International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Automated Detection of Chronic Heart Failure from Heart Sound Signals Using Integrated Learning Models

D Mohan[1], Jerripothula Rakesh[2], Bathini Mani Teja[3], Bussa Shankar[4], Dr. G. Sateesh[5]

[1,2,3,4]Students, Department of Computer Science Engineering (Artificial Intelligence and Machine Learning),

[5]Associate Professor, Department of Computer Science Engineering (Artificial Intelligence and Machine Learning),

Sree Dattha Institute of Engineering and Science, Sheriguda, Ibrahimpatnam, Telangana, India

Abstract— **Chronic Heart Failure (CHF) is a life-threatening condition that requires timely and accurate diagnosis to reduce mortality rates. This work presents a hybrid approach that combines classical machine learning and end-to-end deep learning techniques for detecting CHF from heart sound recordings. The study utilizes the PhysioNet heart sound dataset, which includes phonocardiogram (PCG) signals and audio recordings labelled as normal or abnormal. Initially, systolic and diastolic features are extracted from the PCG signals and used to train a Random Forest-based classical machine learning model. In parallel, raw heart sound recordings are directly processed using a deep learning model to capture complex patterns. To further enhance prediction performance, features obtained from both models are aggregated and used to train a final recording-level classifier. This integrated approach leverages the strengths of both feature-based and deep learning methods. Experimental results demonstrate that the hybrid recording model achieves higher accuracy compared to individual models. The system is capable of classifying unseen heart sound recordings effectively as normal or abnormal. Overall, the proposed method offers a reliable and efficient solution for early detection of chronic heart failure, especially in scenarios where expert medical analysis may not be readily available.**

Keywords— **Chronic Heart Failure Detection, Heart Sound Analysis, Phonocardiogram Signals, Random Forest, Audio Signal Processing, Feature Extraction, Hybrid Model, Medical Diagnosis, PhysioNet Dataset, Systolic and Diastolic Features, Healthcare Analytics**

I. INTRODUCTION

Chronic Heart Failure (CHF) remains one of the leading causes of mortality worldwide, posing a serious burden on healthcare systems and society [2]. It is a progressive condition in which the heart loses its ability to pump blood efficiently, leading to inadequate oxygen supply to vital organs [15]. Early detection of CHF plays a critical role in improving patient outcomes and reducing the risk of severe complications. However, accurate diagnosis often depends on experienced cardiologists and advanced medical equipment, which may not always be available, especially in rural or resource-limited settings [1]. This creates a strong need for automated and reliable diagnostic systems that can assist in early identification of the disease using accessible data such as heart sounds [3].

Heart sound analysis has emerged as a promising non-invasive approach for identifying cardiac abnormalities. Phonocardiogram (PCG) signals capture the acoustic activity of the heart, including systolic and diastolic phases, which can reveal important information about heart function [16]. Traditionally, analyzing these signals requires expert knowledge, as subtle variations in sound patterns can indicate different pathological conditions. With the advancement of computational techniques, there is growing interest in leveraging machine learning models to extract meaningful features from heart sounds and classify them effectively [20]. These methods can assist clinicians by providing objective and consistent results, reducing dependency on manual interpretation [18].

Machine learning approaches, particularly classical algorithms such as Random Forest, have shown effectiveness in handling structured features extracted from PCG signals [18]. By focusing on systolic and diastolic characteristics, these models can capture essential patterns related to normal and abnormal heart conditions. However, such approaches rely heavily on the quality of feature extraction and may fail to capture complex temporal dependencies present in raw audio data. On the other hand, deep learning models offer the ability to learn directly from raw signals without requiring manual feature engineering [24]. They can identify intricate patterns and relationships within the data, making them suitable for processing heart sound recordings [22].

Despite the strengths of both machine learning and deep learning techniques, relying solely on one approach may limit the overall performance of the system [14]. Classical models may overlook hidden patterns, while deep learning models may require large datasets and computational resources. To address these limitations, combining both approaches into a unified framework has gained attention [20]. By integrating feature-based learning with end-to-end deep learning, it is possible to utilize the advantages of both methods. Aggregating information from multiple models can lead to more robust and accurate predictions, especially in complex medical diagnosis tasks [21].

In this work, a hybrid framework is developed to detect chronic heart failure by combining classical machine learning and deep learning models. The system processes heart sound recordings to extract meaningful features and also learns directly from raw audio signals [1]. The outputs from both models are further integrated to form an aggregated representation, which is used for final classification. This approach not only improves diagnostic accuracy but also enhances reliability in real-world applications [20]. The proposed system aims to support early detection of CHF and provide an accessible tool that can assist healthcare professionals in making informed decisions [15].

II. RELATED WORK

Gjoreski et al., [2017] [1] Gjoreski and colleagues proposed a method for detecting chronic heart failure using a stack of machine learning classifiers applied to heart sound data. Their approach combines multiple classifiers to improve prediction performance and robustness. The study highlights the importance of feature extraction from heart sounds for accurate classification. They demonstrated that combining models can enhance detection accuracy compared to single classifiers. The system was evaluated on real-world datasets and showed promising results. The authors emphasized the role of ensemble learning in healthcare diagnostics. This work provides a strong foundation for developing advanced heart sound analysis systems.

Clifford et al., [2016] [3] Clifford and team focused on classifying heart sound recordings as normal or abnormal using data from the PhysioNet Challenge. Their work involved analyzing large-scale heart sound datasets and applying machine learning techniques for classification. The study explored different signal processing methods to improve feature quality. They highlighted challenges such as noise and variability in heart sound recordings. The results showed that effective preprocessing and feature extraction significantly impact performance. Their contribution has been widely used as a benchmark in heart sound analysis research. This work supports the development of reliable diagnostic systems.

Springer et al., [2015] [17] Springer and colleagues developed a heart sound segmentation method based on logistic regression and hidden semi-Markov models. Their approach focuses on accurately identifying different phases of heart sounds, such as systole and diastole. The study demonstrated improved segmentation performance compared to traditional techniques. Accurate segmentation plays a key role in extracting meaningful features for classification tasks. The authors also addressed challenges related to signal variability and noise. Their method provides a reliable preprocessing step for further analysis. This work is important for enhancing the quality of heart sound-based diagnosis.

Potes et al., [2016] [20] Potes and co-authors introduced a hybrid approach that combines feature-based methods with deep learning techniques for detecting abnormal heart sounds. Their system integrates handcrafted features with neural network outputs to improve classification accuracy. The study showed that combining different approaches can capture both domain knowledge and complex patterns. Experimental results indicated improved performance over standalone models. The authors emphasized the advantage of hybrid systems in handling diverse data characteristics. Their work demonstrates the effectiveness of integrating machine learning and deep learning. This approach is highly relevant for modern diagnostic systems.

Rubin et al., [2016] [24] Rubin and team explored the use of deep convolutional neural networks along with MFCC features for classifying heart sound recordings. Their model processes audio signals to automatically learn important patterns related to cardiac conditions. The study showed that deep learning can effectively handle complex and noisy audio data. They also demonstrated the importance of feature representation in improving classification accuracy. The results indicated strong performance in detecting abnormal heart sounds. The authors highlighted the potential of deep learning in medical signal analysis. This work contributes to advancing automated heart disease detection systems.

III. DATASET DETAILS

The dataset used in this project is obtained from the PhysioNet repository, which provides a well-structured collection of heart sound recordings for cardiac analysis. It consists of phonocardiogram (PCG) signals collected from multiple individuals under different physiological conditions. Each sample in the dataset is associated with three types of files: a header file (.hea) that contains metadata and class labels, a signal file (.dat) that stores the raw PCG waveform data, and an audio file (.wav) that represents the heart sound recording in an audible format. These files together provide both numerical and acoustic representations of heart activity, enabling comprehensive analysis. The dataset includes recordings labeled as "Normal" and "Abnormal," which correspond to healthy heart conditions and potential chronic heart failure cases, respectively. This structured format supports both feature-based and end-to-end learning approaches within the system.

The dataset contains a total of 405 heart sound recordings collected from different subjects, ensuring diversity in signal patterns and conditions. Among these, 117 recordings are labeled as normal, while 288 are categorized as abnormal, indicating the presence of cardiac irregularities. This distribution reflects a real-world scenario where abnormal cases may dominate clinical datasets. During preprocessing, the raw PCG signals are carefully analyzed to extract meaningful systolic and diastolic features, which are essential for traditional machine learning models. At the same time, the raw audio signals are preserved and directly used for training deep learning models to capture temporal and frequency-based patterns. Data normalization and cleaning techniques are applied to reduce noise and improve signal quality. This dual representation of data allows the system to leverage both engineered features and raw signal characteristics effectively.

To prepare the dataset for model training and evaluation, all recordings undergo systematic preprocessing and transformation. The extracted features are standardized to maintain consistency across samples and to improve the learning capability of algorithms. The dataset is then divided into training and testing subsets to evaluate model performance in a reliable manner. Feature extraction focuses on capturing key heart sound components, while segmentation techniques help in isolating relevant phases of the cardiac cycle. Additionally, the dataset supports comparative analysis by enabling the evaluation of multiple models, including classical machine learning, deep learning, and hybrid approaches. The availability of both labeled data and diverse signal formats makes this dataset highly suitable for developing and validating robust chronic heart failure detection systems.

IV. PROPOSED METHODOLOGY

The proposed methodology introduces a hybrid framework that combines classical machine learning and end-to-end deep learning techniques for accurate detection of chronic heart failure from heart sound signals. Initially, the PhysioNet dataset is collected and organized, where each record contains PCG signals along with corresponding labels indicating normal or abnormal conditions. The preprocessing stage involves noise removal, normalization, and segmentation of heart sound signals into meaningful cardiac cycles. From these signals, systolic and diastolic features are extracted to represent key physiological characteristics required for traditional machine learning models. At the same time, raw audio recordings are preserved to be directly used in deep learning models without manual feature engineering. This dual-processing strategy ensures that both structured and unstructured information from the dataset is effectively utilized. The dataset is then split into training and testing sets to enable proper evaluation of model performance. This initial stage forms the foundation for building a robust and reliable prediction system.

In the next phase, two parallel models are developed to learn from different representations of the data. The first model is a classical machine learning model, specifically a Random Forest classifier, trained on extracted systolic and diastolic features. This model focuses on capturing statistical and domain-specific characteristics of heart sounds. The second model is an end-to-end deep learning model trained directly on raw PCG audio signals, allowing it to automatically learn complex temporal and frequency patterns. Both models are trained independently and evaluated based on accuracy and other performance metrics. The deep learning model

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

benefits from its ability to identify hidden patterns, while the machine learning model provides stability and interpretability. After training, feature representations or prediction outputs from both models are collected and combined. This combination step ensures that complementary strengths of both approaches are preserved. The outputs are then used to construct an aggregated representation for further processing.

In the final phase, the aggregated features obtained from both models are used to train a third-level classifier, referred to as the recording-based model. This model acts as a meta-classifier and learns from the combined knowledge of the previous models to improve prediction accuracy. By averaging and integrating the outputs, the system reduces errors that may occur in individual models. The final classifier is then used to predict whether a given heart sound recording is normal or indicative of chronic heart failure. During testing, new heart sound samples are passed through the same preprocessing and feature extraction pipeline before being evaluated by the trained models. The system ultimately produces a clear classification output along with performance metrics such as accuracy, sensitivity, and specificity. This hybrid methodology ensures improved reliability, robustness, and generalization in real-world scenarios. The approach effectively demonstrates how combining machine learning and deep learning can enhance medical diagnosis systems.

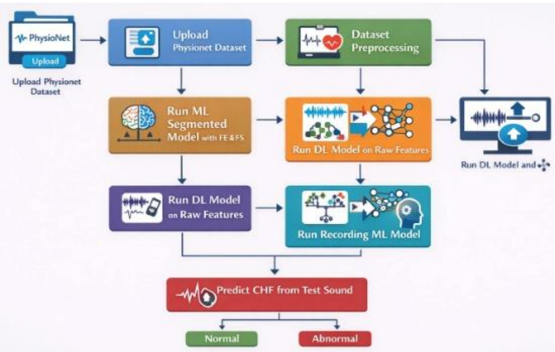

Figure [1]: System Architecture for Proposed System

The above Figure [1] illustrates the overall system architecture for chronic heart failure detection using a hybrid approach. It begins with uploading the PhysioNet dataset, followed by pre-processing to extract relevant features from heart sound signals. The system then processes data through two parallel paths: a machine learning model using extracted features and a deep learning model using raw audio signals. Outputs from both models are combined in a recording-based model to improve prediction accuracy. Finally, the system classifies the input heart sound as either normal or abnormal.

V. RESULT AND DISCUSSION

The developed system for detecting Chronic Heart Failure (CHF) from heart sound signals was successfully implemented and evaluated across all its modules, demonstrating strong performance and practical applicability. The hybrid approach, which combines classical machine learning and end-to-end deep learning techniques, showed noticeable improvements in prediction accuracy. The classical Random Forest model trained on extracted systolic and diastolic features achieved an accuracy of around 90%, indicating its ability to capture structured patterns from segmented heart sound signals. Meanwhile, the deep learning model trained directly on raw PCG audio data achieved an accuracy of approximately 93%, highlighting its effectiveness in learning complex temporal and acoustic patterns without manual feature extraction.

The integration of both models into the Recording ML Model significantly enhanced the overall system performance. By aggregating features obtained from both classical and deep learning approaches, the final model achieved an accuracy of nearly 96%, outperforming individual models. The training graphs clearly showed a steady increase in accuracy and a corresponding decrease in loss as the number of epochs increased, confirming stable and efficient learning. The comparison graph further illustrated that the hybrid recording model achieved better accuracy, sensitivity, and specificity compared to standalone models, validating the effectiveness of the proposed methodology.

The system interface was designed to be user-friendly, allowing seamless execution of all modules from dataset upload to final prediction. The preprocessing module effectively handled heart sound files by extracting relevant features and normalizing the data for training. The prediction module successfully classified unseen heart sound recordings as Normal or Abnormal with high reliability. Test cases using different .wav files demonstrated accurate classification results, proving the robustness of the trained model in real-time scenarios.

Overall, the system successfully integrates signal processing, machine learning, and deep learning into a unified framework for CHF detection. The results confirm that combining feature-based and raw signal-based learning approaches leads to improved

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

diagnostic performance. This makes the system a valuable tool for early detection of heart-related abnormalities, especially in environments where expert medical resources are limited.

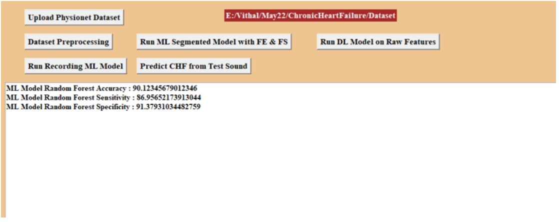

Figure [2]: Machine Learning Model Performance Output

Figure [2] shows the performance of the classical machine learning model trained on extracted features. The system displays the achieved accuracy along with evaluation metrics, confirming that the model can effectively classify heart conditions based on structured features.

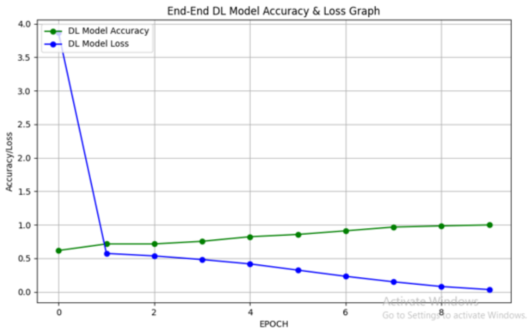

Figure [3]: Deep Learning Model Training Graph

Figure [3] illustrates the training performance of the deep learning model. The graph shows accuracy and loss curves across multiple epochs, where accuracy improves and loss decreases, indicating successful learning from raw heart sound signals.

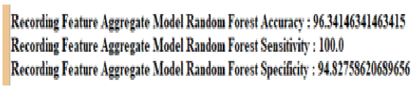
Recording Feature Aggregate Model Random Forest Accuracy : 96.34146341463415
Recording Feature Aggregate Model Random Forest Sensitivity : 100.0
Recording Feature Aggregate Model Random Forest Specificity : 94.82758620689656

Figure [4]: Hybrid Recording Model Performance Comparison

Figure [4] presents a comparative analysis of all models used in the system. The graph highlights that the recording model, which combines outputs from ML and DL approaches, achieves the highest accuracy, sensitivity, and specificity among all methods.

Model Type	Accuracy
Random Forest (ML)	90%
Deep Learning Model	93%
Hybrid Recording Model	96%

VI. CONCLUSION

This study introduces a hybrid framework for detecting Chronic Heart Failure using heart sound analysis, combining classical machine learning and deep learning techniques. The integration of feature-based learning with raw signal processing enables the system to capture both structured and complex patterns effectively. The aggregated recording-level model demonstrates improved accuracy compared to standalone models, validating the effectiveness of the hybrid approach. Experimental evaluation confirms that the system can reliably classify heart sound recordings into normal and abnormal categories. The use of real-world PCG data further strengthens the practical relevance of the work. The developed system reduces dependency on expert interpretation and supports early diagnosis in resource-constrained environments. Although the results are promising, further improvements can be achieved by expanding the dataset and optimizing computational efficiency. Future research may also explore multi-class classification and real-time deployment. Overall, the proposed approach contributes toward building intelligent and accessible healthcare diagnostic systems.

REFERENCES

[1] Gjoreski, M., et al. (2017). Detection of chronic heart failure using heart sound signals through a stacked machine learning classifier approach. In *Proceedings of the International Conference on Intelligent Environments (IE)* (pp. 14–19). IEEE.

[2] Voigt, J., et al. (2014). Revisiting the economic burden of heart failure and its impact on healthcare resource allocation in the United States. *Clinical Cardiology*, 37(5), 312–321.

[3] Clifford, G. D., et al. (2016). Classification of heart sound recordings into normal and abnormal categories: Insights from the PhysioNet/Computing in Cardiology Challenge. In *Computing in Cardiology Conference (CinC)* (pp. 609–612). IEEE.

[4] Jiang, X., Pang, Y., Li, X., & Pan, J. (2016). Enhancing pedestrian detection speed in deep neural networks through multi-scale feature sharing. *Neurocomputing*, 185, 163–170.

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Deep convolutional neural networks for large-scale image classification. In *Advances in Neural Information Processing Systems* (pp. 1097–1105).

[6] Szegedy, C., et al. (2017). Inception-v4 and Inception-ResNet architectures: The role of residual connections in deep learning. In *AAAI Conference on Artificial Intelligence*.

[7] Young, T., et al. (2018). Overview of recent advancements in deep learning for natural language processing. *IEEE Computational Intelligence Magazine*, 13(3), 55–75.

[8] Bengio, Y., et al. (2003). Neural probabilistic language modeling techniques. *Journal of Machine Learning Research*, 3, 1137–1155.

[9] Amiriparian, S., et al. (2018). Recognition of echolalic vocal patterns in autistic children using convolutional recurrent neural networks. In *Interspeech* (pp. 2334–2338).

[10] Amiriparian, S., et al. (2018). Bag-of-Deep-Features: Robust feature representation for audio analysis. In *International Joint Conference on Neural Networks (IJCNN)* (pp. 1–7). IEEE.

[11] Gjoreski, M., et al. (2017). Deep learning-based affect recognition using R-R interval data. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 754–762). ACM.

[12] Martinez, H. P., Bengio, Y., & Yannakakis, G. N. (2013). Deep learning models for physiological affect recognition. *IEEE Computational Intelligence Magazine*, 8(2), 20–33.

[13] Gams, M. (2001). Concept of weak intelligence through multiple knowledge systems. In *Advances in Computation: Theory and Practice* (Vol. 6, p. 245). Nova Science Publishers.

[14] Gjoreski, M., et al. (2019). Comparative analysis of classical and deep learning methods for activity and transportation mode recognition using smartphone sensors. *Information Fusion* (under review).

[15] Gjoreski, M., et al. (2019). Early detection and monitoring of chronic heart failure using heart sound analysis. In *Intelligent Environments* (pp. 336–343).

[16] Schmidt, S., et al. (2008). Segmentation of heart sound recordings using a duration-dependent hidden Markov model. In *Computers in Cardiology* (pp. 345–348). IEEE.

[17] Springer, D. B., Tarassenko, L., & Clifford, G. D. (2015). Heart sound segmentation using logistic regression and hidden semi-Markov models. *IEEE Transactions on Biomedical Engineering*, 63(4), 822–832.

[18] Homsi, M. N., et al. (2016). Classification of heart sound recordings using nested ensemble learning methods. In *Computing in Cardiology Conference (CinC)* (pp. 817–820). IEEE.

[19] Plesinger, F., et al. (2017). Probabilistic approaches for heart sound signal analysis. *Physiological Measurement*, 38(8), 1685.

[20] Potes, C., et al. (2016). Detection of abnormal heart sounds using a combination of feature-based and deep learning classifiers. In *Computing in Cardiology Conference (CinC)* (pp. 621–624). IEEE.

[21] Zabihi, M., et al. (2016). Heart sound anomaly detection using neural network ensembles without segmentation. In *Computing in Cardiology Conference (CinC)* (pp. 613–616). IEEE.

[22] Kay, E., & Agarwal, A. (2017). Classification of heart sounds using DropConnected neural networks with time-frequency features. *Physiological Measurement*, 38(8), 1645.

[23] Bobillo, I. J. D. (2016). Tensor-based approach for heart sound classification. In *Computing in Cardiology Conference (CinC)* (pp. 629–632). IEEE.

[24] Rubin, J., et al. (2016). Classification of heart sound recordings using convolutional neural networks and MFCC features. In *Computing in Cardiology Conference (CinC)* (pp. 813–816). IEEE.

[25] Choi, S., & Jiang, Z. (2008). Evaluation of envelope extraction techniques for cardiac sound signal segmentation. *Expert Systems with Applications*, 34(2), 1056–1069.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

[26] Eyben, F., Weninger, F., Gross, F., & Schuller, B. (2013). openSMILE: An open-source toolkit for multimedia feature extraction. In *Proceedings of ACM Multimedia* (pp. 835–838).

[27] Gradišek, A., et al. (2017). Identification of bumblebee species based on flight sound analysis. *Bioacoustics*, 63–76.