International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 6, December 2025, pp. 5782~5791
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i6.pp5782-5791

Prediction of peripheral arterial disease through non-invasive diagnostic approach

Sobhana Mummaneni[1], Lalitha Devi Katakam[1], Pali Ramya Sri[1], Mounika Lingamallu[1], Smitha Chowdary Ch[2], D.N.V.S.L.S. Indira[3]
[1]Department of Computer Science and Engineering, Faculty of Engineering, Velagapudi RamaKrishna Siddhartha Engineering College, Vijayawada, India
[2]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India
[3]Department of Information Technology, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, India

Article Info

Article history:

Received Jun 17, 2025
Revised Jul 23, 2025
Accepted Sep 16, 2025

Keywords:

Electrocardiography
Laser doppler flowmetry
Machine learning
Peripheral arterial disease
Photoplethysmography

ABSTRACT

Peripheral arterial disease (PAD) is a cardiovascular condition caused by arterial blockages and poor blood circulation, increasing the risk of severe complications such as stroke, heart attack, and limb ischemia. Early and accurate detection is essential to prevent disease progression and improve patient outcomes. This study introduces a non-invasive diagnostic method using laser doppler flowmetry (LDF), electrocardiography (ECG), and photoplethysmography (PPG) to assess vascular health. LDF measures microvascular blood flow, ECG evaluates heart rate variability, and PPG captures pulse waveform characteristics. Key physiological features such as blood flow variability, pulse transit time, and hemodynamic responses are extracted and analyzed using machine learning. Random forest and XGBoost models are employed and combined using ensemble learning to classify individuals into non-PAD, moderate PAD, and severe PAD categories. A comparative evaluation shows that the ensemble model delivers superior classification accuracy. This integrated system offers a fast, reliable screening tool that supports early PAD detection and intervention. By combining multimodal signal analysis with machine learning, the approach enhances diagnostic precision and provides a scalable solution for preventive cardiovascular care.

Corresponding Author:

Sobhana Mummaneni
Department of Computer Science and Engineering, Faculty of Engineering, Velagapudi RamaKrishna Siddhartha Engineering College
Vijayawada, 520007, India
Email: sobhana@vrsiddhartha.ac.in

1. INTRODUCTION

Peripheral arterial disease (PAD), a widespread vascular condition primarily affecting the lower extremities, impacts over 200 million people worldwide [1]. Primarily caused by atherosclerosis, PAD has become a major public health concern due to its rising prevalence, driven by aging populations and risk factors such as diabetes, metabolic abnormalities, and tobacco use [2]. In advanced stages, PAD can lead to critical limb ischemia, resulting in non-healing ulcers or even limb amputation if left untreated, and significantly increases the risk of serious cardiovascular complications like myocardial infarction, stroke, and overall mortality [3]. Despite its severity, PAD remains widely underdiagnosed, particularly in the early stages, due to the limited sensitivity of the ankle brachial index (ABI), which is widely regarded as the standard screening tool especially in patients with diabetes-related arterial calcification [4]. Conventional

statistical approaches are generally inadequate in modeling the intricate nonlinear dependencies between risk factors, leading to reduced predictive performance. This study addresses the following question: Can a multimodal approach that integrates electrocardiography (ECG), photoplethysmography (PPG), and laser doppler flowmetry (LDF) signals with machine learning techniques enhance early detection and severity classification of PAD compared to conventional diagnostic methods?

Recent advancements in non-invasive diagnostic technologies offer promising solutions to these challenges. ECG, a routine clinical tool, provides insights into heart rhythm and variability, aiding cardiovascular risk assessment [5]. PPG, widely used in pulse oximeters, is an affordable and portable method for evaluating vascular health, suitable for integration into wearable devices [6], [7]. LDF measures skin blood perfusion and has shown high sensitivity in detecting compromised lower-limb circulation, particularly in high-risk groups such as hemodialysis patients [8]. Machine learning (ML) enhances these tools by identifying intricate, nonlinear data patterns that conventional models miss [1]. Prior studies demonstrate that ML-driven analyses of electronic health records and PPG data outperform traditional methods in PAD prediction [9]. Techniques for dimensionality reduction, such as principal component analysis (PCA), help retain the most significant variance in the dataset while reducing the number of features. [10] Similarly, ensemble methods like random forest and XGBoost enhance model accuracy and reduce the risk of overfitting. [11].

This study addresses these gaps by proposing a novel, non-invasive diagnostic framework that synergistically combines LDF, ECG, and PPG signals to assess microvascular blood flow, cardiac dynamics, and vascular health. Physiological features, including blood flow variability, pulse transit time, and hemodynamic responses, are extracted and analyzed using an ensemble machine learning model comprising random forest and XGBoost algorithms, optimized via GridSearchCV to classify PAD severity into Non-PAD, Moderate PAD, and Severe PAD categories with 93% accuracy. To maximize clinical utility, the model is deployed through a Flask-based web application, enabling rapid, user-friendly PAD screening in diverse healthcare settings. This approach facilitates early detection, supports preventive care, and paves the way for personalized management of PAD, ultimately aiming to reduce its clinical and economic burden. According to prior research, this work is one of the earliest attempts to combine ECG, PPG, and LDF signals with ensemble models to classify the severity of PAD. This integration offers a novel, non-invasive, and accurate diagnostic framework for early detection and stratification of PAD.

The objective of this study is to propose a non-invasive diagnostic framework for the early detection and stratification of PAD severity. This is achieved through a multimodal methodology that incorporates signals from electrocardiogram, photoplethysmography, and LDF. Essential physiological indicators such as variations in blood flow, pulse transit times, and hemodynamic responses are extracted and processed. These features are subsequently evaluated using ensemble algorithms, namely random forest and XGBoost, with hyperparameters optimized through GridSearchCV. The resulting model stratifies PAD into three categories: Non-PAD, Moderate PAD, and Severe PAD. For real-world applicability, the trained model is integrated into a Flask-based web platform, offering an accessible and real-time screening tool. This approach is intended to enhance diagnostic precision, support early interventions, and enable individualized management strategies for PAD.

The structure of this paper is organized as: section 2 discusses the state-of-the-art research and existing methodologies related to the diagnosis of PAD. Section 3 outlines the identified research gaps along with the main contributions of this work. Section 4 details the adopted methodology and the data utilized. Section 5 discusses the experimental findings. Finally, section 6 concludes the study and suggests directions for future research.

2. LITERATURE REVIEW

Allen *et al.* [12] proposed a deep learning approach utilizing photoplethysmography (DLPPG) was employed to identify PAD through the analysis of toe-based PPG signals. The objective was to evaluate the effectiveness of a convolutional neural network, specifically AlexNet with transfer learning, applied to continuous wavelet transform (CWT) spectrograms. The model achieved 86.6% sensitivity, 90.2% specificity, and 88.9% accuracy with a Cohen's Kappa of 0.76 using 5-fold cross-validation. This approach requires minimal signal preprocessing and prioritizes toe PPG, which is more clinically relevant for PAD detection than finger-based signals. The study highlighted challenges such as managing movement artifacts and signal noise. It also noted that the dataset was not fully balanced and certain health factors like diabetes were not incorporated.

Kim *et al.* [13] explored PAD detection and severity assessment using deep learning on arterial pulse waveforms. A synthetic dataset from a transmission line model simulated various PAD severities. Brachial and ankle waveforms were analyzed using a modified AlexNet CNN, achieving 97% sensitivity, 99% specificity, and accuracy—surpassing the traditional ABI method. This approach better captured

waveform morphology and individual variability. Key challenges included use of virtual data and limited real-world generalization. The study highlights deep learning's potential for accurate, non-invasive PAD screening, with future efforts focused on clinical validation and localization.

McBane *et al.* [14] introduced a model utilizing the inception time architecture to detect PAD from resting arterial Doppler waveforms. Trained on data from 3432 patients and validated on 151, the model predicted abnormal ABI values with high accuracy (rest ABI: 0.89, AUC 0.95; postexercise ABI: 0.85–0.89). While the method reduces the need for exercise testing, it depends on high-quality waveform acquisition. Limitations include exclusion of certain patient groups and limited generalizability. The approach demonstrates strong potential for scalable, non-invasive PAD screening.

Stansby *et al.* [15] conducted a prospective diagnostic study to assess the accuracy of multi-site photoplethysmography (MPPG) in identifying PAD within primary care. Using duplex ultrasound as the reference standard, MPPG demonstrated a sensitivity of 79.8% and specificity of 71.9%, comparable to the traditional ankle-brachial pressure index (ABPI), which showed 80.2% sensitivity and 88.6% specificity. Unlike ABPI, MPPG was faster, automated, and required less operator training. However, the study faced challenges such as an 8.4% test failure rate due to signal quality and prototype device limitations. Despite these constraints, the research highlights MPPG's potential as a scalable, non-invasive diagnostic tool for early PAD detection in primary care settings.

Forghani *et al.* [16] proposed DeepPAD, a novel deep learning framework for identifying PAD using Oscillo metric pulse waveforms recorded at different cuff pressures. The system employed an attention-enhanced bidirectional LSTM model to analyze raw Oscillo metric pulses and extracted features. Evaluated on data from 33 individuals, the model achieved up to 94.8% accuracy, 90.0% sensitivity, and 97.4% specificity, outperforming the conventional ABI and a genetic algorithm-based neural network (GA-NN). Despite its high performance, limitations included a small sample size and lack of PAD severity classification.

Sonderman *et al.* [17] introduced a machine learning model aimed at identifying individuals at high risk for peripheral artery disease by analyzing electronic health record (EHR) data. Unlike traditional screening approaches, this method combined ABI measurements with a broad set of patient features to improve prediction accuracy. The researchers applied a random forest algorithm to select key variables, followed by a logistic regression model to classify PAD risk. The model showed consistent performance with an AUC around 0.68 across internal and external datasets, and slightly higher accuracy (AUC 0.72) on a national sample, outperforming simpler age-based predictions. Despite these strengths, challenges remain in handling the variability and completeness of EHR data. Limitations include moderate predictive power and lack of validation in real-world clinical workflows. Future research could enhance model robustness and assess its impact on patient care**.**

3. RESEARCH GAPS AND PROPOSED CONTRIBUTIONS

Recent studies [12]–[17] on non-invasive PAD detection rely on single modalities such as PPG [12], [15], [16], Doppler waveforms [14], or EHR data [17], limiting their ability to capture PAD's complex microvascular, cardiac, and vascular dynamics. Key limitations include small sample sizes (*e.g.*, 33 individuals in [16]), synthetic datasets with poor real-world generalizability [13], absence of PAD severity classification [12], [14], [16], [17], and lack of scalable deployment mechanisms [14], [15]. The proposed study addresses these gaps by integrating LDF, ECG, and PPG signals within an ensemble machine learning framework (random forest and XGBoost), achieving 93% accuracy in classifying PAD severity (non-PAD, moderate PAD, severe PAD) on a robust 1,000 sample dataset [18], [19]. Deployed via a Flask-based web interface, this approach offers a scalable, accurate, and clinically accessible solution for early PAD detection and management, advancing preventive cardiovascular care.

4. METHODOLOGY

This study presents a robust classification framework aimed at detecting and assessing the severity of PAD using physiological signals and ensemble machine learning. The methodology prioritizes accuracy, interpretability, and effective preprocessing to ensure clinical applicability. The system leverages an ensemble classification model integrating random forest and XGBoost with soft voting, enabling multi-class classification of PAD into Non-PAD, Moderate PAD, and Severe PAD categories. The methodology is detailed in the following subsections. Data acquisition and preprocessing strategies are discussed in subsections 4.1 and 4.2. The architecture and training of the base classifiers are described in subsections 4.3 and 4.4. The ensemble approach is outlined in the subsection 4.5, followed by evaluation protocols in subsection 4.6. Figure 1 illustrates the complete PAD classification pipeline.

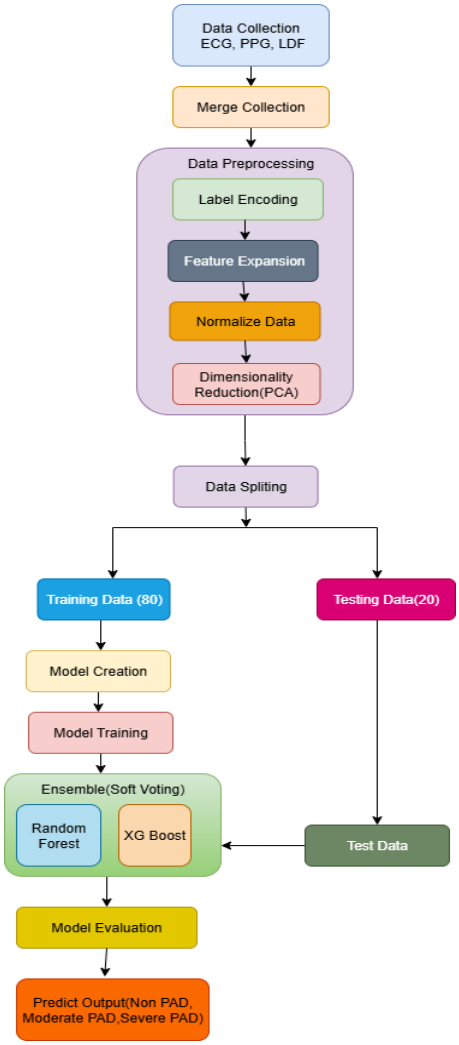

Figure 1. Proposed methodology for PAD classification

4.1. Data collection and preprocessing

The data collection and preprocessing phase, was conducted under carefully controlled conditions, ensured the integrity and relevance of the dataset, which includes essential physiological signals used for the classification of PAD. Two primary sources were considered, both publicly accessible. The first dataset, obtained from Kaggle [18], focuses on ECG and PPG signals and is organized into recordings that capture various cardiovascular conditions. The second dataset consists of LDF measurements sourced from a published medical study [19], which provides microvascular blood flow data for both healthy individuals and PAD patients. The dataset includes high-resolution LDF values indicative of tissue perfusion levels. To ensure the reliability of multimodal signal analysis, ECG, PPG, and LDF data were combined through time synchronization techniques, aligning them to a common temporal window. This enabled accurate cross-signal correlation and robust feature extraction. The final dataset is categorized into three classes: Non-PAD, Moderate PAD, and Severe PAD, supporting effective classification of PAD severity.

4.2. Data loading and processing

The data preprocessing phase begins with importing the PAD dataset, which includes physiological signals such as (LDF; blood flow in mL/min), PPG, and ECG, using the Pandas library for efficient data manipulation and exploration. The target labels ("non-PAD", "moderate PAD", and "severe PAD") are encoded using LabelEncoder to enable supervised learning. To capture nonlinear relationships among features, a second-degree polynomial expansion is applied using polynomial. Features, balancing model complexity with enhanced feature expressiveness while avoiding the computational cost of higher-degree terms. Feature standardization is then applied through StandardScaler, which normalizes the data to zero mean and unit variance, ensuring uniform contribution of all features during model training and improves the

convergence speed and stability of the machine learning algorithms. To further optimize the feature set and reduce redundancy, PCA was applied, as discussed in the following subsection.

4.2.1. Principal component analysis

To address the high dimensionality and multicollinearity introduced by polynomial feature expansion, PCA was employed to compress the feature space. PCA transforms the standardized feature space into a set of uncorrelated principal components that capture the maximum variance in the dataset, thereby simplifying the dataset while maintaining its core structural characteristics. In this study, components were retained such that 98% of the total variance was preserved, implemented via PCA($n\_components$=0.98). This strategy effectively minimizes redundancy, accelerates model training, and maintains critical physiological signal patterns. PCA was selected over alternative methods like linear discriminant analysis (LDA) and t-distributed stochastic neighbor embedding (t-SNE) due to its unsupervised nature, emphasis on variance retention, and computational efficiency for continuous biomedical data. LDA was excluded to avoid potential overfitting to class labels, while t-SNE was considered unsuitable given its computational demands and focus on data visualization rather than predictive modeling. After optimizing the feature space, the next step focused on designing machine learning models for PAD classification using the processed input signals.

4.3. Model creation

The model development phase was centered on establishing a classification framework to identify the severity of PAD, categorized into non-PAD, moderate PAD, and severe PAD, using preprocessed physiological signals. Two machine learning models such as random forest and XGBoost were selected due to their effectiveness in managing multi-class classification, particularly with medical datasets. Random forest, introduced by Breiman, is an ensemble method based on decision trees [20], builds numerous decision trees and predicts the class based on the majority vote across these trees. This method is particularly effective in capturing non-linear dependencies within diverse physiological data. For this study, random Forest was fine-tuned using GridSearchCV, set to build 400 trees, with a maximum depth of ten, minimum samples split of two, and balanced class weighting to address class imbalance and improve generalization. XG Boost, a gradient boosting-based model, sequentially constructs trees to correct errors made by prior ones and incorporates regularization to reduce overfitting [21]. For our task, XGBoost was configured with a learning rate of 0.1, maximum depth of 6, and 300 boosting rounds, using a multi-class log-loss objective. Hyperparameter tuning was also performed via GridSearchCV to enhance classification accuracy. All input features were derived from three physiological signal modalities: LDF, PPG, and ECG. Preprocessing steps included label encoding, polynomial feature expansion, standardization, and dimensionality reduction using PCA to improve learning performance. Once model architectures and hyperparameters were finalized, both classifiers were trained independently on the refined dataset

4.4. Model training

In the training phase, two ensemble learning algorithms random forest and XGBoost were used to create predictive models. Each was trained independently on a preprocessed and dimensionally reduced dataset. The random forest algorithm builds an ensemble of decision trees whose combined output enhances prediction performance and mitigates overfitting, making it suitable for handling noisy and complex data. XGBoost constructs models iteratively, where each new tree is designed to address the mistakes made by the preceding trees, resulting in improved accuracy. To fine-tune the models, a 5-fold cross-validation method was utilized within GridSearchCV, ensuring optimal selection of hyperparameters. The training set included 800 samples, split in an 80/20 ratio, allowing both models to learn relationships between input signals (LDF, PPG, ECG) and output labels while maintaining generalization to unseen data. To improve classification reliability and generalization, predictions from both trained models were integrated through an ensemble approach.

4.5. Ensemble strategy

To enhance predictive accuracy and model stability, an ensemble strategy was employed by combining predictions generated from both random forest and XGBoost models, which were optimized using GridSearchCV. This strategy leverages the individual advantages offered by each algorithm, reducing overfitting and improving generalization to unseen data. A soft voting mechanism, implemented using scikit-learn's VotingClassifier, was utilized, where equal weights were assigned to both models predicted class probabilities, and the averaged probabilities determined the final classification. This method ensures balanced contributions based on each model's confidence, leading to reliable outcomes. The ensemble model achieved superior performance in contrast to using either model independently. Following ensemble integration, the final model was evaluated using comprehensive performance metrics to validate its effectiveness.

4.6. Model evaluation

The effective performance of random forest, XGBoost, and their ensemble was evaluated on an independent test dataset based on commonly used metrics like accuracy, precision, recall, and F1-score. Accuracy measures the frequency of correct predictions, while precision and recall examine the capability of the model to correctly detect relevant cases while reducing the chance of missing them. F1-score, defined as the harmonic mean of precision and recall, provides a unified metric capturing both precision and sensitivity in classification. Additionally, confusion matrices were analyzed to examine classification errors in more detail. The ensemble model consistently surpassed the performance of the individual models across all metrics, highlighting improved reliability and robustness by capitalizing on the combined advantages from the strengths of both random forest and XGBoost algorithms.

5. RESULTS AND DISCUSSION

This research evaluates the classification performance of random forest, XGBoost, and their ensemble for assessing PAD severity using preprocessed physiological features, including LDF, PPG, and ECG signals. The dataset comprised 1,000 instances, with 80% allocated for training and 20% for testing. Model tuning was conducted using GridSearchCV. Classification accuracy was employed as the primary evaluation metric, complemented by precision, recall, and F1-score for a detailed performance analysis. As summarized in Table 1, the ensemble achieved the highest classification accuracy of 93%, surpassing random forest (90%) and XGBoost (91%) in distinguishing among non-PAD, moderate PAD, and severe PAD categories. Figure 2 illustrates these performance differences through comparative accuracy visualization. The confusion matrix in Figure 3 further highlights the ensemble's effectiveness, particularly in identifying non-PAD and severe PAD cases. Slightly reduced accuracy in moderate PAD classification may be due to overlapping physiological patterns within this category. These findings demonstrate that the combined framework enhances predictive accuracy by leveraging random forest's capability to manage high-dimensional features and reducing overfitting with XGBoost's ability to refine predictions through gradient boosting. By combining these models, the ensemble mitigates individual model weaknesses, achieving a robust balance of sensitivity and specificity critical for clinical applications.

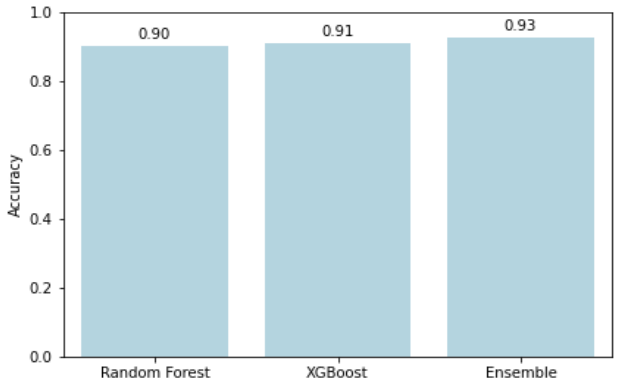

Figure 2. Comparison of classification accuracy of different models for PAD prediction

Figure 3. Confusion matrix for PAD prediction using the ensemble model

Table 1. Accuracy of models for PAD prediction

Model	Accuracy
XG Boost	91.0%
Random forest	90.0%
Ensemble	93.0%

5.1. PAD prediction system

The practical utility of these findings is embodied in the PAD prediction system, depicted in Figure 4. Built using Flask, the system processes real-time physiological inputs (LDF, PPG, ECG) through a Python backend and employs the ensemble model to classify PAD severity into Non-PAD, moderate PAD, or severe PAD. The web interface, designed for minimal user interaction complexity, ensures efficient data entry and delivers clear text outputs (*e.g.*, "severe PAD"). This architecture supports rapid deployment in

clinical and field settings, enabling timely decision support for healthcare providers. The system's accessibility and ease of use address a critical gap in PAD diagnostics, where early detection is essential to prevent complications such as limb ischemia or amputation [22].

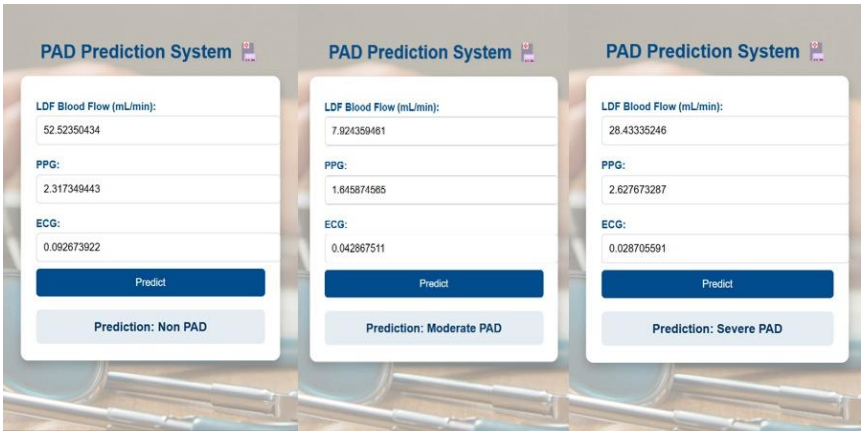

Figure 4. Peripheral artery disease prediction webpage

5.2. Comparative analysis of PAD detection models

Table 2 illustrates a comparative analysis of various approaches for PAD detection, highlighting differences in data sources, algorithms, and classification accuracies. Earlier models utilized convolutional neural networks (CNNs) on arterial pulse waveforms, structured clinical data with XGBoost, and biomarker-based Cox proportional-hazards models, achieving accuracies ranging from 72.0% to 90.0%. In contrast, the proposed model integrates multimodal physiological signals-LDF, Electrocardiography, and Photoplethysmography with Ensemble strategies, including random forest and XGBoost. With an accuracy of 93%, demonstrating its robustness and effectiveness in PAD severity classification.

Table 2. Comparative performance of PAD detection models

Author	Dataset	Methodology	Accuracy
Opitz and Maclin [11]	Arterial pulse waveforms	Deep learning (CNNs)	90.0%
Li *et al.*[23]	Vascular quality initiative (VQI) registry	Machine learning (extreme gradient boosting – XGBoost)	86.0%
Amrock and Weitzman [24]	CRP, NLR, homocysteine, UACR	Cox proportional-hazards modeling and multimarker score analysis	85.0%
Moussa *et al.* [25]	ABI	Statistical analysis using logistic regression	72.0%
Proposed model	LDF, ECG, PPG	Random forest, XG Boost, ensemble model	93.0%

6. CONCLUSION

This study developed a machine learning model for detecting PAD by leveraging physiological signals (LDF, PPG, and ECG) to classify patients into Non-PAD, Moderate PAD, and Severe PAD categories. Utilizing pre-processing techniques such as polynomial feature transformation, standardization, and PCA-based dimensionality reduction, the dataset was optimized for an ensemble model combining random forest and XGBoost, with hyperparameters fine-tuned through GridSearchCV. Deployed within a Flask web application, the model achieved a 93% classification accuracy, demonstrating its potential for real-time clinical diagnostics and remote health monitoring. These results highlight the model's capability to accurately distinguish PAD severity levels, enabling early diagnosis and facilitating personalized patient care. However, the model faced challenges in generalizing across diverse patient populations due to variations in physiological signal patterns. Its performance may be limited when applied to datasets with atypical signal characteristics or smaller sample sizes Future work will focus on addressing these limitations by incorporating larger, more diverse datasets and exploring advanced feature engineering techniques to enhance model robustness. Additionally, expanding the model to accommodate a broader range of physiological variations. These enhancements aim to strengthen the model's applicability in diverse healthcare environments, support early intervention strategies, and inform data-driven health policy decisions.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mummaneni Sobhana		✓				✓	✓		✓	✓		✓	✓	
Lalitha Devi Katakam	✓			✓				✓						
Ramya Sri Pali			✓						✓	✓				
Mounika Lingamallu	✓									✓				
Smitha Chowdary Ch		✓		✓	✓	✓		✓			✓			
D.N.V.S.L.S. Indira			✓				✓				✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest

DATA AVAILABILITY

The ECG and PPG datasets used in this study are publicly available on Kaggle, while the LDF data were sourced from a published medical study. All datasets were integrated using time-synchronization techniques and are accessible under their respective data use terms.

REFERENCES

[1] E. H. Weissler *et al.*, "Model-based algorithms for detecting peripheral artery disease using administrative data from an electronic health record data system: algorithm development study," *JMIR Medical Informatics*, vol. 8, no. 8, 2020, doi: 10.2196/18542.

[2] J. S. Hiramoto, M. Teraa, G. J. de Borst, and M. S. Conte, "Interventions for lower extremity peripheral artery disease," *Nature Reviews Cardiology*, vol. 15, no. 6, pp. 332–350, Jun. 2018, doi: 10.1038/s41569-018-0005-0.

[3] A. N. Lasota, K. Overvad, H. H. Eriksen, A. Tjønneland, E. B. Schmidt, and M.-L. M. Grønholdt, "Validity of peripheral arterial disease diagnoses in the danish national patient registry," *European Journal of Vascular and Endovascular Surgery*, vol. 53, no. 5, pp. 679–685, May 2017, doi: 10.1016/j.ejvs.2016.12.031.

[4] Y. Zhang, J. Huang, and P. Wang, "A prediction model for the peripheral arterial disease using NHANES data," *Medicine*, vol. 95, no. 16, p. e3454, Apr. 2016, doi: 10.1097/MD.0000000000003454.

[5] C.-H. Lin *et al.*, "A multitask deep learning model utilizing electrocardiograms for major cardiovascular adverse events prediction," *npj Digital Medicine*, vol. 8, no. 1, pp. 1–10, Jan. 2025, doi: 10.1038/s41746-024-01410-3.

[6] R. Ferizoli, P. Karimpour, J. M. May, and P. A. Kyriacou, "Arterial stiffness assessment using PPG feature extraction and significance testing in an in vitro cardiovascular system," *Scientific Reports*, vol. 14, no. 1, pp. 1–10, Jan. 2024, doi: 10.1038/s41598-024-51395-y.

[7] M. Elgendi *et al.*, "The use of photoplethysmography for assessing hypertension," *npj Digital Medicine*, vol. 2, no. 1, pp. 1–11, Jun. 2019, doi: 10.1038/s41746-019-0136-7.

[8] T. Ishii *et al.*, "Laser doppler blood flowmeter as a useful instrument for the early detection of lower extremity peripheral arterial disease in hemodialysis patients: an observational study," *BMC Nephrology*, vol. 20, no. 1, pp. 1–11, Dec. 2019, doi: 10.1186/s12882-019-1653-y.

[9] A. M. Flores, F. Demsas, N. J. Leeper, and E. G. Ross, "Leveraging machine learning and artificial intelligence to improve peripheral artery disease detection, treatment, and outcomes," *Circulation Research*, vol. 128, no. 12, pp. 1833–1850, Jun. 2021, doi: 10.1161/CIRCRESAHA.121.318224.

[10] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 1–16, Apr. 2016, doi: 10.1098/rsta.2015.0202.

[11] D. Opitz and R. Maclin, "Popular ensemble methods: an empirical study," *Journal of Artificial Intelligence Research*, vol. 11, pp. 169–198, Aug. 1999, doi: 10.1613/jair.614.

[12] J. Allen, H. Liu, S. Iqbal, D. Zheng, and G. Stansby, "Deep learning-based photoplethysmography classification for peripheral arterial disease detection: a proof-of-concept study," *Physiological Measurement*, vol. 42, no. 5, pp. 1–8, May 2021, doi: 10.1088/1361-6579/abf9f3.

[13] S. Kim, J.-O. Hahn, and B. D. Youn, "Detection and severity assessment of peripheral occlusive artery disease via deep learning analysis of arterial pulse waveforms: proof-of-concept and potential challenges," *Frontiers in Bioengineering and Biotechnology*,

vol. 8, pp. 1–11, Jun. 2020, doi: 10.3389/fbioe.2020.00720.
[14] R. D. McBane *et al.*, "Artificial intelligence for the evaluation of peripheral artery disease using arterial Doppler waveforms to predict abnormal ankle-brachial index," *Vascular Medicine*, vol. 27, no. 4, pp. 333–342, Aug. 2022, doi: 10.1177/1358863X221094082.
[15] G. Stansby *et al.*, "Prospective assessment of the diagnostic accuracy of multi-site photoplethysmography pulse measurements for diagnosis of peripheral artery disease in primary care," *Angiology*, vol. 74, no. 9, pp. 859–867, Oct. 2023, doi: 10.1177/00033197221121614.
[16] N. Forghani, K. Maghooli, N. J. Dabanloo, A. V. Farahani, and M. Forouzanfar, "DeepPAD: detection of peripheral arterial disease using deep learning," *IEEE Sensors Journal*, vol. 22, no. 16, pp. 16254–16262, Aug. 2022, doi: 10.1109/JSEN.2022.3188810.
[17] M. Sonderman *et al.*, "Identifying patients with peripheral artery disease using the electronic health record," *JACC: Advances*, vol. 2, no. 7, pp. 1–8, Sep. 2023, doi: 10.1016/j.jacadv.2023.100566.
[18] A. Achraf, "Signal processing: ECG-PPG beat detectors," *Kaggle*. 2022, Accessed: Oct. 08, 2022. [Online]. Available: https://www.kaggle.com/datasets/elbaronahmedashraf/signal-processing-ecg-ppg-beat-detectors.
[19] K. F. Ma *et al.*, "Laser doppler flowmetry combined with spectroscopy to determine peripheral tissue perfusion and oxygen saturation: a pilot study in healthy volunteers and patients with peripheral arterial disease," *Journal of Personalized Medicine*, vol. 12, no. 6, p. 853, May 2022, doi: 10.3390/jpm12060853.
[20] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–23, 2001, doi: 10.1023/A:1010950718922.
[21] T. Chen and C. Guestrin, "XGBoost: a scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[22] M. H. Criqui and V. Aboyans, "Epidemiology of peripheral artery disease," *Circulation Research*, vol. 116, no. 9, pp. 1509–1526, Apr. 2015, doi: 10.1161/CIRCRESAHA.116.303849.
[23] B. Li *et al.*, "Machine learning to predict outcomes of endovascular intervention for patients with PAD," *JAMA Network Open*, vol. 7, no. 3, p. e242350, Mar. 2024, doi: 10.1001/jamanetworkopen.2024.2350.
[24] S. M. Amrock and M. Weitzman, "Multiple biomarkers for mortality prediction in peripheral arterial disease," *Vascular Medicine*, vol. 21, no. 2, pp. 105–112, Apr. 2016, doi: 10.1177/1358863X15621797.
[25] I. D. Moussa *et al.*, "Prevalence and prediction of previously unrecognized peripheral arterial disease in patients with coronary artery disease: the peripheral arterial disease in interventional patients study," *Catheterization and Cardiovascular Interventions*, vol. 73, no. 6, pp. 719–724, May 2009, doi: 10.1002/ccd.21969.

BIOGRAPHIES OF AUTHORS

Sobhana Mummaneni is currently working as an associate professor in the Department of Computer Science and Engineering, V. R. Siddhartha Engineering College, Vijayawada, India. She received Ph.D. degree in computer science and engineering in 2018 from Krishna University. She has 16 years of teaching experience. Her research interests lie in areas such as artificial intelligence, machine learning, data analytics, cyber security, and software engineering. She published 35 papers in National and International journals and published 7 patents. She can be contacted at email: sobhana@vrsiddhartha.ac.in.

Lalitha Devi Katakam is a third-year B.Tech. student specializing in CSE at V. R. Siddhartha Engineering College, Vijayawada, India. She is passionate about biomedical signal processing, machine learning, and healthcare-related intelligent systems. She can be contacted at email: lalithakatakam2112@gmail.com.

Pali Ramya Sri is a third-year B.Tech. student specializing in CSE at V. R. Siddhartha Engineering College, Vijayawada, India. Her interests lie in the fields of biomedical signal processing, healthcare-related intelligent systems machine learning, and web development. She can be contacted at email: ramyasripali@gmail.com.

Mounika Lingamallu is a third-year B.Tech. student specializing in CSE at V.R. Siddhartha Engineering College, Vijayawada, India. Her academic interests focus on biomedical signal processing, machine learning, and web development. She can be contacted at email: lingamallumounika@gmail.com.

Smitha Chowdary Ch has 18 years of experience in teaching field with more than 35 publications and currently working as professor in KL Deemed to be University. Her research interests include software engineering, AI and machine learning, pattern recognition, block chain and deep learning-based applications, time-series data analysis and prediction, image processing, image applications and recommendation systems She can be contacted at email: smithacsc@gmail.com.

D.N.V.S.L.S. Indira is currently working as professor and head in the Department of Information Technology, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, AP, India. She completed B.Tech. (CSE) in 2006, and M.Tech. (CSE) in 2009 from JNTUK, Kakinada. Received Ph.D. from Krishna University, Machilipatnam in 2018. Her research interest covers sentiment analysis, data analytics, cyber security, artificial NN, WSN and image/vudio processing. She published 47 papers in national/ international journals and presented 17 papers in international conferences. She received an amount of one lakh from AICTE for SPICES program. She organized 2 E&ICT FDPs in collaboration with NIT Warangal. She has 19 years of teaching experience. She can be contacted at email: indiragamini@gmail.com.