Hybrid Stacked Autoencoder - Based Optimization Framework for Cardiovascular Disease Risk Factor Prediction Using Photoplethysmography Signal Data

Deepak P. Franklin[1]* , Krishnamoorthi Murugasamy[2]

[1] Department of Electronics and Communication Engineering, S. A. Engineering College, Chennai 600077, India
[2] Department of Information Technology, Dr N. G. P Institute of Technology, Coimbatore 641048, India

Corresponding Author Email: deepakfranklinp@gmail.com

ABSTRACT

The high level of blood pressure (BP), heart rate, and blood glucose level are the cardiovascular disease (CVD) risk factors that commonly cause death and disability in the world. The photoplethysmogram (PPG) data is considered for monitoring BP, heart rate, and blood glucose levels. In the presented methodology, a deep learning based hybrid stacked autoencoder approach is presented to accurately classify the risk levels as pre-hypertension, hypertension, and normal. Initially, the input PPG data is taken for predicting the risk factors. This PPG data comprises BP levels, heart rate, and the glucose level in the blood. The input data is initially filtered with a Fast Normalized Least Mean Square (FNLMS) filtering approach to remove the unnecessary data in the input. Then, the effective features such as Gaussian Fitting, Spectral Entropy feature, statistical features, and demographic features are extracted from the filtered data. Afterwards, the rain optimization (RO) approach is utilized to optimally select the features and reduce the features' dimensionality. Finally, a hybrid stacked autoencoder with sparrow search optimization (HSAE-SSO) methodology is utilized to accurately predict the risk condition of the patients. This framework predicts the risk factor by classifying the data into normal, pre-hypertension, or hypertension classes. The presented methodology is implemented in the PYTHON platform. The experimental results are examined with different performance metrics in terms of accuracy (95.43%), precision (92.26%), F-measure (92.62%), mean absolute error (MAE) (0.0456), and RMSE (Root mean squared error) (0.0523).

1. INTRODUCTION

Health monitoring systems are enhanced with wireless communication technology, which is proficient in communicating wirelessly. Abnormalities in the human body are diagnosed with this kind of development in devices [1]. Continuous monitoring with this system is essential for emergency purposes in the medical field, and it provides an alert or warning to the patients. Moreover, the patient's health condition is continuously monitored with the development of a cost-effective monitoring system [2]. Generally, cardiovascular disease (CVD) is considered one of the major reasons for death worldwide. Due to the different food styles, stress, environmental pollution, and other health related factors, mortality with CVD is estimated to increase in the near future [3]. Also, CVD patients are affected by multiple risk factors like high blood pressure (BP), obesity, and stress. So, it is essential to predict the risk of BP and heart rate, and high glucose levels in the population to prevent people from CVD [4]. In general, it should be noted that continuous monitoring of human BP and pulse rate is necessary, as these are considered important parameters in CVD. High BP increases more risky health related problems in the human body, creating a large amount of stress on the heart. This causes the

arteries to be less stretchy over a period of time [5]. When the arteries are flexible, the lumen becomes narrower, increasing the chance of clotting. This clot may create many hazardous diseases like stroke, heart attack, dementia, and kidney diseases [6].

Diabetes, like hypertension, is considered a lifelong metabolic condition due to decreased physiologic effects and insulin secretion. Nowadays, no effective treatments are available for diabetes, and continuous blood glucose level screening is also difficult [7]. Diabetes is controlled with self-monitoring at regular intervals. Normally, the glucose level of the human body is measured by getting a blood sample [8, 9]. This technique does not allow real-time monitoring and causes pain due to frequent blood collection from the patient's fingertip. So, this kind of shortcoming is exceeded by developing a non-invasive blood glucose technology in smart healthcare research. Conversely, self-monitoring with the help of non-invasive techniques is not in the advanced stage of improvement and needs enhancement by utilizing deep learning based approaches [10].

Generally, PPG data is utilized to estimate blood volume variations in particular areas of the body [11-13]. PPG is a methodology to attain CVD data based on the deviations of blood volume in veins [14, 15]. For providing information

about the tissue, a PPG device typically contains a light source [16]. In the circulatory system, the quantity of absorbed light changes as per the blood volume variations. The PPG signal includes data related to the circulatory system, breathing, and heartbeat beat and blood flow. Many researchers developed a non-invasive BP measurement and performed effective HR and respiratory rate monitoring [17, 18]. Particular elements in the PPG signals are analyzed to perform the abovementioned monitoring. For personal health monitoring, PPG signal processing is a new technology that reveals information about the human body's blood parameters and hemodynamic properties [19]. Also, PPG is considered a prominent technology utilized for the early prediction of several health diseases like arrhythmias, heart rate, cardiac output, coronary heart disease, systolic BP, and diastolic BP [20]. Some of the existing approaches used in the prediction of these risk levels are SVM (support vector machine), ANN (Artificial neural network) [21], ResNet, and WaveNet [22].

In order to extract clinically significant, noise-robust features from PPG data for CVD risk prediction, this study is new in that it combines a hybrid stacked autoencoder with sparrow search optimization (HSAE-SSO). Methodology closely combines representation learning and hyperparameter optimization, in contrast to previous research that either uses autoencoders for compression or metaheuristics for feature selection separately. This expands the use of PPG-based screening to practical, affordable health monitoring by producing interpretable latent biomarkers, increasing convergence efficiency, and improving prediction accuracy in noisy, device-heterogeneous environments.

The organization of this manuscript is summarized as follows: Section 2 provides a survey of various recent related works, Section 3 describes the detailed explanation of the presented methodology, Section 4 analyzes the experimental results and their discussions, and the manuscript is concluded in Section 5.

2. LITERATURE REVIEW

Panwar et al. [23] developed a deep learning system for PPG based BP and HR evaluation. The PP-Net model was developed in this article to calculate physiological parameters like HR, DBP (Diastolic blood pressure), and SBP (Systolic blood pressure) concurrently from a similar framework utilizing PPG data. The deep learning framework was designed with the LRCN (Long-term Recurrent Convolutional Network) framework, which exhibits the inherent ability of feature extraction. This feature extraction step excludes the cost-effective processes as well as the selection to make a less complex deployment on mobile platforms. The performance of the proposed methodology was evaluated with the publically available MIMIC-II database.

Chowdhury et al. [24] developed machine learning techniques to calculate BP from the PPG signal and demographic features. Hypertension leads to different health complications in the human body, and it is essential to monitor BP continuously. The general BP measures were considered discrete, and they were not comfortable for the patient. In order to overcome this issue, a continuous, non-invasive, and cuff-less based BP measurement framework was developed using extracted features with ML (machine learning) techniques. From PPG and their derivative signals, frequency, time, and time-frequency domain features were extracted by

feature extraction. The computational complexity of the whole framework was reduced with the feature selection method, which diminishes the over-fitting issues in ML techniques. Afterwards, the extracted features were utilized to estimate these predictions.

Esmaelpoor et al. [25] developed a multiple-stage DNN (deep neural network) model for BP prediction utilizing PPG signals. Difficulties in cuff-based and invasive BP measuring methods were overcome with easy access to bio-signals. To calculate SBP and DBP using the PPG signal, this article developed an enhanced DNN framework. Initially, the first phase of this framework was utilized to extract morphological features from every PPG signal and calculate SBP and DBP separately. In the second stage, another deep learning network is to attain the temporal variations. Also, this methodology integrates the dynamic utilized for the connection among the SBPs and DBPs to enhance the accuracy.

Tjahjadi and Ramli [26] developed an innovative BP classification framework that depends on the PPG signal data. The presented methodology utilized a KNN framework for the prediction of the risk status of the patient. This developed framework attained improved accuracy than the existing deep learning, Adaboost, and bagged tree approaches. But this approach considered only the BP data for predicting risk conditions. However, an enhanced methodology is needed for further improvement in prediction accuracy.

Zhang et al. [27] developed a non-invasive blood glucose screening based on PPG signal data processing with a smartphone and ML approaches. This framework can distinguish the blood glucose levels of a user into normal, borderline, and warning, depending on PPG data, which were developed to enable daily care at home. In this article, PPG signals were collected by smartphone camera videos, and the segmentation of the signal into a single period was done by the FSW (fitting-based sliding window) approach; the feature extraction was done with the help of the Gaussian fitting function. Finally, ML algorithms were applied to classify the valid samples into three glucose levels. Bharti et al. [28] demonstrate that combining deep learning with machine learning improves heart disease prediction accuracy. Katarya and Meena [29] gives a comparative analysis of different machine learning techniques. Elgendi et al. [30] present a dataset combining psychological scales and physiological signal collected with portable device.

The most critical illness that causes heart failure is CVD; therefore, prompt identification can reduce premature mortality and damage. Sinha et al. [31] have introduced filter-based feature selection and machine learning approaches to achieve an efficient identification. Four feature sets, each with 5, 10, 15, and 20 features, were selected using LASSO (Least Absolute Shrinkage and Selection Operator), RFE (Recursive Feature Elimination), Chi-square, and tree-based algorithms. Finally, the SVM was used with the selected 15 features and achieved efficient accuracy. A total of 112 models were trained to select the best feature set. This architecture showed its effective analysis in anticipating the CVD.

Effective spectral analysis of the PPG signal to monitor the heart rate has gained huge attention because the available algorithms have failed to show effective performance for resting persons due to the motion artifacts. To overcome such issues, the benefits of deep learning were leveraged by Ismail et al. [32], which exploits the time-frequency and spectral perspectives of the signal. The signal from the subjects who were performing exercise was collected and processed for

heart rate estimation. A set of features was extracted from the signal and fed into a hybrid CRNN (convolutional-recurrent neural network) for rate estimation from the heart signal. It achieves less error rate of 3.8±2.3 bpm for subject-independent and 2.41±2.90 bpm for subject-dependent.

Early detection and diagnosis of CVD are mandatory for reducing the disease severity. Therefore, Anny Leema and Jothiaruna [33] developed an image augmentation process that maintains a balance between the ECG image data classes. A balanced dataset with 6322 images showing 5 different classes, such as Myocardial infarction (MI), previous details about MI, COVID-19 ECG images, abnormal and normal heartbeat images. This balanced dataset was trained using ResNet-50, DenseNet-161, and VGG-16 models. Hyperparameters such as optimizer, dropout, and learning rate were introduced over the models to improve the CVD classification accuracy. DenseNet has obtained 93.33% accuracy, which was found to be more efficient than VGG-16 and ResNet-50.

A deep learning based arrhythmia classification was analyzed by Karri et al. [34], which undergoes QRS/PT detection, SDM (Sigma-Delta Modulation), and 1D-CNN (One-Dimensional Convolution Neural Networks) for achieving high performance. The local maximum and minimum algorithms were developed with SDM and 1D-CNN for QRS/PT wave detection. The LSTM (Long Short-Term Memory) based RNN (recurrent neural network) with a blend classifier was developed for arrhythmia classification, which utilizes features extracted by SDM and 1D-CNN. The MIT-BIH arrhythmia database was utilized for performance analysis. It consumes 1.5µw power and achieves efficient accuracy.

2.1 Problem statement

Nowadays, stress or hypertension plays a substantial role in every person's life, and BP is considered a main health indicator in human beings. One-sixth of the total population is affected by hypertension, a chronic disease. Fluctuations in the BP level form different kinds of critical conditions like heart attack and stroke. This BP level is related to HR and glucose levels in the blood. The increase in BP can reduce the heart rate slightly. In existing works, different approaches are utilized to predict these factors. However, the existing approaches have not obtained an enhanced performance in predicting CVD risk factors. Hence, the presented methodology utilized an effective deep learning based approach for accurately classifying CVD risk factors as normal, hypertension, and pre-hypertension.

2.2 Motivation

The higher BP level, glucose levels, and heart rate are the major reasons for the CVD issue in humans. Besides, the currently utilized glucose estimation techniques still depend on invasive approaches, which increases the risk of infection. Deep learning based approaches have recently been developed and have also enhanced the prediction performance. Moreover, it is essential to know the optimal health indicator for the earliest diagnosis of disease. A major disadvantage of the traditional method is lower accuracy in prediction. Therefore, it is essential to present a deep learning based framework to accurately predict the stages of the disease. This research presents the best deep learning based prediction framework with a hybrid stacked autoencoder. The advantage over the existing method is the accurate prediction of normal, hypertension, and pre-hypertension classes. The main contributions of this research are described as,

➤ To achieve the fastest convergence and the lowest complexity of PPG signal data, effective FNLMS filtering is presented.

➤ An efficient and effective feature extraction that deeply analyzes the patient data for accurate CVD detection at an early stage is required to improve the detection rate.

➤ An enhanced and recent optimization approach is presented to optimally select the relevant features from a set of features. This process decreases the dimensionality of the features.

➤ A hybrid deep learning technique with an optimization algorithm is used to maximize the performance of CVD risk factor prediction. Here, a hybrid stacked autoencoder framework is presented to enhance the prediction of CVD risk conditions as normal or pre-hypertension, or hypertension.

➤ The optimal parameter tuning must be performed over the introduced detection algorithm to achieve reasonable accuracy with fewer errors. For that, an SSO algorithm is introduced, which avoids overfitting issues.

3. PROPOSED METHODOLOGY

In this work, the CVD risk levels are classified into pre-hypertension, hypertension, and normal using an enhanced deep learning based framework. Initially, PPG signal data is taken and pre-processed using an FNLMS filtering approach. The input data comprised BP, glucose, and the patient's heartbeats. Then, the different features are extracted from the pre-processed data. Here, the Gaussian fitting feature, statistical feature with spectral entropy, and demographic features are extracted. The RO methodology is then used to optimally select the features and decrease the dimensionality of the features. At the end of the process, a hybrid stacked autoencoder framework with SSO methodology is utilized to classify the risk conditions accurately. This framework classifies the data into normal, pre-hypertension, or hypertension classes. The schematic representation of the presented methodology is depicted in Figure 1.
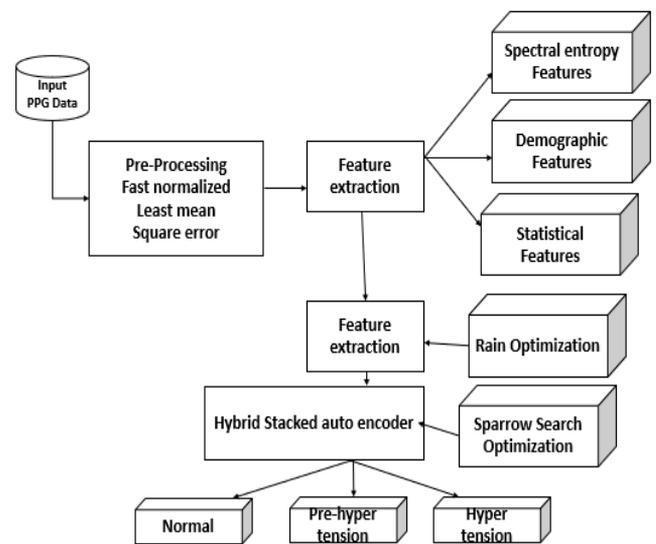

Figure 1. Schematic diagram of the presented methodology

3.1 Pre-processing

Initially, the input PPG data with heart rate, BP, and blood glucose level data are considered for predicting CVD risk conditions. The input data is pre-processed with the FNLMS processing and is described in the subsequent sub-sections.

3.1.1 Fast normalized least mean square (FNLMS) filtering

Initially, the input signal data is pre-processed using FNLMS filtering. The input signal data comprises heartbeat, BP, and glucose level. It is an enhanced filtering approach with the fastest convergence and the lowest complexity. This algorithm's normalization process normalizes the signal data's power with filtering. In this filtering approach, the simplified gain is updated and described in Eq. (1):

$$\begin{bmatrix} \bar{Q}_{k,m} \\ \bar{P}_m \end{bmatrix} = \begin{bmatrix} \dfrac{E_m}{\bar{\lambda}\bar{\alpha}_{m-1} + b_0} \\ \bar{Q}_{k,m-1} \end{bmatrix} \tag{1}$$

where, $\bar{Q}_{k,m}$ signifies adaptation gain, E_m signifies the error, $\bar{P}_m$ signifies the normalizing power quantity, $\bar{\lambda}\bar{\alpha}_{m-1}$ signifies the term to avoid the division with the least values, b_0 signifies the constant parameter, and the least square error is computed by the subsequent Eq. (2):

$$E_m = \bar{y}_m - c_m \bar{y}_{m-1} \tag{2}$$

where, $\bar{y}_m$ signifies the input signal data, the parameter c_m is utilized to reduce the cost function $Exp(E_m)^2$, and the parameter c_m is evaluated by Eq. (3).

$$c_m = \frac{L_{1y,m}}{L_{0y,m} + b_c} \tag{3}$$

where, $L_{1y,m}$ signifies the first-order lag correlation function of data y_m, $L_{0y,m}$ signifies the input signal power. b_c represents the constant utilized to remove the possibility of dividing by the element that is nearest to zero. Then the forward based prediction of error is described in Eq. (4).

$$\bar{\alpha}_m = \lambda \bar{\alpha}_{m-1} + \left(E_m\right)^2 \tag{4}$$

where, the constant b_0 is summed with the division component $\bar{\lambda}\bar{\alpha}_{m-1}$ to resolve the issues of least values in the NLMS approach to get fast convergence and enhanced filtering. This process enhances the filtering process in the PPG signal data.

3.2 Feature extraction

This section extracts an effective set of statistical, demographic, Gaussian fitting, and spectral entropy features. An effective set of feature extraction is important to accurately predict the risk factor of CVD. These feature extractions are described in the subsequent sub-sections.

3.2.1 Gaussian fitting feature

Gaussian fitting based feature extraction is described in the subsequent Eq. (5):

$$\bar{F}_t(m) = \bar{H}_t * Exp\left(-\frac{2(m - m_k)^2}{\bar{W}_m^2}\right), k = 1,2,3\ldots N; m = 1,2,3\ldots N \tag{5}$$

where, $\bar{F}_t(m)$ signifies the Gaussian fitting feature, $\bar{H}_t$ signifies the peak amplitude value, m_k signifies the position of peak time, and $\bar{W}_m^2$ signifies the half-width of every Gaussian wave.

3.2.2 Spectral entropy feature

In this spectral entropy feature extraction process, an entropy feature is extracted to attain the peak of the spectrum and its location. Here, the individual frequency elements of the spectrum are divided by the whole sum of all the elements to convert the spectrum into a PMF (probability mass function). It is described in the subsequent Eq. (6):

$$\bar{y}_k = \frac{\bar{Y}_k}{\sum\limits_{k=1}^{M} \bar{Y}_k}, \qquad k = 1,2,3\ldots M \tag{6}$$

where, $\bar{Y}_k$ signifies the energy of kth element of the spectrum, $y = \{y_1, y_2, y_3 \ldots y_M\}$ signifies the spectrum PMF, and M signifies the number of points in the spectrum. This feature ensures that the area below the normalized spectra is added to 1, and this normalized spectral data is considered a PMF for estimating the entropy feature. Moreover, the entropy evaluated for every frame is described in the subsequent Eq. (7):

$$\bar{E}_S = -\sum_{k=1}^{M} \bar{y}_k \log \bar{y}_k \tag{7}$$

where, $\bar{E}_S$ signifies the spectral entropy feature.

3.2.3 Demographic features

In the presented methodology, six different demographic feature data are considered for the prediction of CVD. The demographic features are gender, height, width, age, BMI, and heart rate. These demographic features vary from person to person, which is important to accurately predict a person's health condition.

3.2.4 Statistical features

This section extracts statistical features like mean, median, and standard deviation from the pre-processed PPG signal data.
(1) Mean
The mean feature is estimated by the ratio of the sum of all the data considered to the total quantity of the data. It is computed through the subsequent Eq. (8):

$$M_{\bar{Y}} = \frac{\sum \bar{y}_k}{k} \tag{8}$$

where, $M_{\bar{Y}}$ signifies the mean of the considered data, $\bar{y}_k$ signifies each feature data, and k signifies the total quantity of the considered data.
(2) Median
The median feature is the centre feature of the considered data group. This median feature is important to signify the data variations.
(3) Standard deviation
This measure estimates the variability and the consistency of the considered group of data. It is computed by the subsequent Eq. (9),

$$\bar{S}_D = \sqrt{\frac{\sum y_k - \bar{y}_k}{k - 1}} \tag{9}$$

where, $\bar{S}_D$ signifies the standard deviation, y_k signifies the pre-processed data.

(4) Percentile

This measure provides the percent value of the considered data. It is computed through the subsequent Eq. (10).

$$P(k^{th}) = \left(\frac{k}{100}\right)m \tag{10}$$

where, $P(k^{th})$ signifies the percentile data, m signifies the data in the dataset.

(5) Skewness

This measure evaluates whether the data is symmetrical. It is computed by the subsequent Eq. (11):

$$\hat{S}_w = \frac{\frac{\sum_{k=1}^{M}(M_{\bar{Y}}(k) - M_{\bar{Y}})^3}{M}}{(\bar{S}_D)^3} \tag{11}$$

where, $\hat{S}_w$ signifies the skewness, $M_{\bar{Y}}$ signifies the mean, and $\bar{S}_D$ signifies the standard deviation.

(6) Kurtosis

This measure is used to evaluate the peak value in the considered data, and it is computed through the subsequent Eq. (12):

$$\bar{K}_s = \frac{\sum_{k=1}^{M}(M_{\bar{Y}}(k) - M_{\bar{Y}})^4 \Big/ M}{(\bar{S}_D)^4} - 3 \tag{12}$$

where, $\bar{K}_s$ signifies the kurtosis measure, M signifies the total number of data.

(7) Mean absolute deviation (MAD)

It measures the average distance variation between the original and the predicted data. It is computed through the subsequent Eq. (13):

$$\bar{M}_{AD} = \frac{\sum_{k=1}^{M}|\bar{y}_k - \bar{y}|}{m} \tag{13}$$

where, $\bar{M}_{AD}$ represents the absolute mean distance. The flowchart for the overall proposed architecture is shown in Figure 2.
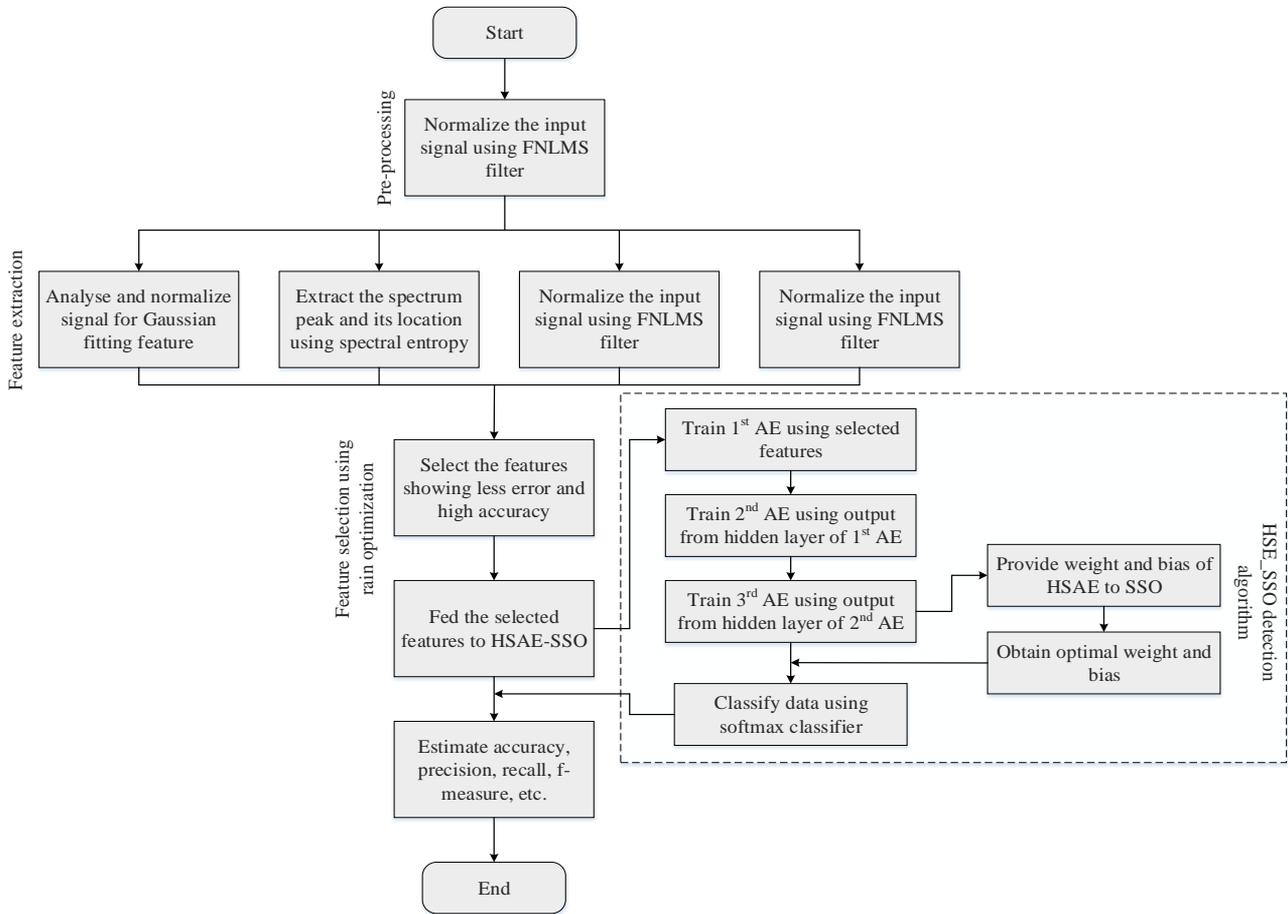

Figure 2. Flowchart for proposed processing steps

3.3 Feature selection using rain optimization

The behavior of raindrops from the highest position to the lowest position inspires the rain optimization approach. This optimization is initialized with the extracted features set, and it is represented as $\bar{Z}$, and the number of raindrops described in the subsequent Eqs. (14) and (15):

$$\bar{M}_d = \left\{\bar{f}_{m,1}, \bar{f}_{m,2}, \bar{f}_{m,3} .... \bar{f}_{m,k}\right\} \tag{14}$$

$$m \in \left\{1, 2, 3, ... \bar{F}\right\} \tag{15}$$

where, $\bar{M}_d$ signifies the feature set data and $\bar{f}$ signifies the feature elements. In the presented optimization approach, the

most important features are selected from the feature variables to reduce the dimensionality of the feature size, and it is described in the subsequent Eq. (16):

$$\bar{f}_m = U_d(\hat{u}_m, \hat{l}_m) \tag{16}$$

where, U_d signifies the uniform distribution, $\hat{u}_m$ signifies the upper bound, and $\hat{l}_m$ signifies the lower bound. The position of the optimized neighboring data is arbitrarily expressed in the subsequent Eqs. (17) and (18):
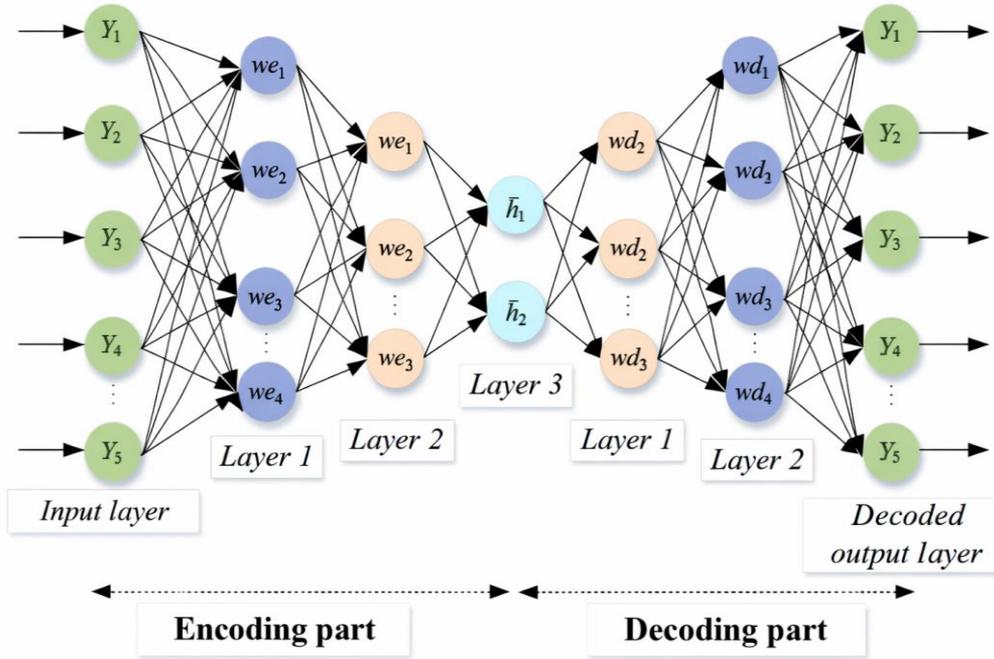
$$\left\| (M_d - M_p) * \bar{f}_m \right\| \leq \bar{P}_R * \bar{f}_m \tag{17}$$

$$\bar{P}_R = \bar{P}_R (Start * Iteration) \tag{18}$$

where, $\bar{P}_R$ signifies the positive real number that describes the nearest point $\bar{n}_p$, $(\cdot)$ function describes the unit vector size. The dominant position of the nearest point $(\overline{df})$ is utilized for the chosen optimal set of features, and it is described in the subsequent Eq. (19):

$$O_F \left(\overline{df}^m_{pk} \right) < O_F \left(\bar{f}^m_{pk} \right), \qquad m = 1, 2, 3, ... M_p \tag{19}$$

where, O_F signifies the optimized feature function based on the priority of important features from highest to lowest. Then the feature ranking is expressed in the subsequent Eq. (20):

$$F_{rank} = O_F \left| Max(i) - O_F \right| \tag{20}$$

where, $Max(i)$ signifies the maximum iteration, O_F signifies the optimal features, and F_{rank} signifies the selected features through ranking. The ranked feature is the feature attained above the threshold (T_h), and it is represented as $F_{rank} > T_h$. Here, the higher priority data is considered for further processing. This optimization process selects the features optimally to reduce the dimensionality of the feature set. This ranking criterion is utilized to optimally choose the features and reduce the dimensionality. In this way, features are optimally selected to reduce the dimensionality.

3.4 Classification using a hybrid stacked autoencoder

An autoencoder is a neural network framework comprised of three layers: input layer, hidden layer, and output layer. These layers comprised the encoder and the decoder to learn the features and make the output decision. In the presented methodology, hybrid stacked auto encoders are utilized for the classification. The schematic diagram of the presented hybrid stacked autoencoder is depicted in Figure 3.

Figure 3. Schematic diagram of hybrid SAE

The presented hybrid SAE framework comprises the stacked form of autoencoders. This framework performs both the encoding as well as the decoding process of the extracted features for feature learning and provides the final output decision. This process retains the essential information from the input feature representation for accurate output. Here, the encoder is utilized for mapping the input information to the hidden information, and the decoder part reconstructs the input information from the hidden representation. The encoding process of the stacked autoencoder is described in the subsequent Eq. (21):

$$H_k = \bar{f}(\bar{w}_1 y_k + \bar{B}_1) \tag{21}$$

where, $\bar{f}$ signifies the encoding operation, $\bar{w}_1$ signifies the encoder's weight matrix, and $\bar{B}_1$ signifies the vector function of bias. Similarly, the decoder operation is described in the subsequent Eq. (22):

$$\tilde{y}_k = \bar{g}(\bar{w}_2 H_k + \bar{B}_2) \tag{22}$$

where, $\bar{g}$ signifies the decoding operation, $\bar{w}_2$ signifies the

decoder's weight matrix, and $\bar{B}_2$ signifies the bias vector. The training procedure exhibited by hybrid SAE is discussed below:

Initially, the first autoencoder is trained to reduce errors between the original and reconstructed input.

Then, the output from the hidden layer of 1st autoencoder is taken as input for 2nd autoencoder, and training is performed.

Then, the procedure is repeated for subsequent autoencoders.

After that, the output from the last layer of the autoencoder is provided as an input to the softmax layer, and training is performed.

Finally, the weight and bias parameters are fine-tuned using the SSO algorithm.

3.4.1 Softmax regression classifier

In this, softmax regression is used for the final classification. It utilizes the output from the hidden layer of the final autoencoder. Based on the input x, a hypothesis function is designed by estimating the probability $P(y = j|x)$ for several categories j. The hypothesis is defined below Eq. (23):

$$h_\theta\left(x^{(i)}\right) = \begin{matrix} p\left(y^{(i)} = 1 \middle| x^{(i)}; \theta\right) \\ p\left(y^{(i)} = 2 \middle| x^{(i)}; \theta\right) \\ p\left(y^{(i)} = k \middle| x^{(i)}; \theta\right) \end{matrix}$$

$$= \frac{1}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}} \begin{matrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ e^{\theta_k^T x^{(i)}} \end{matrix} \qquad (23)$$

where, $\theta_1, \theta_2, .., \theta_k$ are model parameters, $\frac{1}{\left(\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}\right)}$ normalizes the probability distribution, and ensures the probability sum is 1. The loss function is defined in Eq. (24):

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{k} 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{j=1}^{k} e^{\theta_j^T x^{(i)}}}\right] + \frac{\lambda}{2}\sum_{i=1}^{k}\sum_{j=0}^{n} \theta_{ij}^2 \qquad (24)$$

The optimized weights are attained through the SSO approach for updating the hybrid SAE framework.

3.4.2 Weight optimization with Sparrow search optimization

This optimization approach is motivated by the foraging and the anti-predation habit of sparrows. This comprises one detector and one connector process. Here, the initial detector is the head of the population, and the connector follows the population guide to reach the data. The population head finds the priority of the optimal data, and it is described in the subsequent Eq. (25):

$$Y_{l,m}^{\bar{T}+1} = \begin{cases} Y_{l,m}^{\bar{T}} * Exp\left(\frac{-l}{\beta * \bar{I}_{max2}}\right) & if \quad W < S_w \\ Y_{l,m}^{\bar{T}} + A * Q & if \quad W \geq S_w \end{cases} \qquad (25)$$

where, $\bar{T}$ signifies the number of current iterations, $\bar{I}_{max}$ signifies the maximum number of iterations, $Y_{l,m}^{\bar{T}}$ signifies the position of lth data with dimension (m), the arbitrary number β range is between 0 and 1. A signifies the arbitrary number follows the normal distribution, Q signifies the matrix with all 1 components, and the dimensionality is $1 * \bar{d}$. W signifies the prior warning rate with the range of [0, 1], and S_w signifies the

safety esteem in the range of [0.5, 1]. If $W < S_w$, then there is no predator identified in the searching area. Similarly, if the condition of $W \geq S$ is satisfied, then the predator is found in that area, and this information is transmitted to others with the prior warning factor. Then, this particular position is further updated as per the subsequent Eq. (26):

$$Y_{l,m}^{\bar{T}+1} = \begin{cases} A * Exp\left(\frac{Y_{worst}^{\bar{T}} - Y_{l,m}^{\bar{T}}}{l^2}\right) & if \quad l < {}^m\!/_2 \\ Y_{best}^{\bar{T}+1} + \left|Y_{l,m}^{\bar{T}} - Y_{l,m}^{\bar{T}+1}\right| * B^* * M & else \end{cases} \qquad (26)$$

where, $Y_{worst}^{\bar{T}}$ signifies the current worst position in global, B^* signifies that the matrix $1 * \bar{d}$ comprises arbitrarily considered values in the range of -1 to 1. $B^* = BT(BBT) - 1$ and $l < \frac{m}{2}$ signifies that due to lower fitness, l^{th} the connector has no optimal data and needs to search in another updated position. If any unnecessary data is found, then the data group of the population performs the predation behavior, and their position is updated through the subsequent Eq. (27):

$$Y_{l,m}^{\bar{T}+1} = \begin{cases} Y_{best}^{\bar{T}+1} + \gamma * \left|Y_{l,m}^{\bar{T}} - Y_{best}^{\bar{T}+1}\right| & if \quad \bar{F}_l > \bar{F}_g \\ Y_{l,m}^{\bar{T}} + P * \left(\frac{\left|Y_{l,m}^{\bar{T}} - Y_{best}^{\bar{T}+1}\right|}{\left(\bar{F}_l - \bar{F}_w\right) + \bar{e}}\right) & if \quad \bar{F}_l = \bar{F}_g \end{cases} \qquad (27)$$

where, γ signifies the step length control. This is a normal distribution arbitrary number with an average value of 0 and a variance of 1, $P \in [-1,1]$ represents the arbitrary number of the direction of data selection, and this signifies the step length control variable. $\bar{e}$ signifies the minimum constant to avoid the denominator going to zero, $\bar{F}_l$ signifies the fitness value of lth data, $\bar{F}_g$, $\bar{F}_w$ signifies the best and worst fitness values correspondingly. If $\bar{F}_l > \bar{F}_g$ signifies the data at the edge of the population and it is easily affected by the predator, $\bar{F}_l = \bar{F}$ signifies the data in the centre of the population, and it is considered the data is safe. This optimization approach is utilized to update the optimized weights in the hybrid stacked autoencoder framework. After the update of optimized weights, the encoder operation is improved by minimizing the loss, and it is described in the subsequent Eq. (28):

$$\phi(\Theta) = \arg\min_{\theta, \theta''} \frac{1}{M}\sum_{k=1}^{m} l\left(y^k, \tilde{y}^k\right) \qquad (28)$$

where, l signifies the loss function $l(y, \tilde{y}) = \|y - \tilde{y}\|^2$. The architecture of the SAE framework comprises of M number of encoders with M a number of hidden layers through unsupervised learning of layers. Further, weight optimization uses the Sparrow search optimization approach. The presented SAE framework comprised subsequent steps. Initially, the input autoencoder layer attains the feature vectors, and these learned feature data are utilized in the subsequent layers for further processing.

4. RESULTS AND DISCUSSION

This section analyses the experimental results with the presented hybrid stacked autoencoder. The presented work is implemented in the PYTHON platform. Moreover, the dataset

considered for the analysis of the presented approach is the PPG signal figshare dataset. The experimental results are compared with the different existing approaches in terms of accuracy, sensitivity, specificity, F-measure, MAE, precision, and RMSE. The hyperparameter details for the proposed algorithm are shown in Table 1.

Table 1. Hyperparameter details

Method	Hyperparameter	Values
	Batch size	16
	Encoder unit	100, 70, 50, 30
	Decoder unit	50, 70, 100
	Activation	ReLU
SAE	Output activation	softmax
	Learning rate	0.001
	L1/L2 regularization	0.001
	Dropout	0.1
	Epochs	100
	Num-dimension	5
	Max-iter	100
	Num-droplet	50
	Hy	-1
ROA	Hx	-1
	w	0.5
	C1	1
	C2	2
	K	5
	Population size	30
	Max-iteration	100
SSO	Proportion of procedures	0.2
	Safety threshold	0.1
	Safety threshold constant	0.8

4.1 Dataset description

PPG signal figshare dataset (https://figshare.com/articles/dataset/PPG-BP_Database_zip/5459299) [30]:

It is a health database for the prediction of CVD. The 19 volunteers, among whom 5 are females and 14 males, whose age ranges from 26.15± 8.69 years (males) and 18-56 years (females), from different backgrounds, are participating in the data collection process. This dataset comprised 657 information records. Moreover, the information records are comprised of diabetes and hypertension. The presented methodology utilized the BP data, glucose level information, and the heart rate of the individuals.

The dataset used for processing contains 219 waveform folders with 3 files having a 2nd length infrared PPG signal txt file, recorded files with physiological data, and 1 table file. The PPG signal comprised 2.1-second length 12-bit AD raw value data. Three segments of waveform data are separately available in ID_1, ID_2, and ID_3 text files. The "PPG-BP database.xlsx" table file includes the details about the disease and the physiological information of subjects. Information record contains disease, diastolic pressure, Body mass index (BMI), systolic pressure, sex, age, weight, ID, weight, height, and heart rate records. Before aggregating the patients' records, the information is subjected to initial screening, and then the waveform quality evaluation is performed to generate a high-quality dataset.

The majority of individuals in the Figshare PPG dataset are young-to-middle-aged people (mean age 34.6±9.2 years; range 18-65), with 38% being female and 62% being male. Only a small percentage of respondents had a diagnosis of CVD; about 74% of subjects were categorized as healthy. High-risk patients and the elderly (over 70) are underrepresented.

4.2 Performance metrics

This section analyses performance metrics like accuracy, F-measure, sensitivity, specificity, precision, MAE, and RMSE [28] with the existing approaches. These different performance metrics are described in the subsequent sub-sections.

4.2.1 Accuracy
This measure is utilized to predict the percentage of accuracy. It is evaluated through the subsequent Eq. (29):

$$\bar{A}_Y = \frac{\bar{t}_{(positive)} + \bar{t}_{(negative)}}{\bar{t}_{(positive)} + \bar{t}_{(negative)} + \bar{f}_{(positive)} + \bar{f}_{(negative)}} \tag{29}$$

where, $\bar{t}_{(positive)}$ signifies the true positive, $\bar{t}_{(negative)}$ signifies the true negative, $\bar{f}_{(positive)}$ signifies the false positive, and $\bar{f}_{(negative)}$ signifies the false negative.

4.2.2 Sensitivity ($\bar{S}_Y$) or recall
This measure computes the accurate prediction of positive classes in the total considered data. This measure is also known as recall. This measure is estimated through the subsequent Eq. (30):

$$\bar{S}_Y = \frac{\bar{t}_{(positive)}}{\bar{t}_{(positive)} + \bar{f}_{(negative)}} \tag{30}$$

4.2.3 Specificity ($\bar{S}_P$)
This measure evaluates the accurately predicted negative classes amongst the total data. It is estimated through the subsequent Eq. (31):

$$\bar{S}_P = \frac{\bar{t}_{(negative)}}{\bar{t}_{(negative)} + \bar{f}_{(positive)}} \tag{31}$$

4.2.4 F-measure
This measure computes the harmonic mean of precision and sensitivity. The enhanced performance of this estimation provides a better balance in the prediction of sensitivity and precision estimations. It is estimated through the subsequent Eq. (32):

$$\bar{F}_{measure} = 2 * \frac{\bar{S}_Y * \bar{P}_N}{\bar{S}_Y + \bar{P}_N} \tag{32}$$

4.2.5 Precision
This measure evaluates the number of positives predicted accurately from the total number of positives. It is computed through the subsequent Eq. (33):

$$\bar{P}_N = \frac{\bar{t}_{(positive)}}{\bar{t}_{(positive)} + \bar{f}_{(positive)}} \tag{33}$$

4.2.6 Root mean squared error (RMSE)

This performance measure is used to predict the variation between the predicted and the original values. The reduced value of this error proved that the presented approach attains enhanced performance than the existing works. It is assessed through the Eq. (34):

$$\overline{RMSE} = \sqrt{\frac{\sum_{j=1}^{k}|b_j - c_j|^2}{k}} \qquad (34)$$

where, $\overline{RMSE}$ signifies the RMSE measure, b_j signifies the original data, and c_j signifies the predicted data.

4.2.7 Mean absolute error (MAE)

The MAE performance metric computes the mean variation of the absolute values in the entire original and the predicted data. This measure evaluates the predicted output error value. This performance computation is described in the subsequent Eq. (35):

$$\overline{MAE} = \frac{1}{K}\sum_{j=1}^{k}|b_j - c_j| \qquad (35)$$

where, $\bar{M}_{AE}$ signifies the mean absolute error. This minimized error value proved that the presented approach is effective in accurate prediction.

4.3 Performance analysis

In this section, the performance of the presented methodology is compared with the different existing methodologies. The developed confusion matrix of the presented approach in predicting CVD risk factors is depicted in Figure 4.

Figure 4 shows the generated confusion matrix with three different classes: normal, hypertension, and pre-hypertension.

The performance metrics of the proposed are examined with the different existing approaches like ABETN-Attention based Bidirectional Encoder Transformer Network (94.2%), DDLSA-Trans-Dual Decoder Locality Self Attention Based Transformer (84.8%), VIT-Vision Transformer (82.3%), GLST-Global Local Synchronous Transformer (83.2%), CNN -convolution neural network (83.3%), and k-nearest neighbor (80.3%) [35, 36]. It proves that the presented approach attained enhanced accuracy (97.05%) than the existing methods. Moreover, the performance examination of the presented approach in terms of accuracy performance is depicted in Figure 5.

In Figure 5, the comparative accuracy analysis is provided with existing methodologies like ABETN, DDLSA, VIT, GLST, CNN, and kNN. This proved that the presented approach attains enhanced accuracy (97.05%) performance. The performance examination of various performance metrics for three different output classes is compared with the existing approaches. The performance comparisons on accuracy with every three classes are depicted in Figure 6.

In Figure 6, the performance of the presented approach is examined with the predicted three different classes separately. This proved that the presented approach attained enhanced performance over the existing KNN approach [28] in terms of different performance metrics in all the classes.

Figure 7 depicts the performance comparison on sensitivity. The presented approach attains enhanced performance in terms of sensitivity (97%) than the existing approaches. This proved that the presented approach attains a significant improvement over the existing approaches like ABETN (82.3%), DDLSA (85%), VIT (78.5%), GLST (78.2%), CNN (86.3%), and kNN (78.2%).
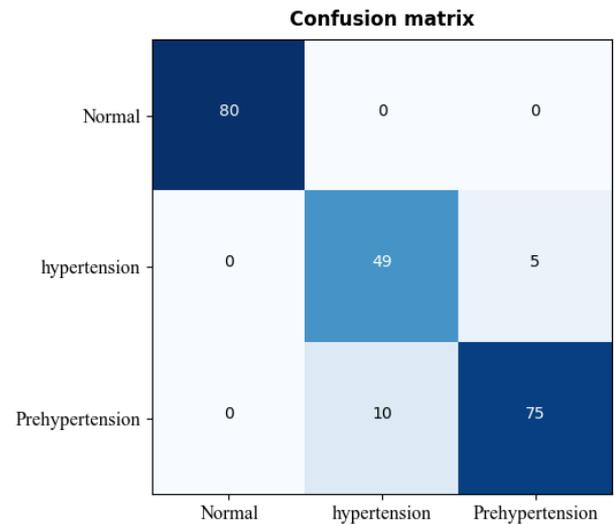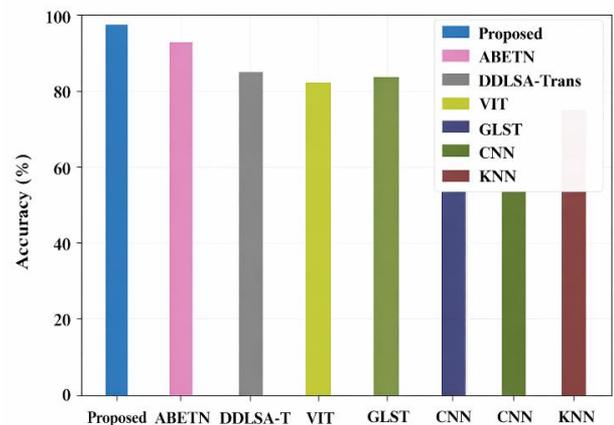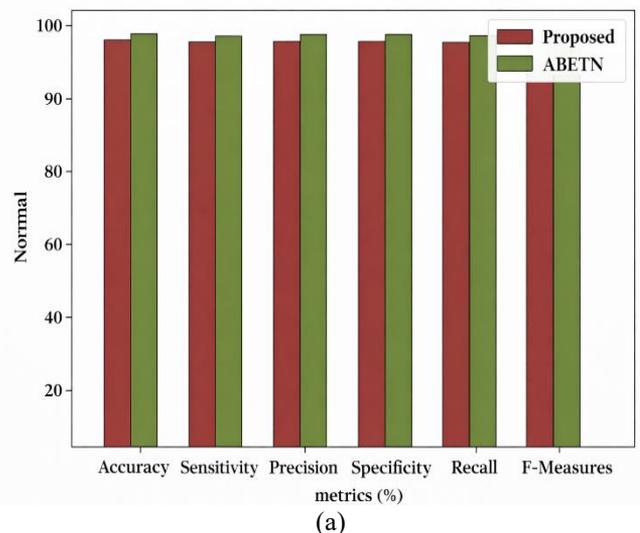

Figure 4. Confusion matrix

Figure 5. Comparison analysis of the accuracy
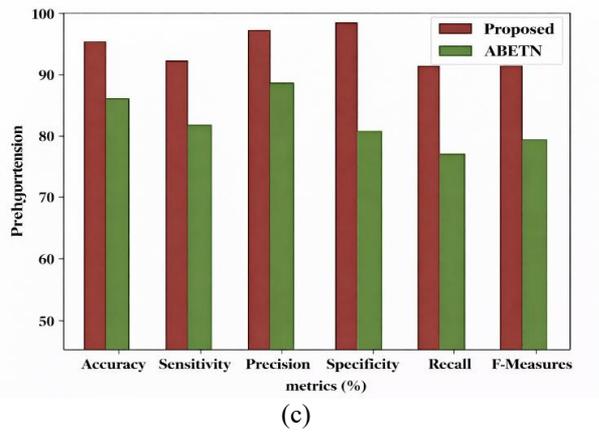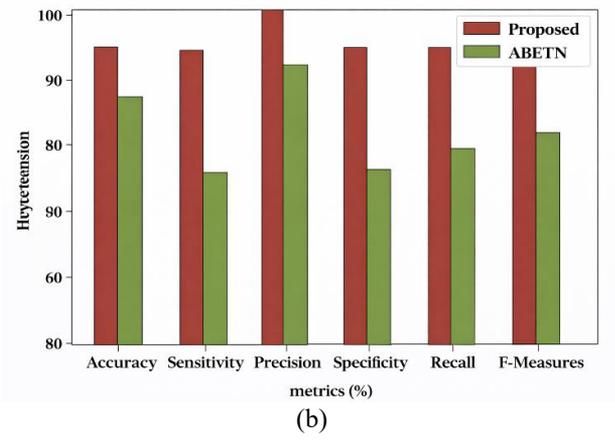

(a)

(b)

(c)

Figure 6. Performance analysis of classes: (a) normal, (b) hypertension, and (c) pre-hypertension
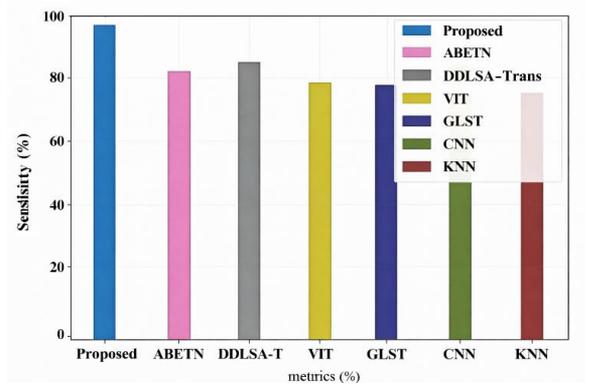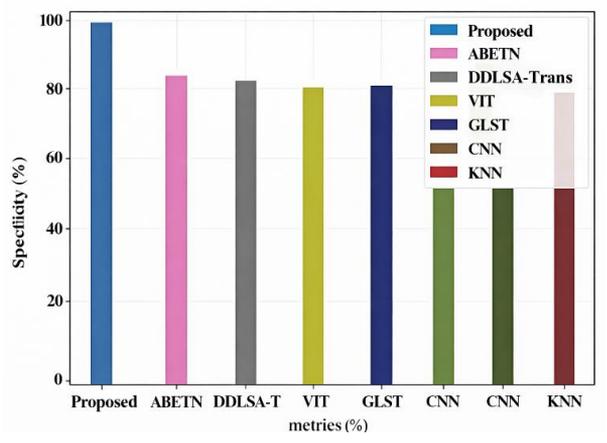

Figure 7. Comparison analysis on sensitivity

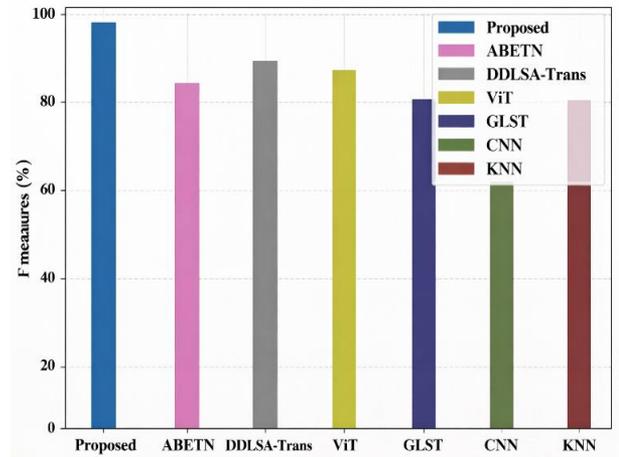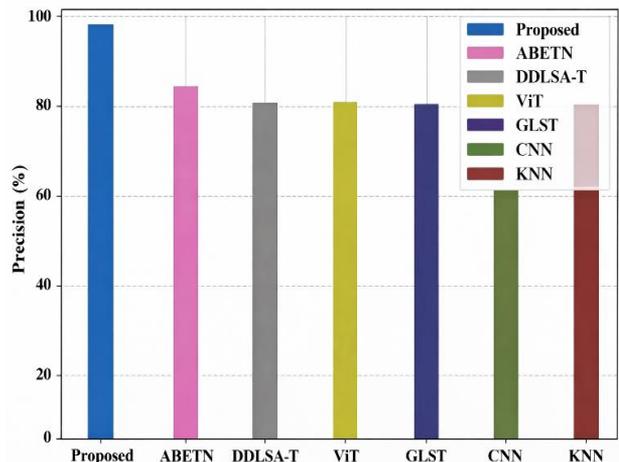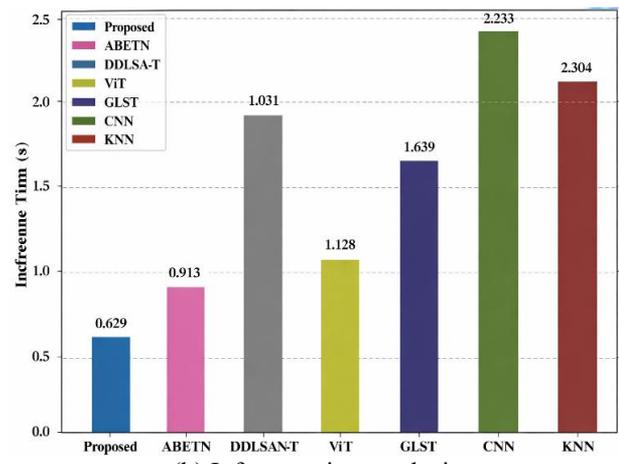**Figure 8.** Comparison analysis on specificity

Figure 9. Comparison analysis on F-measure

(a) Precision

(b) Inference time analysis

Figure 10. Comparison analysis of precision and inference time

In Figure 8, the performance on specificity is illustrated. The performance of the presented approach in terms of specificity is 97%. This clearly shows that the approach presented achieves improved performance than the various existing approaches, such as ABETN (83.1%), DDLSA (77.7%), VIT (78.9%), GLST (78.7%), CNN (82.3%), and kNN (78.7%). This demonstrated that the performance of the presented approach attained an improved performance than the different existing methodologies. The presented approach

attained a performance analysis of F-measure is 97%. The F-measure performance of the presented and the existing approaches is ABETN (82.95%), DDLSA (90.8%), VIT (88.22%), GLST (79.71%), CNN (79.11%), and kNN (78.03%). Moreover, this proved that the presented approach obtained enhanced performance over the other mentioned approaches. Moreover, the performance comparison on F-measure is depicted in Figure 9.

Figure 9 examines the F-measure performance with the various existing approaches. This demonstrates that the presented approach attains a significant improvement compared to the existing methodologies [29]. Here, the attained performance of the presented approach in regard to precision is 97%. The precision performances on existing approaches are ABETN (82.8%), DDLSA (77.9%), VIT (79.5%), GLST (77.5%), CNN (82.8%), and kNN (79.6%). Moreover, the performance analysis on precision and inference time is depicted in Figure 10.
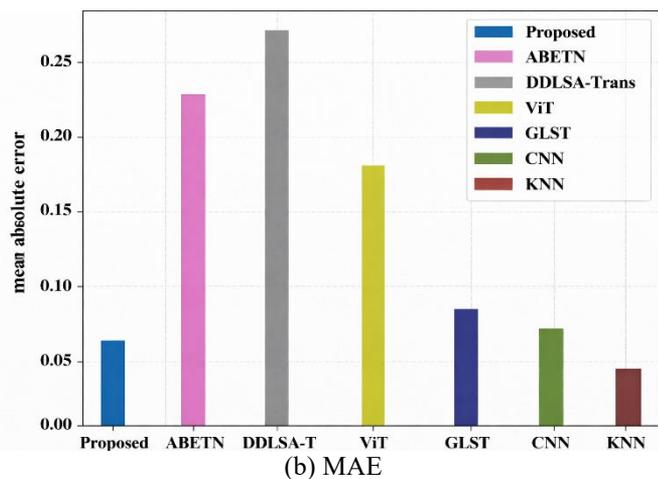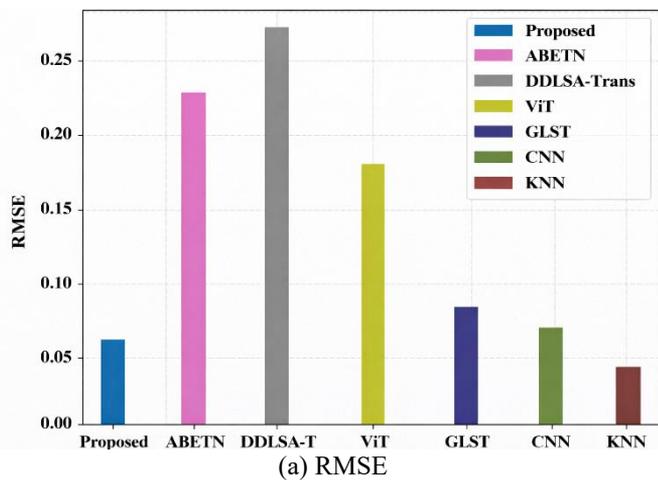

(a) RMSE

(b) MAE

Figure 11. RMSE and MAE comparison

In Figure 10, the comparison of precision performance is illustrated. This figure demonstrates that the presented approach attains an improved precision (97%) performance than the existing methodologies. Figure 11 depicts the RMSE and MAE comparison.

The RMSE performance is compared with the different existing approaches. This demonstrates that the presented approach significantly improved RMSE (0.05). The RMSE value is significantly decreased than the existing methodologies like ABETN (0.235), DDLSA (0.29), VIT (0.187), GLST (0.08), CNN (0.068), and kNN (0.048). The

MAE performance is examined with the various existing approaches. The performance of the presented methodology on the MAE is 0.0456. But the existing approaches attained a higher error value than the presented performance. The MAE performances on other existing approaches are ABETN (0.22), DDLSA (0.275), VIT (0.187), GLST (0.069), CNN (0.062), and kNN (0.047). This proved that the presented approach significantly improved over the existing methodologies.
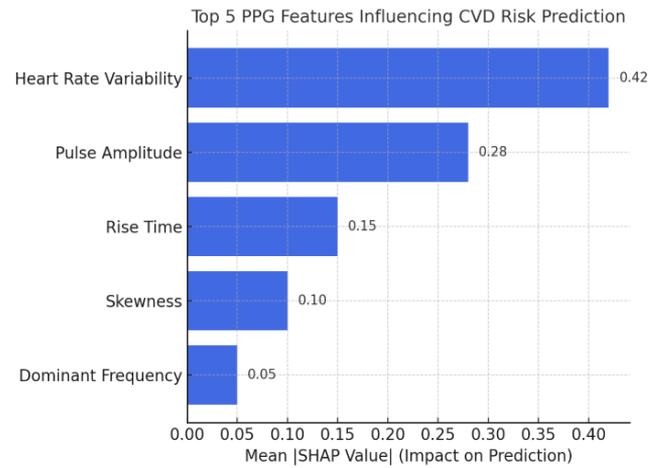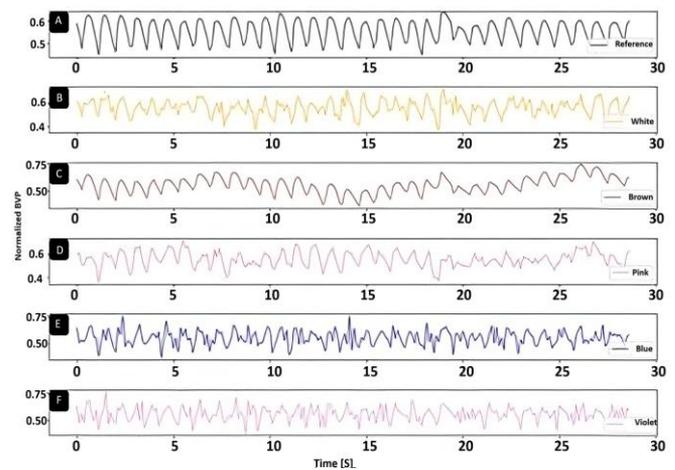

Figure 12. SHAP analysis based feature rating

Figure 13. Signal output for noised and denoised samples
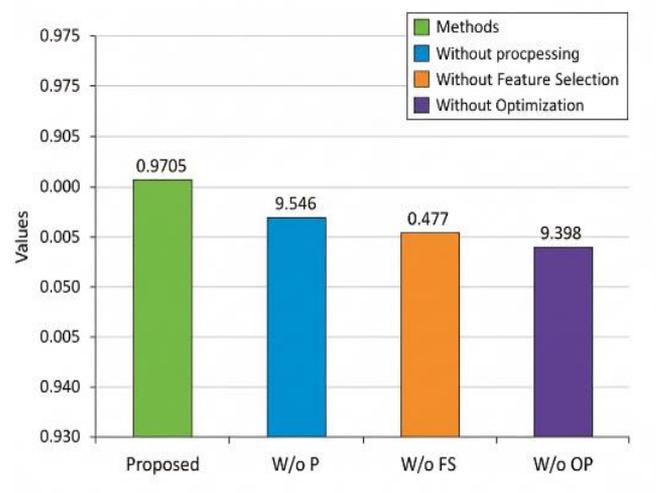

Figure 14. Ablation study output

The features that are high rated based on SHAP interpretation are shown in Figure 12. The proposed framework supplemented the hybrid framework with patient-level explanation cards and SHAP-based feature importance analysis to improve clinical interpretability. These analyses showed that the most important predictors of the model, including pulse amplitude and Heart rate variability (HRV) measurements, match clinically recognized cardiovascular indicators, increasing confidence and the likelihood that medical professionals will embrace the model. The model's robustness under varying noise levels or data distribution conditions is shown in Figure 13. This is to assess the practical applicability of the proposed model in abnormal conditions (e.g., high noise or incomplete data). The overall performance achieved by the proposed model with noise is accuracy-95.6%, sensitivity-95.434%, specificity-95.421%, precision-95.492%, f-measures-95.425%, MAE-0.161, RMSE-0.211, and inference time-0.2223.

The performance comparison between the proposed and existing methods is shown in Table 2. The comparative analysis has shown that the proposed algorithm has achieved a lower error rate performance and higher accuracy than other existing algorithms. This is because the proposed approach has used an efficient SSO algorithm for weight parameter selection. SSO can obtain optimal parameters over a large area. Due to this, the error rate of classification gets reduced. Finally, efficient accuracy is achieved with a lower error rate during CVD detection.

The ablation study output is shown in Figure 14. The proposed model's predictions might not apply to older populations or individuals with several comorbidities due to the dataset's narrow age range and lack of health diversity. In a similar vein, biases specific to particular devices (such as fingertip versus wrist PPG) were not adequately reflected, which might have limited their use. Future research will include subgroup performance analysis, external validation cohorts, and fairness-aware training techniques to reduce these biases and guarantee that the model's predictions hold true for various demographic groups.

Table 2. Performance comparison between proposed and existing methods

Ref no	Technique	Dataset	Performance Metric
Proposed	HSAE-SSO	PPG–BP figshare database	Accuracy (95.43%), specificity (96.73%), precision (92.26%), MAE (0.0456), and RMSE (0.0523)
Panwar et al. [23]	LRCN	UCI repository database (MIMIC-II (Multi-parameter Intelligent Monitoring in Intensive Care II) database)	MAE-2.32±0.095, NMAE-0.059, NRMSE-0.009
Chowdhury et al. [24]	Gaussian process regression (GPR)	PPG-BP database	SBP-MAE (3.02), MSE (45.49), RMSE (6.74), DBP-MAE (1.74), MSE (12.89), RMSE (3.59)
Esmaelpoor et al. [25]	DNN	MIMIC-II database	Mean error-+1.91±5.55mmHg (SBP), and +0.67±2.84mmHg (DBP)
Tjahjadi and Ramli [26]	KNN	PPG–BP figshare database	Accuracy-86.7%, F1-score-90.8%
Zhang et al. [27]	GSVM, BT, KNN	Real-time data	Accuracy-71.27% (Blood glucose), 90.57% (invalid sample)
Sinha et al. [31]	SVM	UCI repository database	Accuracy-92.31%
Ismail et al. [32]	CRNN	IEEE signal processing cup dataset	error rate-3.8±2.3 bpm (subject-independent) and 2.41±2.90 bpm (subject-dependent)
Anny Leema and Jothiaruna [33]	ResNet-50, DenseNet-161, and VGG-16	ECG Images Dataset (EID)	Accuracy-93.33%, precision-94.23%, F1-score-93.3%
Karri et al. [34]	LSTM-RNN	MIT-BIH arrhythmia database	positive predictivity-96.5%, sensitivity-93.87%, and F1-score-95.18%
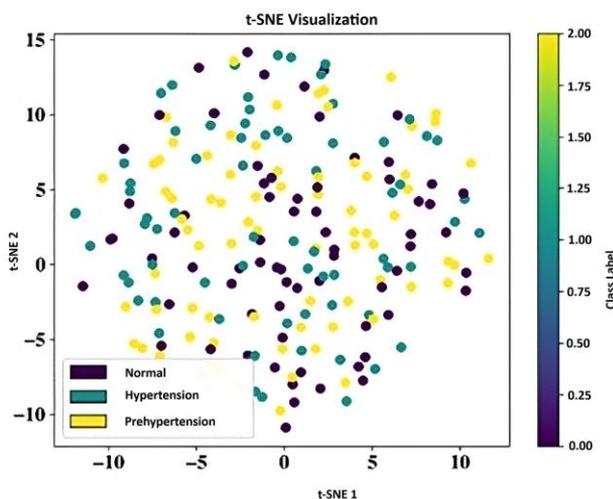
Figure 15. t-SNE plot

The t-SNE plot for used PPG dataset used is shown in Figure 15. The accuracy and inference time obtained by the hybrid SAE and SAE are shown in Figure 16. Moreover, the SAE is hybrid with PSO and GA, then the obtained results are also illustrated in Figure 16. The Sparrow Search Optimization (SSO) is integrated directly with the autoencoder's latent space. This coupling reduces irrelevant/collinear features with fewer dimensions retained compared to PCA or GA-based feature selection baselines. By searching over compressed representations rather than raw features, convergence is faster and less prone to local minima. SSO requires fewer iterations than classical optimizers. Its adaptive exploration–exploitation balance yields more stable hyperparameter tuning (learning rate, hidden size) in the hybrid autoencoder.

The proposed HSAE-SSO achieves 95.43% higher accuracy than benchmark models. Importantly, the framework improves specificity for high-risk patients, which is clinically more valuable than overall accuracy alone. By combining unsupervised representation learning (SAE) with evolutionary optimization (SSO), the model highlights which compressed health indicators most influence CVD risk. This hybrid design is not just a technical stacking, but addresses a real medical challenge: balancing predictive accuracy with transparent, feature-driven insights.
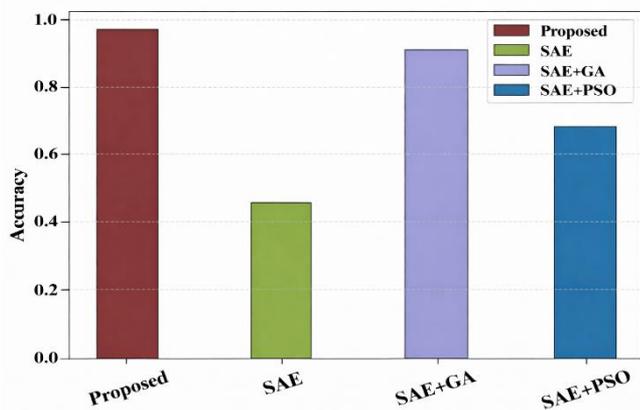
The inference time, energy efficiency, and cost efficiency

analysis for the proposed model are discussed in Table 3. On an NVIDIA RTX 3090 (PyTorch 2.0), the inference delay, computational cost, and energy consumption are assessed using the suggested framework. A total of 1000 timed inferences and 200 warm-up iterations are conducted for each model version, yielding 96.2 ± 0.05 latency and energy per inference (0.00264 ± 0.00005). Model weights were used to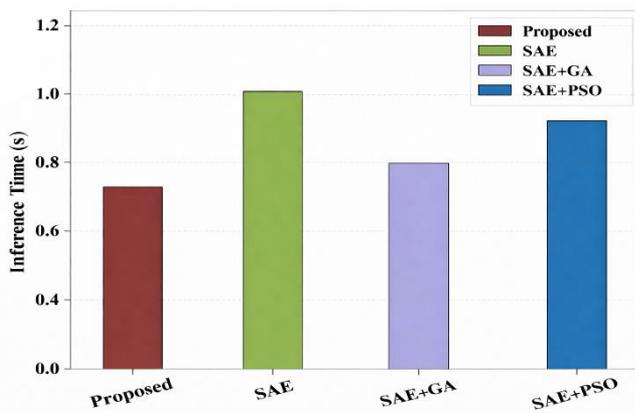 calculate parameter counts, and FLOPs were measured using thop. For validation, energy measurement used an external power meter in conjunction with NVML sampling. It also profiled improved variants (INT8 TensorRT and FP16 mixed precision). The TensorRT INT8 variation delivers 2.5× greater throughput at a minor AUC reduction ($\Delta AUC = 0.8\%$), while the entire HSAE-SSO model achieves AUC 95.6% with latency 18.4ms (batch = 1) and energy 1.70 mJ per inference, according to the results (Table 3).

Table 3. Inference time and energy efficiency analysis

Model Variant	Params (M)	Model Size (MB)	FLOPs (G)	Throughput (Samp/s)	Mean Power (W) ± std	Energy Efficiency (J) ± std	AUC (%)
HSAE-SSO (full)	4.1	15.6	11.9	64	93 ± 5	0.00264 ± 0.00005	96.2
HSAE-SSO (FP16)	4.1	7.6	11.9	89	88 ± 2.9	0.00079 ± 0.00009	93.9
HSAE-SSO (int8 TRT)	3.9	5.2	11.9	2429	85 ± 3.5	0.00030 ± 0.00001	94.5
SAE baseline	2.8	9.8	7.6	111	79 ± 2	0.00076 ± 0.00002	92.2

(a) Accuracy

(b) Inference time

Figure 16. Accuracy and inference time for proposed hybrid and solitary models

methodology is utilized for selecting the features optimally. Finally, a hybrid stacked autoencoder framework is utilized to accurately predict CVD conditions like normal, pre-hypertension, and hypertension. Here, this prediction performance is further improved with the SSO methodology. The SSO methodology is utilized to update the optimized weights in the HSAE framework. The experimental results of the presented approach are analyzed with the different performance metrics. Moreover, the performance of the presented approach achieved significant enhancement in terms of accuracy (95.43%), F-measure (92.62%), precision (92.26%), MAE (0.0456), and RMSE (0.0523) than the different existing approaches.

Compared to low-risk/healthy individuals, high-risk CVD cases are generally underrepresented in large-scale clinical databases. The model may be biased to anticipate "low risk" as a result. Therefore, to artificially enhance minority (high-risk) classes by avoiding class imbalance in the future, employ sophisticated resampling techniques like Adaptive Synthetic (ADASYN), Synthetic Minority Oversampling Technique (SMOTE) algorithm, or generative adversarial networks (GANs).

Variations in sample rate, noise, wavelength, and calibration of PPG signals obtained from various devices (such as fingertip sensors, hospital monitors, and smartwatches) may limit their generalizability. Device variability will be lessened, though, if domain adaptation strategies (such as CORAL and Domain-Adversarial Neural Networks) are used to align feature distributions across devices. In order to do this, the suggested model will use domain adaptation techniques, synthetic augmentation, and cost-sensitive loss functions. Future multi-device cross-validation will also be carried out. When applied to bigger, actual clinical datasets, these procedures will guarantee the suggested framework's scalability and resilience.

5. CONCLUSION

This paper effectively predicted CVD risk factors using the hybrid stacked autoencoder framework. At first, input PPG signal data is taken and pre-processed with an effective FNLMS filtering methodology. Afterwards, an effective set of features, like statistical features with spectral entropy features, demographic features, and Gaussian features, was extracted from the pre-processed data. Subsequently, the RO

REFERENCES

[1] Kim, Y.J., Jang, H.S., Byun, C.S., Choi, B.S. (2018). Development of u-health monitoring system using PPG sensor. In 2018 International Conference on Electronics, Information, and Communication (ICEIC), Honolulu, USA, pp. 1-2. https://doi.org/10.23919/ELINFOCOM.2018.8330688

[2] Reddy, G.N.K., Manikandan, M.S., Murty, N.N. (2020).

On-device integrated PPG quality assessment and sensor disconnection/saturation detection system for IoT health monitoring. IEEE Transactions on Instrumentation and Measurement, 69(9): 6351-6361. https://doi.org/10.1109/TIM.2020.2971132

[3] Prabhakar, S.K., Rajaguru, H., Lee, S.W. (2019). Metaheuristic-based dimensionality reduction and classification analysis of PPG signals for interpreting cardiovascular disease. IEEE Access, 7: 165181-165206. https://doi.org/10.1109/ACCESS.2019.2950220

[4] Liu, W., Fang, X., Chen, Q., Li, Y., Li, T. (2018). Reliability analysis of an integrated device of ECG, PPG and pressure pulse wave for cardiovascular disease. Microelectronics Reliability, 87: 183-187. https://doi.org/10.1016/j.microrel.2018.06.008

[5] Mousavi, S.S., Firouzmand, M., Charmi, M., Hemmati, M., Moghadam, M., Ghorbani, Y. (2019). Blood pressure estimation from appropriate and inappropriate PPG signals using a whole-based method. Biomedical Signal Processing and Control, 47: 196-206. https://doi.org/10.1016/j.bspc.2018.08.022

[6] Riaz, F., Azad, M.A., Arshad, J., Imran, M., Hassan, A., Rehman, S. (2019). Pervasive blood pressure monitoring using Photoplethysmogram (PPG) sensor. Future Generation Computer Systems, 98: 120-130. https://doi.org/10.1016/j.future.2019.02.032

[7] Reddy, V.R., Choudhury, A.D., Jayaraman, S., Thokala, N.K., Deshpande, P., Kaliaperumal, V. (2017). PerDMCS: Weighted fusion of PPG signal features for robust and efficient diabetes mellitus classification. In Special Session on Smart Medical Devices-From Lab to Clinical Practice, Scitepress, 6: 553-560. https://doi.org/10.5220/0006297205530560

[8] Golap, M.A.U., Raju, S.T.U., Haque, M.R., Hashem, M.M.A. (2021). Hemoglobin and glucose level estimation from PPG characteristics features of fingertip video using MGGP-based model. Biomedical Signal Processing and Control, 67: 102478. https://doi.org/10.1016/j.bspc.2021.102478

[9] Islam, T.T., Ahmed, M.S., Hassanuzzaman, M., Bin Amir, S.A., Rahman, T. (2021). Blood glucose level regression for smartphone ppg signals using machine learning. Applied Sciences, 11(2): 618. https://doi.org/10.3390/app11020618

[10] Habbu, S., Dale, M., Ghongade, R. (2019). Estimation of blood glucose by non-invasive method using photoplethysmography. Sādhanā, 44(6): 135. https://doi.org/10.1007/s12046-019-1118-9

[11] Tsai, C.W., Li, C.H., Lam, R.W.K., Li, C.K., Ho, S. (2019). Diabetes care in motion: Blood glucose estimation using wearable devices. IEEE Consumer Electronics Magazine, 9(1): 30-34. https://doi.org/10.1109/MCE.2019.2941461

[12] Islam, M., Biswas, T., Saad, A.M., Haque, C.A., Salah Uddin Yusuf, M. (2019). A non-invasive heart rate estimation approach from photoplethysmography. In Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018, Singapore: Springer Nature Singapore, pp. 383-394. https://doi.org/10.1007/978-981-13-7564-4_33

[13] Rastegar, S., GholamHosseini, H., Lowe, A. (2020). Non-invasive continuous blood pressure monitoring systems: Current and proposed technology issues and challenges. Physical and Engineering Sciences in Medicine, 43(1): 11-28. https://doi.org/10.1007/s13246-019-00813-x

[14] Deepakfranklin, P., Krishnamoorthi, M, Kalamani, M. (2018). Contact and non-contact methods of photo plethysmography. International Journal of Engineering and Advanced Technology, 8(2): 378-383.

[15] Deepakfranklin, P., Krishnamoorthi, M., Hemanand, D., Chembian, W.T., Kumar, G.M., Jayalakshmi, D.S. (2021). Survey on methods of obtaining biomedical parameters from PPG signal. Turkish Journal of Computer and Mathematics Education, 12(10): 2684-2692.

[16] Priyadarshini, R.G., Kalimuthu, M., Nikesh, S., Bhuvaneshwari, M. (2021). Review of PPG signal using machine learning algorithms for blood pressure and glucose estimation. In IOP Conference Series: Materials Science and Engineering, IOP Publishing, 1084(1): 012031. https://doi.org/10.1088/1757-899X/1084/1/012031

[17] Shin, H., Min, S.D. (2017). Feasibility study for the non-invasive blood pressure estimation based on ppg morphology: Normotensive subject study. Biomedical Engineering Online, 16(1): 10. https://doi.org/10.1186/s12938-016-0302-y

[18] Pankaj, Kumar, A., Komaragiri, R., Kumar, M. (2022). A review on computation methods used in photoplethysmography signal analysis for heart rate estimation. Archives of Computational Methods in Engineering, 29(2): 921-940. https://doi.org/10.1007/s11831-021-09597-4

[19] Wu, C.M., Chuang, C.Y., Chen, Y.J., Chen, S.C. (2016). A new estimate technology of non-invasive continuous blood pressure measurement based on electrocardiograph. Advances in Mechanical Engineering, 8(6): 1687814016653689. https://doi.org/10.1177/1687814016653689

[20] Lo, F.P.W., Li, C.X.T., Wang, J., Cheng, J., Meng, M.Q.H. (2017). Continuous systolic and diastolic blood pressure estimation utilizing long short-term memory network. In 2017 39th Annual International Conference of The IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea (South), pp. 1853-1856. https://doi.org/10.1109/EMBC.2017.8037207

[21] Tarvirdizadeh, B., Golgouneh, A., Tajdari, F., Khodabakhshi, E. (2020). A novel online method for identifying motion artifact and photoplethysmography signal reconstruction using artificial neural networks and adaptive neuro-fuzzy inference system. Neural Computing and Applications, 32(8): 3549-3566. https://doi.org/10.1007/s00521-018-3767-8

[22] Paviglianiti, A., Randazzo, V., Villata, S., Cirrincione, G., Pasero, E. (2022). A comparison of deep learning techniques for arterial blood pressure prediction. Cognitive Computation, 14(5): 1689-1710. https://doi.org/10.1007/s12559-021-09910-0

[23] Panwar, M., Gautam, A., Biswas, D., Acharyya, A. (2020). PP-Net: A deep learning framework for PPG-based blood pressure and heart rate estimation. IEEE Sensors Journal, 20(17): 10000-10011. https://doi.org/10.1109/JSEN.2020.2990864

[24] Chowdhury, M.H., Shuzan, M.N.I., Chowdhury, M.E., Mahbub, Z.B., Uddin, M.M., Khandakar, A., Reaz, M.B.I. (2020). Estimating blood pressure from the photoplethysmogram signal and demographic features

using machine learning techniques. Sensors, 20(11): 3127. https://doi.org/10.3390/s20113127

[25] Esmaelpoor, J., Moradi, M.H., Kadkhodamohammadi, A. (2020). A multistage deep neural network model for blood pressure estimation using photoplethysmogram signals. Computers in Biology and Medicine, 120: 103719. https://doi.org/10.1016/j.compbiomed.2020.103719

[26] Tjahjadi, H., Ramli, K. (2020). Non-invasive blood pressure classification based on photoplethysmography using k-nearest neighbors algorithm: A feasibility study. Information, 11(2): 93. https://doi.org/10.3390/info11020093

[27] Zhang, G., Mei, Z., Zhang, Y., Ma, X., Lo, B., Chen, D., Zhang, Y. (2020). A non-invasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning. IEEE Transactions on Industrial Informatics, 16(11): 7209-7218. https://doi.org/10.1109/TII.2020.2975222

[28] Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning. Computational Intelligence and Neuroscience, 2021(1): 8387680. https://doi.org/10.1155/2021/8387680

[29] Katarya, R., Meena, S.K. (2021). Machine learning techniques for heart disease prediction: A comparative study and analysis. Health and Technology, 11(1): 87-97. https://doi.org/10.1007/s12553-020-00505-7

[30] Elgendi, M., Galli, V., Ahmadizadeh, C., Menon, C. (2022). Dataset of psychological scales and physiological signals collected for anxiety assessment using a portable device. Data, 7(9): 132. https://doi.org/10.3390/data7090132

[31] Sinha, N., Jangid, T., Joshi, A.M., Mohanty, S.P. (2022). iCARDO: A machine learning based smart healthcare framework for cardiovascular disease prediction. arXiv Preprint arXiv: 2212.08022. https://doi.org/10.48550/arXiv.2212.08022

[32] Ismail, S., Siddiqi, I., Akram, U. (2022). Heart rate estimation in PPG signals using convolutional-recurrent regressor. Computers in Biology and Medicine, 145: 105470. https://doi.org/10.1016/j.compbiomed.2022.105470

[33] Anny Leema, A, Jothiaruna, N. (2023). A deep learning framework for automatic cardiovascular classification from electrocardiogram images. Research Square. 1-15. https://doi.org/10.21203/rs.3.rs-2413127/v1

[34] Karri, M., Annavarapu, C.S.R., Pedapenki, K.K. (2023). A real-time cardiac arrhythmia classification using hybrid combination of delta modulation, 1D-CNN and blended LSTM. Neural Processing Letters, 55(2): 1499-1526. https://doi.org/10.1007/s11063-022-10949-9

[35] Darbhasayanam, S.H., Mohammad, H.A.A., Komalla, A.R. (2024). Photoplethysmogram (PPG)-based blood pressure estimation using vision transformer networks: A deep learning approach. In 2024 3rd International Conference on Artificial Intelligence for Internet of Things (AIIoT), Vellore, India, pp. 1-6. https://doi.org/10.1109/AIIoT58432.2024.10574539

[36] Divya, N.J., Kumar, N.S., Devi, R.K. (2025). An optimized progressive attention-based bidirectional encoder enclosed transformer network-based cardiovascular disease detection. AI, Computer Science and Robotics Technology. 4(1): 1-32. https://doi.org/10.5772/acrt.20250009