https://doi.org/10.46335/IJIES.2024.9.7.3 e-ISSN: 2456-3463
Vol. 9, No. 7, 2024, PP. 10 -15

International Journal of Innovations in Engineering and Science, www.ijies.net

Harnessing Machine Learning and Artificial Intelligence for Electrocardiogram Analysis: A Comprehensive Review

Pulkeshini Vishal Taksande[1], Abhinav Muley[2]

[1]*UG Student,* [2]*Asst. Professor,Dept. of CSE(DS), SVPCET, Nagpur,Maharashtra,India.*
pulkeshinitaksande@gmail.com

***Received on**: 5 May,2024* ***Revised on**: 29 June,2024* ***Published on**: 02 July ,2024*

Abstract – The main focus of this paper is to review and analyze the various techniques used to analyze Electrocardiogram, ECG using different Machine Learning Techniques. Electrocardiogram is pivotal in the diagnosis and management of cardiovascular diseases, which is the leading cause of mortality. The emergence of machine learning technologies has transformed the field of biomedical signal processing attempting promising avenues for enhancing ECG analysis accuracy and efficiency. This review paper provides a comprehensive study of the various machine-learning models applied in ECG analysis, highlighting their methodologies, strengths, challenges, and applications. The models generally used for ECG analysis are Convolutional Neural Network (CNN), deep neural networks, Long-Short Term Memory (LSTM), and Recurrent Neural Networks(RNNs).

Keywords – ECG, ECG analysis, CNN, deep neural network, LSTM, RNN, disease prediction.

I. INTRODUCTION

ECG, an electrocardiogram, is a medical test that measures the electrical activity of the heart over some time using electrodes placed on the skin. These electrodes detect the tiny electrical changes on the skin that arise from the heart muscles' electrophysiologic pattern of depolarizing and repolarizing during each heartbeat. It is a non-invasive, painless test with quick results, widely used in the medical field to check the heart's condition.

The ECG paper is marked with a grid of small and large squares. Each small square represents 40 milliseconds(ms) in time along the horizontal axis and each large square contains 5 small squares, collectively representing 200 milliseconds (ms). Standard paper speeds and square markings allow easy measurement of cardiac timing intervals. Image 1 describes the sample of standard ECG paper showing the scale of voltage, measured on the vertical axis, against time on the horizontal axis.

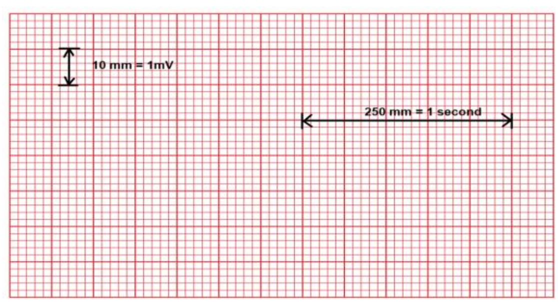

Figure 1: sample of standard ECG paper

The first electrical signal on a normal ECG originates from the atria and is known as the P wave. Although there is only one P wave in most leads of an ECG, the P wave is the sum of the electrical signals from the two atria, which are usually superimposed. Then, there is a short delay as the atrioventricular node slows the electrical depolarization before it proceeds to the ventricles. This delay is responsible for the PR interval, a short amount of time where no electrical activity

https://doi.org/10.46335/IJIES.2024.9.7.3 **e-ISSN: 2456-3463**
Vol. 9, No. 7, 2024, PP. 10 -15

is seen on the ECG, which is shown by a straight horizontal(isoelectric) line.

Depolarisation of the ventricles results in the largest part of the ECG signal and this is known as the QRS complex. The Q wave is the first initial downward deflection. The R wave is then the next upward deflection. The S wave is then the next downward deflection provided that it crosses the horizontal line to become briefly negative before returning to the horizontal baseline.

In the case of the ventricles, there is also an electrical signal reflecting the repolarisation of the myocardium. This is shown as the ST segment and the T wave. The ST segment is normally isoelectric, and the T wave in most leads is an upright deflection of variable amplitude and duration

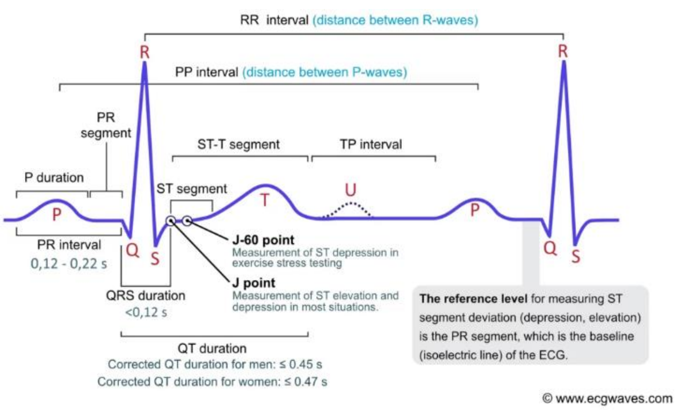

Figure 2 describes the various intervals in an ECG wave [1]

ECG analysis is the study and analysis of Electrocardiograms, used to detect diseases and take precautionary measures against it. By implementing machine learning in this field, the task that generally takes doctors a considerable amount of time will be reduced significantly, along with increased accuracy and efficiency. Now, various machine learning models have been used for this purpose, out of which, the most commonly used are: Artificial Neural Networks, Convolutional Neural Networks, Deep learning models, Support Vector Machines, Autoencoders, Ensemble methods, and Transfer Learnings.

II. LITERATURE REVIEW

1. Advancing Cardiovascular Diagnostics: Integrating Machine Learning and 3D Simulation in ECG Analysis for Enhanced Arrhythmia Detection and Disease Prediction [2]

This study emphasizes the crucial role of ECG analysis in diagnosing and predicting cardiovascular disorders, the significant contribution of computational methods like machine learning in accurate heartbeat classification for arrhythmia detection and disease diagnosis, and the potential of 3D computer simulations in interpreting ECG data and generating synthetic datasets for training machine learning classifiers. The paper also emphasizes the significance of ECG analysis in diagnosing cardiovascular disorders, discusses the role of computational techniques like machine learning in advancing medical discoveries, and highlights the recent progress and growing interest in the field. The methodology in the paper involves the use of machine learning techniques and 3D computer simulations for ECG analysis, including supervised and unsupervised learning for dataset analysis, with a focus on heartbeat classification and patient diagnosis. The study also discusses the challenges of dealing with medical data for clinical applications and the role of 3D computer simulations in addressing these challenges. Intervention effects include the classification of heartbeats, SVM specificity:99.72%, AF event identification, Sensitivity: ~97%, Specificity: ~97%, Real-time analysis using deep learning, Computer simulation models: Provided insights into the effects of diseases and treatments on the ECG.

2. Deciphering the Heart's Rhythm: An Explainable Deep Learning Approach for Atrial Fibrillation Detection from ECGs [3]

This research discusses the development of an explainable artificial intelligence model to detect atrial fibrillation using an electrocardiogram, emphasizing the importance of interpretability in deep learning models for accurate detection. The study's objectives were to develop an explainable deep learning model to detect atrial fibrillation using an electrocardiogram, validate its performance using diverse ECG formats, and enhance its transparency for clinical application. The main findings in this paper were the successful detection of AF using diverse ECGs with an explainable DLM, enhanced transparency of the DLM for clinical application, and the outperformance of previous models in detecting AF The methodology involved a retrospective study using the Sejong ECG dataset for internal validation, development of two feature modules for DLM decisions, and external validation using datasets from PTB-XL, Chapman, and PhysioNet.

https://doi.org/10.46335/IJIES.2024.9.7.3 e-ISSN: 2456-3463
Vol. 9, No. 7, 2024, PP. 10 -15

The intervention effects in the study show that the DLM had high accuracy in detecting atrial fibrillation (AF) using different formats of ECG, with AUC values ranging from 0.990 to 0.999. The sensitivity of the DLM ranged from 0.982 to 0.999, and the specificity ranged from 0.970 to 0.999. The explainability of features such as rhythm irregularity and absence of P-wave also showed high AUC values, ranging from 0.961 to 0.993 and 0.983 to 0.993, respectively. These results indicate that the explainable artificial intelligence methodology used in the DLM was effective in accurately detecting AF and providing insights into the reasons for its decisions.

3. Precision in Prediction: Evaluating Supervised Machine Learning Algorithms for Differential Diagnosis in ECG Report [4]

The study aims to use supervised machine learning to distinguish between normal and abnormal ECG reports and predict specific heart diseases, evaluating the performance of six different algorithms. The methodology in the paper involves designing a model using supervised machine learning to find anomalies in ECG reports, applying six supervised machine learning algorithms to distinguish between normal and abnormal ECG, dividing the dataset for training and testing, using Cross Validation and Random Train-Test Split for accuracy, normalizing and scaling data, creating different datasets for specific diseases, training classifiers with different algorithms, and using sample data to predict heart diseases. The main findings of this research were Logistic Regression performed well with a score above 0.90 for all diseases, Myocardial Infarction had the highest accuracy score across all algorithms, and different algorithms were effective for different diseases.

The intervention effects in the study include: Coronary Artery Disease (CAD): Logistic Regression: 90%, Naive Bayes: 83% Myocardial Infarction (MI): Logistic Regression: 100%, Naive Bayes: 91% Sinus Tachycardia (ST): Decision Tree: 97%, Nearest Neighbor: 70% Sinus Bradycardia (SB): Support Vector Machine: 96%, Nearest Neighbor: 69% Right Bundle Branch Block (RBBB): Logistic Regression: 96%, Naive Bayes: 81%, Overall, all the algorithms gave relatively very good results for all diseases.

4. Decoding Heartbeats: Comparative Analysis of ANN Models for Cardiac Arrhythmia Diagnosis from ECG Data [5]

This research proposes an ANN-based system for diagnosing cardiac arrhythmia using ECG signal data, evaluates three ANN models based on classification performance measures, and discusses the importance of accurate detection of arrhythmias for clinical purposes. The MLP model exhibited high accuracy and sensitivity in diagnosing cardiac arrhythmia, while the MNN model showed superior classification specificity. The performance of the ANN models varied based on the dataset used.This research discusses the performance of different ANN models in classifying cardiac arrhythmia cases, highlighting the strengths of Multilayer Perceptron in accuracy and sensitivity, and Modular Neural Network in specificity, with variations in performance across different datasets.

The methodology involved using Artificial Neural Network models to diagnose cardiac arrhythmia from ECG signal data, training the models with the backpropagation algorithm, evaluating performance using various measures, and cleaning the dataset by replacing missing values. The study objectives were to identify cardiac arrhythmias automatically from ECG recordings and to classify diseases into normal and abnormal classes using three different neural network models.

5. Enhancing Early Intervention: AI Methods in Identifying Shockable ECG Rhythms and Overcoming Database Challenges [6]

The study discusses the importance of early detection of shockable ECG rhythms, the incorporation of AI methods into CAAC systems, the challenges related to the use of deep learning methods, and the need for large databases for accurate classification. It emphasizes the significance of accurate identification of shockable ECG rhythms and the role of feature extraction in achieving high detection accuracy. The paper also highlights the limitations of small databases in achieving optimal classification performance. It also emphasizes the importance of detecting shockable ECG rhythms, the role of AI in improving accuracy, and the need for advanced systems to enhance real-time detection. It also highlights the significance of accurate ECG diagnosis in designing automated external defibrillators (AEDs) and the increasing use of AI in computer-aided arrhythmia classification systems. The introduction sets the stage for the paper's focus on reviewing machine and deep learning methods for detecting shockable ECG signals. The future research recommendations include applying state-of-

https://doi.org/10.46335/IJIES.2024.9.7.3 e-ISSN: 2456-3463
Vol. 9, No. 7, 2024, PP. 10 -15

the-art methods on big data, collecting new high-quality and long-duration ECG data, using compression techniques to simplify deep learning models, employing data augmentation methods, utilizing cross-validation techniques, optimizing algorithms for feature selection, balancing data, exploring lightweight transfer learning models, and considering new loss functions.

6. Advancing ECG Analysis: Deep Multi-Task Learning for Enhanced Accuracy and Transferability [7]

This research introduces a deep multi-task learning scheme for ECG data analysis to improve the accuracy and transferability of models, addressing challenges faced by traditional deep learning algorithms. The proposed deep multi-task learning scheme for ECG data analysis improves accuracy by up to about 5.1% and shows potential for generalizability to other datasets.
The paper discusses the application of a deep multi-task learning scheme for ECG data analysis, emphasizing the division of tasks, dataset construction, and the design of a deep parameter-sharing network to improve accuracy by up to 5.1% using the MIT-BIH database.
The methodology involved proposing a deep multi-task learning scheme for ECG data analysis, converting the problem into a multi-task learning one, constructing multiple datasets for each task, designing a deep parameter-sharing network, and conducting experiments using the MIT-BIH database.
The hypotheses tested in the study are that the deep multi-task learning scheme for ECG data analysis can improve the accuracy of ECG data analysis by up to about 5.1% compared to traditional deep learning algorithms and that the proposed scheme requires limited efforts to fine-tune the network and can enable the trained model to be well applied to other datasets.

7. Revolutionizing ECG Arrhythmia Detection: A Novel Inter-Patient Paradigm with SVM Classification Superiority [8]

This study aims to design and investigate an automatic classification system for ECG arrhythmia detection using a new comprehensive ECG database inter-patient paradigm separation without performing any feature extraction. The study focuses on improving the detection of minority arrhythmical classes using machine learning techniques. The research emphasizes the importance of the proposed inter-patient paradigm separation method for ECG classification without the need for complex data preprocessing for feature engineering. The study evaluates four supervised machine learning models (SVM, KNN, Random Forest, and ensemble) for classifying Normal Beat, Left Bundle Branch Block Beat, Right Bundle Branch Block Beat, Premature Atrial Contraction, and Premature Ventricular Contraction using real ECG records from MIT-DB. The results show that the SVM classifier outperforms the other methods in terms of accuracy, precision, recall, and f1-score, achieving an accuracy of 0.83. The paper concludes that the SVM model is more realistic in a clinical environment for classifying various types of ECG signals collected from different patients.
The main findings of the study indicate that the SVM classifier outperformed other methods in classifying ECG arrhythmias using the proposed inter-patient paradigm, achieving an accuracy of 0.83. The study also showed that the separation approach used in the study was more effective and promising for all classes, including the minority class PAC. Overall, the SVM model demonstrated the best performance compared to other methods, highlighting its importance as a classification tool for ECG arrhythmia.
The methodology in the study involved testing four supervised machine learning models (SVM, KNN, Random Forest, and Ensemble) on inter-patient real ECG records from MIT-DB for classifying five types of beats, without the need for feature extraction. The study included segmentation and normalization of the data, followed by evaluating the performance using metrics like accuracy, precision, recall, and f1-score.

8. Automated ECG Signal Analysis: Machine Learning Approaches for Distinguishing Cardiovascular Health [9]

This study discusses the importance of accurate categorization of cardiovascular diseases using ECG data and machine learning models, highlighting challenges in distinguishing between healthy and sick individuals, emphasizing the significance of automated ECG signal identification, and mentioning the Physio Net initiative. The conclusion addresses the advantages and disadvantages of different machine learning methods and declares no competing interests.The study focuses on using machine learning algorithms for the automated categorization of ECG data, utilizing open-source ECG data banks to develop classification methods, aiming to accurately distinguish between healthy individuals and those with cardiovascular disorders.The methodology involved using machine learning models for ECG data categorization, performing cross-validation to improve evaluation metrics and prevent bias, specifically using 10-fold cross-validation, and

https://doi.org/10.46335/IJIES.2024.9.7.3 e-ISSN: 2456-3463
Vol. 9, No. 7, 2024, PP. 10 -15

adjusting hyperparameters using a specific formula. The study tests hypotheses related to the automated categorization of ECG data using machine learning algorithms, including Gaussian NB, Random Forest, Logistic Regression, Linear Discriminant Analysis, and Dummy Classifier. The research aims to classify ECG data, including single-lead ECGs, by leveraging open-source ECG data banks.

III. SUMMARY TABLE

Sr. no.	Research paper	Model used	Limitations found
1.	[2]	Random Forest, Bayesian Network, Neural Networks	-Challenges in digitizing ECG data. -lack of digital ECG clinical acquisition.
2.	[3]	Deep Learning Model combined with Neural Network and Decision Trees	-Further studies are needed to assess the clinical significance of the new technology for application in clinical practice.
3.	[4]	Logistic Regression, Decision Trees, Nearest Neighbours	-Increased accuracy when focusing on specific diseases rather than normal and abnormal ECG
4.	[5]	Artificial Neural Network, specifically the Multilayer Perceptron (MLP), Generalized Feedforward Neural Network (GFFNN), and Modular Neural Network (MNN) models.	None
5.	[6]	Support Vector Machine (SVM)	-Imbalanced datasets affecting DL model efficiency -Need for big data for better performance
6.	[7]	Deep Multi-task Learning Scheme with a Deep Parameter-Sharing Neural Network	-The study mentions the challenge of different diseases having similar effects on normal ECG signals, leading to a potential decline in accuracy when applying a model trained on one group of patients to another.
7.	[8]	Support Vector Machine (SVM)	-Unexpected performance of the ensemble voting method -Lack of significant improvement from the proposed inter-patient paradigm separation approach.
8.	[9]	Various ML algorithms including Gaussian NB, Random Forest, Logistic Regression, Linear Discriminant Analysis, and Dummy Classifier.	-Reliance on a specific dataset -Lack of generalizability -Need for further research on larger and more diverse datasets

Table 1: Analysis of various techniques used for ECG analysis.

IV. CONCLUSION AND FUTURE SCOPE FOR RESEARCH

The paper discusses the advantages and disadvantages of different machine learning methods in ECG data analysis, the focus on using machine learning algorithms for automated categorization of ECG data, and suggests a potential future research focus on improving accuracy in distinguishing between healthy and sick individuals using machine learning and deep learning methods in cardiovascular disease classification.

The future scope of ECG analysis includes applying advanced methods on big data, collecting high-quality ECG data, using compression techniques, data augmentation, cross-validation, optimizing algorithms, balancing data, pattern mining from different sources [10], exploring transfer learning models, and considering new loss functions. Future research could also focus on designing deep learning schemes for ECG data analysis carefully and improving the accuracy of distinguishing between healthy and sick individuals using machine learning and deep learning methods in cardiovascular disease classification.

References

[1] ECG interpretation: Characteristics of the normal ECG (P-wave, QRS complex, ST segment, T-wave) – Cardiovascular Education (ecgwaves.com)

[2] Lyon, Aurore; Mincholé, Ana; Martínez, Juan Pablo; Laguna, Pablo; Rodriguez, Blanca (2017). Supplementary material from "Computational techniques for ECG analysis and interpretation in light of their contribution to medical advances". The Royal Society. Collection. https://doi.org/10.6084/m9.figshare.c.3956569.v1

[3] Y.-Y. Jo, Y. Cho, S.Y. Lee, et al., Explainable artificial intelligence to detect atrial fibrillation using electrocardiogram, International Journal of Cardiology, https://doi.org/10.1016/j.ijcard.2020.11.053

[4] Sushmita Roy Tithi, Afifa Aktar, Fahimul, 2019 IEEE Region 10 Symposium (TENSYMP), https://doi.org/10.1109/tensymp46218.2019.8971374

[5] Shivajirao M. Jadhav, Sanjay L. Nalbalwar, Ashok A. Ghatol . Artificial Neural Network Models based Cardiac Arrhythmia Disease Diagnosis from ECG Signal Data. International Journal of Computer Applications. 44, 15 (April 2012), 8-13. DOI=10.5120/6338-8532

[6] M. Hammad, Rajesh N.V.P.S. Kandala, A. Abdelatey et al., Automated detection of shockable ECG signals: A review, Information Sciences, https://doi.org/10.1016/j.ins.2021.05.035

[7] J. Ji, X. Chen, C. Luo and P. Li, "A deep multi-task learning approach for ECG data analysis," 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Las Vegas, NV, USA, 2018, pp. 124-127, doi: 10.1109/BHI.2018.8333385. keywords: {Task analysis; Electrocardiography; Neural networks; Data analysis; Feature extraction; Heart beat; Machine learning},

[8] Sraitih, M.; Jabrane, Y.; Hajjam El Hassani, A. An Automated System for ECG Arrhythmia Detection Using Machine Learning Techniques. *J. Clin. Med.* 2021, *10*, 5450. https://doi.org/10.3390/jcm10225450

[9] Seyed Matin Malakouti, Heart disease classification based on ECG using machine learning models, Biomedical Signal Processing and Control, Volume 84, 2023, 104796, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2023.104796.

[10] Muley, A. (2020). Global Data Fusion versus Local Pattern Fusion in Mining Multiple Databases: A Comparative Review. Journal of Computational and Theoretical Nanoscience, 17(9-10), 3844-3849.