Vol. 07, No. 01, pp. 124 –136 (2026)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Deep Learning Based Early Detection of Atherosclerosis for Stroke Prevention using Multi-Sensor Data Integration

Akshaya D. Shetty[1,2,3*], Usha Desai[1,3], Akash Saxena[4]

[1]Department of Electronics and Communication Engineering, S.E.A College of Engineering & Technology, Bengaluru 560049, Karnataka, India, akshaya926@gmail.com, dr.ushadesai@seaedu.ac.in
[2]Department of Artificial Intelligence and Machine Learning, Vivekananda College of Engineering & Technology, Puttur 574203, Karnataka, India
[3]Visvesvaraya Technological University, Belagavi 590018, Karnataka, India
[4]Department of Electrical Engineering, School of Engineering and Technology, Central University of Haryana 123031, Haryana, India, drakash@cuh.ac.in

*Correspondence: akshaya926@gmail.com

Abstract

Atherosclerosis is a progressive cardiovascular condition where arteries narrow due to plaque buildup, significantly increasing the risk of heart attacks and strokes. This study presents a non-invasive, Deep Learning-architecture based system for the timely diagnosis of atherosclerosis by means of real-time physiological and clinical data. This system integrates wearable sensors namely, Electrocardiogram (ECG), Photoplethysmography (PPG), Galvanic Skin Response (GSR), and Blood Pressure (BP) to continuously monitor heart-related parameters. Also, incorporates the clinical indicators namely blood glucose and cholesterol levels to enhance predictive accuracy. Data set comprises 226 records of the subjects having signs of atherosclerosis and 180 records of healthy subjects. The feature extraction involves total 8 original and 5 engineered features. Collected data undergoes preprocessing and is analyzed using various Deep Learning architectures including LSTM, BiLSTM, GRU, CNN-LSTM, and Transformer networks. These models are trained and evaluated using stratified K-fold cross-validation, ensuring consistent and generalized performance. The assessment metrics involves accuracy, precision, recall and F1 score. Among these, CNN-LSTM and Transformer models achieved superior accuracy and robustness in classifying individuals as healthy or at risk of atherosclerosis. The best model is chosen as CNN-LSTM with highest weighted score of 0.98 in comparison with other individual models. The final model is deployed in a user-friendly Streamlit interface, which helps users to input physiological data and receive real-time health predictions. The system provides a diagnostic output, confidence score, and highlights of any abnormal parameters. This solution addresses limitations of traditional diagnostics such as high cost, invasiveness, and lack of real-time feedback by offering a portable, affordable, and continuous monitoring tool. It empowers users, especially in remote or underserved areas, to take proactive measures for stroke prevention and cardiovascular health management.

Keywords: *Atherosclerosis, Deep Learning, CNN-LSTM, Transformer, Wearable Sensors.*

Received: October 16th, 2025 / Revised February 12th, 2026 / Accepted: February 24th, 2026 / Online: February 26th, 2026

I. INTRODUCTION

Heart diseases or cardiovascular disorders are becoming common now a days, which may be deadly [1]. To avoid the sudden heart issues it is primarily important to track the real time signals from the human body. The atherosclerosis is also a significant factor among the cardiovascular diseases. Atherosclerosis is a disease that advances over time, causing arteries to narrow and harden from plaque accumulation, which may result in heart attacks, strokes, or other significant health issues. This disease becomes unnoticed over time; it is dangerous due to its asymptomatic nature and its sudden evident is either heart attack or stroke. To prevent the life threatening threats early detection of atherosclerosis is very important. However the conventional diagnostic techniques are very expensive, and are not always accessible for real time regular monitoring. This work bridges the gap through the development of a non-invasive, sensor based system that helps in early detection of atherosclerosis using physiological real time data.

We have integrated several sensors- including Electrocardiogram (ECG), Photoplethysmography (PPG), Galvanic Skin Response (GSR), and Blood Pressure (BP) sensors - to collect various bio-signals and health parameters. In addition, clinical indicators like glucose and cholesterol levels from blood test reports are incorporated to enhance prediction accuracy. The data is then processed and analyzed using Deep Learning techniques to identify the potential cases of atherosclerosis.

doi: 10.38094/jastt71583

The system is made to be affordably priced, versatile and suitable for continuous monitoring, especially in remote or resource-constrained environments. By combining sensor technology with intelligent algorithms, this work takes a step towards smarter, more accessible healthcare, where early signs of vascular disease can be identified significantly.

In the current healthcare landscape, the detection of atherosclerosis heavily relies on clinical examinations and diagnostic imaging techniques. Methods like Angiography, Doppler ultrasound, CT (Computed Tomography) scans, and MRI (Magnetic Resonance Imaging) are widely utilized to assess arterial health and detect blockages or plaque buildup. These techniques provide detailed and accurate insights into the cardiovascular system, helping doctors identify the presence and severity of atherosclerosis. Additionally, routine blood tests are performed to keep close watch on risk factors like cholesterol levels, blood sugar, and inflammatory indicators such as CRP (C-reactive protein). These indicators help in estimating a person's risk of developing atherosclerosis, even before symptoms appear. Even though the existing methods are clinically reliable, they come with a few major challenges such as high cost and limited accessibility, invasiveness and it would be too late for preventive action.

However, these limitations, signifies the growing needs for development of easily available, affordable, and non-invasive methods of detecting atherosclerosis, mainly systems which can provide real-time information about person's cardiovascular well-being. This leads to exploration of wearable sensors, AI-driven analysis, and home-based monitoring systems that can detect subtle physiological changes and alert users before a serious issue develops. A ring sensor that is flexible serves as the biomarkers for atherosclerosis. This device integrates two techniques namely, micro fluid technology and the electrochemical bio sensing. A biosensor named FredJewel is introduced for measuring C-reactive protein, cholesterol, potassium, pH level in sweat, which was validated in early diagnosis for ten samples. Absolutely, there is a need to design a system involving wearable sensors like ECG monitors, PPG device, glucose monitors for the real time monitoring of physiological signals. The wearable sensor that is being integrated onto the monitoring systems helps in identifying the early cardiac risks. Complete review of signal processing systems indicted in the monitoring of cardiovascular disease is explored [1, 2].

In a clinic the atherosclerosis can be accessed by blood lesion level, blood flow and blood viscosity. An electrochemical sensor is developed by integrating the multiple combinations like cholesterol, potassium, and transferrin, which are biomarkers for the atherosclerosis. Another sensor is PPG that tracks daily activities like sleep and exercise and measures heart rate and rhythm. Moreover, PPG has the high capability for the health and wellbeing application of individual [3, 4]. In the prediction of the cardiovascular disease, use of the smart wearable for the diagnosis and treatment of fatal diseases has gained the increased attention due to medical significance.

The wearable devices application in the cardiovascular healthcare includes arrhythmias early detection, measurement of blood pressure, and detection of diabetes. Cardiovascular events can be promptly and early detected with wearable sensors. The movement of the body, skin contact, noise, and motion artifacts throughout daily life can contaminate wearable technology data.

The algorithms used to process the raw data have a significant impact on accuracy and dependability. Multiple machine learning and deep learning models such as Convolutional Neural Networks (CNN), Support Vector Machine (SVM), Long Short Term Network (LSTM), and Decision Trees (DT) are used for cardiovascular disease prediction using wearable sensors [5, 6]. To identify the subclinical atherosclerosis, several ECG abnormalities are combined into ECG integrated risk score. Conversely, in a young adult apparently the Coronary Artery Calcium (CAC) is measured, where the individuals considered for measurement were from Korea. Authors measured five abnormalities such as heart rate, QRS duration, T-wave inversion, prolonged QTc and left ventricular hypertrophy. All these are summed into an ECG risk score. Via the CT scan the CAC is computed. The ECG risk score higher which is correlated with greater prevalence for the value of CAC. The CAC is the direct marker of the plaque buildup so finding of ECG abnormalities directly reflect the early atherosclerosis [7]. Also, other major risk factor is cholesterol.

A new technique is identified to measure the cholesterol from skin, where based on fluorescent chemical staining one portable skin cholesterol detector is designed, which works on absorption technique spectrography. Increase of cholesterol in the skin reflects the high blood cholesterol level and early signs of the atherosclerosis. Through the detection reagent the light passes it reaches the epidermis of the skin and then it reflected back to the detector [8]. In the clinical practice using PPG signal the arterial blood volume information's are precisely captured. The PPG systems are included with wearable components like sensor design, acquisition of signals, and methods for signal processing. To achieve sustainable accuracy, reliability and robustness of the measurement of cardiovascular related wearable PPG sensors plays an important role in the daily life. To the smart phone a three channel PPG device is connected and it can measure the pulse transit and pulse wave velocity using the probes that is clipped into the big toe, index finger and ear. Using 100 participants of aged between 20 to 77 years the validation is done at the Nagpur hospital in India. The subjects are divided into three groups that are healthy, coronary arterial disease and hypertensive [9, 10], using an ECG 12 leads the Global Electrical Heterogeneity (GEH) derived, correlated with the functional and structural abnormalities of heart.

This research highlighted the potential of the ECG derived markers in the cardiovascular health. There is a high challenge in treating the subjects in hospital emergency rooms and critical care unit with Acute Hypotensive Episode (AHE) [11]. Furthermore, a study involving GSR or Electro Dermal Activity (EDA), which is a measure of activation of the nervous system, this helps to identifying the danger of AHE. The low cost wearable device is designed and tested to predict the blood pressure and circulatory dynamics. Relation between the two parameters viz. GSR values and four indexes related to

BP like pulse pressure, arterial pressure, diastolic BP, systolic BP is observed. The medical decision support system is conceptualized in detection of Atherosclerosis. The Cleveland heart disease dataset is taken as an input to the 4 models such as Artificial Neural Network (ANN), K Nearest Neighbor (KNN), K-means, and K-medoids algorithms. The dataset included 270 records with 13 clinical features per patient. ANN, with Levenberg–Marquardt back propagation, achieved the highest accuracy (96%) and Matthews Correlation Coefficient (MCC) of 0.93, outperforming other models.

Studies reports stress leads to gradual increase of the cardiac problems. Using the dataset from the 16 adults, a bio sensor to acquire the GSR on Zigbee communication standard is designed, which achieves the success rate of 76.5%, and the GSR, is able to detect each user's diverse states. The machine learning Bayesian networks achieved a maximum accuracy of the value 90.97%, for the atherosclerosis detection, [12, 13]. Atherosclerosis leads to cardiovascular events like stroke and heart attacks primarily due to thrombosis following the plaque rupture. Due to the factors like rise in blood pressure and Low Density Lipoprotein (LDL) cholesterol the blood vessels inner layer becomes dysfunctional. Due to this the lipoproteins enters the arterial wall. The plaque rupture is the primary reason of the heart related events [14].

According to World Health Organization (WHO), the cardiovascular diseases are the main factor of death. The cardiovascular risk factor such as BP is predicted using the advanced machine learning methods like ANN, SVM, DT, Naïve Bayes (NB) and KNN. A medical dataset comprising of 558 patients with atherosclerosis is deliberated as input, where subjecting ANN accuracy obtained was 96.67% [15]. In order to detect the cardiovascular diseases the conventional machine learning algorithms such as KNN, SVM and Random Forests applying ECG, PPG and Phonocardiograms (PCG) signals are applied. Outcome of these algorithms are compared with deep learning architectures such as CNN, LSTM and Transfer Learning models. The significance of denoising, data compression, and digitization techniques are clearly illustrated in order to improve the quality of the acquired signals. The summary of the multiple datasets for ECG, PPG and PCG are tabulated with the description of the dataset and the applications [16].

A new dataset called PPG-DaLiA is presented and published for the heart rate estimation based on PPG signal. In which, the sensor data information using the device worn on the chest and wrist is recorded with sampling rate of 700 Hz and data was collected from 15 subjects, age group of 21-55 years, where eight were female and seven male participants. The performance of CNN was compared with several other datasets such as IEEE_training, PPG_motion, PPG_Bruse and WESAD [17]. Deep learning approach for discriminating the classes of data is challenging. In addition, two-dimensional feature analysis is medically significant and the bio signals extracted are of non-stationary and low amplitude which needs nonlinear conversion to extract the important transform domain coefficients. In literature over a decade for ECG, EMG, EEG and other physiological signal analysis several nonlinear methods are applied with machine learning and neural network classifiers [18]. The qualities and quantitative analysis of deep learning models with key parameter extraction on the publicly dataset such as NinaProDB database, BioPatRec sub-database and MYO is conducted [19].

One of the major parameter for the evaluation of the circulatory systems is blood pressure. Variation in the blood pressure will results in cardiovascular diseases. So there is a requirement of evaluation of hypertension risk. Deep learning method is used for evaluation of the individuals' hypertension using PPG signals. Wavelet transform and pre trained CNN network are the main foundation of this work. Total 121 data is collected from an individual using the intensive care database. Each signal in the database consists of PPG and arterial blood pressure signals. The proposed deep learning models required the high quality PPG signal [20]. Using two bio signals such as ECG and PPG the blood pressure are being predicted by the ResNet LSTM model. Hypertension affects the millions of individuals in developing the cardiovascular devices.

The MIMIC (Medical Information Mart for Intensive Care) dataset is the primary source of this research work. This model performance was also validated with the Pulse Transit Time PPG dataset. The deep learning used cardiovascular risk score based on PPG is designed. Combined with the basic demographics the PPG signals are obtained using the smart devices. Considering the data obtained using UK Biobank specifically participant data and PPG waveforms the model is trained and validated [21, 22]. The main factors associated with risk in cardiovascular events are almost identical in both men and female gender. Some of the independent cardiovascular disease risk factors are exclusively evident only in the females due to the disorders in pregnancy such as hypertension, diabetes; endocrine disorders and early menopause accelerated the risk of cardiovascular diseases.

In the cardiovascular disease main risk factor prevalence and its consequence in both men and women is being studied and examined. Studies suggest more consideration in terms of awareness and lifestyle changes. There is a more need of appropriate attention in both men and women in cardiovascular care [23]. The glucose level monitoring is done among the Korean individuals with fasting glucose levels. The relation between the glucose levels with fasting followed the j-shaped curves. The glucose level lower risks range is of 85–99 mg/dL. After fasting the glucose level increased more than 100 mg/dL.

The main predictors of the risk for coronary heart disease and stroke are the variations in glucose level. The research provided the guidance in terms of predicting the baseline of glucose levels [24]. One more major parameter in coronary artery disease prediction is the cholesterol lipoproteins. The optimal range which includes: LDL cholesterol: less than 100 mg/dL, Non- High Density Lipoprotein (HDL) cholesterol: less than 130 mg/dL, HDL cholesterol: greater than 40 mg/dL in men and greater than 50 mg/dL in women, Triglycerides: less than 150 mg/dL, Lipoprotein (a): less than 50 mg/dL and total cholesterol desirable below 200 mg/dL. The lipid parameter range is the important factor in predicting atherosclerosis risk. Elevations in total cholesterol, LDL, low HDL, high triglycerides, and high lipoprotein (a) levels are linked with the buildup of atherosclerosis plaques and it contributes to cardiovascular risk [25]. Managing and monitoring is very essential and is of high concern. There exists relationship between the heart rate variability and cardiovascular disease risk in lifetime. The heart rate variability is evaluated using the metrics such as standard deviation of RR interval, root mean square of successive differences, mean of all normal RR interval, Low Frequency (LF) power, High Frequency (HF) power, and LF/HF ratio. The lower HRV is linked with the higher life time risk of cardiovascular disease [26].

This paper mainly focuses on improving cardiovascular health monitoring and providing a system for early detection of atherosclerosis, aiming for a practical, user friendly and cost-effective solution. The multiple factors necessary for the cardiovascular healthcare is studied from the literature review study. Further, there is crucial requirement of monitoring the vital body parameters. However, this work contributes in development of multiple sensor integration with the user interface to access the health state of an individual. Inspired from the literature review in this study following research objectives are framed:

i. To provide a real-time cardiovascular health monitoring framework, with low cost and less complexity, ensuring that any changes in heart health are detected immediately.

ii. To integrate multiple health parameters, including the sensor data (ECG, PPG, GSR, and BP) along with clinical values (cholesterol, glucose) to offer a comprehensive evaluation of a person's cardiovascular health.

iii. To demonstrate the cardiovascular health related data on deep learning algorithms to achieve sustainable accuracy and reliability in detecting early signs of atherosclerosis, reducing false positives and negatives.

This paper is further organized as follows: Section II explains methodology of proposed study and Sections III and IV explores results and conclusion, respectively.

II. Methodology

The fundamental objective of this paper is to facilitate early atherosclerosis detection and risk stratification. This is initiated by determining the target population with suspected cardiovascular risk factors that include diabetes, high blood pressure, high cholesterol, or a family history of cardiovascular disease. The second step is to define technical specifications. This entails defining the system's main functionalities, including the recording of biological signals such as ECG, PPG, GSR, and BP information.

A. Hardware Setup

The suggested system in Fig. 1, offers a unique approach of early atherosclerosis detection by combining biomedical sensor technology with advanced data processing and Deep Learning techniques. The system consists of 4 biosensors such as PPG, ECG, GSR and BP for the real time physiological signal extraction. Proposed work is extended by clinical values such as glucose and cholesterol levels. The hardware setup consists of controller Arduino based on ATmega328P chip. The Arduino Integrated Development Environment (IDE) by Arduino.cc is be used to write program for the controller. The code is written in the C language. The ECG sensors play a major role in monitoring heart activity. This sensor detects the electrical signal generated by heart beats. Each time the heart contracts to pump blood, a wave of electrical impulses travels through the heart muscles. ECG sensors carefully placed on the skin, usually around the chest, arms, and legs, can detect these subtle signals through the body's tissues.

These signal pattern, timing, and strength provide crucial insights into heart rhythm, rate, and overall cardiac health, allowing doctors to identify conditions such as arrhythmias, heart attacks, and other abnormalities. However, the body temperature regulator, sweat plays a very important role. When the stress, excitement, or anxiety is experienced the sweat glands becomes more active. This leads to increase in moisture and also the conductivity in the skin surface that is detected by the electro dermal sensor or GSR [27, 28]. The electrical conductivity of the skin is measured by GSR. When the person is highly stressed reflects the change in sympathetic nervous system, the change in stress causes changes in skin conductivity, which are perceived in the biosensor as electrical signals.

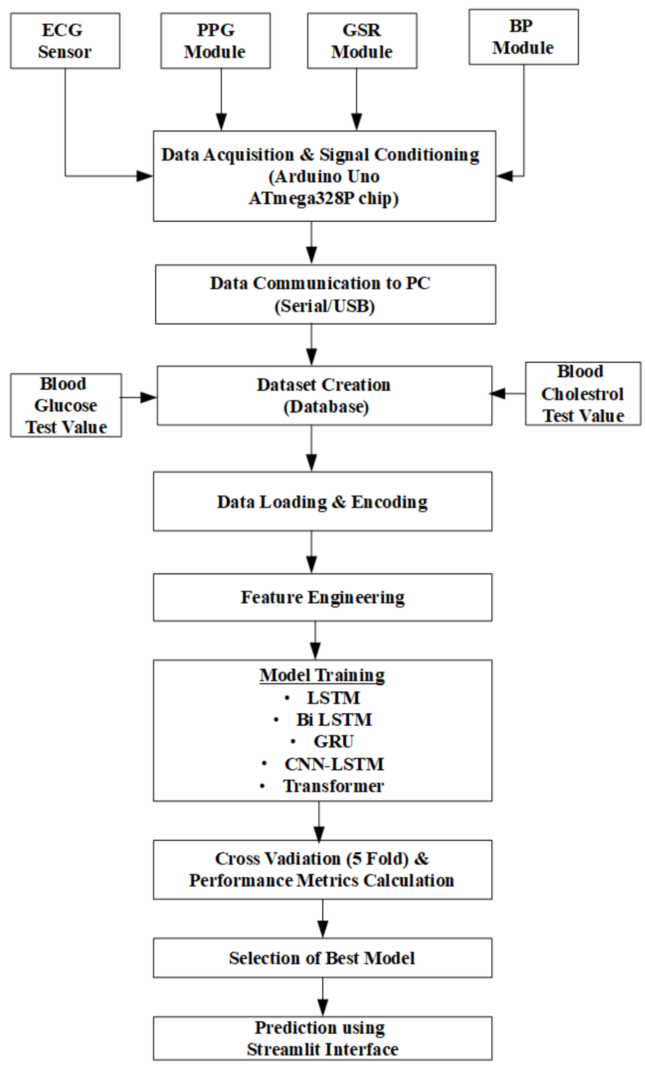

Fig. 1. Block diagram for the atherosclerosis detection, using biosensor signals and deep learning model.

GSR alone cannot be used as a biomarker for the cardiovascular events; moreover, this sensor is used in this research to access the state of the individual relaxed or non-relaxed. The PPG sensor measures variations in blood volume to monitor the physiological rhythm. Ratherthan relying on external signs of health, the PPG sensor uses light to reveal the inner body mechanisms, which shines a gentle beam, usually red or infrared, onto the skin, often at the fingertip or wrist. As blood pulses with each heartbeat, it absorbs and reflects light in tiny but detectable ways. The PPG sensor reads these variations, like finely tuned radar, to measure heart rate and blood flow. These signals offer insight not just into physical well-being but also into emotional and mental states, as heart rate can rise with stress, fear, or excitement. By translating light reflections into meaningful data, PPG sensors provide a non-invasive real-time window into cardiovascular health. They are especially valuable in wearable devices, enabling continuous monitoring of a person's heart activity with remarkable simplicity and precision.

The BP sensor acts as a vital window into cardiovascular health, allowing non-invasive monitoring of the force generated by circulating blood on the walls of arteries. Unlike traditional manual cuffs, modern BP sensors are compact, automated, and ideal for continuous or periodic monitoring. These sensors typically operate using the oscillometric method, where small pressure variations are detected as the cuff inflates and deflates. By analyzing these oscillations, the sensor determines both systolic and diastolic pressure values, along with pulse rate, which helps to detect early signs of hypertension, stress related changes, or irregular heart rhythms. Integrated with microcontroller, BP sensors can store, process, and transmit real-time data for medical analysis.

For analyzing the collected sensor data and early detection of atherosclerosis, deep learning is employed. Deep Learning models, particularly CNNs and Recurrent Neural Networks (RNNs), analyze complex patterns in the sensor data. These models can identify subtle trends and correlations within the physiological signals that may indicate the early onset of atherosclerosis. The atherosclerosis detection system's user-friendly interface is created using Streamlit.

Users can enter newly acquired clinical and physiological data through the interface, including blood pressure, cholesterol, and heart rate. Based on this input, the model provides a prediction along with the corresponding confidence level, assisting users in determining the likelihood of atherosclerosis. This simple, interactive setup ensures efficient and accessible health assessment in real-time.

Fig. 2 represents integration of sensor module to Arduino microcontroller for real-time acquisition. In which PPG, ECG, and GSR using MAX30102, AD8232, and SEN0208 sensor modules, and BP monitoring device are integrated on the Arduino Uno ATmega328P chip for acquiring the required biomedical signals and analysis.

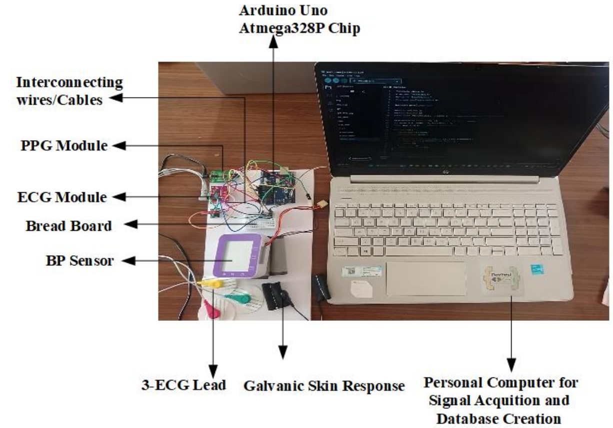

Fig. 2. Biosensors signal acquisition and processing framework.

B. Data Acquisition, Preprocessing and Feature Selection

The arrangement is carefully designed, with the Arduino Uno serving as the crucial interface between the sensor modules and the serially obtained data. Firstly the PPG sensor module is placed. This module is designed to measure blood volume changes and monitor the heart rate, essential for assessing cardiovascular health and detecting potential issues like atherosclerosis.

The ECG sensor module is equipped with a plug-and-play electrode circuit for recording the electrical response of the heart, which aids in identifying irregular heartbeats and contributing valuable insights into cardiovascular conditions.

The digital BP sensor module positioned is responsible for measuring the blood pressure, a vital sign for determining the likelihood of atherosclerosis and other related diseases. Additionally, the GSR sensor module is placed to measure skin conductance based on sweat gland activity, providing insights into autonomic nervous system responses that can complement the health data collected from other sensors.

The data acquisition is done using the ECG, GSR, PPG and BP sensor modules interfaced with Arduino Uno and the preprocessing is done following the biomedical signal processing standards. ECG signal at a sampling rate of 250 Hz is band pass filtered between 0.5-40 Hz, PPG signal is sampled at 100 Hz and filtered between 0.5-8 Hz, GSR signal sampling frequency is 10 Hz and is low pass filtered below 1 Hz. Beat wise basis the BP measurements are obtained and mean average filter is used for smoothing. The recording of the signals using the bio sensors was held for 5-10 minutes, the abnormal amplitude variations due to motion artifacts are removed and linear interpolation is performed for reconstructing the short missing segments.

The outlined steps provide a structured approach for collecting real-time biological data through four sensors with an Arduino microcontroller. This integrated code facilitates synchronized data acquisition from the PPG, ECG, GSR, and BP sensors. Each sensor is assigned to specific analog or digital pins, and data is transmitted through serial communication. The loop section continuously reads and averages the sensor values, and displays them on the serial monitor. This integration is crucial for developing AI-based diagnostic systems that analyze multi-sensor biomedical signals.

Using the set up represented in Fig. 2 the data acquisition is done. To create a model for detecting atherosclerosis, the system is trained using two sets of numeric data stored in CSV files. The data includes key medical parameters relevant to arterial health such as glucose, cholesterol, heart rate, SpO_2 and other circulatory system indicators. These measurements serve as the key input features to the model. The dataset is collected from the subjects at the Pragathi Multi Specialty Hospital, Puttur, Dakshina Kannada, Karnataka in India, consisting of 226 records of individuals (121 males and 105 females) of age 20 to 54 years diagnosed with signs of atherosclerosis. The data is collected over the duration of three months. The healthy dataset includes 180 records (110 males, and 70 females) from the individuals of age varying from 20 to 35 years showing no signs of arterial disease. The ethical approval is obtained from the medical hospital for collection of dataset, and the medical experts have validated the dataset. The aim of the model is to accurately classify each record into one of the two categories that is healthy (no signs of atherosclerosis) and unhealthy (presence of atherosclerosis conditions).

The healthy and unhealthy data saved as CSV files for which the label encoding is performed. The healthy is encoded to be 0 (zero) and unhealthy as 1 (one). Missing values are handled in the dataset. This step is followed by extraction of significant features using deep learning models LSTM, BiLSTM, GRU, CNN-LSTM and Transformer, respectively. In this study, in total thirteen features are collected using individual model, where eight are original and five engineered features. The original features are systolic BP, diastolic BP, LDL, HDL, total cholesterol, heart rate, SpO_2 and gender encoded by a numeric value (1=male, and 0=female). The engineered features which includes pulse pressure; mean arterial pressure, ratio of bad to good cholesterol, cholesterol ratio (cholesterol/HDL), and heart rate SpO_2 ratio. However, the 8 features are basically clinical measurements and 5 features mainly contribute the respiratory and cardiovascular health. Further, the input features are separated from target labels, where all the features are scaled to 0-1 range using min max normalization and are reshaped for the deep learning models.

The code is implemented in Python 3.9 using Jupiter notebook. Deep learning is implemented using the TensorFlow 2.12 library is utilized, where Scikit-learn for preprocessing and evaluation, and data handling using numpy and pandas. Also, for visualization the matplotlib and seaborn libraries are utilized.

C. Deep Learning Models

Relevant features are selected from the dataset to enhance the model performance. The feature which includes pulse pressure, mean arterial pressure, cardiovascular risk marker, cholesterol ratio, heart rate relative to oxygen saturation, respctively. However, in literature atherosclerosis detection has been studied by employing the deep learning algorithms [17, 19] such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Transformer, Convolutional Neural Network - Long Short-Term Memory (CNN-LSTM), and Bidirectional Long Short Term Memory (BiLSTM) for classification of bio sensor signals, and the clinical data. These models are designed efficiently to capture the temporal dependencies in physiological signals such as ECG, PPG, GSR, blood pressure, glucose, and cholesterol levels. By leveraging these recurrent architectures, our proposed research aims to accurately classify the presence or absence of atherosclerosis, contributing to the development of a non-invasive and intelligent healthcare diagnostic tool. Techniques such as correlation analysis or statistical testing are used to retain, only the most medical significant (probability (p) <0.05) parameters. Multiple deep learning models including CNN-LSTM, CNN-LSTM, GRU, BiLSTM, and Transformers are trained using the preprocessed dataset. These models learn spatial and temporal features to classify between healthy and atherosclerotic cases. The overall model training and pipeline is represented in Fig. 3.

Algorithm: Model training and evaluation pipeline

1. Load two datasets (healthy and unhealthy).
2. Assigning class labels and combining datasets.
3. Label encoding encodes categorical values (gender, label).
4. Apply feature engineering to compute features.
5. Min max scalar to normalize all numerical features.
6. For deep learning models reshape feature matrix into (samples, features, 1)
7. Initialization of 5 fold stratified cross validation.
8. For each model named M in {LSTM, GRU, BiLSTM, CNN-LSTM, Transformer}:
    a. Split data into training and testing folds.
    b. Using architecture definitions build a model M.
    c. Train using Adam optimizer of batch size 32 and epochs of 25 including Early Stopping and ReduceLROnPlateau.
    d. Evaluation metrics: accuracy, precision, recall and F1-score.
    e. Average metrics and store per fold.
9. Selection of model with highest F1-score.
10. Retrain the best model on entire dataset with validation split of 20%.
11. Final Trained model is saved

Fig. 3. Deep learning models training, and evaluation pipeline.

a) LSTM: This is type of RNN architecture, particularly designed to learn long-term dependencies in sequential data [30]. Unlike regular RNNs, LSTM networks are able to retain the information for longer time steps by utilizing a memory cell, which can store information and regulate its flow through three gates viz. input gate, forget gate, and output gate. The input gate determines what values from the input should be written into the memory cell. The forget gate determines what must be forgotten in the past state of memory, and finally the output gate determines what gets passed to the next hidden state. These procedures help to pick up relevant features efficiently over time. LSTM is particularly useful in learning physiological signal sequence patterns where patterns evolve over time. In fact for detecting the atherosclerosis, LSTM helps to process continuously the sensor data and learns predictive features of arterial stiffness or abnormality classes. Gates and cell states are represented using the Equations (1)-(6), in which the LSTM model remembers the long term patterns [30].

Forget Gate: $f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f)$ (1)

Input Gate: $i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i)$ (2)

Candidate cell state: $\tilde{c}_t = tanh(W_c.[h_{t-1}, x_t] + b_c)$ (3)

New cell state: $c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$ (4)

Output Gate: $o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o)$ (5)

Hidden state: $h_t = o_t \odot \tanh(c_t)$ (6)

where, x_t is the input vector at the time step t, h_t is the hidden state, W_f, W_i, W_c, W_o are the weight matrices for the forget, input, candidate and output gates, respectively. Further, b_f, b_i, b_c, b_o are the bias terms for each gate, σ is the sigmoid function, and $\odot$ is the element wise Hadamard multiplication. Fig. 4 represents the pseudo code of LSTM model.

Algorithm: LSTM model architecture
1. Input: x ∈ R^{1xF}
2. Layer 1: LSTM layer (16 units, L2 regularization with λ = 0.001)
3. Layer 2: Dropout layer (rate=0.3)
4. Layer 3: Dense layer (8 units, ReLU activation)
5. Output layer: Dense (1 unit, sigmoid activation function)

Fig. 4. Pseudo code of LSTM model

b) GRU: This is a reduced form of the LSTM network that also learns temporal dependencies in sequential data [19, 30, 31]. This model merges forget, and input gates into one update gate and combines the cell state and retired state, reducing the architecture. Two primary rudiments of a GRU are update gate, which deals with the extent to which former information must be encouraged to future, and reset gate conveys how to incorporate new input together with the memory of the history. GRU networks use smaller parameters than LSTM, making them hastily to train with analogous performance. In this study, GRU learns patterns from ECG and GSR signals, and assists in the overall bracket delicacy at a lower computational cost. Architecture using Equations (7)-(10) uses the five layers namely, input layer, GRU layer, dropout layer, dense (fully connected) layer and output dense layer [19, 30, 31].

Update gate: $z_t = \sigma(W_z.x_t + U_z h_{t-1} + b_z)$ (7)

Reset Gate: $r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r)$ (8)

Candidate state: $\tilde{h}_t = tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h)$ (9)

New Hidden state: $h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$ (10)

where, x_t represents the input vector at the time step t, h_t being the hidden state, W, U, b are weight matrices and bias, r_t reset gate and $\tilde{h}_t$ candidate hidden state. The pseudo code of the GRU model architecture is shown in Fig. 5.

Algorithm: GRU model architecture
6. Input: x ∈ R^{1xF}
7. Layer 1: GRU layer (16 units, L2 regularization with λ = 0.001)
8. Layer 2: Dropout layer (rate=0.3)
9. Layer 3: Dense layer (8 units, ReLU activation)

Output layer: Dense (1 unit, sigmoid activation function)

Fig. 5. Pseudo code for the implementation of GRU learning model

c) Transformer: Transformer is a deep learning model developed for processing sequential data without the need for recurrence, employs a self-attention mechanism to calculate the relative importance of various positions in the input sequence, which facilitates parallelization and improved long-range dependence handling. The model has encoder and decoder stacks, each of which includes multi-head attention layers, position-wise feedforward networks, and layer normalization. Subjecting the physiological sensor data, Transformers provide improved performance in capturing the complex interdependencies among time steps compared to standard RNNs, and hence, they are appropriate for atherosclerosis detection tasks that involve high-dimensional time-series data. In addition, ability of model in providing dynamic weights to various segments of the sequence, this improves the feature interpretability and classification performance [29]. Input layer is of shape (T, F) where, T is the time step and F is the features per step. The input is passed to the 1D convolutional layer to extract the local patterns defined in Equation (11) as:

$$Conv1D(x) = \sum_{i=1}^{k} w_i.x_{t+i} + b \quad (11)$$

where, *k=3* is the kernel size and output is of shape *(T, 32)*. The convolution will capture the short term interaction between the features.

The multi head attention shown in Equation (12)-(13) is the core component of the transformer [29]. This attention mechanism allows the model to weigh the relevance of all features for each time step or vector. Here each input vector is linearly transformed into three vectors viz. Query (Q), a key (K), and a Value (V). However, for a given input, theta belongs to $R^{Txd_{model}}$.

$$Q = XW^Q, K = XW^K, V = XW^V \quad (12)$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (13)$$

$\sqrt{d_k}$ is the scaling factor in order to prevent large gradients.

Considering the multi head attention this process is performed in parallel, represented in Equation (14):

$$MultiHead(X) = Concat(head_{1 \ldots\ldots} head_h\ )W^O \quad (14)$$

W^O is output projection matrix.

A residual connection adds the input back to the attention output (Equation 15), which helps the gradient flow and stabilizes the training.

$$X_{attn} = LayerNorm(X + Attention(\mathrm{Q,K,V}) \quad (15)$$

Each feature vector is passed through the same two layer feed forward network (Equation (16)):

$$\mathrm{FFN(x)} = \max(0, \mathrm{x}W_1 + b_1)W_2 + b_2 \quad (16)$$

$W_1 \epsilon R^{d_{model} x\ d_{ff}}$, $W_2 \epsilon R^{\ d_{ff} x\ d_{model}}$ and d_{ff} is the dimension of the hidden layer (taken as 64).

To summarize the feature information across the time, the samples (time step) shown in Equation (17):

$$\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_{ff}^{(t)} \quad (17)$$

where, $\bar{X}$ is reduced into single vector and $X \epsilon R^{d_{model}}$

Finally, the model passes the vector through the fully connected layers represented in Equation (18):

$$h = RELU(\bar{X}W_3 + b_3) \quad (18)$$

$$\hat{y} = \sigma(hW_4 + b_4)$$

where, $\hat{y}$ is the probability of being unhealthy, W_3 is the weight matrix of the dense layer, b_3 is the bias vector, W_4 is the final output weight, h is the hidden representation and b_4 is the final output bias. The pseudo code of Transformer model is described in Fig. 6.

Algorithm: Transformer model architecture
1. Input: x ∈ R^{1xF}
2. Layer 1: 1D convolutional layer having 32 filters of kernel size 3, ReLU activation
3. Layer 2: Two heads-Multihead attention layer with (key_dim=16)
4. Layer 3: Residual connection layer with layer normalization
5. Layer 4: Feedfoward network with:
    a. Dense layer of 64 units (ReLU unit)
    b. Dense layer of 32 units
6. 1D global average pooling layer
7. Dense layer of 16 units with ReLU activation unit
8. Output layer: Dense layer of 1 unit and sigmoid activation function.

Fig. 6. Pseudo code for the Transformer model

d) BiLSTM: This is a sophisticated version of LSTM networks that processes the input sequence in forward and backward directions [18]. The two way processing allows the network to gain a thorough understanding of the sequence by taking into account past and future context. The BiLSTM model consists of two distinct LSTM layers: one to process the sequence in left-to-right (forward) direction and the other in right-to-left (backward) directions, respectively. These outputs are concatenated at every time step, allowing the model to capture contextually richer representations. In an atherosclerosis detection application, BiLSTM enhances the accuracy of the prediction by analyzing the sequential health information, detecting trends, and outliers that are not visible in a unidirectional model.

The BiLSTM consists of two LSTM layers in parallel; one processes the sequence in the forward direction and other processes in the backward direction. The outputs are concatenated at each time steps so the model captures the information from the input in both directions. In Equation (19), $\vec{h}_t$ is the hidden state from the forward LSTM and $\overleftarrow{h}_t$ is the hidden state from backward LSTM, respectively. Also, h_t is the final hidden state for the time step t [19].

$$h_t = [\vec{h}_t\ ;\ \overleftarrow{h}_t\ ] \quad (19)$$

where, the symbol [.;.] used to indicate the concatenation.

Equations (20)-(25) represent the architectural representation of the BiLSTM network.

Forget Gate:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (20)$$

Input Gate:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (21)$$

Cell candidate:

$$\bar{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (22)$$

Cell State:

$$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t \quad (23)$$

Output Gate:

$$o_t = \sigma(W_0 x_t + U_0 h_{t-1} + b_0) \quad (24)$$

Hidden State:

$$h_t = o_t \odot \tanh(c_t) \quad (25)$$

All these equations are applied separately for the forward as well as backward direction and then it is concatenated. In this, interdependencies learned between the vital features like SpO2, heart rate, LDL/HDL ratio, etc., regardless of any order. Here W_*, U_*, b_* are the trainable weights and bias for each gate. Fig. 7 depicts the pseudo code of Bi-LSTM model. L2 regularization is used in Layer 1.

Algorithm: Bi-LSTM model architecture
1. Input: x ∈ R^{1xF}
2. Layer 1: 16 units LSTM Bidirectional layer
3. Layer 2: Dropout layer having dropout rate=0.3.
4. Layer 3: 8 units dense layer with ReLU activation.
5. Output layer of 1 unit with sigmoid activation function.

Fig. 7. Pseudo code of Bi-LSTM model

e) CNN-LSTM: This is a deep learning architecture that is a combination of CNNs to extract spatial features and LSTM networks for temporal modeling [19, 29-30]. In this system, CNN layers initially process subsequences of sensor signals to detect localized patterns or features. Extracted features are then fed to LSTM layers to capture the time-dependent relations over sequences. CNN-LSTM is very strong for multivariate time series with both spatial (e.g., between several sensors or windows) and temporal dependence. This model is a potential application to tap into both local trends as well as the sequential patterns in sensor values for sustainable classifying the atherosclerosis cases.

The CNN-LSTM is hybrid architecture using Equations (26)-(29), which takes the CNN output as the input to LSTM. The CNN-LSTM combines the convolutional feature extraction along with the temporal modeling [19, 29, 30].

$$X \in \mathrm{R}^{Txd} \quad (26)$$

This the time series input data, T is the number of time steps, d is the number of features per time step.

$$\mathrm{Z}^{(2)} = Conv1D_2(Conv1D_1(X)) \,\epsilon R^{TxF_2} \quad (27)$$

where, $\mathrm{Conv1D_1}$ and $\mathrm{Conv1D_2}$ are the first, and second one-dimensional (1D) convolutional layers, $\mathrm{Z}^{(2)}$ is the output of convolutional layer, F_2 is the filters count in the second $\mathrm{conv1D_2}$ layer, i.e. the CNN layer uses the sliding window scans the input sequences using filters to extract local patterns like sharp rises in the heart rate.

$$h_T = LSTM(\mathrm{Z}^{(2)}) \in \mathrm{R}^{n} \quad (28)$$

where, $\mathrm{Z}^{(2)}$ is the high level CNN feature now taken as input to LSTM, h_T is the hidden state at the final step T, n is the hidden size that is number of LSTM units.

$$\hat{y} = \sigma(Wh_T + b)\epsilon(0,1) \quad (29)$$

where, W is the weight matrix in the last dense layer, b is the bias term, and $\hat{y}$ is the final prediction of the model (probability). The final hidden state will pass through the dense layer having sigmoid as the activation function, to provide the likelihood probability of a patient being unhealthy.

The pseudo code of CNN-LSTM hybrid architecture is designed, and represented in Fig. 8.

Algorithm: CNN-LSTM model architecture
1. Input: x ∈ R^{1xF}
2. Layer 1: 1D convolution layer with 64 filters, kernel size =5, activation function = 'ReLU', and padding = 'same'.
3. Layer 2: The Dropout layer having Dropout out rate of 0.3
4. Layer 3: 1D convolution layer with 64 filters, kernel size of 3, activation function = 'ReLU', and padding = 'same'.
5. Layer 4: Dropout layer with Dropout rate of 0.3.
6. Layer 5: LSTM layer of 32 units, return sequences of 'False', L2 regularization =0.001.
7. Layer 6: Dropout layer with Dropout rate of 0.3.
8. Layer 7: Dense layer of 16 units with activation function 'ReLU'.
9. Output layer with dense layer of 1 unit and activation function 'Sigmoid'.

Fig. 8. Pseudo code of CNN-LSTM model

In this paper, five deep learning architectures such as LSTM, GRU, Bi-LSTM, CNN-LSTM, and a light weight Transformer are implemented. The dataset is created where each patient record is a single snapshot comprising 13 features. However, the LSTM, BiLSTM, CNN-LSTM, GRU, and Transformer architectures are designed for the temporal or the sequential data are applied in an exploratory manner for these static patient snapshots, to evaluate the model performance on the static measurements. The input to each model is shaped as $(1\ x\ F)$; where, F is the total number of features per patient.

Letting the physiological dataset each model is adapted, where 1D convolution is applied to CNN-LSTM to capture the local signal patterns prior sequence modeling with the LSTM layers. On the feature sequence the Transformer model applies the multi head self-attention. Binary cross entropy loss with Adam optimizer of learning rate 0.001 model is trained, with 25 epochs and batch size of 32. To prevent over fitting drop out (0.3), and L2 regularization (0.001) is applied. To optimize convergence early stopping and ReduceLROnPlateau is used. Considering an average accuracy, precision, recall, and F1 score as metrics, the performance is evaluated using stratified 5-fold cross-validation process.

The dataset consists of total 406 unique subjects' data. The healthy data is of 180 records and remaining 226 with the signs of atherosclerosis. Each patient contributes exactly single record, such that the stratified 5-fold cross-validation is applied at the record level, and between the training and test folds so data leakage is simplified. In addition, no subject appeared in both training and test folds, for each epoch training and validation is recorded, with the validation split of 0.2 the best performing model is retrained on full dataset.

III. Results And Discussion

This section represents the results of the proposed deep learning models for the detection of atherosclerosis using the sensor and clinical data. The evaluation was performed through a rigorous 5-fold stratified cross-validation approach to ensure robustness and generalization across varied data subsets. Each model is assessed using key classification metrics: Accuracy, Precision, Recall, and F1 Score.

The performance comparison in TABLE I, which includes five deep learning architectures: LSTM, BiLSTM, GRU, CNN-LSTM (hybrid model), and Transformer (attention-based model). The reported performance metrics reflect the capacity of the model on patient static snapshot.

Considering the results obtained in TABLE I, a weighted scoring system in Equation (30) is employed,

$$F1_{weighted} = \frac{\sum_{c=1}^{N} n_c F1_c}{\sum_{c=1}^{N} n_c} \quad (30)$$

where, N is the number of classes, n_c is the total number of true samples in class 'c' and $F1_c$ is the F1 score for a class 'c'.

$$Score = \frac{w_1.A + w_2.P + w_3.R + w_4.F1_{score}}{w_1 + w_2 + w_3 + w_4} \quad (31)$$

where, w_1, w_2, w_3 and w_4 are the important weights being assigned to each metric. If all metrics are equally important then w_1, w_2, w_3 and w_4 equals to 1.

TABLE I. PERFORMANCE OF DEEP LEARNING MODELS, AVERAGED OVER THE 5 FOLDS CROSS VALIDATION WITH VALUES PRESENTED AS MEAN ± STANDARD DEVIATION, IN 5 RUNS.

Deep Learning Models	Accuracy (A)	Precision (P)	Recall (R)	F1 Score
LSTM	90.80% ± 0.45%	92.80% ± 0.45%	91.00% ± 0.00%	92.00% ± 0.00%
BiLSTM	95.80% ± 0.45%	97.00% ± 0.00%	95.60% ± 0.55%	96.00% ± 0.00%
GRU	95.00% ± 0.00%	96.00% ± 0.00%	95.00% ± 0.00%	96.00% ± 0.00%
CNN-LSTM	98.00% ± 0.00%	98.00% ± 0.00%	99.00% ± 0.00%	98.00% ± 0.00%
Transformer	97.00% ± 0.00%	97.00% ± 0.00%	97.00% ± 0.00%	97.00% ± 0.00%

The composite score is calculated for each model as per the Equation (31) and is represented in TABLE II.

TABLE II. FORMULATION OF COMPOSITE SCORE WITH MEAN ± STANDARD DEVIATION FOR DEEP LEARNING MODELS.

Deep Learning Models	Composite Score
LSTM	91.60% ± 0.55%
BiLSTM	96.00% ± 0.00%
GRU	95.60% ± 0.55%
CNN-LSTM	98.00% ± 0.00%
Transformer	97.00% ± 0.00%

However, the data set is balanced so F1 score calculated, in Table I is effectively the $F1_{weighted}$ score represented in Equation (30). Overall, the CNN-LSTM model achieved the best weighted score of 98.00% ± 0.00% in comparison with individual LSTM (91.60% ± 0.55%), GRU (95.60% ± 0.55%), BiLSTM (96.00% ± 0.00%) or Transformer (97.00% ± 0.00%) models, and which was thus selected as final predictive model. This model was saved in CNN-LSTM_best_model.h5 format for compatibility with deployment environments.

Fig. 9 presents the training and validation accuracy and loss for the models considered in this study. The CNN-LSTM and Transformer models outperform others, achieving the best highest accuracy, closely followed by GRU and BiLSTM. LSTM records the lowest accuracy, aligning with the higher loss. These results reinforce the effectiveness of hybrid and attention-based models in capturing the complex patterns in atherosclerosis-related physiological signals.

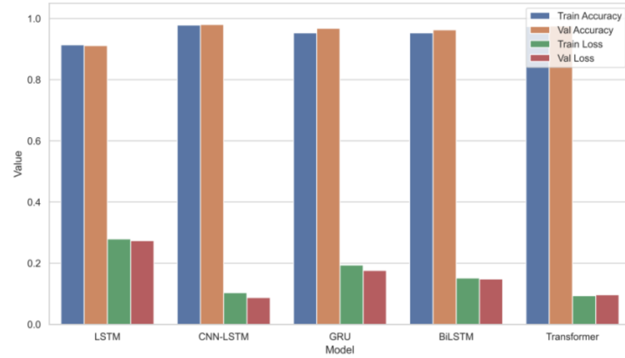

Fig. 9. Comparison of models training and validation using accuracy and loss.

In final phase, the saved CNN-LSTM model is loaded into a Streamlit-web based user interface, where it accepts real-time user inputs such as age, gender, glucose, cholesterol, LDL-HDL lipoproteins, GSR, systolic BP, diastolic BP, heart rate, and the ECG signal physiological parameters. The input data undergoes feature engineering and normalization before being passed to model. The model executes outputs: the prediction label: Healthy or Unhealthy (at-risk), and the model confidence score which is a health risk analysis based on number of parameters falling outside standard medical ranges. Table III represents the clinical ranges of the health parameter as per the reference [32, 33, 34, 35, 36].

Fig. 10 shows model's prediction for a healthy class based on inputs namely, age, BP, glucose, cholesterol levels, and others.

TABLE III. CLINICAL REFERENCE DATA FOR ATHEROSCLEROSIS RISK ASSESSMENT PARAMETERS.

Parameters	Specifications
Gender	Male = higher baseline risk, Female = lower until menopause
Glucose (mg/dL)	70–99 mg/dL (normal), 100–125 mg/dL (prediabetes), ≥126 mg/dL (diabetes)
Cholesterol (mg/dL)	<200 mg/dL (optimal)
LDL (mg/dL)	<100 mg/dL (optimal), 100–129 mg/dL (near optimal)
HDL (mg/dL)	>40 mg/dL (men), >50 mg/dL (women), ≥60 mg/dL protective
GSR (kΩ)	10–200 kΩ (normal physiological range; varies with stress & skin condition)
Systolic BP (mmHg)	<120 mmHg (normal), 120–129 mmHg (elevated), ≥130 mmHg (hypertension)
Diastolic BP (mmHg)	<80 mmHg (normal), 80–89 mmHg (elevated), ≥90 mmHg (hypertension)
SpO_2 (%)	95–100 % normal
Heart Rate (bpm)	60–100 bpm resting normal
HRV (RMSSD, ms)	≥20 ms healthy, 15–20 ms moderate risk, <15 ms high risk

ECG (bpm)	60–100 bpm normal
Weight (kg)	Depends on height → BMI 18.5–24.9 kg normal

Fig. 10. The Streamlit web based user interface showing the output of healthy individual typical details.

The model processes and normalizes these (TABLE III) physiological parameters, before generating a prediction. The HRV is being quantified using RMSSD (Root Mean Square of Successive Differences); which is expressed in milliseconds (ms). HRV value in web based interface is the normalized ECG derived HRV index value. In this case, the prediction label is "Healthy," with a model confidence score. The health risk analysis confirms that all input parameters are within standard medical ranges, suggesting no immediate risk of atherosclerosis. This output serves as an example for the model's ability in accurately identifying the individuals without signs of the abnormal condition, providing a clear baseline for further evaluation of the at-risk patients. The interpretation of atherosclerosis risk is illustrated using user interface.

Fig. 11 illustrates the model's prediction of a subject with abnormal physiological pattern (at risk). Analyzing the output, the cholesterol 228 mg/dL, blood pressure is more than the normal range shown in Table II. Total of 7 parameters (Fig. 11) linked closely to cardiovascular health which is outside the optimal range. The risky parameters include LDL-HDL, glucose, and cholesterol, systolic BP, SpO2, and ECG where, predicts the label as "Unhealthy" (at-risk) with a confidence score. The health risk analysis indicates the major parameters of human body with elevated range such as blood pressure and cholesterol levels fall out of standard medical ranges. This output shows the model's ability to detect potential atherosclerosis cases, aiding early identification in at-risk individuals. The model predictions for the atherosclerosis risk are predicted and the prediction values are evaluated and tested using the user interface.

IV. CONCLUSION

The proposed research work successfully demonstrates an effective integration of biomedical sensor data and the deep learning algorithms for the early detection of atherosclerosis, which is a major contributor to stroke and other cardiovascular healthcare conditions. Considering non-invasive bio sensors such as ECG, PPG, GSR, and BP, the system captures essential physiological signals that are indicative of cardiovascular health. The proposed approach is tested with the individual subjects from the real-time human physiological parameters. In this study, the acquired bio signals are subjected to applied deep learning models, including LSTM, BiLSTM, GRU, CNN-LSTM, and Transformer architectures, to extract the subtle patterns associated with early atherosclerotic changes. However, CNN-LSTM model provided a best weighted score of 0.98 using 5-fold cross-validations. The novelty of this work lies in the application of the clinical dataset with multiple parameters combining engineered cardiac features.

The proposed method offers a scalable and cost-effective alternative to traditional diagnostic methods, enhancing the accessibility and timeliness of atherosclerosis screening. The system demonstrates promising accuracy in distinguishing between healthy and at-risk individuals, thereby underscoring the potential of predictive healthcare solutions. Overall, this work bridges the gap among raw physiological data and meaningful clinical insights, paving the way for intelligent, real-time, and preventive cardiovascular care. By facilitating early intervention, the designed system has the capability to significantly lower the incidence of stroke and improve patient outcomes, highlighting its relevance and impact in the realm of modern healthcare.

Fig. 11. The Streamlit web based user interface showing the output of an abnormal physiological pattern in an individual .

Future work involves expansion of the dataset to involve more subjects and design of mobile application platform to provide instant predictions, visual analysis, and track the health status of a subject continuously. Moreover, the sensors can be integrated with respiration sensor to provide efficient health state monitoring system.

Acknowledgment

The Authors are thankful to under graduate students and Head of Department of Artificial Intelligence and Machine Learning, Vivekananda College of Engineering & Technology, Puttur, Dakshina Kannada, Karnataka, India. The clinical data is self-collected from the Pragathi Multi Specialty Hospital, Puttur, 574201, Dakshina Kannada, Karnataka, India for this research work. The data collection is approved; ethical standard including the patient confidentiality is maintained.

"Ethical Approval: Approved by Pragathi Specialty Hospital Ethics Committee, Dakshina Kannada, Puttur-574201, Dakshina Kannada, Karnataka, India, (Approval #: PSH-EC-2025-001, dated May 10, 2025). Individual subject confidentiality and ethical standards are strictly maintained. All participants provided written informed consent. Study complies with Declaration of Helsinki."

References

[1] J. Fu, Y. Wang, Y. Ding, J. Wang, S. Deng, Z. Jiang, “Wearable ring sensor for monitoring biomarkers of atherosclerosis in sweat,” Talanta, vol. 287, pp. 127608, Jan. 2025, doi: 10.1016/j.talanta.2025.127608.

[2] R. U. S. Ahmad, W. U. Khan, M. S. Khan, P. Cheung, “Emerging rapid detection methods for the monitoring of cardiovascular diseases: current trends and future perspectives,” Mater Today Bio, vol. 32, pp. 101663, Mar. 2025, doi: 10.1016/j.mtbio.2025.101663.

[3] J. Wei, X. Zhang, Q. Chang, S. M. Mugo, Q. Zhang, “An integrated sweat sensor for synchronous detection of multiple atherosclerosis biomarkers,” Anal Chem, vol. 95, no. 42, pp. 15786–15794, Oct. 2023, doi: 10.1021/acs.analchem.3c03310.

[4] P. H. Charlton, P. A. Kyriacou, J. Mant, V. Marozas, P. Chowienczyk, J. Alastruey, “The 2023 wearable photoplethysmography roadmap,” Physiol Meas, vol. 44, no. 11, pp. 111001, Jul. 2023, doi: 10.1088/1361-6579/acead2.

[5] M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, A. Raad, “Smart wearables for the detection of cardiovascular diseases: a systematic literature review,” Sensors, vol. 23, no. 2, pp. 828, Jan. 2023, doi: 10.3390/s23020828.

[6] F. Miao, D. Wu, Z. Liu, R. Zhang, M. Tang, Y. Li, “Wearable sensing, big data technology for cardiovascular healthcare: current status and future prospective,” Chin Med J, vol. 136, no. 9, pp. 1015–1025, Sep. 2022, doi: 10.1097/cm9.0000000000002117.

[7] M. Kang, Y. Chang, J. Kang, Y. Kim, S. Ryu, “Electrocardiogram risk score and prevalence of subclinical atherosclerosis: a cross-sectional study,” J Pers Med, vol. 12, no. 3, pp. 463, Mar. 2022, doi: 10.3390/jpm12030463.

[8] J. Ni, X. Zhang, Y. Liu, Z. Chen, “Development of a non-invasive method for skin cholesterol detection: pre-clinical assessment in atherosclerosis screening,” Biomed Eng Online, vol. 20, no. 1, Jun. 2021, doi: 10.1186/s12938-021-00889-1.

[9] P. H. Charlton, P. A. Kyriacou, J. Mant, V. Marozas, P. Chowienczyk, J. Alastruey, “Wearable photoplethysmography for cardiovascular monitoring,” Proc IEEE, vol. 110, no. 3, pp. 355–381, Mar. 2022, doi: 10.1109/jproc.2022.3149785.

[10] V. Ouyang, J. Li, M. Chen, S. Wang, “The use of multi-site photoplethysmography (PPG) as a screening tool for coronary arterial disease and atherosclerosis,” Physiol Meas, vol. 42, no. 6, pp. 064006, Aug. 2020, doi: 10.1088/1361-6579/abad48.

[11] T. Biering-Sørensen, J. M. Smith, R. K. Johnson, "Global ECG measures and cardiac structure and function," Circ Arrhythm Electrophysiol, vol. 11, no. 3, Mar. 2018, doi: 10.1161/circep.117.005961.

[12] O. Terrada, B. Cherradi, A. Raihani, O. Bouattane, "A novel medical diagnosis support system for predicting patients with atherosclerosis diseases," Inform Med Unlocked, vol. 21, pp. 100483, Jan. 2020, doi: 10.1016/j.imu.2020.100483.

[13] M. V. Villarejo, B. G. Zapirain, A. M. Zorrilla, "A stress sensor based on galvanic skin response (GSR) controlled by ZigBee," Sensors, vol. 12, no. 5, pp. 6075–6101, May 2012, doi: 10.3390/s120506075.

[14] E. Falk, "Pathogenesis of atherosclerosis," J Am Coll Cardiol, vol. 47, no. 8 Suppl, pp. C7–C12, Apr. 2006, doi: 10.1016/j.jacc.2005.09.068.

[15] A. El-Ibrahimi, O. Terrada, O. E. Gannour, B. Cherradi, A. E. Abbassi, O. Bouattane, "Optimizing machine learning algorithms for heart disease classification and prediction," Int J Online Biomed Eng, vol. 19, no. 15, pp. 61–76, Oct. 2023, doi: 10.3991/ijoe.v19i15.42653.

[16] Y. N. Fuadah, K. M. Lim, "Advances in cardiovascular signal analysis with future directions: a review of machine learning and deep learning models for cardiovascular disease classification based on ECG, PCG, and PPG signals," Biomed Eng Lett, vol. 15, no. 4, pp. 619–660, Apr. 2025, doi: 10.1007/s13534-025-00473-9.

[17] A. Reiss, I. Indlekofer, P. Schmidt, K. Van Laerhoven, "Deep PPG: large-scale heart rate estimation with convolutional neural networks," Sensors, vol. 19, no. 14, pp. 3079, Jul. 2019, doi: 10.3390/s19143079.

[18] Desai, Usha, Roshan Joy Martis, C. Gurudas Nayak, G. Seshikala, K. Sarika, and R. A. N. J. A. N. SHETTY K. "Decision support system for arrhythmia beats using ECG signals with DCT, DWT and EMD methods: a comparative study." Journal of Mechanics in Medicine and Biology 16, no. 01 (2016): 1640012. doi: 10.1142/S0219519416400121

[19] B. Rim, N.-J. Sung, S. Min, M. Hong, "Deep learning in physiological signal data: a survey," Sensors, vol. 20, no. 4, pp. 969, Feb. 2020, doi: 10.3390/s20040969.

[20] Y. Liang, Z. Chen, R. Ward, M. Elgendi, "Photoplethysmography and deep learning: enhancing hypertension risk stratification," Biosensors, vol. 8, no. 4, pp. 101, Oct. 2018, doi: 10.3390/bios8040101.

[21] C. K. K. Reddy, J. Smith, M. Lee, "Detecting anomalies in smart wearables for hypertension: a deep learning mechanism," Front Public Health, vol. 12, Jan. 2025, doi: 10.3389/fpubh.2024.1426168.

[22] W. H. Weng, P. Zhao, L. Chen, "Predicting cardiovascular disease risk using photoplethysmography and deep learning," PLOS Glob Public Health, vol. 4, no. 6, pp. e0003204, Jun. 2024, doi: 10.1371/journal.pgph.0003204.

[23] Y. Appelman, B. B. Van Rijn, M. E. T. Haaf, E. Boersma, S. A. E. Peters, "Sex differences in cardiovascular risk factors and disease prevention," Atherosclerosis, vol. 241, no. 1, pp. 211–218, Jan. 2015, doi: 10.1016/j.atherosclerosis.2015.01.027.

[24] C. Park, E. Guallar, et al., "Fasting glucose level and the risk of incident atherosclerotic cardiovascular diseases," Diabetes Care, vol. 36, no. 7, pp. 1988–1993, Feb. 2013, doi: 10.2337/dc12-1577.

[25] A. K. Pancholia, N. K. Kabra, R. Gupta, "Laboratory evaluation of lipid parameters in clinical practice," Indian Heart J, vol. 76, pp. S29–S32, Feb. 2024, doi: 10.1016/j.ihj.2024.02.002.

[26] Y. Kubota, L. Y. Chen, E. A. Whitsel, A. R. Folsom, "Heart rate variability and lifetime risk of cardiovascular disease: the Atherosclerosis Risk in Communities Study," Ann Epidemiol, vol. 27, no. 10, pp. 619–625.e2, Oct. 2017, doi: 10.1016/j.annepidem.2017.08.024.

[27] U. Desai, A. D. Shetty, "Electrodermal activity (EDA) for treatment of neurological and psychiatric disorder patients: a review," in Proc. 7th Int. Conf. on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1424–1430, doi: 10.1109/ICACCS51430.2021.9441808.

[28] U. Desai, A. D. Shetty, T. M., A. S. A., G. Nekar, S. B., "Detection of anxiety in psychiatric patients using physiological signals," in Proc. IEEE 19th India Council Int. Conf. (INDICON), Kochi, India, 2022, pp. 1–5, doi: 10.1109/INDICON56171.2022.10040142.

[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," Adv. Neural Inf. Process. Syst., vol. 30, 2017, pp. 5998–6008. doi: 10.5555/3295222.3295349.

[30] I. Goodfellow, Y. Bengio, A. Courville, Deep learning. Cambridge, MA: MIT Press; 2016.

[31] F. Lazzeri, Machine learning for time series forecasting with Python. Sebastopol, CA: O'Reilly Media; 2020.

[32] American Diabetes Association, "Standards of medical care in diabetes—2024," Diabetes Care, 2024.

[33] National Cholesterol Education Program (NCEP) Expert Panel, Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). NIH Publication No. 02-5215; Sep. 2002.

[34] P. K. Whelton et al., "2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines," Hypertension, vol. 71, no. 6, pp. 1269–1324, 2018, doi: 10.1161/HYP.0000000000000066.

[35] F. Shaffer, J. P. Ginsberg, "An overview of heart rate variability metrics and norms," Front Public Health, vol. 5, pp. 258, 2017, doi: 10.3389/fpubh.2017.00258.

[36] W. Boucsein, Electrodermal activity, 2nd ed. New York, USA: Springer, 2012. doi: 10.1007/978-1-4614-1126-0.