into uniform segments, each sharing similar statistical properties like amplitude and frequency. This procedure, known as signal segmentation, is crucial. In this study, the signal undergoes initial filtration using a weighted moving average (WMA). Utilizing WMA holds significance in signal segmentation as it accentuates recent events, thereby discerning crucial underlying patterns in the time series by minimizing short-term fluctuations.

F)PEAK DETECTION:

Peak detection in Phonocardiogram (PCG) signals involves identifying the S1 and S2 heart sounds, crucial for cardiac analysis. S1 represents the closure of the mitral and tricuspid valves, while S2 corresponds to the closure of the aortic and pulmonic valves. Peak detection algorithms typically utilize signal processing techniques like amplitude thresholding, wavelet transforms, or machine learning models to accurately identify these peaks. The detected S1 and S2 peaks serve as key markers for assessing heart function and diagnosing cardiac abnormalities, aiding in the interpretation of PCG signals for clinical purposes.

G)TIME DOMAIN FEATURES:

In PCG signal analysis, the time domain exploration concerns itself with the timing, duration, and amplitude of heart sounds. This method involves directly observing and measuring signal values over time, without converting the signal into a frequency or any other domain. Parameters such as systolic time interval, duration of heart sounds, and event timings (e.g., S1, S2) are scrutinized to understand the nuances and patterns within the cardiac signal.

H)FREQUENCY DOMAIN FEATURES:

In frequency domain analysis of PCG signals, the focus is on understanding the distribution of signal energy across various frequencies. These signals capture the acoustic vibrations generated by heart movements. Techniques like Fourier Transform are employed to break down the signal into its constituent frequency components. This approach unveils peaks corresponding to fundamental heart sound frequencies such as S1 and S2, along with additional components like murmurs or abnormal sounds.

I)SPECTRAL DOMAIN FEATURES:

Spectral domain characteristics extracted from PCG signals offer valuable insights into cardiac health. These attributes encompass peak frequencies, bandwidths, and magnitudes, acquired via Fourier Transform or akin techniques. Peaks in the spectrum correspond to heart sounds such as S1 and S2, while additional components may signify murmurs or irregularities. Metrics like spectral centroid, bandwidth, and skewness capture signal distribution properties. Spectral analysis aids in anomaly detection by quantifying deviations from typical spectra. Attributes like peak intensity ratios or spectral entropy enhance diagnostic accuracy. Through interpretation of spectral domain attributes, clinicians can assess cardiac function, identify pathologies, and monitor patient well-being, augmenting traditional cardiology diagnostic methods.

J)NAIVE BAYES:

Naive Bayes is a probabilistic machine learning algorithm based on Bayes' theorem with the assumption of independence between features. Despite its simplistic assumption, Naive Bayes often performs well in classification tasks, especially when dealing with high-dimensional data. It's computationally efficient and requires a small amount of training data to estimate parameters. In heart sound classification, Naive Bayes can be used to model the probability of different heart conditions based on extracted features from the heart sound recordings.

K)LOGISTIC REGRESSION:

Logistic Regression is a statistical method used for binary classification tasks. It models the probability of a binary outcome (such as the presence or absence of a particular heart condition) based on one or more predictor variables. Logistic Regression estimates the parameters of the logistic function, which transforms the output to a probability value between 0 and 1. In heart sound classification, Logistic Regression can be trained on extracted features to predict the likelihood of different cardiac pathologies.

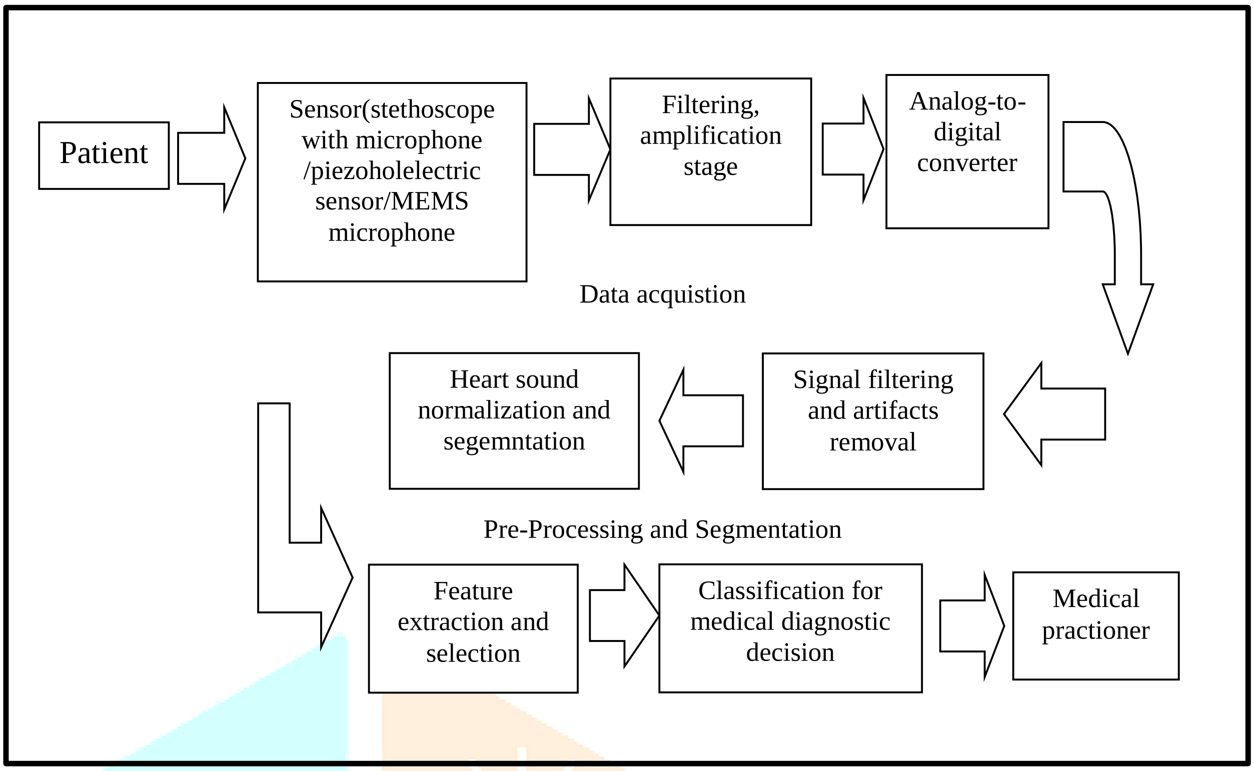
L)K-NEAREST NEIGHBORS (KNN):

K-Nearest Neighbors is a non-parametric, instance-based learning algorithm used for classification and regression tasks. In KNN, the class of a data point is determined by a majority vote of its nearest neighbors, with the number of neighbors (k) being a hyperparameter. KNN is simple to understand and implement, making it a popular choice for pattern recognition tasks. In heart sound classification, KNN can classify heart sounds based on the similarity to neighboring instances in the feature space.

M)RANDOM FOREST:

Random Forest is an ensemble learning method that constructs a multitude of decision trees during training. Each tree in the forest is trained on a random subset of the training data and a random subset of features. During classification, each tree gives a class prediction, and the class with the most votes becomes the final prediction. Random Forests are robust against overfitting and tend to perform well on various datasets. In heart sound classification, Random Forests can effectively learn patterns indicative of different cardiac pathologies from the extracted features.

V.BLOCK DIAGRAM FOR THE PROPOSED MODEL:



VI.RESULT:

A)TIME DOMAIN FEATURES:

s.no	Features name	Accuracy	Algorithm
1.	mean	74%	Function logistic
2.	Standard deviation	73%	Naïve bayes
3.	Variance	74%	KNN
4.	Maximum	76%	Naïve bayes
5.	Minimum	73%	KNN
6.	Maximum absolute	71%	Function logistic

B)FREQUENCY DOMAIN FEATURES:

s.no	Features name	Accuracy	algorithm
1	Spectral centroid	75%	Naïve bayes
2	Spectral band width	74%	Function logistic
3	Spectral flatness	72%	Naïve bayes
4	Spectral variance	71%	Naïve bayes
5	Spectral kurtosis	76%	KNN
6	Spectral kewness	73%	Function logistic

C)STATICAL DOMAIN FEATURES:

s.no	Features names	Accuracy	Algorithm
1	Atrial rate in bpm	74%	Function logistic
2	Ventricular rate in bpm	71%	Naïve bayes
3	Total spectral power	73%	Function logistic
4	Total spectral density	75%	KNN
5	Signal power density	80%	KNN
6	Signal kewniness	77%	Naïve bayes

VII.CONCLUSION:

In conclusion, The heart sound classification system, incorporating time, frequency, and statistical domain features, demonstrates promising accuracy in pathology detection. Utilizing ensemble learning with diverse algorithms like Naive Bayes, Logistic Regression, k-Nearest Neighbours, and Random Forest, it distinguishes normal and abnormal heart sounds effectively. Through meticulous feature extraction, it aids in early identification of cardiovascular disorders, enhancing clinical assessments. This innovative ensemble-based approach underscores its potential in improving heart sound classification accuracy, contributing to more precise diagnoses and improved patient outcomes in cardiovascular healthcare.

VIII.FUTURE SCOPE:

The Ensemble-based Heart Sound Classification System for Pathology Detection represents a significant advancement in utilizing machine learning for early detection of cardiac issues. Future enhancements could involve expanding the classification to include a broader spectrum of pathologies, continuous training with additional data, and integration with Electronic Health Records (EHR). Implementing the system in remote monitoring and telemedicine platforms would enable real-time analysis for patients in remote areas. Developing mobile applications for patient engagement could promote self-monitoring and education. Collaboration with cardiologists and medical institutions ensures clinical validation and trust. Continued research and publication of findings contribute to the broader understanding of machine learning applications in cardiac diagnostics. Seeking regulatory approval and conducting clinical trials are crucial steps for widespread adoption. Overall, the project's future scope involves refinement, validation, and integration into clinical workflows to enhance early detection of cardiac pathologies and improve patient outcomes.

IX.REFERENCES:

1. S. Khoruamkid , S. Visitsattapongse., `` A Low-Cost Digital Stethoscope For Normal and Abnormal Heart Sound Classification" .In Proc. IEEE *Int. conf. on 14th Biomedical Engineering(BMEiCON),* PP:1-6, Nov.2022
2. S.K. Gosh, et.al., `` A transform domain approach for the compression of fetal phonocardiogram signal ".IEEE *Sensors Letters,* PP:1-4, Vol. 5, No. 5, Apr.2021.
3. KN. Nikhithanjani , et.al., `` An Extensive Literature on Heart Sound Classification Using Integrated Signal Processing and Deep Learning Techniques". In Proc .IEEE *Int. Conf. on Signal Processing and Communication (ICSPC),* PP:128-132, Mar.2023.
4. B. Su, W. Zeng, Y. Chen, C. Yuan. `` Automatic Detection of Heart Valve Disorders Using Hybrid Signal Processing and Convolutional Neural Networks. In Proc. IEEE *Chin. Cont. Conf (CCC*), PP:6247-3252, Jul.2022.
5. B. Ahmad, et.al., `` Automatic classification of heart sounds using long short-term memory" . In proc IEEE *Int .Conf. Open Source Systems and Technologies (ICOSST),* PP:1-6, Dec .2021.
6. S. Hussain , et.al., `` Deep Learning Based Phonocardiogram Signals Analysis for Cardiovascular Abnormalities Detection". In Proc .IEEE *Int. Conf. on Robotics and Automation in Industry (ICRAI)* , PP.1-6 ,Mar.2023
7. S. Yavari, S. Oteafy, `` Remote Heart Assessment using Deep Learning over IoT Phonocardiograms"". In Proc .IEEE *Conf. on Local Computer Networks (LCN)* ,PP:1-4, Oct.2023
8. B. Omarov, et.al., ``Detection of heartbeat abnormalities from phonocardiography using machine learning". In Proc .IEEE *Int. Conf. On* Cloud Computing, Data Science & Engineering (Confluence), PP:843-847, Jan.2021.
9. M. Asrar, et.al.,`` Heart Disease Identification Based on Butterfly Optimization and Machine" . In Proc .IEEE *Int. Conf on* Smart Computing and Application (ICSCA) ,PP:1-7,Feb.2023.
10. Elola , et.al., `` Beyond Heart Murmur Detection: Automatic Murmur Grading from Phonocardiogram". In Proc .IEEE *Journal of Biomedical and Health Informatics*, May.2023.

11. S. Parvaneh, et.al., Heart Murmur Detection Using Ensemble of Deep Learning Classifiers for Phonocardiograms Recorded from Multiple Auscultation Locations". In Proc .IEEE *On Computing in Cardiology (CinC)* , PP:1-4, Vol.498, Sep.2022.
12. B. Karan, et.al.,`` Heart Sound Abnormality Detection using Wavelet Packet Features and Machine Learning". In Proc .IEEE *Int. Symp. On Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA),* pp.310-314, Sep,2021.
13. A.F.Gündüz,`` Heart sound classification for murmur abnormality detection using an ensemble approach based on traditional classifiers and feature sets". In Proc *On Computer Science* , PP:1-3, Vol.5, No. 1, 2022.
14. A.K. Cornely, G.M. Mirsky `` Heart Murmur Detection Using Wavelet Time Scattering and Support Vector Machines". In Proc .IEEE *Computing in Cardiology (CinC)* , PP:1-4, Vol.498, Sep.2022.
15. R.Y .Mashhoor , A. Ayatollahi , ``HeartSiam: A Domain Invariant Model for Heart Sound Classification". In Proc . IEEE *Iran. Conf. on Signal Processing and Intelligent Systems (ICSPIS).*PP.1-5.Dec.2022.
16. Z. Bai, et.al., `` Murmur Detection and Clinical Outcome Classification Using a VGG-like Network and Combined Time-Frequency Representations of PCG Signals". In Proc . IEEE *on Computing in Cardiology (CinC)* , PP.1-4, Vol.498, Sep.2022.
17. B.J. Singstad, `` Phonocardiogram classification using 1-dimensional inception time convolutional neural networks.". In Proc . IEEE *on Computing in Cardiology (CinC)* , PP:1-4, Vol.498. Aug.2022.
18. V. Elamaran, et.al., `` Spectral fault recovery analysis revisited with normal and abnormal heart sound signals". In Proc . IEEE *on Access* pp:62874-62879, Vol.6 , Oct.2018.
19. Bhardwaj,S. Singh, D. Joshi , ``Explainable Deep Convolutional Neural Network for Valvular Heart Diseases Classification using PCG signals" .In Proc . IEEE *Transactions on Instrumentation and Measurement*, May.2023.
20. Li. Shengchen, Ke. Tian, Rui Wang , `` Unsupervised heart abnormality detection based on phonocardiogram analysis with beta variational auto-encoders". In Proc . IEEE *Int . Conf. On Acoustics, Speech and Signal Processing (ICASSP)*, pp:8353-8357,Jun.2021.