AI-Assisted Heart Disease Prediction System

Nalamothu Vamsi[1], Nunna Sai Venkata Satya Siva[2], Manikonda Charan Raj[3], Manivannan[4]
[1,2,3,4]*Dhanalakshmi Srinivasan college*

***Abstract*—This project focuses on developing an AI-assisted Heart Disease Prediction System using machine learning techniques. The system analyses patient health parameters such as age, cholesterol level, blood pressure, heart rate, and other medical attributes to predict the presence of heart disease. The dataset used for this project contains multiple patient records and medical features, which are processed and analysed using Python and machine learning algorithms. The proposed system uses the Logistic Regression algorithm to train the model and make predictions. Data preprocessing, feature scaling, and model evaluation techniques are applied to improve prediction accuracy and reliability. The system generates diagnostic results that help healthcare professionals in faster and more consistent decision-making. The developed AI-based prediction system aims to reduce diagnosis time, minimize human errors, and provide reliable decision support for doctors. It can be useful in hospitals, telemedicine services, and rural healthcare centers where specialist availability is limited. Overall, this project demonstrates how Artificial Intelligence can enhance healthcare services and improve patient outcomes through early disease detection and intelligent data analysis.**

***Index Terms*—Artificial Intelligenc, Heart Disease Prediction, Machine Learning, Logistic Regression, Healthcare Analytics, Medical Data Analysis, Disease Diagnosis, Predictive Modelin, Data Preprocessing, Clinical Decision Support.**

I. INTRODUCTION

Heart disease is one of the leading causes of death worldwide and continues to be a major public health concern. According to the World Health Organization, cardiovascular diseases account for a significant percentage of global mortality each year, affecting millions of people across different age groups. Early detection and timely medical intervention play a crucial role in reducing complications and improving survival rates. However, traditional diagnostic methods often depend on manual evaluation of medical reports, clinical experience, and time-consuming laboratory tests, which may sometimes lead to delayed or inconsistent results. With the rapid advancement of Artificial Intelligence and Machine Learning, the healthcare industry has witnessed transformative changes in disease diagnosis and patient care. Machine learning algorithms can analyse large volumes of medical data, identify hidden patterns, and provide predictive insights with high accuracy. By leveraging patient health parameters such as age, cholesterol levels, blood pressure, heart rate, and other clinical attributes, intelligent systems can assist doctors in identifying the risk of heart disease at an early stage. This project presents an AI-assisted Heart Disease Prediction System that utilizes the Logistic Regression algorithm to analyse medical datasets and predict the likelihood of heart disease. The system involves data preprocessing, feature scaling, and model evaluation techniques to ensure accurate and reliable predictions. By integrating predictive analytics into healthcare decision-making, the proposed system aims to support medical professionals, reduce diagnosis time, and minimize human errors. Furthermore, it can be effectively implemented in hospitals, telemedicine platforms, and rural healthcare centers where access to specialized cardiologists may be limited.

Overall, the introduction of Artificial Intelligence in heart disease prediction demonstrates the potential of data-driven healthcare solutions in enhancing early diagnosis, improving treatment planning, and ultimately saving lives.

II. LITERATURE REVIEW

This section provides an overview of the existing literature involving the various approaches employed to diagnose cardiovascular diseases based on heart sound analysis. The analysis determines the correlations between heart sounds and normal or abnormal heart conditions. Typically, conventional heart sound auscultation remains ineffective for diagnosing certain heart conditions due to the limitations of human hearing nd the transient nature of heart sounds. The clinical diagnosis of heart sounds is also problematic, requiring much practice and experience. Recent advances in portable electronic medical devices and PCG signal processing using ML techniques have facilitated the creation of intelligent, non-invasive, low-cost, and user-friendly tools. These tools can assist cardiologists in diagnosing cardiovascular diseases and their prediction at early stages. Hence, PCG signals are crucial for describing the heart's mechanical activities, which can be undetectable under ECG signals. Table tabulates the general differences between ECG and PCG. Numerous studies have employed different dimension reduction techniques to reduce similar and unnecessary data features, such as ReliefF, Linear Discriminant Analysis (LDA), and Particle Swarm Optimisation (PSO). These studies have reported high accuracy. A study by Tang et al. extracted 515 features from the PhysioNet Challenge 2016 dataset. The study demonstrated that the SVM classifier yielded the most optimal results, with an accuracy of 88%. Likewise, Alshamma et al. compared SVM and KNN algorithms. The study denoted that the KNN algorithm achieved significant results with an accuracy of 93.50%. Nonetheless, accurately identifying heart sounds remains a critical unresolved issue in ML. Figure depicts that applying ML for diagnosing heart sounds in the reviewed studies can be divided into three main processes: pre-processing, feature extraction and selection, and classification.

A. Pre-Processing of The Pcg Signal

The data pre-processing procedure is crucial in training ML-based models for analysing PCG signals. This process encompasses filtering, segmentation, and normalisation methods. The PCG signals are typically recorded at low frequencies, which can be easily affected by various power disturbances. For example, noise can degrade the quality of the features, complicating the diagnosis process. A study by Prasad and Thalluri, observed that PCG signals possessed high sensitivity and low strength, which rendered them susceptible to noise interference. Therefore, filtering methods are necessary to remove unwanted noise from PCG signals. Generally, analogue (or digital)- based signal processing techniques are utilised to remove the noise. This feature suggests that differential amplifiers and filters should be used to prevent interference. For example, the Butterworth filters are the most prevalent techniques employed in literature for noise suppression, such as low bandpass, high band-pass, and band-pass filter. A segmentation process involves dividing PCG signals into four segments: S1, systole, S2, and diastole (see Figure 2). This process extracts specific information for aiding the detection of normal and abnormal heart sounds. Systole is the cardiac cycle phase when the heart muscle contracts, pushing blood out of the chambers into the arteries. The ventricles contract during this process, forcing blood into the pulmonary artery from the right ventricle and into the aorta from the left ventricle. This phase is also measured from the beginning of ventricular contraction to the closure of the aortic and pulmonary valves. Alternatively, diastole is the cardiac cycle phase when the heart muscle relaxes and fills with blood. The ventricles expand during this process and are filled with blood from the atria. This phase is calculated from the closure of the aortic and pulmonary valves to the beginning of ventricular contraction

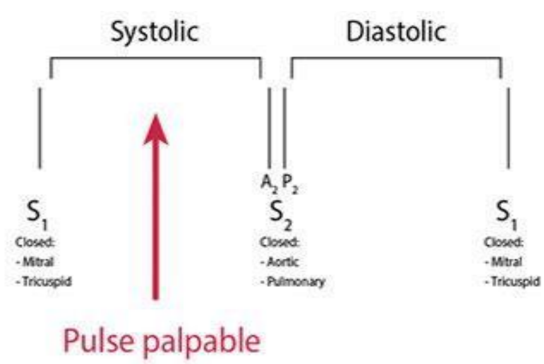

FIGURE. The working principles of S1 and S2.

B. Feature Extraction

The PCG features are independent variables of the heart sounds, functioning as input data for training ML models. Two processes are used for the data reduction process: feature extraction and feature selection. Feature extraction can improve PCG analysis procedures by providing more precise feature characterisation and lowering computing complexity. This strategy also obtains essential and specific information about the functionality and changes in

PCG signals. Given that the use conditions differ, the appropriate features are used depending on the system when analysing PCG signals. A study by Rhif et al. described a signal processing method consisting of two wavelet transform (WT) types: continuous wavelet transform (CWT) and discrete wavelet transform (DWT). Considering that the CWT method generates many wavelet coefficients for each scale, applying the CWT method to non-linear signals is challenging and time-consuming. The DWT method is advantageous due to its smaller number of coefficients. This method can effectively analyse non-stationary signals compared to the CWT method. A study by Touahria et al. suggested that the DWT method was suitable as it provided high classification performance and a more comprehensive representation of the features than the Mel-frequency cepstral coefficient (MFCC) method. According to Potdar utilised the DWT method to extract 152 features from the PhysioNet Challenge 2016 dataset. The extracted features included standard deviation, mean, and energy. Likewise, Son and Kwon investigated the DWTand MFCC methods to extract multiple features from their dataset and classified them using the SVM algorithm. The study achieved an accuracy of 97.90%, a sensitivity of 98.20%, and a specificity of 99.40%. Meanwhile, Li et al. proposed a method for extracting deep features using the wavelet scattering transform method (WST). The study revealed that the technique could extract 498 features from 409 signals collected from the PhysioNet Challenge 2016 dataset. Another study by Li et al. established that extracting deep features by using the CNN extractor was focused on several recent studies. Nevertheless, CNNs could extract abstract features and possess limited explanatory power. However, the study of Li et al. utilised the PhysioNet Challenge 2016 dataset, comprising 3,153 signals, and extracted 497 features using an ensemble feature method that has been presented in. A global average pooling (GAP) layer was then integrated into the CNN algorithm to improve the classification performance for obtaining a global search for features. Consequently, the CNN classifier achieved an accuracy of 86.80%, a sensitivity of 87%, and a specificity of 86.60%. A recent study by Guven and Uysal utilized the MFCC method to extract 33 timefrequency domains from the PhysioNet Challenge 2016 dataset, which comprised 2,435 PCG signals. This method integrates both short (5-second) and long-time-duration features (entire PCG signals) to capture more features related to heart cycle changes over the entire duration of the PCG signals. The study highlights the significance of extracting comprehensive information about heart sounds, aiding in the precise identification and measurement of the qualitative and quantitative features of PCG signals.

C. Feature Selection

Feature selection refers to selecting important features from the original features of the PCG signals. This process reduces the total number of features by eliminating unnecessary features. Given that feature selection eliminates redundant features, it is significant in feature processing. A study by Prasad and Thalluri applied the DWT method to extract features from PCG signals. The study presented that the extracted features included peaksignal-to-noise ratio (PSNR), variance, mean square, entropy, and energy. Global features were selected based on information gain (IG) and mutual information (MI). Another study by Milani et al. employed the MFCC method to extract features from PCG signals collected from the PhysioNet Challenge 2016 dataset. This dataset consisted of approximately 3,126 PCG signals and 130 extracted features. The features also included 30 and 100 features from the time (prominent peaks, amplitude, cycle duration, time interval, intensity, and area) and frequency domains (filter bank energy features), respectively. Various statistical values of these features were then considered, such as median, standard deviation, mean, minimum and maximum. Subsequently, the critical features were selected using the linear discriminant analysis (LDA) method to reduce redundant features and improve the classification performance. A study by Li et al. conducted a comparison with the multi-dimensional scaling (MDS) and principal component analysis (PCA). This selection process resulted in the selection of 12 features. The researchers used the Twin SVM (TWSVM) to improve the classification model further. Consequently, the MDS approach was significantly superior to PCA, achieving an accuracy of 98.58%, a sensitivity of 98.58%, a specificity of 98.57%, and an F1 score of 99%. The study by Potdar utilised the PCA method is used to select the most optimal features. Er performed local binary and ternary methods to extract features from the PhysioNet Challenge 2016 dataset

(2,868 PCG signals). The study then employed the ReliefF method for feature selection. Certain studies have also mentioned that genetic algorithm (GA) and particle swarm optimization (PSO) are exciting methods that are reported with PCG signals to reduce unrelated features and improve performance accuracy. Moreover, the study presented the PSO method with PCG signals of the PhysioNet Challenge 2016 to select global features and mitigate optimization challenges. A recent study by Touahria et al. applied the feature selection method using Joint Mutual Information (JMI).

D.Dataset

Several PCG datasets have been collected and employed in previous cardiac ultrasound analysis studies over the past five years. This review highlighted three datasets: PhysioNet/Computing in Cardiology (CinC) Challenge 2016 dataset consisting of only two classes (normal and abnormal signals), Yaseen Khan 2018 dataset containing two classes (normal and abnormal signals) and several valvular heart disease classes, and PhysioNet Challenge 2022 limited dataset comprising two classes (normal and abnormal signals) and several heart issue classes, such as mitral regurgitation murmurs. Table tabulates the details of the three datasets in detail. The primary issues identified in the literature for the three datasets were noise and class imbalance, especially since these challenges existed in the PhysioNet Challenge 2016 and 2022.

Typically, the class imbalance issue in a dataset occurs when there are significantly more samples for certain classes than others. Hence, the training algorithms under these circumstances are likelier to be overwhelmed by larger classes while neglecting smaller classes. Multiple studies have also investigated the Pascal dataset, and specific methodologies have involved merging different PCG signal datasets, such as the PhysioNet 2016 and College of Michigan Heart Sound and Murmur Library datasets. Moreover, Li et al. employed three datasets: the PhysioNet 2016, Pascal, and Yaseen Khan 2018 datasets.

III. PROPOSED SYSTEM

The proposed system is an AI-based clinical decision support solution designed to predict the likelihood of heart disease using machine learning techniques. The system collects patient health parameters such as age, cholesterol level, blood pressure, heart rate, and other relevant medical attributes. These inputs are processed and analyzed using a supervised machine learning model based on Logistic Regression to determine whether a patient is at risk of heart disease. The architecture of the system consists of four main modules: Data Collection Module, Data Preprocessing Module, Model Training and Prediction Module, and Result Visualization Module. In the Data Collection Module, patient medical data is gathered either from hospital databases, electronic health records, or manual input through a user interface. The Data Preprocessing Module handles missing values, removes inconsistencies, performs feature scaling, and converts categorical data into numerical format to ensure high-quality input for the model. In the Model Training and Prediction Module, the Logistic Regression algorithm is trained using historical patient data. The model learns patterns and relationships between health parameters and heart disease outcomes. Once trained, the system can analyse new patient data and generate predictions indicating the probability of heart disease. Model evaluation techniques such as accuracy, precision, recall, and confusion matrix analysis are used to measure performance and reliability. The Result Visualization Module presents the prediction outcome in a clear and understandable format. It displays whether the patient is at low or high risk along with probability scores. This output assists doctors in making faster and more informed clinical decisions. The system can be deployed as a web-based application or integrated into hospital management systems, making it suitable for hospitals, telemedicine platforms, and rural healthcare centers. Overall, the proposed AI-assisted Heart Disease Prediction System enhances early detection, reduces diagnosis time, minimizes human error, and provides reliable decision support to healthcare professionals.

IV. METHODOLOGY

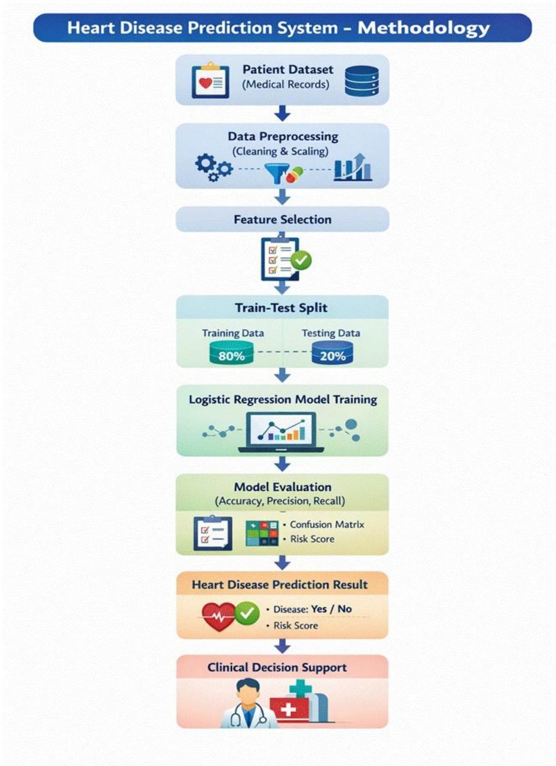

FIG: Heart Disease Prediction system

The proposed Heart Disease Prediction System follows a structured machine learning workflow to ensure accurate and reliable prediction of heart disease. The first step involves data collection, where a structured dataset containing patient medical records is obtained. The dataset includes important health attributes such as age, cholesterol level, blood pressure, heart rate, and other clinical parameters that influence heart disease diagnosis. After data collection, data preprocessing is performed to improve data quality and prepare it for model training. In this stage, missing values are handled, duplicate records are removed, and categorical values (if any) are converted into numerical format. Feature scaling techniques such as normalization or standardization are applied to ensure that all input attributes are on a similar scale, which improves the performance of the Logistic Regression algorithm. Next, the dataset is divided into training and testing sets. The training dataset is used to train the machine learning model, while the testing dataset is used to evaluate the model's performance. The proposed system uses the Logistic Regression algorithm because it is effective for binary classification problems such as predicting whether a patient has heart disease or not. During the model training phase, the Logistic Regression algorithm learns the relationship between input health parameters and the target output (presence or absence of heart disease). The algorithm calculates probability scores and applies a decision threshold to classify patients into risk categories. After training, the model undergoes performance evaluation using metrics such as accuracy, precision, recall, and F1-score. These evaluation metrics help in measuring the reliability and consistency of the prediction system. If necessary, hyperparameter tuning is performed to improve model performance.

Finally, the trained model is integrated into a user-friendly prediction interface, where healthcare professionals can enter patient details and obtain instant prediction results. The system generates diagnostic outputs that support doctors in early detection and decision-making. This structured methodology ensures improved prediction accuracy, reduced diagnosis time, and reliable clinical decision support.

V. MODEL AND IMPLEMENTATION

The proposed Heart Disease Prediction System is developed using a supervised machine learning approach to classify whether a patient is likely to have heart disease or not. The system is built using the Logistic Regression algorithm, which is suitable for binary classification problems in medical diagnosis. The model takes various medical parameters as input and processes them to generate a prediction result. The entire system is designed to assist healthcare professionals by providing faster and more consistent decision support. The model begins with data collection, where patient health records are gathered from a structured dataset. The dataset contains several important medical attributes that influence heart disease prediction. These features serve as input variables for training the machine learning model.

Some of the important input parameters include:

- Age
- Gender
- Blood Pressure
- Cholesterol Level
- Chest Pain Type

- Fasting Blood Sugar
- ECG Results
- Maximum Heart Rate Achieved
- Oldpeak (ST depression)
- Thalassemia

These medical attributes are strongly related to cardiovascular health and are used to determine the probability of heart disease. After collecting the dataset, the next step is data preprocessing. In real world medical data, there may be missing values, inconsistent entries, or duplicate records. Therefore, preprocessing is performed to clean and prepare the dataset before feeding it into the model. This step improves the reliability and accuracy of the prediction system.

The preprocessing stage includes:

- Removing missing or null values
- Eliminating duplicate records
- Converting categorical data into numerical format
- Splitting the dataset into training and testing sets
- Applying feature scaling

Feature scaling is an important step because medical parameters like cholesterol and age have different value ranges. If scaling is not applied, features with larger values may dominate the model. Therefore, Standard Scaler is used to normalize the data so that all features contribute equally to the Logistic Regression algorithm. Once preprocessing is completed, the dataset is divided into two parts: training data and testing data. Generally, 80% of the data is used for training and 20% for testing. The training dataset is used to teach the model patterns between input features and the target variable, while the testing dataset is used to evaluate the model's performance. The Logistic Regression algorithm is then applied to train the model. Logistic Regression works by calculating the weighted sum of input features and passing it through a sigmoid function to generate a probability value between 0 and 1. This probability represents the likelihood of the presence of heart disease.

The working principle of the model can be explained as follows:

- The model receives patient medical inputs.
- It calculates a weighted combination of all input features.
- The sigmoid function converts the result into a probability value.
- If the probability is greater than 0.5, the system predicts "Heart Disease Present."
- If the probability is less than 0.5, the system predicts "Heart Disease Not Present."

After training, the model is tested using unseen data to measure its accuracy and performance. Evaluation metrics such as accuracy score and confusion matrix are used to determine how well the model predicts correctly. High accuracy indicates that the system can reliably assist doctors in diagnosis. Finally, when implemented in a real-time environment, the system allows users to enter patient details through an interface. The trained model processes the input instantly and provides a prediction result along with the probability score. This helps healthcare professionals make faster decisions, especially in hospitals, telemedicine platforms, and rural healthcare centers where specialist doctors may not always be available. Overall, the implemented AI-based heart disease prediction model works as an intelligent decision support system. It reduces diagnosis time, minimizes human errors, and improves early detection of heart disease by analysing medical data efficiently and accurately.

VI. RESULTS AND DISCUSSION

demonstrate that the studies have extensively applied multiple ML and DL-based methods for improving cardiovascular disease detection results using heart sound signals. also reveals that the studies have primarily employed SVM, CNN, and RNN. Subsequently, summarises the criteria for the appropriate model selection process.

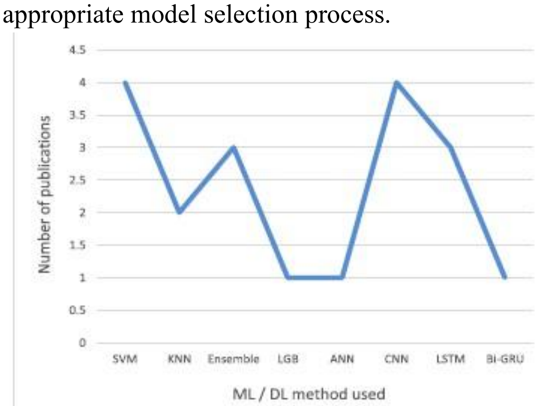

FIGURE. The ML or DL techniques practised in the investigated studies for this review.

The performance metrics were utilised to measure and observe the performance of a model in diagnosing heart sounds during the test phase. These metrics were accuracy, sensitivity, specificity, and F1 score. mathematically defines "True Positive", "False Positive", "True Negative", and "False Negative" as TP, FP, TN, and FN, respectively. Meanwhile, the accuracy rate was compared in the
previous studies

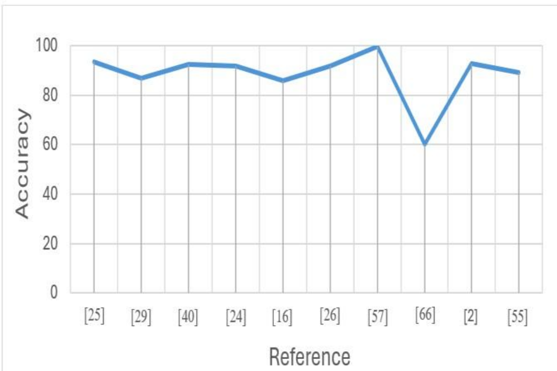

FIGURE. The accuracy rate for the previous studies in this review

Higher accuracy in ML models resulted in more precise predictions or data classifications (see Table 4), which could be highly valuable in various applications (such as medical diagnosis). The imbalanced dataset was also either dominated by positive instances or negative instances. Therefore, the specificity and sensitivity of the model were highly significant when addressing an imbalanced data set. The higher the sensitivity, the higher the TP predictions of the model and the lower the FNs. The F1 score is calculated by integrating precision and sensitivity into a single metric to assess model performance comprehensively. This metric was essential in diagnosing positive cases while minimising FPs and FNs to address imbalanced dataset issues. The class imbalance problem (CIP) is the situation in ML where the classes in the dataset are not equally represented. This phenomenon often occurs when one class (minority class) is significantly less frequent than another class (majority class). Several techniques were presented to effectively handle class imbalance, including data resampling methods like oversampling minority classes, such as synthetic minority over-sampling technique (SMOTE) and under-sampling majority classe].

VII. CONCLUSION

In conclusion, the proposed Heart Disease Prediction System successfully demonstrates the practical application of Artificial Intelligence in the healthcare domain. By utilizing machine learning techniques, particularly the Logistic Regression algorithm, the system effectively analyses key medical parameters such as age, cholesterol level, blood pressure, and heart rate to predict the likelihood of heart disease. The implementation of data preprocessing, feature scaling, and model evaluation methods enhances the accuracy, consistency, and reliability of predictions. The developed system provides a supportive diagnostic tool that assists healthcare professionals in making faster and more informed decisions. It reduces manual errors, saves time in the diagnosis process, and offers early detection support, which is crucial for preventing severe cardiac complications. Moreover, the system has significant potential for use in hospitals, telemedicine platforms, and rural healthcare centers where access to specialized cardiologists may be limited. Overall, this project highlights the transformative role of Artificial Intelligence in modern healthcare by improving clinical decision-making and contributing to better patient outcomes through intelligent data-driven prediction models.

REFERENCES

[1] World Health Organization. (2023). World Health Statistics 2023: Monitoring Health for the SDGs, Sustainable Development Goals. World Health Org. [Online]. Available: https://books.google.com/books?id=0RDAEAAAQBAJ

[2] M. Guven and F. Uysal, "A new method for heart disease detection: Long short-term feature extraction from heart sound data," Sensors, vol. 23, no. 13, p. 5835, Jun. 2023, doi: 10.3390/s23135835.

[3] R. Jaros, J. Koutny, M. Ladrova, and R. Martinek, "Novel phonocardiography system for heartbeat detection from various locations," Sci. Rep., vol. 13, no. 1, pp. 1–16, Sep. 2023, doi: 10.1038/s41598-023-41102-8.

[4] R. Kotha, C. Streitmatter, A. Serdiuk, N. N. Aldawoodi, and R. S. Ackerman, "Cardiac remote monitoring devices and technologies: A

review for the perioperative physician and telemedicine providers," Cureus, vol. 16, no. 2, pp. 1–9, Feb. 2024, doi: 10.7759/cureus.53914.

[5] Q. Liu, X. Wu, and X. Ma, "An automatic segmentation method for heart sounds," Biomed. Eng. OnLine, vol. 17, no. 1, pp. 1–22, Dec. 2018, doi: 10.1186/s12938-018-0538-9.

[6] J. Chen, Z. Guo, X. Xu, G. Jeon, and D. Camacho, "Artificial intelligence for heart sound classification: A review," Expert Syst., vol. 41, no. 4, pp. 1–20, Apr. 2024, doi: 10.1111/exsy.13535.

[7] S. K. Ghosh, P. R. Nagarajan, and R. K. Tripathy, "Heart sound data acquisition and preprocessing techniques: A review," in Handbook of Research on Advancements of Artificial Intelligence in Healthcare Engineering. Hershey, PA, USA: IGI Global, 2020, pp. 244–264, doi: 10.4018/978-1-79982120-5.ch014.

[8] J. Li, L. Ke, Q. Du, X. Ding, X. Chen, and D. Wang, "Heart sound signal classification algorithm: A combination of wavelet scattering transform and twin support vector machine," IEEE Access, vol. 7, pp. 179339–179348, 2019, doi: 10.1109/ACCESS.2019.2959081.

[9] Y. Arjoune, T. N. Nguyen, R. W. Doroshow, and R. Shekhar, "A noise-robust heart sound segmentation algorithm based on Shannon energy," IEEE Access, vol. 12, pp. 7747–7761, 2024, doi: 10.1109/ACCESS.2024.3351570.

[10] S. Aziz, M. U. Khan, M. Alhaisoni, T. Akram, and M. Altaf, "Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features," Sensors, vol. 20, no. 13, p. 3790, Jul. 2020, doi: 10.3390/s20133790.

[11] B. Karan and T. Padhi, "Hilbert domain characterizations of wavelet packets for automated heart sound abnormality detection," Biomed. Signal Process. Control, vol. 90, Apr. 2024, Art. no. 105793, doi: 10.1016/j.bspc.2023.105793.

[12] O. Alshamma, F. H. Awad, L. Alzubaidi, M. A. Fadhel, Z. M. Arkah, and L. Farhan, "Employment of multi-classifier and multi-domain features for PCG recognition," in Proc. 12th Int. Conf. Develop. eSystems Eng. (DeSE), Oct. 2019, pp. 321– 325, doi: 10.1109/DESE.2019.00066.

[13] J. A. M. Sidey-Gibbons and C. J. Sidey-Gibbons, "Machine learning in medicine: A practical introduction," BMC Med. Res. Methodol., vol. 19, no. 1, pp. 1–18, Dec. 2019, doi: 10.1186/s12874-019-0681-4.

[14] M. Chowdhury, K. Poudel, and Y. Hu, "Detecting abnormal PCG signals and extracting cardiac information employing deep learning and the Shannon energy envelope," in Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB), Dec. 2020, pp. 1–4, doi: 10.1109/SPMB50085.2020.9353624.

[15] Q. Abbas, A. Hussain, and A. R. Baig, "Automatic detection and classification of cardiovascular disorders using phonocardiogram and convolutional vision transformers," Diagnostics, vol. 12, no. 12, p. 3109, Dec. 2022, doi: 10.3390/diagnostics12123109.

[16] D. Chen, W. Xuan, Y. Gu, F. Liu, J. Chen, S. Xia, H. Jin, S. Dong, and J. Luo, "Automatic classification of normal–abnormal heart sounds using convolution neural network and long-short term memory," Electronics, vol. 11, no. 8, p. 1246, Apr. 2022, doi: 10.3390/electronics11081246.

[17] S. K. Ghosh, R. N. Ponnalagu, R. K. Tripathy, and U. R. Acharya, "Automated detection of heart valve diseases using Chirplet transform and multiclass composite classifier with PCG signals," Comput. Biol. Med., vol. 118, Mar. 2020, Art. no. 103632, doi: 10.1016/j.compbiomed.2020. 103632.

[18] W. Chen, Q. Sun, X. Chen, G. Xie, H. Wu, and C. Xu, "Deep learning methods for heart sounds classification: A systematic review," Entropy, vol. 23, no. 6, p. 667, May 2021, doi: 10.3390/e23060667.

[19] Q. Zhao, "Deep learning for heart sound analysis: A literature review," medRxiv, pp. 1–17, Sep. 2023, doi: 10.1101/2023.09.16.23295653.

[20] S. M. Debbal and C. Hamza, "Heart sounds analysis using the three wavelet transform versions the continuous wavelet transforms (CWT), the discrete wavelet transforms (DWT) and the wavelet packet transforms (PWT)," J. Cardiol. Interv., vol. 1, no. 1, pp. 1–9, Jan. 2021.