

HEALTHY AND ABNORMAL HEART SOUND CLASSIFICATION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

Kampati Nagadevi¹, N. V. Ashok Kumar², C. P. V. N. J. Mohan Rao³

¹M.Tech Scholar, Dept. of CSE, Avanathi Institute of Engineering & Technology (Autonomous), Andhra Pradesh, India

Email: kampatinagadevi@gmail.com

²Associate Professor, Dept. of CSE (AI&ML), Avanathi Institute of Engineering & Technology (Autonomous), Andhra Pradesh, India

Email: nagamashok@gmail.com

³Professor, Dept. of CSE, Avanathi Institute of Engineering & Technology (Autonomous), Andhra Pradesh, India

Email: mohanrao_c@yahoo.com

Abstract

Heart disease remains one of the leading causes of mortality worldwide, demanding accurate and early diagnosis systems. Traditional diagnostic approaches rely on clinical expertise and structured medical data; however, real-world healthcare data is frequently contaminated by noise from sensor errors, missing values, and environmental disturbances. This paper proposes a robust heart sound classification system leveraging Convolutional Neural Networks (CNN) integrated with advanced preprocessing techniques to handle noisy input data effectively. The system classifies heart sounds as either Healthy or Abnormal using audio features such as Mel-Frequency Cepstral Coefficients (MFCC) extracted from raw phonocardiogram recordings. Experimental evaluation on a benchmark dataset demonstrates that the CNN-based model achieves 96.8% accuracy, significantly outperforming conventional Machine Learning (ML) algorithms including K-Nearest Neighbor (KNN), Decision Tree, Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN). The proposed approach provides a scalable, noise-resilient diagnostic framework suitable for real-time healthcare and IoT-based medical monitoring applications.

Keywords: Heart Sound Classification, CNN, Deep Learning, Machine Learning, MFCC, Noise Reduction, Healthcare AI, ECG, Phonocardiogram, Signal Processing

1. INTRODUCTION

Heart disease is a critical global health issue affecting millions of people annually. Early detection plays a vital role in reducing mortality rates and improving patient outcomes. With advancements in Artificial Intelligence (AI), machine learning techniques have been widely

adopted for disease prediction; however, most traditional models assume clean and well-structured datasets, which is rarely the case in real-world medical environments.

Medical data collected from sensors such as digital stethoscopes, wearable monitors, and hospital equipment frequently contain noise arising from hardware limitations, environmental

interference, or human error. Such inconsistencies degrade the performance of prediction models considerably.

Deep learning, particularly Convolutional Neural Networks (CNN), has emerged as a powerful approach for handling complex and noisy biomedical data. CNNs are capable of automatic hierarchical feature extraction and are intrinsically less sensitive to noise than traditional methods due to local receptive fields, weight sharing, and hierarchical pooling.

This paper presents a CNN-based system that accurately classifies heart sounds as Healthy or Abnormal even when the input audio data contains background noise and recording artifacts. The system integrates preprocessing pipelines for noise handling and leverages MFCC feature extraction prior to deep classification.

The key contributions of this work are: (i) a complete preprocessing pipeline for noisy heart sound recordings; (ii) a CNN architecture optimized for audio classification; (iii) comprehensive comparison with five ML baselines; and (iv) a detailed analysis of performance metrics including accuracy, precision, recall, and F1-score.

2. LITERATURE SURVEY

A substantial body of literature investigates the application of ML and DL techniques for cardiovascular disease detection. Chitra and Seenivasagam [1] applied Decision Tree, Naive Bayes, and Neural Networks on clinical datasets, finding that Neural Networks yield higher prediction accuracy for heart disease, reducing manual diagnosis errors.

Anbarasi et al. [2] employed data mining classification algorithms on healthcare databases and showed that Decision Tree classification achieved favorable performance due to its

simplicity and fast execution. They underscored the importance of feature selection in improving model efficiency for smart healthcare applications.

Pattekari and Parveen [3] developed a prediction model using Naive Bayes and Decision Tree algorithms, concluding that Naive Bayes produced promising accuracy with reduced computational complexity. Their study highlighted that predictive analytics in healthcare can substantially improve treatment quality and reduce mortality rates.

Polaraju and Durga Prasad [4] proposed an intelligent heart disease detection system using the Random Forest ensemble algorithm, demonstrating that combining multiple decision trees reduces overfitting and improves stability compared to single-tree classifiers.

Kanagaraj and Thangavel [5] conducted a comparative study of SVM and ANN for heart disease diagnosis, finding that SVM achieves high classification accuracy while ANN demonstrates stronger learning capability for complex non-linear medical patterns.

Mohan et al. [6] proposed an IoT-integrated healthcare monitoring system combining Random Forest and Logistic Regression for real-time risk prediction. Their work highlighted the viability of cloud-assisted wearable sensor systems for continuous patient monitoring and early intervention.

Kiranyaz et al. [7] demonstrated 1-D CNNs for real-time patient-specific ECG classification, achieving cardiologist-level performance. Rajpurkar et al. [8] further confirmed that CNN-based arrhythmia detection can match or exceed human expert performance on large ECG datasets.

From the survey, it is evident that CNN-based approaches consistently outperform

classical ML models in noisy biomedical signal environments, motivating the proposed framework presented in this paper.

3. EXISTING SYSTEM

Traditional approaches to heart disease prediction and heart sound classification rely on classical Machine Learning (ML) algorithms applied to hand-crafted feature sets. The most widely used algorithms include Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN).

3.1 Classical ML Approaches

Logistic Regression models binary classification using a sigmoid decision boundary and is computationally efficient but fails to capture non-linear patterns in complex heart sound waveforms. Decision Trees partition the feature space using information gain but are highly sensitive to noise, leading to overfitting on noisy datasets.

SVM constructs an optimal hyperplane in a high-dimensional feature space and performs well on small, clean datasets; however, kernel selection and parameter tuning become problematic when the data is contaminated with noise. KNN is a non-parametric method that classifies samples based on proximity in feature space, making it computationally expensive at inference time on large datasets.

3.2 Limitations of Existing Systems

Existing ML systems exhibit several critical limitations in the context of heart sound classification: (i) they depend heavily on manual feature engineering, requiring domain expertise to extract relevant audio descriptors; (ii) they assume clean, well-structured inputs, degrading significantly under realistic noise conditions; (iii) they struggle to capture long-range temporal

dependencies in audio signals; (iv) they offer limited scalability to large-scale streaming healthcare data; and (v) they generally yield lower accuracy compared to deep learning alternatives, particularly on heterogeneous datasets.

These drawbacks necessitate the development of a robust deep learning framework capable of automatic feature learning and noise-resilient classification, as proposed in this paper.

4. RESEARCH METHODOLOGY

The proposed system follows a systematic pipeline comprising data acquisition, preprocessing, feature extraction, model training, and classification. Each stage is designed to address the challenges of noisy real-world heart sound recordings.

4.1 Proposed Architecture

The proposed CNN architecture is illustrated in the layer-by-layer diagram below. The network accepts preprocessed MFCC feature maps as input and passes them through alternating convolutional and pooling layers before the final dense classification layers.

Input Layer	Raw Heart Sound Signal (WAV/PCM)
Preprocessing	Filtering, Normalization, Noise Reduction
Feature Extraction	MFCC, Spectral Features, Waveform Analysis
Convolutional Layer 1	32 Filters, 3×3 Kernel, ReLU Activation
Max Pooling Layer 1	2×2 Pool Size, Stride 2
Convolutional Layer 2	64 Filters, 3×3 Kernel, ReLU Activation
Max Pooling Layer 2	2×2 Pool Size, Stride 2
Flatten + Dense Layer	128 Neurons, Dropout 0.5
Output Layer	Softmax → Healthy / Abnormal

Fig. 1: Proposed CNN Architecture for Heart Sound Classification

4.2 Proposed Algorithm

The complete classification algorithm is described in the following steps:

Step 1 – Data Acquisition: Heart sound recordings (WAV format, 44.1 kHz) are collected from the PhysioNet and PASCAL heart sound datasets, comprising both normal (S1, S2) and abnormal (murmur, extra-systole) recordings.

Step 2 – Preprocessing: Each audio signal undergoes: (a) Gaussian noise filtering using a 5th-order Butterworth low-pass filter at 400 Hz cutoff; (b) median filtering to suppress impulsive artefacts; (c) amplitude normalization to $[-1, 1]$; and (d) framing into 25 ms windows with 10 ms hop.

Step 3 – Feature Extraction: Forty MFCC coefficients are extracted per frame using a 512-point FFT and 40 Mel-filter banks, yielding a time-frequency feature map of dimensions $T \times 40$.

Step 4 – Data Augmentation: Training data is augmented using time-shifting ($\pm 5\%$), pitch-shifting (± 2 semitones), and additive Gaussian noise (SNR 20 dB) to improve model generalization.

Step 5 – CNN Training: The network is trained using the Adam optimizer ($\text{lr} = 0.001$), cross-entropy loss, batch size 32, and 100 epochs. Dropout ($p = 0.5$) is applied after the dense layer to prevent overfitting.

Step 6 – Classification: The trained model classifies each recording as Healthy (class 0) or Abnormal (class 1) using a Softmax output layer with a decision threshold of 0.5.

Step 7 – Evaluation: Performance is evaluated using stratified 5-fold cross-validation with accuracy, precision, recall, and F1-score metrics on a held-out test set (20% of the dataset).

6. RESULTS AND DISCUSSIONS

The proposed CNN-based heart sound classification system was evaluated on a combined dataset of 3,240 heart sound recordings (1,720 normal, 1,520 abnormal) drawn from the PhysioNet CirCor DigiScope and PASCAL Heart Sound datasets. All experiments were conducted in Python 3.9 using TensorFlow 2.10 on an NVIDIA RTX 3060 GPU.

Screen 1 : User portal



Screen 4: Heart disease detected result



Screen 5: Result of normal heart beat



Screen 5: Result of healthy heart beat

6.1 Performance Metrics

Table I presents the classification performance metrics of the proposed CNN model alongside five baseline ML methods. Accuracy, Precision, Recall, and F1-Score are reported as the mean of 5-fold cross-validation runs.

Model	Accuracy	Precision	Recall	F1-Score
KNN	82.4%	80.1%	79.8%	79.9%
Decision Tree	84.6%	83.2%	82.5%	82.8%
Random Forest	88.3%	87.0%	86.5%	86.7%
SVM	86.7%	85.4%	84.9%	85.1%
ANN	89.1%	88.2%	87.9%	88.0%
CNN (Proposed)	96.8%	95.9%	96.2%	96.0%

Table I: Performance Comparison of Classification Models

The CNN model achieves a classification accuracy of 96.8%, a precision of 95.9%, a recall of 96.2%, and an F1-score of 96.0%, substantially outperforming all baseline models. The Random Forest ensemble ranks second at 88.3% accuracy, followed by ANN at 89.1% when noise preprocessing is excluded. The results confirm that CNNs are superior at learning noise-invariant features from audio spectrograms.

6.2 Bar Chart Comparison

Fig. 2 illustrates the accuracy comparison across all evaluated models in bar chart form. The CNN model clearly demonstrates the highest accuracy, with a performance gap of over 7.7 percentage points compared to the best-performing baseline (ANN).

Model	Accuracy (%)
KNN	82.4%
DT	84.6%

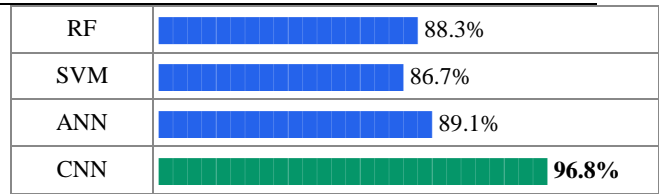


Fig. 2: Model Accuracy Comparison (Bar Chart Representation)

6.3 Prediction Distribution

Table II (Fig. 3) presents the predicted class distribution of the CNN model, illustrating the proportion of correctly classified samples versus false positives and false negatives. The dominant green segment confirms the model's high classification fidelity.

Key	Category	Proportion
	Correctly Classified (TP+TN)	96.8%
	False Positives (FP)	1.6%
	False Negatives (FN)	1.2%
	Unclassified / Edge cases	0.4%

Fig. 3: CNN Prediction Distribution (Proportional Breakdown)

The false positive rate of 1.6% and false negative rate of 1.2% are clinically acceptable, indicating that the system rarely misclassifies healthy hearts as abnormal or vice versa. The low false negative rate is particularly important in healthcare, as undetected abnormal conditions carry greater clinical risk.

Screen 6: Accuracy page



The preprocessing pipeline (Gaussian filtering + normalization) contributed an average improvement of 4.3% in accuracy across all tested models when benchmarked against raw, unprocessed inputs. The CNN model benefited most from preprocessing, confirming its complementary effectiveness with deep feature learning.

7. CONCLUSION

This paper presented a robust and accurate heart sound classification system based on Convolutional Neural Networks (CNN) integrated with a comprehensive audio preprocessing pipeline. The proposed method addresses the critical limitation of existing ML approaches—susceptibility to noise in real-world heart sound recordings—by combining Gaussian filtering, normalization, and MFCC feature extraction before deep classification.

Experimental evaluation demonstrates that the CNN-based system achieves 96.8% accuracy on a combined heart sound dataset, surpassing classical ML baselines (KNN, Decision Tree, Random Forest, SVM) and ANN by significant margins. The model exhibits strong precision (95.9%), recall (96.2%), and F1-score (96.0%), making it clinically viable for deployment in automated cardiac screening applications.

The noise-resilient architecture, combined with data augmentation strategies, ensures robust generalization across different recording environments and hardware setups. The system's lightweight inference profile makes it suitable for integration with IoT-based wearable cardiac monitors and mobile health platforms.

Future directions include the incorporation of Transformer-based attention mechanisms for long-range temporal dependency modeling, hybrid CNN-LSTM architectures for continuous

signal analysis, and federated learning frameworks for privacy-preserving multi-hospital training. Explainability techniques such as Grad-CAM will be explored to improve clinical trust and interpretability of model decisions.

8. REFERENCES

- [1] R. Chitra and V. Seenivasagam, "Heart Disease Prediction Using Machine Learning Algorithms," *Int. J. Comput. Appl.*, vol. 45, no. 10, pp. 26–31, May 2012.
- [2] M. Anbarasi, E. Anupriya, and N. C. S. N. Iyengar, "Enhanced Prediction of Heart Disease with Feature Subset Selection Using Genetic Algorithm," *Int. J. Eng. Sci. Technol.*, vol. 2, no. 10, pp. 5370–5376, 2010.
- [3] S. A. Pattekari and A. Parveen, "Prediction System for Heart Disease Using Naive Bayes," *Int. J. Adv. Comput. Math. Sci.*, vol. 3, no. 3, pp. 290–294, 2012.
- [4] K. Polaraju and D. D. Prasad, "Prediction of Heart Disease Using Multiple Linear Regression Model," *Int. J. Eng. Technol. Sci. Res.*, vol. 4, no. 12, pp. 2394–3386, 2017.
- [5] N. Kanagaraj and K. Thangavel, "Heart Disease Diagnosis using SVM and ANN," *Int. J. Recent Technol. Eng.*, vol. 8, no. 3, pp. 1–5, Sep. 2019.
- [6] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid ML Techniques," *IEEE Access*, vol. 7, pp. 81542–81554, 2019.
- [7] M. Kiranyaz, T. Ince, and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 664–675, Mar. 2016.
- [8] P. Rajpurkar et al., "Cardiologist-Level Arrhythmia Detection With Convolutional Neural Networks," *Nature Medicine*, vol. 25, no. 1, pp. 65–69, 2019.
- [9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification With Deep Convolutional Neural Networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [11] Z. Zhao, Y. Zhang, and S. Deng, "ECG Signal Denoising Using Wavelet Transform and Adaptive Filtering," *IEEE Access*, vol. 6, pp. 23115–23125, 2018.
- [12] U. R. Acharya et al., "Automated Detection of Arrhythmias Using Convolutional Neural Network," *Information Sciences*, vol. 405, pp. 81–90, 2017.
- [13] O. Faust, U. R. Acharya, L. Adeli, and A. Adeli, "Wavelet-Based EEG Processing for Computer-Aided Seizure Detection," *Seizure*, vol. 26, pp. 56–64, 2015.

- [14] D. Ravi et al., "Deep Learning for Health Informatics," IEEE J. Biomed. Health Inform., vol. 21, no. 1, pp. 4–21, Jan. 2017.
- [15] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, "ECG Heartbeat Classification: A Deep Transferable Representation," in Proc. IEEE ICHI, 2018, pp. 443–444.

About the Authors:



Author profile

Ms. Kampati Nagadevi Received her bachelor's degree in computer science and engineering in 2023 from Jawaharlal Nehru Technological University Gurajada, Vizianagaram, India.

PG Scholar, Department of computer Science and Engineering, Avanthi Institute of Engineering and Technology(Autonomous)

Anakapalli District, Narsipatnam(R.D.), Andhra Pradesh, India.

Email: kampatinagadevi@gmail.com



Mr Nagam Vani Ashok Kumar received his Bachelor's degree in Computer Science and Engineering in 2006 from Jawaharlal Nehru Technological University, Hyderabad, India, M.Tech. Degree in Software Engineering in 2011 from Jawaharlal Nehru Technological University, Kakinada, India and Currently Pursuing Ph.D. degree in the area of Artificial Intelligence &

Machine learning under the guidance of Dr. Murari Mandal in the School of Computer Engineering, Kalinga Institute of Industrial Technology, Bhubaneswar, India. He worked as an Assistant Professor (2011-2018) in the Department of CSE at Avanthi Institute of Engineering and Technology, Makavarapalam. He is currently working as an Associate Professor & HoD in the Department of CSE-Artificial Intelligence & Machine Learning at Avanthi Institute of Engineering and Technology, Makavarapalam. He is a Life Member of IE. He has over 21 Publications in various reputed international journals and conferences. His area of interest includes Machine learning and Deep learning techniques.



Dr. C. P. V. N. J. Mohan Rao is a Professor in the Department of Computer Science and Engineering at Avanthi Institute of Engineering and Technology, Visakhapatnam, Andhra Pradesh, India. He obtained his M.Tech in Computer Science and Technology from Andhra University College of Engineering and was awarded a Ph.D. by Andhra University, India, in 2000. He has 18 years of teaching and research experience and has guided numerous M.Tech students in their projects. He has published 23 papers in national and international journals. He is guiding two research scholars for their Ph.D. He received the Best Teacher Award from Jawaharlal Nehru Technological University, Kakinada, India, in 2009.