Heart Sound Classification using Weka for Machine Learning

Tanner Mergist
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
tanner.mergist1@louisiana.edu

T. Maheswara Reddy Yenumula
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
C00538416@louisiana.edu

Rafeeq Muhammad
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
rafeeq.muhammad1@louisiana.edu

Tytiana James
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
tytiana.james1@louisiana.edu

Ali Harimi
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
ali.harimi1@louisiana.edu

Kasem Khalil
Electrical and Computer Engineering
University of Mississippi
Mississippi, USA
kmkhalil@olemiss.edu

Ashok Kumar
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, USA
ashok.kumar@louisiana.edu

SoutheastCon 2025 | 979-8-3315-0484-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/SOUTHEASTCON56624.2025.10971716

***Abstract*—The accurate and efficient classification of heart sounds represents a crucial challenge in cardiac care, with significant implications for early diagnosis and timely intervention. This study investigates the application of machine learning techniques, utilizing the Weka framework, to address this problem and used PhysioNet dataset, which comprises 3,240 heart sound recordings, and extracted 238 features to capture the essential spectral and temporal characteristics of heartbeats. Considering the high dimensionality of the dataset, various feature selection methods were evaluated, including information gain ranking, Principal Component Analysis (PCA), and Recursive Feature Elimination (RFE). Ultimately, the InfoGain algorithm was applied, selecting the 50 most significant features from the initial set of 238. The performance of five machine learning models—Multilayer Perceptron, Random Forest, Decision Tree, k-Nearest Neighbor (k-NN), and Support Vector Machine (SVM) — on both the full and reduced feature sets, using 80-20 split and 10-fold cross-validation was evaluated. The results demonstrate that the Random Forest model achieved the highest accuracy of 92.1% on the full feature set, while the k-NN model exhibited superior computational efficiency on the reduced set, with an accuracy of 89.4% and significantly faster training times. These findings highlight the trade-offs between model complexity and efficiency, providing valuable insights for clinicians and researchers in selecting appropriate algorithms for real-world heart sound classification applications.**

***Index Terms*—Heartbeat Classification , Feature Selection, Machine Learning, PhysioNet Dataset, Medical Diagnostics, Signal Processing,**

I. Introduction

Cardiovascular diseases are the leading global cause of mortality, responsible for an estimated 17.9 million deaths annually [1]. Early diagnosis of heart conditions is crucial to effective treatment and prevention of these life-threatening diseases [2]. Conventional diagnostic methods, such as electrocardiograms and echocardiograms, often require invasive procedures, a large investment in time, and specialized medical expertise, making them less accessible and efficient for widespread cardiac evaluation [3]. In this context, the analysis of heart sounds has emerged as a promising, non-invasive, and cost-effective approach to cardiac assessment.

Recent advances in machine learning have demonstrated the remarkable potential of automated heart sound classification techniques, allowing the detection of various heart conditions, including murmurs, arrhythmias, and valve abnormalities [4]. A popular open-source machine learning framework, has been extensively utilized in the field of biomedical analysis due to its extensive collection of robust algorithms and user-friendly interface, which makes it a particularly attractive choice for researchers and clinicians seeking to develop and deploy effective cardiac diagnostic tools [5].

This study stands out through its comprehensive assessment and comparison of five widely-adopted machine learning algorithms - Multilayer Perceptron, Random Forest, Decision Tree, k-NN, and Support Vector Machine - for the task of heart sound classification using the Weka framework. The models were assessed based on their classification accuracy and computational efficiency, with the goal of identifying the most suitable algorithm for accurate and efficient heart sound classification. Ultimately, this research provides a robust and optimized machine learning-based approach for the early and accurate detection of heart conditions, which can lead to timely treatment and prevention of cardiovascular disease.

Additionally, feature selection techniques were leveraged to optimize the models' performance and efficiency, offering valuable insights into the balance between model complexity and effectiveness.

The dataset used in this study was obtained from the PhysioNet repository, which contains 3,240 heart sound recordings from both healthy individuals and those with various cardiac conditions. This research involved a structured pipeline of preprocessing, feature extraction, feature selection, and classification, utilizing Weka's comprehensive tool set to evaluate the performance of the models.

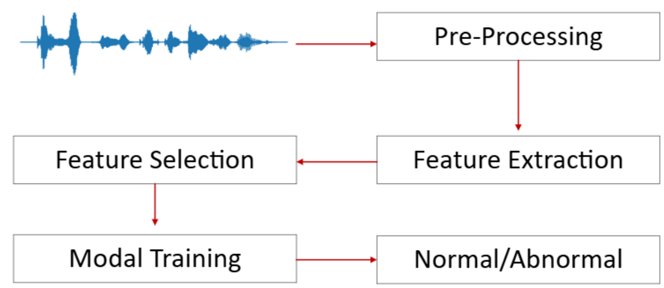

Fig. 1. Classification Pipeline

II. Literature Review

Recent advancements in medical big data and artificial intelligence have driven significant progress in the development of machine learning and deep learning methods for heart sound classification. Traditional machine learning algorithms, such as Support Vector Machine, k-Nearest Neighbors, Decision Tree, and Logistic Regression, have been extensively explored in previous studies to analyze heart sounds and detect cardiac abnormalities [6]. These studies leveraged a variety of time-domain, frequency-domain, and wavelet-based features to train and evaluate model performance, demonstrating the potential of ML-based classification systems in cardiovascular disease diagnosis [7][8].

Feature selection plays a critical role in optimizing ML models by reducing dimensionality, improving generalization, and minimizing computational costs. Prior research has applied various feature selection techniques, including PCA [9], Recursive Feature Elimination [10], and Information Gain [11]. PCA has been widely employed to transform high-dimensional data into lower-dimensional components while preserving maximum variance, while RFE iteratively removes less relevant features to enhance model performance. Information Gain, on the other hand, ranks features based on their relevance to classification, making it particularly useful for interpretability in medical applications. However, comparative studies analyzing the impact of different feature selection techniques on heart sound classification remain limited.

Beyond traditional ML, deep learning models have also shown promising results in heart sound classification. Convolutional Neural Networks and Recurrent Neural Networks have been developed to effectively distinguish between normal and abnormal heart sounds by automatically extracting high-level features from raw waveforms [12]. CNNs, with their spatial feature extraction capabilities, have been successfully applied to Mel spectrograms and time-frequency representations, while RNN-based architectures, such as Long Short-Term Memory networks, have demonstrated strength in capturing sequential dependencies in heart sound signals. Despite their high classification performance, deep learning models often require large amounts of labeled data and significant computational resources, making their deployment challenging in resource-constrained environments.

For instance, The PhysioNet Challenge 2022 presented a multi-task learning model that detected heart murmurs and predicted clinical outcomes. This approach divided heart sound recordings into 3-second segments and extracted features in the time and frequency domains. While the model achieved a weighted accuracy of 69.4%, it exhibited strong performance in cost-sensitive scenarios[13]. Another study developed a real-time phonocardiogram classification system using a convolutional neural network and long short-term memory network. This model processed raw signals directly and attained 86% accuracy, 87% sensitivity, and 89% specificity, making it suitable for real-time applications[14]. Additionally, a hybrid method combining feature extraction and deep learning achieved an accuracy of 86.8%[15].

Furthermore, hybrid approaches combining feature selection and deep learning have gained attention in recent years. Studies have explored integrating PCA with CNNs, or using mutual information-based feature selection before training ML models to enhance efficiency [16]. However, limited research has systematically compared different feature selection techniques in combination with traditional ML models.

The existing literature underscores the potential of Weka-based machine learning platforms for effective and scalable heart sound classification, while highlighting the need for a comprehensive evaluation of feature selection strategies and their impact on model performance.

III. Methodology

A. Dataset

The dataset used in this study was obtained from the PhysioNet repository[17], which contains 3,240 heart sound recordings from both healthy individuals and those with various cardiac conditions. The diversity of the dataset supports the development of models that can generalize effectively, which is crucial in healthcare applications.

B. Data Preprocessing

Several preprocessing steps were applied to ensure the recordings were uniform and ready for feature extraction:

- **Normalization:** The audio files were amplitude-normalized to ensure consistent loudness across recordings, reducing variability that could otherwise mask critical acoustic characteristics.
- **Resampling:** To maintain consistency, all recordings were resampled to a common rate of 22,050 Hz. This

sample rate was chosen to balance the need for detailed frequency information with computational efficiency, ensuring compatibility with standard audio-processing tools.
- **Noise Reduction:** Background noise can interfere with subtle acoustic details. To mitigate this, filtering was applied to reduce unwanted noise, ensuring that essential sound patterns remained intact for analysis.

C. Feature Extraction

The extraction of features is a critical component in machine learning-based heart sound classification. Using Python's Librosa library, we extracted 238 features from each recording, encompassing both spectral and temporal characteristics of the heartbeat sounds.

1) Spectral Features:

- **Mel-Frequency Cepstral Coefficients**: These coefficients capture the spectral properties of the audio, reflecting the perception of the human ear, and are particularly useful for distinguishing the timbre and quality of heart sounds.
- **Formants**: These resonance frequencies represent the shape and quality of the sound within specific frequency ranges, providing insights into potential abnormalities.

2) Temporal Features:

- **Pitch**: Pitch is a valuable indicator of the tonal qualities of heartbeats and can highlight differences between normal and abnormal patterns.
- **RMS Energy**: This feature measures the intensity or power of each heartbeat, which can be a crucial distinguishing factor in identifying structural or functional abnormalities.

To further enrich the feature set, 14 statistical descriptors were computed for each feature, including metrics such as maximum, minimum, mean, range, and percentiles. This comprehensive approach aimed to capture the nuances within the heart sound data, enabling the models to detect subtle variations and enhance the accuracy of the classification process.

D. Feature Selection

The selection of relevant features is crucial for optimizing machine learning models in heart sound classification tasks. This study explored three feature selection methods: Principal Component Analysis, Recursive Feature Elimination, and Information Gain.

- **PCA**: PCA is a widely-used dimensionality reduction technique that transforms features into a new set of uncorrelated principal components. However, PCA may not be ideal for medical applications, as it can obscure the interpretability of the original features, which is important for understanding the significance of the selected features [18].
- **RFE**: RFE is an iterative approach that removes the least important features while retraining the model at each step. While effective in many scenarios, RFE can be computationally expensive, particularly when dealing with high-dimensional datasets like heart sound signals[19].

Attribute	Info Gain Score
mfcc_3_mean	0.79202
mfcc_2_std	0.79202
pitch_kurtosis	0.70105
mfcc_2_min	0.70105
formant_f3_rms	0.70105
pitch_rms	0.43199
mfcc_2_range	0.43199
formant_f1_mean	0.43199
formant_f3_mean	0.43199
mfcc_4_mean	0.34668
mfcc_2_rms	0.34668
formant_f2_90th_percentile	0.34668
formant_f1_std	0.34668
formant_f1_rms	0.34668
formant_f3_50th_percentile	0.32334
mfcc_3_rms	0.32334
formant_f1_max	0.32334
mfcc_3_min	0.29564
formant_f1_75th_percentile	0.29564
mfcc_3_max	0.27737
mfcc_10_percentile_50	0.20657
mfcc_10_mean	0.20657

Fig. 2. Info Gain Scores of Some Features

- **Information Gain**: In contrast, Information Gain ranks features based on their contribution to reducing entropy in the classification task. This approach allows for better interpretability and efficiency [20].

The Information Gain for a feature F is given by:

$$IG(F) = H(Y) - H(Y|F)$$

where $H(Y)$ represents the entropy of the class labels and $H(Y|F)$ is the conditional entropy given feature F.

The entropy $H(Y)$ is defined as:

$$H(Y) = -\sum_{i=1}^{n} P(y_i) \log_2 P(y_i)$$

where $P(y_i)$ represents the probability of class y_i in the dataset.

After comparing different methods, the Information Gain approach was chosen as the best way to select features. IG was able to keep features that are easy to understand from a medical perspective, while still maintaining the model's performance. Plus, a test running PCA, RFE, and IG showed that IG led to only a small drop in accuracy, but a big reduction in how long it takes to train and use the model. This makes IG the most practical choice for real-time applications.

This feature selection step trimmed the number of features from 238 down to 50 and some of the features shown in figure 2, without much impact on accuracy. This made the model a lot faster to train and use, which is important for real-world use.

E. Model Training and Evaluation

This study evaluates the effectiveness of traditional machine learning models in classifying heartbeat sounds. The selected models—Random Forest, k-Nearest Neighbors, Support Vector Machine, Decision Tree, and Multilayer Perceptron—were chosen based on their established success in biomedical signal processing and their ability to balance accuracy and computational efficiency. Deep learning models were not considered due to their high computational cost and the need for large datasets, which may not be feasible in resource-constrained environments.

1) Experimental Setup: The dataset was partitioned into training and testing subsets, with an 80:20 ratio allocation. The models were trained on the training subset and their performance was evaluated using the held-out testing subset. To further enhance the robustness of the evaluation, a 10-fold cross-validation approach was employed. This involved dividing the dataset into 10 equal subsets, where each subset was used as the testing set once, while the remaining 9 subsets were utilized for training. This cross-validation procedure helps to mitigate biases arising from data distribution and ensures that the results are generalizable across different partitions of the dataset.

Performance Metrics: The metrics used were

- Accuracy: the overall proportion of correct predictions, served as the primary evaluation metric.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

  where TP and TN are the correctly classified positive and negative samples, respectively, while FP and FN are the misclassified samples.
- Precision: the ratio of true positive predictions to total positive predictions, gauged the models' ability to correctly identify positive cases.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

- Recall: the ratio of true positive predictions to actual positive instances, measured the models' sensitivity in detecting positive cases.

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

- F-measure: the harmonic mean of precision and recall, provided a balanced metric that accounted for both precision and recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

This combination of metrics offered a thorough assessment of the models' performance, capturing both their overall accuracy and their capacity to handle the nuances of the heart sound classification task.

IV. Results and Discussion

The results of this study demonstrate the impact of feature selection and dimensionality reduction on the performance of machine learning models for heartbeat sound classification. The models were evaluated based on accuracy, precision, recall, F1-score, training time, and inference time to assess their classification performance and computational efficiency.

Among the evaluated models from table II and IV, Random Forest and Multilayer Perceptron achieved the highest accuracy on the full feature set, reaching 92.1% and 90.5%, respectively. These outcomes suggest that Random Forest and the complex pattern recognition capabilities of Neural Networks are advantageous for handling high-dimensional data, allowing these models to effectively distinguish between normal and abnormal heartbeats.

TABLE I
Model Accuracies Using 80-20 Train-Test Split.

Model	Precision	Recall	F1-Score	Accuracy
Random Forest	90.7%	89.6%	90.1%	91.2%
k-NN	87.1%	85.4%	86.2%	87.6%
SVM	86.5%	84.2%	85.3%	86.4%
Decision Tree	82.3%	80.1%	81.2%	82.7%
MLP	88.2%	86.5%	87.3%	88.7%

TABLE II
Accuracies Using 80-20 Train-Test Split for Full and Reduced Feature Sets

Model	Accuracy (Full Set 238)	Accuracy (Reduced Set 50)
Random Forest	91.2%	89.9%
k-NN	87.6%	86.7%
SVM	86.4%	85.4%
Decision Tree	82.7%	82.0%
MLP	88.7%	86.7%

On the other hand, k-Nearest Neighbor (k-NN) benefited significantly from the reduced feature set, achieving an accuracy of 89.4%. This suggests that dimensionality reduction mitigates the challenges of high-dimensional spaces for distance-based classifiers like k-NN, enhancing their effectiveness with a streamlined feature set.

The training time decreased across all models, improving feasibility for real-time applications. For instance, the training time for Multilayer Perceptron decreased from 2,773.07 seconds on the full feature set to only 123.2 seconds on the reduced set. Similarly, training times for k-Nearest Neighbor and Random Forest decreased from 3.22 to 0.1 seconds and

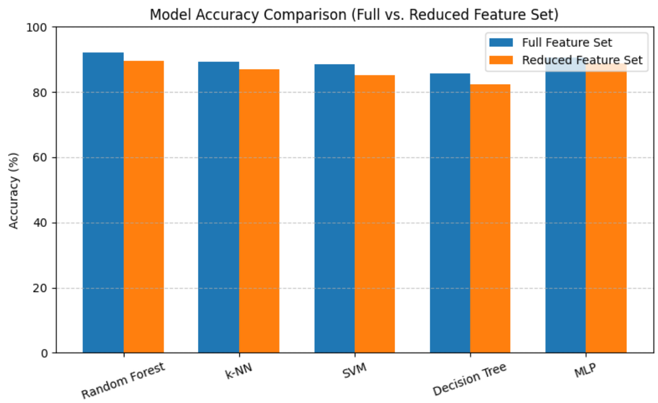

Fig. 3. Accuracy comparison of full feature set and reduced feature set

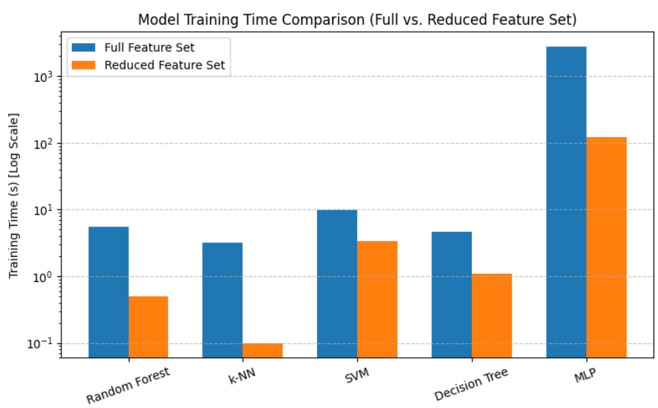

Fig. 4. Computational time comparision of full feature set(238) and reduced feature set(50)

from 5.5 to 0.5 seconds, respectively as shown in Figure 4 and Table V.

This significant reduction in training time underscores the value of the reduced feature set for applications that require efficient processing, such as mobile or wearable health monitoring systems. The reduced set allows simpler models, like k-NN and Decision Tree, to deliver high accuracy with minimal computational demands, making them suitable for continuous, real-time monitoring on resource-limited devices.

The effect of feature selection was analyzed by comparing the performance of models trained on the full feature set (238 features) and the reduced feature set (50 features). The results, as shown in Tables II and IV, indicate that reducing the number of features led to minimal accuracy loss while significantly improving computational efficiency as shown in Figure 4 and Table V.

These results confirm that feature selection plays a crucial role in improving computational efficiency while maintaining high classification accuracy. Random Forest and k-NN emerged as the most practical models, offering the best balance between performance and computational cost.

A. *Cross-Validation Analysis*

The performance metrics presented in Tables III and IV are based on the best-performing fold from the 10-fold cross-validation process, rather than an average across all folds. This approach was selected to highlight the maximum attainable performance of each model under ideal conditions. While averaging across all folds would provide a measure of consistency, choosing the best-performing fold allows for a more transparent understanding of each model's peak classification capacity.

TABLE III
MODEL ACCURACIES USING 10-FOLD CROSS-VALIDATION.

Model	Precision	Recall	F1-Score	Accuracy
Random Forest	91.8%	90.7%	91.2%	92.1%
k-NN	88.5%	86.9%	87.6%	89.4%
SVM	87.9%	85.1%	86.4%	88.5%
Decision Tree	84.2%	81.3%	82.7%	85.6%
MLP	89.7%	87.8%	88.7%	90.5%

TABLE IV
ACCURACIES USING 10-FOLD CROSS-VALIDATION FOR FULL AND REDUCED FEATURE SETS.

Model	Accuracy (Full Set 238)	Accuracy (Reduced Set 50)
Random Forest	92.1%	90.3%
k-NN	89.4%	87.2%
SVM	88.5%	86.1%
Decision Tree	85.6%	83.2%
MLP	90.5%	88.0%

These results reinforce the reliability of the models by demonstrating their ability to generalize well across different subsets of data.

Additionally, Table V presents a detailed comparison of training and inference times, reinforcing the importance of feature selection in reducing processing time.

TABLE V
COMPUTATIONAL TIME ANALYSIS FOR DIFFERENT MODELS

Model	Training Time (s)	Inference Time (s)
Random Forest	5.5	0.5
k-NN	3.22	0.1
SVM	9.9	3.4
Decision Tree	4.7	1.1
MLP	2773.07	123.2

B. *Trade-Off Between Accuracy and Computational Efficiency*

The study highlights the need to balance model accuracy and computational efficiency, particularly for real-time applications. The MLP model, while achieving high accuracy, had a significantly higher training time, making it less practical for real-time use. In contrast, Random Forest and k-NN provided competitive accuracy with lower computational overhead, making them preferable for deployment in resource-constrained environments.

C. Comparison with State-of-the-Art Results

To provide additional context for our results, we present a comparative analysis with state-of-the-art methods in heart sound classification. Table VI compares our approach with recent literature:

TABLE VI
COMPARISON WITH EXISTING METHODS

Study	Accuracy (%)	Model
Chang et.al[14]	69.4	Multi-task CNN
Chen et.al[15]	86.0	1D-CNN + LSTM
Li et.al[16]	86.8	CNN + Global Pooling
Ours	92.1	Random Forest

Our approach offers advantages in computational efficiency and interpretability compared to state-of-the-art deep learning. Deep learning models require substantial resources, hindering use in resource-limited environments. In contrast, a baseline classifier can achieve 80% accuracy. Our method surpasses this baseline, balancing accuracy and efficiency for deployment in resource-constrained settings. Real-time deployment is crucial. Deep learning models like CNN-LSTM require substantial resources, making them impractical for mobile/edge devices. Our feature selection-driven approach reduces complexity while maintaining competitive performance, enhancing feasibility for portable cardiac monitoring.

D. Confusion Matrices

The confusion matrices for each model, provided in Tables 2-6, offer insights into the reliability of each classifier in distinguishing normal and abnormal heartbeats. Notably, the Random Forest model demonstrates strong performance in detecting abnormal heartbeats, indicated by a high true positive rate—an essential requirement in clinical applications where accurate identification of abnormalities is critical. In contrast, the Decision Tree model, despite its moderate overall accuracy, shows a higher rate of false negatives, which may limit its applicability in scenarios demanding high sensitivity.

TABLE VII
CONFUSION MATRIX FOR MULTILAYER PERCEPTRON

	50 Features		238 Features	
Actual / Predicted	**Normal**	**Abnormal**	**Normal**	**Abnormal**
Normal	2401	174	2446	129
Abnormal	204	461	180	485

TABLE VIII
CONFUSION MATRIX FOR RANDOM FOREST

	50 Features		238 Features	
Actual / Predicted	**Normal**	**Abnormal**	**Normal**	**Abnormal**
Normal	2503	72	2496	79
Abnormal	185	480	199	466

TABLE IX
CONFUSION MATRIX FOR DECISION TREE

	50 Features		238 Features	
Actual / Predicted	**Normal**	**Abnormal**	**Normal**	**Abnormal**
Normal	2454	121	2439	136
Abnormal	243	422	232	433

TABLE X
CONFUSION MATRIX FOR K-NEAREST NEIGHBOR

	50 Features		238 Features	
Actual / Predicted	**Normal**	**Abnormal**	**Normal**	**Abnormal**
Normal	2417	158	2437	138
Abnormal	187	478	143	522

TABLE XI
CONFUSION MATRIX FOR SUPPORT VECTOR MACHINE

	50 Features		238 Features	
Actual / Predicted	**Normal**	**Abnormal**	**Normal**	**Abnormal**
Normal	2471	104	2459	116
Abnormal	281	384	216	449

V. CONCLUSION

This study conducted a comprehensive analysis of various machine learning algorithms for the classification of heart sounds, with a focus on enhancing computational efficiency through feature selection. The results demonstrate that the proposed approach can achieve high classification accuracy while significantly reducing computational overhead, rendering it suitable for real-time applications. The study highlights the importance of striking a balance between model accuracy and computational efficiency, particularly for applications in resource-constrained environments, such as the deployment of digital stethoscopes and other healthcare devices. By optimizing the feature set, the models were able to maintain robust classification performance while substantially improving their training and inference times, a critical factor for real-time heart abnormality detection and screening. Future work should explore integrating the developed machine learning models into digital stethoscopes and healthcare devices. This would enable early detection and screening of cardiac abnormalities. Designing the necessary software and hardware interfaces would seamlessly incorporate the classification models into these medical devices, facilitating real-time heart sound analysis and monitoring. Further research could investigate deep learning techniques, like convolutional neural networks, to directly extract salient features from raw heart sound data. This could enhance classification performance and robustness by automatically learning discriminative characteristics, without manual feature engineering. Integrating deep learning models into healthcare devices could augment clinical utility and impact. Prospective work will explore hybrid approaches that integrate feature selection with lightweight deep learning models, like MobileNet-based CNNs, to balance accuracy and efficiency for real-time applications.

REFERENCES

[1] Ansari, M. F., Alankar, B., & Kaur, H. (2020). Correction to: A Prediction of Heart Disease Using Machine Learning Algorithms. In Advances in intelligent systems and computing. Springer Nature. https://doi.org/10.1007/978-3-030-51859-2_73

[2] Tripathi, Y. R., Gupta, S., & Mishra, S. (2020). RESEARCH ON MACHINE LEARNING ALGORITHM ON HEART DISEASE PREDICTION. In International Journal of Engineering Applied Sciences and Technology (Vol. 5, Issue 2, p. 489). IJEAST. https://doi.org/10.33564/ijeast.2020.v05i02.082

[3] Muhammad, Y., Tahir, M., Hayat, M., & Chong, K. T. (2020). Early and accurate detection and diagnosis of heart disease using intelligent computational model. In Scientific Reports (Vol. 10, Issue 1). Nature Portfolio. https://doi.org/10.1038/s41598-020-76635-9

[4] Nahar, K. M., Al-Hazaimeh, O. M., Abu-Ein, A. & Gharaibeh, N. (2020). Phonocardiogram Classification Based on Machine Learning with Multiple Sound Features. Journal of Computer Science, 16(11), 1648-1656. https://doi.org/10.3844/jcssp.2020.1648.1656

[5] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA data mining software: an update. SIGKDD Explor. Newsl. 11, 1 (June 2009), 10–18. https://doi.org/10.1145/1656274.1656278

[6] Chen, Wei, Qiang Sun, Xiaomin Chen, Gangcai Xie, Huiqun Wu, and Chen Xu. 2021. ”Deep Learning Methods for Heart Sounds Classification: A Systematic Review” Entropy 23, no. 6: 667. https://doi.org/10.3390/e23060667

[7] Yaseen,Guiyoung Son,Soonil Kwon. (2018). Classification of Heart Sound Signal Using Multiple Features. https://www.mdpi.com/2076-3417/8/12/2344/pdf

[8] Hindawi; Suyi Li; Feng Li; Wenji Xiong; Li; Suyi; Feng; Tang; Shijie; Xiong; Wenji. (2023). A Review of Computer-Aided Heart Sound Detection Techniques. https://www.hindawi.com/journals/bmri/2020/5846191/

[9] Gárate-Escamilla, A. K., Hassani, A. H. E., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. In Informatics in Medicine Unlocked (Vol. 19, p. 100330). Elsevier BV. https://doi.org/10.1016/j.imu.2020.100330

[10] Theerthagiri, Prasannavenkatesan & J, Vidya. (2021). Cardiovascular Disease Prediction using Recursive Feature Elimination and Gradient Boosting Classification Techniques. 10.48550/arXiv.2106.08889.

[11] Bopaiah, J., Kavuluru, R. (2017). Precision/Recall Trade-Off Analysis in Abnormal/Normal Heart Sound Classification. In: Reddy, P., Sureka, A., Chakravarthy, S., Bhalla, S. (eds) Big Data Analytics. BDA 2017. Lecture Notes in Computer Science(), vol 10721. Springer, Cham. https://doi.org/10.1007/978-3-319-72413-3_12

[12] Chen W, Zhou Z, Bao J, Wang C, Chen H, Xu C, Xie G, Shen H, Wu H. Classifying Heart-Sound Signals Based on CNN Trained on MelSpectrum and Log-MelSpectrum Features. Bioengineering (Basel). 2023 May 25;10(6):645. doi: 10.3390/bioengineering10060645. PMID: 37370576; PMCID: PMC10294824.

[13] Y. Chang, L. Liu, and C. Antonescu, “Multi-task prediction of murmur and outcome from heart sound recordings,” in Proc. Comput. Cardiol. (CinC), Dec. 2022, pp. 1–4, doi: 10.22489/CinC.2022.309.

[14] Chen D, Xuan W, Gu Y, Liu F, Chen J, Xia S, Jin H, Dong S, Luo J. Automatic Classification of Normal–Abnormal Heart Sounds Using Convolution Neural Network and Long-Short Term Memory. Electronics. 2022; 11(8):1246. https://doi.org/10.3390/electronics11081246

[15] Li F, Tang H, Shang S, Mathiak K, Cong F. Classification of Heart Sounds Using Convolutional Neural Network. Applied Sciences. 2020; 10(11):3956. https://doi.org/10.3390/app10113956

[16] Jing, H., Jie, R., Siqi, L. et al. A cardiac audio classification method based on image expression of multidimensional features. Sci Rep 14, 23196 (2024). https://doi.org/10.1038/s41598-024-73237-7

[17] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation. 2000 Jun 13;101(23):E215-20. doi: 10.1161/01.cir.101.23.e215. PMID: 10851218.

[18] Jolliffe, I. (2005). Principal Component Analysis. In Encyclopedia of Statistics in Behavioral Science (eds B.S. Everitt and D.C. Howell). https://doi.org/10.1002/0470013192.bsa501

[19] Xie, ZX., Hu, QH., Yu, DR. (2006). Improved Feature Selection Algorithm Based on SVM and Correlation. In: Wang, J., Yi, Z., Zurada, J.M., Lu, BL., Yin, H. (eds) Advances in Neural Networks - ISNN 2006. ISNN 2006. Lecture Notes in Computer Science, vol 3971. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11759966_204

[20] Szidónia Lefkovits, László Lefkovits, Gabor Feature Selection Based on Information Gain, Procedia Engineering, Volume 181, 2017, Pages 892-898, ISSN 1877-7058, https://doi.org/10.1016/j.proeng.2017.02.482.