Detection of Peripheral Pulse Patterns in Coronary Artery Disease and Hypertension Using Machine Learning

[1]Raksha Amrutkar, [2]Aditya Jadhav, [3]Nishant Patil
[1]Student, [2]Student, [3]Professor
[1]Biomedical Engineering,
[1]Mgm's College of Engineering and Technology, Kamothe, India

Abstract: CVDs especially the coronary artery disease (CAD) and hypertension are the major morbidity and mortality burden worldwide, and engaged a dire need of non-invasive, reliable, diagnostic measures. This study presents a machine-learning-based model of detecting the peripheral pulse pattern of CAD and hypertension using impedance plethysmography. A dataset of fifteen morphologies of pulse was curated, processed, and analyzed using various learning algorithms, such as Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbors (KNN) and Decision Tree (DT) classifiers. An impedance plethysmography-based analyzer was used to record peripheral pulse waveforms of subjects with varying cardiovascular conditions. Strict preprocessing and feature-extraction steps were used to improve signal fidelity and optimize the output of the model. The highest predictive accuracy of 84.62 was obtained with the Random Forest classifier, where the precision, the recall, and the sensitivity were equal to 84.72, 84.62, and 84.62 respectively. The suggested method shows a good prospect of a non-invasive diagnostic means of early detection of cardiovascular abnormalities which would provide the promise of early intervention, better patient recovery, and less healthcare cost.

Keywords: Peripheral Pulse, Machine Learning, Coronary Artery Disease, Hypertension, Impedance Plethysmography, Pattern Recognition, Random Forest, Cardiovascular Diagnostics.

I. INTRODUCTION

Cardiovascular diseases, especially coronary artery disease and hypertension, are major global health burdens. The main cause of CAD is the progressive constriction of coronary arteries under the influence of deposition of atherosclerotic plaques, which can trigger such serious incidents as myocardial infarction and stroke. Hypertension, or chronically high blood pressure, is a significant risk factor without any symptoms that can substantially increase the morbidity and mortality of cardiovascular disease, which develops silently in the early stages. Due to the increasing incidence of these disorders, the requirement of non-invasive diagnostic methods that are easy to use and provide early diagnosis and effective management is dire [1],[2].

The traditional diagnostic methods, which include angiography and echocardiography, are clinically effective but tend to be invasive, expensive and can only be done by a specialist, therefore limiting their universal availability and consistent results in terms of diagnosis [3],[4]. Such a situation highlights the need to implement novel solutions that are both accurate, cost-effective, and clinically useful, especially where resources are limited by the healthcare environment. The peripheral pulse signals present a bountiful source of cardiovascular data as they represent the mechanics between the heart and the arterial system [5]. The change in the morphology of the waveform has the capability of detecting minor abnormalities in the arterial rigidity, blood flow, and the heart contractile force. An effective method to detect these waveform properties is impedance plethysmography (IP) which is a non-invasive method used to record alterations in blood volume in the peripheral vessels based on changes in electrical impedance, [6].

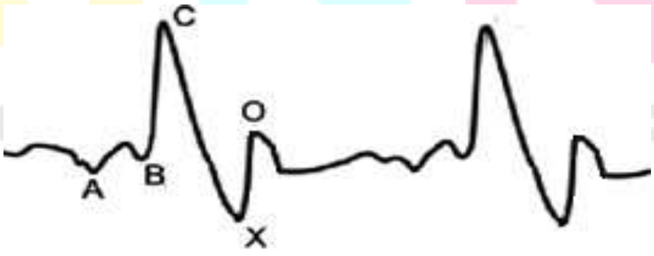


Fig. 1. Normal Radial Pulse waveform with characteristics points

Machine learning (ML) has become a promising approach to the use of cardiovascular diagnostics, which aims to reveal concealed patterns within complex physiological data, as well as improve the accuracy of the diagnostic technique [7]. The interpretation of peripheral pulse waveforms can be automated with the aid of ML algorithms to minimize the subjective bias and facilitate the reliable classification of cardiovascular states [8]. The Peripheral Pulse Analyzer (PPA) is an impedance plethysmography instrument of the Bhabha Atomic Research Centre (BARC) which detects characteristic pulse morphologies [9]. Although past research has characterized up to eight major patterns of waveforms used in cardiovascular research (some of which are associated with particular diseases), little has been done to map out the complete set of patterns based on sophisticated computational methods [10].

It aims to contribute to the development of this area by creating a data-base of fifteen peripheral pulse patterns of CAD, hypertension, and a normal cardiovascular profile. A set of machine learning algorithms are run and their results are systematically compared to select the most successful classifier [11]. Accuracy, precision, recall and sensitivity are some of the key metrics used in assessing model performance. Moreover, a graphical user interface (GUI) is created which will be used in real-time clinical practice [12]. Through developing an automated none-invasive diagnostic system, the proposed research will assist healthcare providers in diagnosing diseases at their early-stage, minimize the need to use invasive methods, and enhance patient outcome using a convenient and cost-effective cardiovascular risks assessment metric.

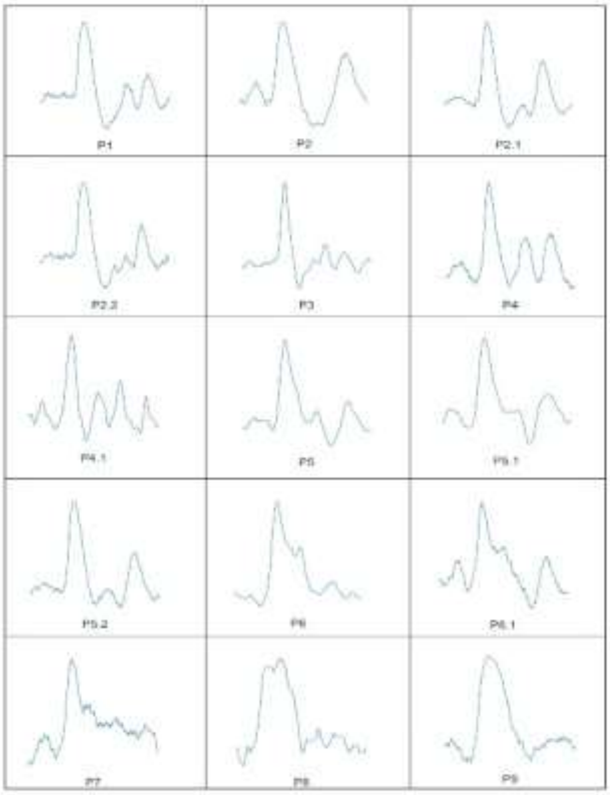


Fig. 2. Fifteen morphological patterns of the peripheral pulse

LITERATURE REVIEW

Diagnostic measurement of peripheral pulse waveforms as a predictor of cardiovascular diseases, and predominantly coronary artery disease (CAD) and hypertension in particular, has received growing interest in recent years. The constant problem with this area is that the morphology of the pulse is highly inter-individually variable, even amongst patients of the same disease type, which makes it difficult to consistently and reliably classify pulses [13]. A number of studies have examined the application of impedance plethysmography (IPG), which is a well-developed non-invasive method that captures changes in electrical impedance as an indicator of variation in peripheral blood volume [14]. The Peripheral Pulse Analyzer (PPA), invented by the Electronics Division of the Bhabha Atomic Research Centre (BARC), has allowed the recording of fine details of peripheral pulse which display characteristic morphological characteristics [15].

Machine learning (ML) algorithms have become potent methods of categorizing such waveform patterns [16]. Supervised algorithms such as Support Vector Machines (SVM), Decision Trees, Random Forests (RF), and K-Nearest Neighbors (KNN) have been used to classify the normal, hypertensive and CAD-affected subjects. An example is the case study of Yadav et al., which used an artificial neural network (ANNs) on Fast Fourier Transform (FFT)-processed information to identify CAD and hypertension cases with classification accuracy of up to 85 percent [17]. The recent events emphasize the increasing application of deep learning in this field. VGG16, VGG19, and especially ResNet50 (enhanced by residual connections to help deep networks with their training) are convolutional Neural Networks capable of finding patterns indicative of a disease on a pulse with classification accuracy exceeding 90% [18]. These models are able to automatically learn rich hierarchical properties on raw waveform data, eliminating the need of handcrafted features, and providing better diagnostic performance.

Previously published research has tended to describe eight major morphological patterns of pulse (P1 to P8) that are associated with different clinical conditions, such as normal cardiovascular functioning to severe arterial disease. Quantitative methods of spectral analysis such as Fourier-based calculation of a Morphology Index (MI) have been employed to measure arterial compliance and to distinguish normal and diseased populations with an assessment of much lower MI values in CAD patients. Although the clinical gold standards of electrocardiography (ECG), angiography, and cardiac computed tomography (CT) prove useful in the treatment process, being invasive, expensive, and requiring a professional analysis, the informativity of non-invasive machine-learning-based diagnostic models as a possible alternative is strengthened. Regardless of these encouraging developments, there are a number of limitations. The limited size of the available datasets, the effect of confounding variables on pulse morphology, e.g. age, comorbidities, and the necessity of combining multi-modal data to increase diagnostic strength are the key challenges. Future research areas focus on annotated database extension, advanced feature-extractors and full clinical validation to make safe and extensive clinical use.

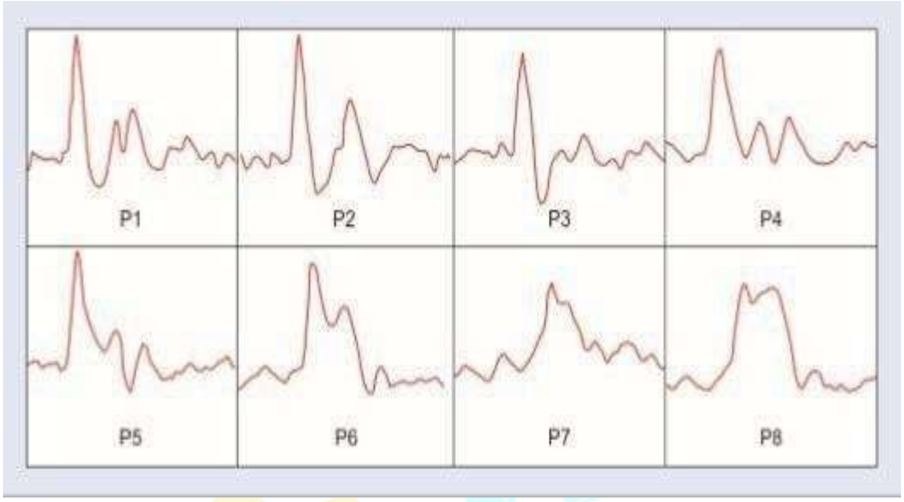


Fig. 3 Eight morphological patterns (P1–P8) of the peripheral pulse.

MATERIALS AND METHODS

The study is devoted to the idea of using machine learning (ML) algorithms to identify coronary artery disease (CAD), hypertension, and normal cardiovascular conditions based on the analysis of peripheral pulse waves. The methodology is modeled to gather, process and categorize physiological indications obtained with non-invasive sensorimotor measurements, such as electrocardiography (ECG) and photoplethysmography (PPG). Once acquired, these signals are preprocessed and features are extracted by the time and frequency domains. The features extracted are then processed through an array of ML algorithms such as classical classifiers and deep learning models to create predictive models that will accurately distinguish between healthy, hypertensive, and CAD-affected people. It also has a graphical user interface (gui) to make it easy to use in clinical environments. The dataset contains 300 peripheral pulse waveforms, with 100 samples each for Normal, Hypertension, and coronary artery disease (CAD) conditions. In this study, 33.3% of the data from each category was used for training and evaluation. Of this data, approximately 80% was used for training the models, 10% for validation to reduce overfitting, and the remaining 10% for testing to evaluate model performance.

3.1 System Architecture

The entire system architecture makes the data acquisition, signal processing, feature extraction, machine learning classification, and result interpretation processes linear.

Data collection: Ecg and PPG sensors were used to measure non-invasive measures of cardiovascular health, both of which are peripheral pulse signals.

Preprocessing: The raw signals were filtered on noise, normalization and data imputation in order to cut any artifact and also to ensure high signal fidelity.

Feature Extraction: Temporal features and spectral frequency domain features were calculated with special attention to Fast fourier Transform (FFT)-based spectral analysis.

The Morphology Index (MI) was one of the most important parameters that were computed to measure arterial compliance and aid in classification [19]. Several machine learning models were then trained using the enriched feature dataset. Classical algorithms were applied including Support Vector Machines (SVM), Decision Trees or K-Nearest Neighbors (KNN), and Random Forests (RF) as well as deep learning models, namely Convolutional Neural Networks (CNNs) using the architecture of VGG16, VGG19, and ResNet50. The problem of small size of clinical dataset was solved with the help of transfer learning techniques to enhance the generalizability of the models. The architecture facilitates the training process, hyperparameter optimization, and validation, the best classification performance is guaranteed. After the training, the system provides the predictions of the inputs as normal, hypertensive, or CAD-positive, so as to assist in clinical decision-making.

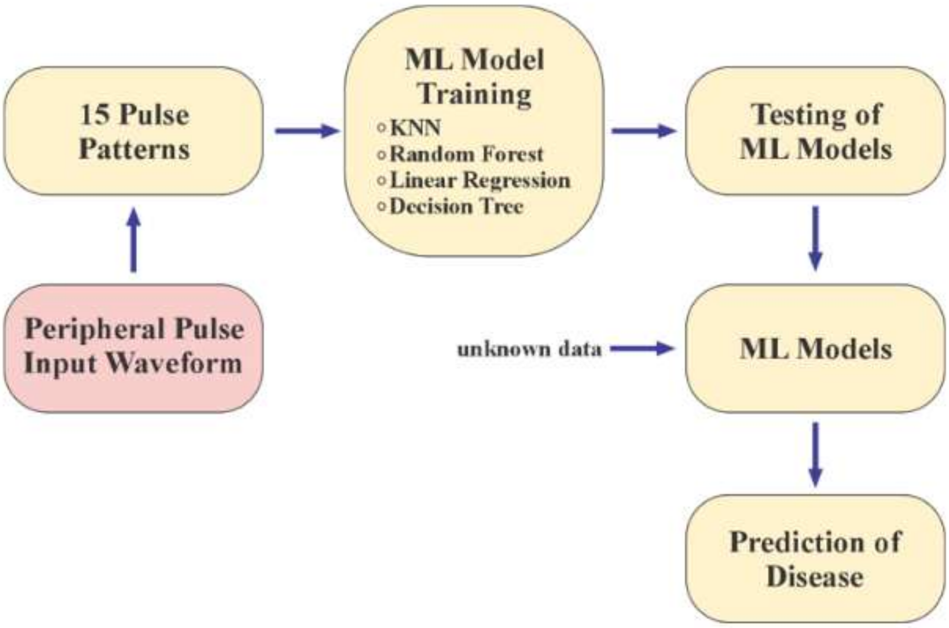


Fig. 4 System Architecture

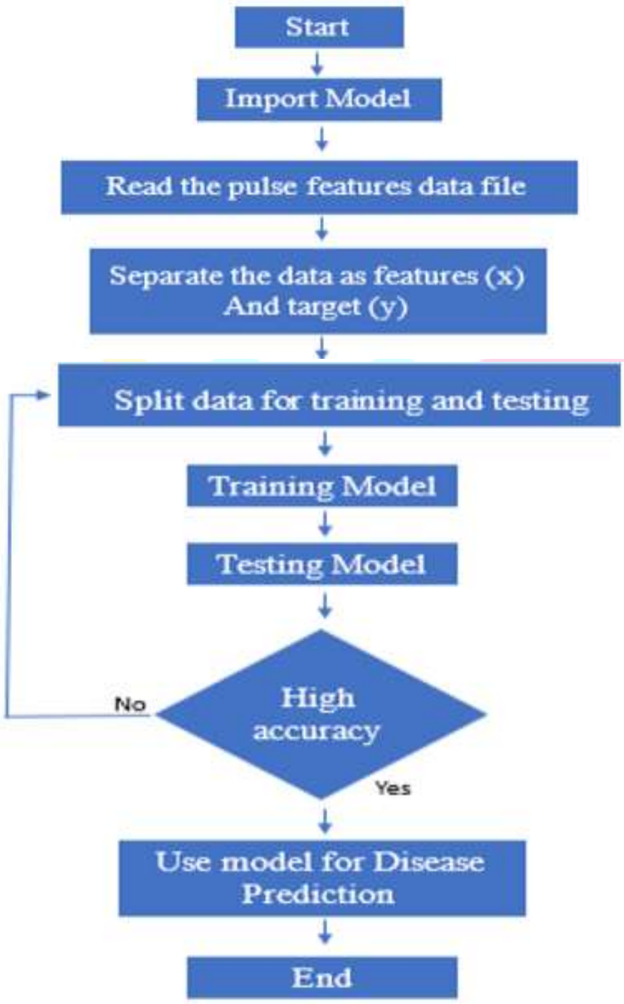


Fig. 5 Flow Chart for training ML mode

3.2 Performance Evaluation

Four main machine learning models (Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Tree) were built with Python because of its strong ecosystem to support machine learning application [20].

Supervised Learning: The models have been trained with well-labeled datasets with a common 70 percent-30 percent training-testing split to ensure sound generalization.

Logistic Regression:Logistic Regression was used as a baseline statistical model with which the likelihood of cardiovascular disease presence condition was estimated as a dependent variable of patient characteristics. In as much as it was good at binary classification, its predictive capability was poor in non-linear physiological relationships.

Support Vector Machine (SVM): The support vector machine (SVM) was employed to detect the best hyperplane to discriminate between cardiovascular groups, which offers good performance in high dimensional space features.

Decision Tree (DT): A Decision Tree classifier was used to build rule based decision boundaries, which provides interpretability and non-linear dependencies.

K-Nearest Neighbors (KNN): KNN is a non-parametric flexible method of test sample classification, in which the majority label of closest feature space neighbors is used to classify the test sample.

In the case of unsupervised learning processes, k-means clustering has been used to investigate natural groupings in unlabelled data. The most important performance measures were Accuracy, Precision, Recall (Sensitivity), Specificity, and F1-Score to represent a more in-depth evaluation of the clinical reliability of each model. The grid search and cross-validation were used to optimize the hyperparameters and decrease the overfitting. Random Forest was the most successful out of all tested methods; it reached a high classification accuracy of 84.62, as compared with both classical machine learning solutions and various deep learning baselines published in the literature.

3.3 Developed Graphical User Interface (GUI)

A specialized Graphical User Interface was created to help establish an interactive interface between the intricate machine learning pipeline and end-user demands to achieve a real-time clinical deployment.
The major aspects of the GUI are,

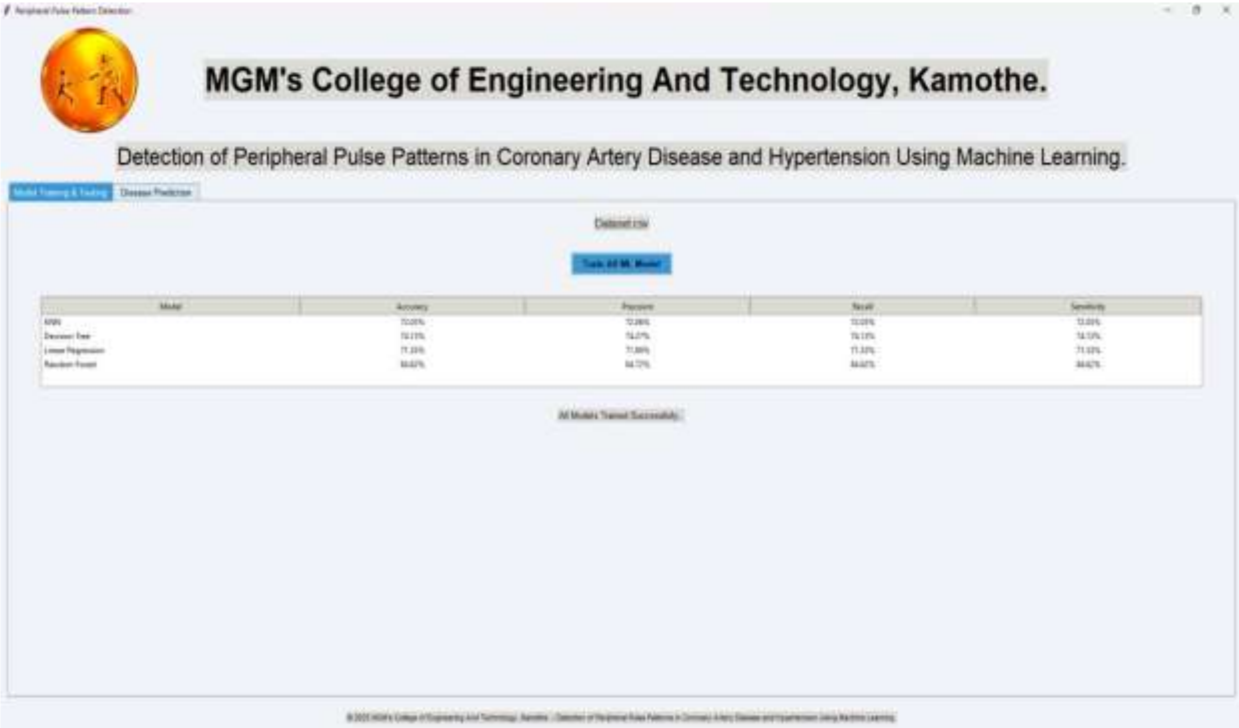
Data Importation: Support of ECG and PPG waveform formats, seamless integration of datasets of patients collected through a variety of sources.

Automated Preprocessing: Noise removal, normalization and missing-data filling to create clean and analysis-ready signals.

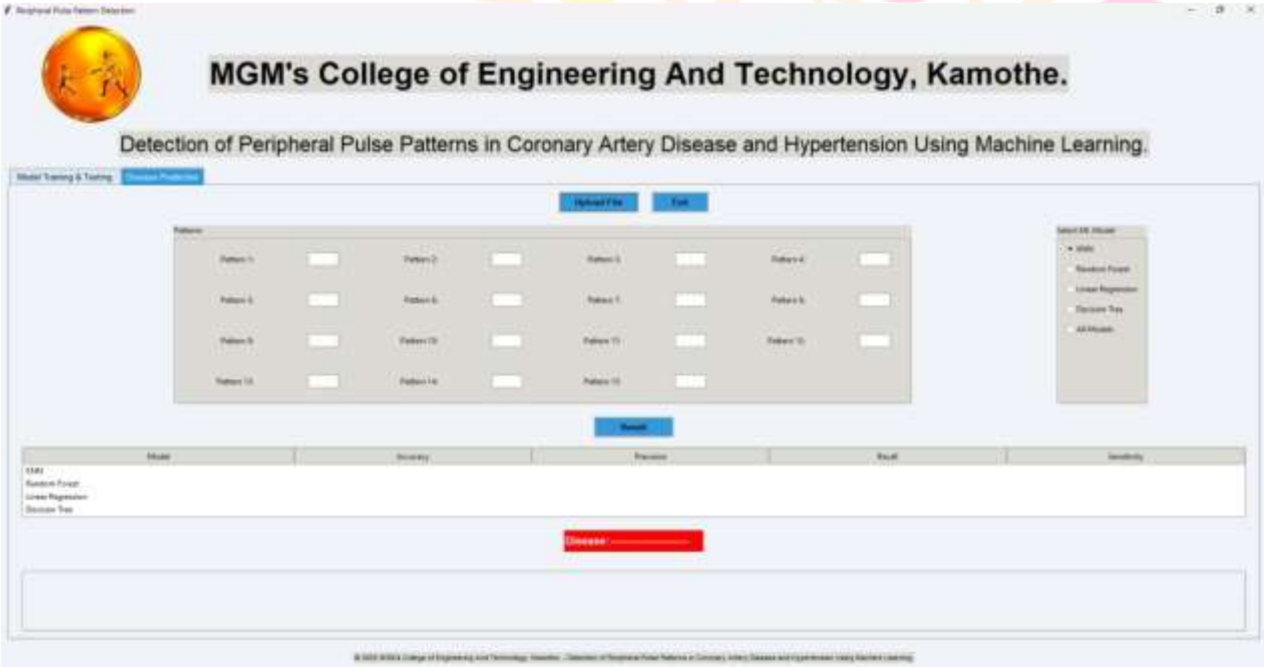
Model Selection: A trained machine learning model menu enables clinicians to make a choice of an algorithm according to the performance metrics they wish.

Real-Time Diagnostics: Immediate classification of the input data as normal, hypertensive or CAD-impacted, with confidence scores to be used in clinical interpretation.

Visualization and Reporting: Waveforms and extracted feature and classification results can be presented in graphical form, making them easier to interpret and enabling the creation of patient-friendly diagnostic reports. This clinician-friendly architecture will make complicated computational models accessible to clinicians to encourage quick adoption and enhance the efficiency of cardiovascular risk assessment.



(a)



(b)

Fig. 6 Interface of GUI Model (a) Model Training & Testing (b) Disease Prediction

RESULT AND DISCUSSION

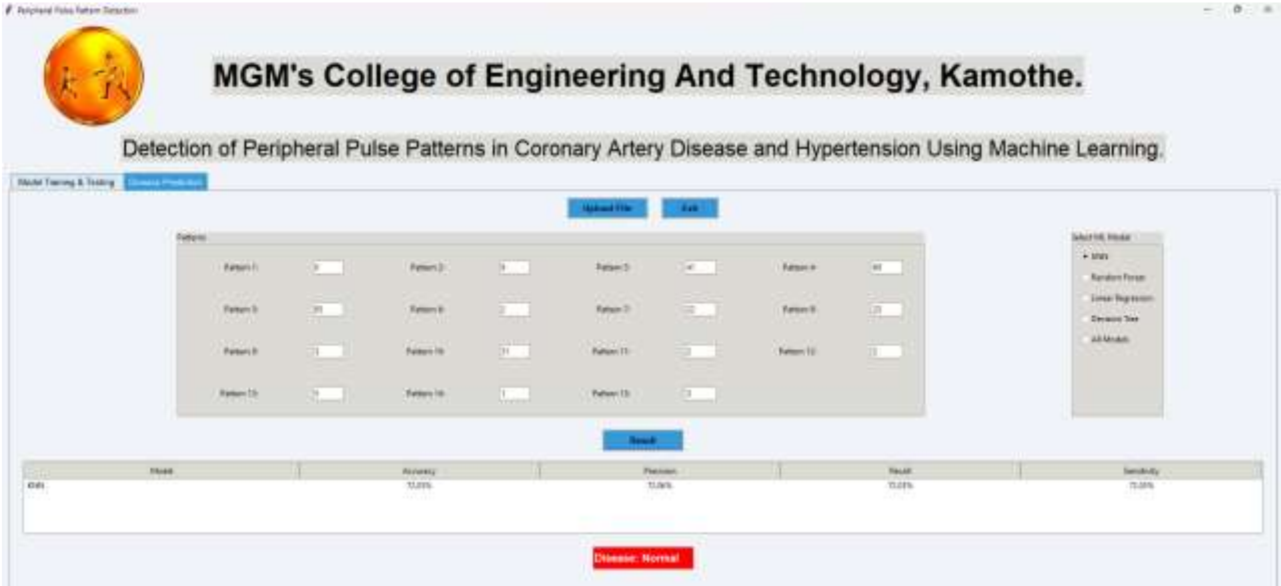
The results of the experiments in this paper indicate that the Random Forest (RF) algorithm is at a better position to predict the type of the peripheral pulse than other machine learning models. The RF model showed the best predictive performance with a classification accuracy of 84.62, precision of 84.72, a recall of 84.62 and the same sensitivity of 84.62 among all considered classifiers. These findings validate potential of RF to efficiently record the morphological changes in peripheral pulse waveforms that are both complicated and to differentiate between normal, hypertensive and CAD-impacted subjects. Other algorithms like the Decision Tree (DT), K-Nearest neighbours (KNN) and Linear Regression yielded lower accuracy rates of 74.13, 72.03 and 71.33 respectively. The existence of this gap in performance hints at the increased stability and strength of ensemble-based models such as Random Forest which harnesses on several decision trees to minimize variance and enhance classification stability. The data gathered in 30 different subjects who represented different cardiovascular conditions was used to validate the model. The RF classifier produced predictions that were highly consistent with clinical diagnoses indicating that the proposed framework is valid in the real world. Notably, the model could identify fifteen distinct pulse morphologies, which provided a broader diagnostic range than prior research that used fewer patterns of the waveforms. Introduction of Graphical User Interface (GUI) made the system much easier to interpret and to use clinically. The GUI did not just show the results of classification, but also offered the probability level, knowledge about the features, and visualization of the pulse signals. This aspect enabled doctors to compare machine predictions with the waveform data and increase the confidence in the diagnosis and provide timely treatment. These results support the possibility that machine learning and especially the ensemble methods can be used as non-invasive diagnostic tools in cardiovascular medicine. Although the deep learning architecture like ResNet50 has shown relatively higher accuracies in other works, the Random Forest model in this work had a better balance in performance, interpretability and clinical adaptability. All the comparative performance of the implemented ML algorithms are summarized in Table 1, and disease predictions results are provided in Table 2 in 30 subjects. Figure 7 and Figure 8 show the representative predictions of the RF model of healthy and hypertensive cases, respectively. In general, the research shows that the Random Forest is a valid and understandable cardiovascular disease classifier. The fact that it can produce high classification measures and the user-friendly nature of the GUI points to a clear direction of producing real-time, clinical grade diagnostic tools.

Table 1. Parameters of ML models (%)

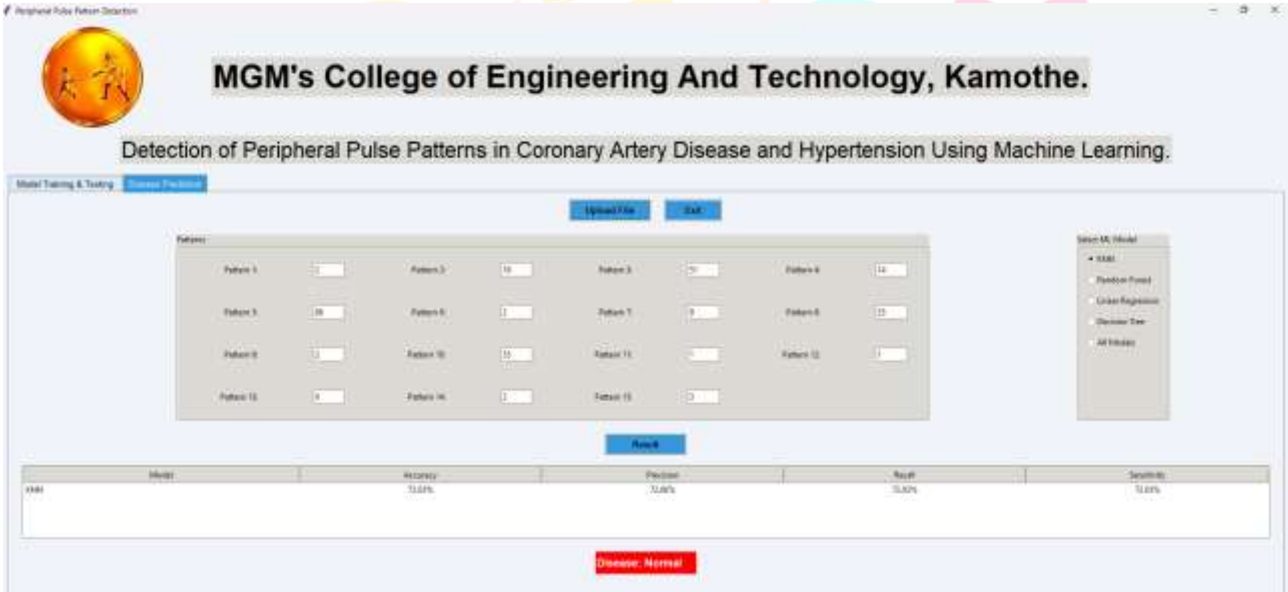
Parameters	**Accuracy**	**Precision**	**Recall**	**Sensitivity**
KNN	72.03%	72.06%	72.03%	72.03%
Random Forest	84.62%	84.72%	84.62%	84.62%
Linear Regression	71.33%	71.33%	71.89%	71.33%
Decision Tree	74.13%	74.27%	74.13%	74.13%

Table 2. Result of disease prediction using ML model for 30 different subjects
(1 - Normal, 2 - Hypertension disease, 3 - CAD disease)

Subject	Input	Probabilities of Patterns numbers															Predicted Output
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	
1	1	1	10	37	41	64	1	13	36	0	8	21	6	3	14	43	1
2	1	0	9	41	60	85	2	22	23	3	31	2	2	5	1	3	1
3	1	2	16	51	14	86	2	9	33	2	35	1	1	0	2	3	1
4	1	1	1	45	23	71	3	8	3	0	85	0	0	9	6	44	1
5	1	20	4	5	17	87	60	0	57	9	7	0	0	0	0	2	1
6	1	5	7	111	81	186	0	3	0	1	0	0	0	1	0	6	1
7	1	3	1	25	101	61	3	26	0	0	18	0	0	0	0	10	1
8	1	6	1	11	96	64	1	36	0	0	18	0	0	0	0	35	1
9	1	14	0	17	18	158	25	33	14	0	3	0	0	0	0	12	1
10	1	105	1	3	21	54	20	21	0	0	9	0	0	1	0	1	1
11	2	0	0	0	0	33	2	66	0	0	0	0	0	150	58	0	2
12	2	2	1	5	11	4	5	9	13	6	100	0	0	6	1	2	2
13	2	0	0	0	0	33	2	66	0	0	0	0	0	150	58	0	2
14	2	6	48	2	0	7	6	1	46	5	4	6	1	3	9	292	2
15	2	3	35	2	1	3	1	5	42	14	4	3	5	8	10	305	2
16	2	5	15	2	5	5	4	3	38	5	5	5	0	2	6	307	2
17	2	11	0	0	0	25	3	24	2	0	0	2	2	107	126	2	2
18	2	14	0	0	0	32	2	14	1	0	0	4	1	104	148	2	2
19	2	0	0	0	0	35	0	18	0	0	0	1	1	143	60	0	2
20	2	10	8	0	0	24	3	1	1	0	3	33	1	93	107	343	2
21	3	1	1	0	1	1	1	0	222	3	0	20	2	19	1	1	3
22	3	0	0	0	0	137	3	110	0	0	5	0	0	1	96	10	3
23	3	0	0	0	0	150	7	88	0	0	9	0	0	0	97	4	3
24	3	1	0	1	0	74	1	187	0	0	0	0	0	2	0	0	3
25	3	0	0	0	0	17	0	238	0	0	6	0	0	0	13	26	3
26	3	0	0	1	0	30	2	212	1	0	0	0	0	0	1	6	3
27	3	0	0	1	0	30	2	212	1	0	0	0	0	0	1	6	3
28	3	3	0	0	0	5	5	168	2	0	8	0	2	39	1	1	3
29	3	4	36	8	63	0	19	2	0	2	124	0	0	8	1	0	3
30	3	4	0	0	2	1	2	5	17	15	221	2	0	14	0	1	3

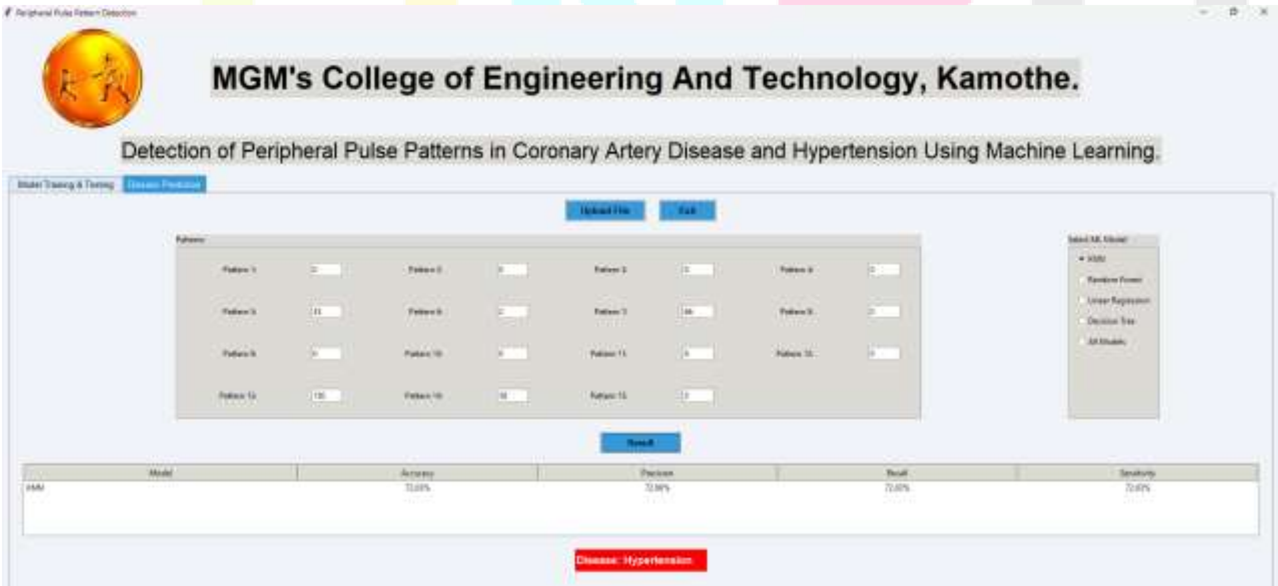


(a)

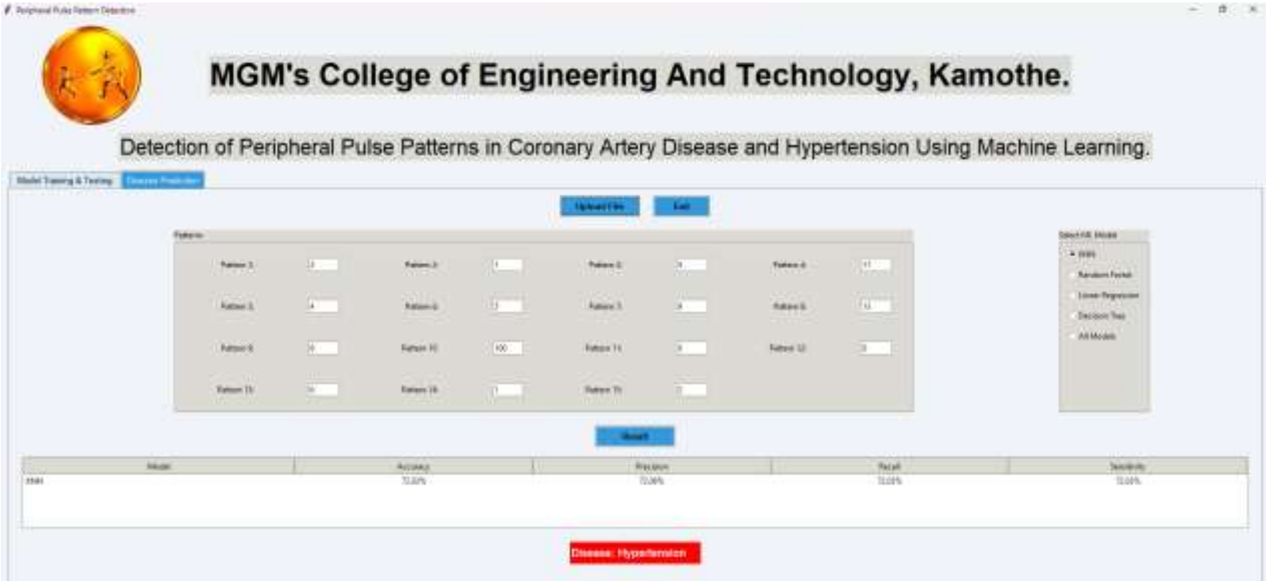


(b)

Fig. 7. Healthy Prediction using RF Model of two different subjects (a), (b)



(a)



(b)

Fig. 8. Hypertension Prediction using RF Model of two different subjects (a), (b)

CONCLUSION

It has also shown that machine learning methods are viable and useful in the non-invasive detection and classification of coronary artery disease (CAD), hypertension and normal cardiovascular conditions using peripheral pulse waveforms. The suggested framework has demonstrated that it has a strong potential to boost the accuracy and reliability of cardiovascular diagnostics through the use of a variety of machine learning algorithms, specifically, Random Forest (RF) and sophisticated deep learning models. The use of frequency domain information, especially the Morphology Index (MI) calculated using Fast Fourier Transform (FFT), was useful in the representation of fine vascular characteristics related to pathological states. Random Forest classifier was the most suitable model to use in clinical practice, among other evaluated models; it exhibited the best balance between accuracy, interpretability, and robustness, which surpasses the traditional statistical methods. Moreover, the computational complexity and real-life demands in healthcare were connected through the creation of a graphical user interface (GUI). The GUI allowed making predictions of the disease in real-time, offered confidence-based outputs, and provided a complex set of visualization tools, which simplified the adoption of the system in clinical processes and made healthcare professionals more timely in making their decisions. The results highlight the promise of machine learning-based diagnostic solutions to aid in early diagnosis and to enhance patient risk-classification and reduce the reliance on expensive and invasive diagnostic methods. These developments are in tandem with the increasing need to have accessible, affordable and patient-friendly solutions in cardiovascular care. Moving forward, future directions should focus on growing the volume and variety of data, unifying multimodal physiological measurements (e.g., ECG, PPG, and imaging data) and confirming them by conducting massive clinical studies. These measures are critical to the reinforcement of the robustness and generalizability of the models to other populations and healthcare settings. The further development of these types of systems can potentially revolutionize the world of cardiovascular diagnostics, as it will allow delivering personalized, efficient, and timely medical care, which, in the end, will lead to a better patient outcome and the optimal use of medical resources.

II. ACKNOWLEDGMENT

The authors would like to thank Dr. Geeta S. Lathkar, Dr. V. G. Sayagavi, and Rev. Fr. Richard Aloysius Coelho for their valuable support and encouragement throughout this study.

REFERENCES

[1] S. Bhat, K. S. Bhat, S. D'sa, and R. Roopa, "Variability Studies and Disease Characterization," in *A Handbook on Physiological Variability: Advanced Applications of Physiological Variability (AAPV-2010)*, G. D. Jindal, K. K. Deepak, and R. K. Jain, Eds. Mumbai: Electronics Division, BARC, 2010, pp. 98–102.

[2] R. Yadav, M. Deshmukh, and R. K. Jain, "Detecting the Patterns of Coronary Artery Disease and Hypertension using Artificial Neural Network," *Int. J. Appl. Eng. Res.*, vol. 13, no. 16, pp. 12843–12845, 2018.

[3] I. Hlimonenko, K. Meigas, and R. Vahisalu, "Waveform Analysis of Peripheral Pulse Wave Detected in the Fingertip with Photoplethysmograph," *Measurement Science Review*, vol. 3, no. 2, pp. 1–4, 2003.

[4] M. Sawant and G. D. Jindal, "Variability in Peripheral Blood Flow and Morphology Index in Controls and Hypertensive Subjects," *Int. J. Sci. Res.*, vol. 7, no. 12, pp. 45–48, 2018.

[5] World Health Organization, "Cardiovascular diseases (CVDs)," WHO Fact Sheet, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[6] D. Mozaffarian et al., "Heart Disease and Stroke Statistics—2016 Update: A Report From the American Heart Association," *Circulation*, vol. 133, no. 4, pp. e38–e360, 2016.

[7] A. Rumberger, "Coronary Artery Imaging Techniques: Angiography, CT, and MRI," *Curr. Cardiol. Rep.*, vol. 7, no. 1, pp. 7–12, 2005.

[8] J. Warrier, A. Deshpande, P. Athavale, U. Bagal, M. Rajput, and G. Jindal, "Peripheral Pulse Morphology for Early Detection of Coronary Artery Disease," *MGM J. Med. Sci.*, pp. 112–116, 2016.
[9] G. D. Jindal et al., "An Introduction to Impedance Cardiovasography," BARC External Report BARC/2001/E/003, 2001.
[10] M. Deshmukh et al., "Machine Learning Approach Combined with Statistical Features in the Classification of Peripheral Pulse Morphology," *Int. J. Comput. Methods Exp. Meas.*, vol. 12, no. 1, pp. 69–75, 2024.
[11] A. Khan, M. Qureshi, M. Daniyal, and K. Tawiah, "A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction," *Health & Soc. Care in the Community*, vol. 2023, Article ID 1406060.
[12] P. Varier, K. Muraleedharan, T. S. Muraleedharan, T. P. Udayakumari, and A. V. Sudheer, "Evaluation of Peripheral Pulse Analyzer from Ayurvedic Point of View," in *A Handbook on Physiological Variability*, G. D. Jindal, K. K. Deepak, and R. K. Jain, Eds. Mumbai: Electronics Division, BARC, 2010, pp. 128–136.
[13] R. Yadav, M. Deshmukh, and R. K. Jain, "Detecting Patterns of Coronary Artery Disease and Hypertension using ANN," *Int. J. Appl. Eng. Res.*, vol. 13, no. 16, pp. 12843–12845, 2018.
[14] S. Mascarenhas and M. Agarwal, "A Comparison Between VGG16, VGG19 and ResNet50 Architecture Frameworks for Image Classification," in *Proc. 2021 Int. Conf. Disruptive Technol.*, 2021.
[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[16] M. Sawant, G. Jindal, and A. Shah, "Fourier-Based Morphology Index for Arterial Compliance," *Int. J. Sci. Res.*, vol. 8, no. 2, pp. 21–25, 2019.
[17] A. Einstein, "Non-invasive Cardiovascular Imaging Modalities," *J. Am. Coll. Cardiol.*, vol. 63, no. 4, pp. 423–437, 2014.
[18] T. Attia et al., "Challenges and Opportunities in Artificial Intelligence for Cardiovascular Diagnostics," *Nature Rev. Cardiol.*, vol. 18, pp. 387–400, 2021.
[19] A. Allen, "Photoplethysmography and its Application in Clinical Physiological Measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.
[20] H. Kim et al., "Clinical Decision Support Systems Based on Machine Learning: A Review," *J. Med. Syst.*, vol. 46, no. 3, pp. 1–13,