Journal of Education for Pure Science
Vol.15, No.2 (2025)
DOI: https://doi.org/10.32792/jeps.v15i2.455
Website: jceps.utq.edu.iq
Email: jceps@eps.utq.edu.iq

Beyond Cuffs and Needles: Exploring Photoplethysmography-Based Machine Learning for Blood Pressure Estimation

Haifaa Dwaish Firas Sabar Miften

haifaadwaish.comp@utq.edu.iq firas@utq.edu.iq

College of Education for Pure Sciences, Computer science department, Thi-Qar University, IRAQ

Received 12 / 6 /2024, Accepted 14 / 8 /2024, Published 1 / 6 /2025

Abstract:

This study explores the use of machine learning techniques to classify blood pressure (BP) based on Photoplethysmography (PPG) signals. Three popular algorithms - decision trees (DT), random forests (RF), and support vector machines (SVM) - were evaluated and compared. The data preprocessing involved extracting relevant features from the PPG signals, including temporal, morphological, and frequency-domain characteristics. The RF model outperformed the DT and SVM models, achieving nearly 98% accuracy and F-score, demonstrating its ability to capture complex non-linear relationships between PPG features and BP. The RF model also showed robust performance in the presence of noise and variations in input PPG signals, making it a promising choice for real-world BP monitoring applications. The evaluation was performed using a publicly available dataset of simultaneous PPG and BP measurements.

Keywords: Photoplethysmography, Feature extraction, Blood Pressure, Classification, Machine learning

1-Introduction

Blood pressure (BP) is crucial for early detection of heart disease due to its association with symptoms of hypertension or hypotension [1]. BP mea- sures the power provided by the heart pump to artery walls when circulating throughout the body [2, 3] . A BP measurement includes three parameters: diastolic BP (DBP), systolic BP (SBP), and mean arterial pressure (MAP), all measured in millimeters of mercury (mmHg).

Blood pressure (BP) can be measured using either invasive or nonin- vasive procedures. Although invasive technologies can accurately and con- stantly measure blood pressure, they are inconvenient to use and can cause infections in patients [4]. Current noninvasive approaches, such as utiliz- ing a cuff, can be uncomfortable for those who are injured, overweight, or have recently given birth [5, 6]. A non-invasive optical method called pho- toplethysmography (PPG) used to measure volumetric variations in blood inside the microvascular bed of tissue. Some important cardiovascular sys- tem information is contained in BP measurements. and other important are contained in PPG signals.

Recent research concentrates on PPG signal-based blood pressure classi- fication as the subject of research since it offers a useful and cost-effective substitute for traditional cuff-based blood pressure monitoring [7, 8]. Ma- chine learning models can categorize an individual's blood pressure as normal.

pre-hypertensive, or hypertensive by analyzing the temporal characteristics and morphological features of the PPG waveform, can be developed [9]. One important consequence of the successful creation of a PPG-based blood pres- sure classification system may be the early identification and treatment of hypertension. It is a key risk factor for cardiovascular disease[10, 1]. Con- tinuous, non-invasive blood pressure monitoring could be made possible by this technology, which would enhance the capacity to identify and react to changes in a person's cardiovascular health.

Creating systems for blood pressure estimate or hypertension risk as- sessment faces multiple issues that needs to be taken into account. Firstly, to properly extract PPG morphological properties, a high-quality waveform captured at a high sample rate was required [9]. Furthermore, it is challeng- ing to extract morphological features since they are prone to drifts, artifacts, and noise [11]. That makes using deep learning models (DL) not acceptable since DL can't produce a high-performance model without a sizable sample set and significant processing power. Configuring deep learning (DL) re- sembles an artistic endeavor, as there is no exact approach for determining parameters like the quantity of neurons, layers, or learning rate, and they vary greatly depending on the specific task at hand [12]. DL employs black- box techniques, which restrict the use of ML or DL in medical applications due to the aforementioned [12, 13]. Finally, it is not possible to continuously monitor cardiovascular health when relying just on the patient's clinical or sociodemographic data.

In this paper, we propose a new classification method for blood pressure (BP) based on photoplethysmography (PPG) signals, utilizing a CNN architecture (specifically DenseNet201) for feature extraction. Our main contributions can be summarized as providing non-invasive monitoring with im- proved accuracy, noise detection and integration with healthcare systems, contributing to advancements in continuous and reliable BP monitoring.

2. ,Related Work

This section covers a few of the investigations and theoretical models that surround PPG-based blood pressure readings. In Figure 1, the existing approaches are categorized.

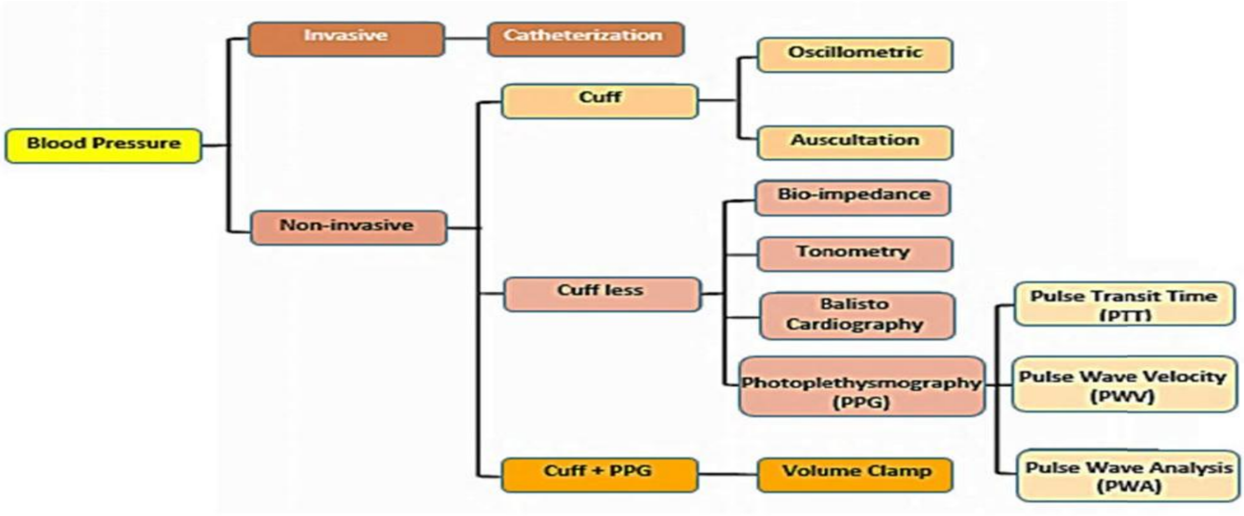

Figure 1:An overview of the body of research on BP measurement techniques[14]

In this paper, we will concentrate on the final category of non invanse type which is cutoff with PPG for volume clamp.Gang et al.[15] presents a two-stage multi-task learning network (ABPMTL) that uses electrocardiogram (ECG) and photoplethysmogram (PPG) signals as input to predict ABP waveforms. The Resnet18 combined domain adver- sarial network is trained to produce class labels in the first stage, which is a classification job. These class labels are considered the auxiliary input in the subsequent stage. The two branch duties in the second stage are the creation of ABP waveforms and the perdition of BP values. To achieve simultaneous preservation of specificity and hierarchical feature sharing between two tasks,a dual attention-based task consistency learning block (TCL) is presented. The suggested approach performs exceptionally well in both BP value prediction and ABP waveform creation, taking into account the correlation of characteristics across several BP tasks for the first time.

Hamza et al. [16] work investigates the novel combination of machine learning (ML) methods and photoplethysmography (PPG) signals, with a particular emphasis on the categorization of aberrant arterial pulse (AAP) patterns—a field that has not received much attention up to this point. We acknowledge the difficulties in this undertaking, chief among them being the dearth of clinically characterized AAP waveform datasets. This lack of avail- ability is due to the challenges in finding volunteers who display a range of disease-related AAPs and the inherent dangers involved with ABP measuring techniques. Moreover, the existing guidelines do not provide enough informa- tion about AAP features, which restricts the use of PPG and ML to identify ABP-related anomalies primarily in situations of hypertension and hypoten- sion. In order to close these gaps, the current work presents a PPG-based categorization system that makes use of the bagged trees (BT) and k-nearest neighbors (KNN) algorithms. These were chosen due to their ability to rep- resent intricate, nonlinear interactions at a lower level of complexity than alternatives such as Support Vector Machines (SVM) or Deep Neural Networks (DNN). Furthermore, new detectors have been created to detect important pulse wave characteristics as dicrotic notches and troughs, which are essential for PPG feature extraction and AAP pattern recognition. A modeling pro- cedure that makes use of diseased cases that are known to exhibit particular AAP patterns is part of the methodology. A comprehensive test with 1,120 PPG and ABP signals produced remarkable 90.9% and 91% accuracy rates for KNN and BT algorithms, respectively. Both algorithms demonstrated strong performance across 11 different classes, highlighting their potential as efficient AAP detectors.

Nassir et al. [17] uses STFT with various neural networks (Convolution Neural Network (CNN), Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (Bi-LSTM), and more to analyze PPG signals from over 200 patients (650+ signal samples) with hypertension. Two cate- gories have been classified: hypertension (which encompasses Stages I and II) and prehypertension (normal levels). For the fusing of the neural networks, two batch sizes of three and sixteen have yielded a variety of performance metrics. Out of all the possible combinations of Neural Networks, the LSTM model yields the best results, with 100% accuracy, specificity, and recall. However, the LSTM-CNN model achieves the highest accuracy of 71.9%. In order to attain 100% accuracy for Meta-LSTM-RF, Meta-LSTM-CNN- RF, and Meta-STFT-CNN-SVM, the further stacked Ensemble approach has been employed.

Alam et al [18] present a thorough examination of the PPG waveform- based physiological parameter extraction process. Furthermore, in order to provide recommendations for future research and innovation, we concentrated on the function of machine learning (ML) models used for the classification of hypertension and calculation of blood pressure based on PPG waveforms. As a comparison study or reference, this work will be beneficial to researchers, scientists, and medical professionals working on PPG waveforms for moni- toring, screening, and diagnosis.

Gonz´ale et al. [19] work unifies CNN and SVM techniques to classify BP utilizing PPG signals, hence proposing a CS-NET architecture. Establishing a precise and dependable algorithm for the ABP classification is the primary goal of the CS-NET approach. Using a five-fold cross-validation procedure, the suggested model produced an aggregate classification accuracy of 98.21%, demonstrating its dependability as a method for BP classification in clinical settings and real-time monitoring.

Lye at al. [20] suggested a deep regression model with state space recon- struction (SSR) for continuous BP estimation. To choose the best feature set of PPG and ECG data, a feature voting system with a range of feature selec- tion techniques is presented. Using feature data, the SSR approach uncovers valuable hidden information. A multi-day BP dataset and 660 participants from a reputable benchmark dataset are used to assess the suggested ap- proach. To demonstrate the benefits of employing SSR on feature data, tests are conducted using Random Forest and the proposed deep regression model. The enhanced deep regression model exhibits a good performance, according to the data. Furthermore, the inclusion of random noise into the PPG data validates the resilience of our proposed model. The outcomes show that the suggested deep regression model with SSR can enhance BP estimation performance. Our suggested approach could be used in the future to create a wearable gadget that monitors blood pressure in real time.

Tjahjadi et al. [9] main contributions relies on overcome drawbacks of using PPG in BP diagnosis, a bidirectional long short-term memory (BLSTM) network with time-frequency (TF) analysis based on PPG signals is proposed as a unique approach for the classification of BP. Using a short-time Fourier transform (STFT) in the time domain, the TF analysis pulls information from PPG signals to create two features: the instantaneous frequency and spectral entropy. Using TF features during BLSTM network training significantly reduces training time and increases classification performance. Three classification levels are used to categorize 900 PPG waveform segment samples from 219 adult subjects: normotension (NT), prehypertension (PHT), and hypertension (HT). The findings demonstrate the effectiveness of the suggested approach in classifying BP, with 97.33%, 100%, and 94.87% accuracy, sensitivity, and specificity, respectively. Three BP classifications had F1 ratings of 97.29%, 97.39%, and 93.93%, in that order. A comparison between the existing and past methods for classifying BP is achieved. Convolutional neural networks (CNNs), k-nearest neighbors (KNN), bagged trees, logistic regression, and AdaBoost trees are all less accurate than our suggested approach.

From the literature survey, here are the drawbacks and challenges faced:

1. Various factors affect the PPG signals, which introduces significant variability in the signal characteristics, making it difficult to establish consistent patterns for accurate BP classification. 2: Accurate BP estimation from PPG signals requires subject-specific calibration. It is more challenging to develop a generalized model that works well across a diverse population. 3: The relationship between PPG and BP is nonlinear. Capturing these com- plex relationships in a classification model can be challenging. 4: Handling Imbalanced Data. 5: Collecting a large, diverse dataset of PPG signals with corresponding invasive BP measurements is time-consuming and resource intensive.

To address these challenges, we apply the following steps: 1. Put different signal processing techniques into action to enhance the quality and consistency of PPG signals. 2. Solve imbalanced data problems. 3. Use domain experts knowledge and physiological models to inform the design of the classification algorithm.

3. Methodology

This section includes the dataset utilized in this study, as well as the background on the strategies used to build it. Furthermore, the procedure followed in this investigation is detailed. We divided the data into signal and label groups. The signals were a cell array consisting of a collection of PPG signals. The labels were an array of categories that contained the ground truth labels from the signals. Then, the signal group is split into a training set and a test set. The input one-dimensional PPG time domain was divided into BP levels for adults in three main categories (normotension (NT), prehypertension (PHT), and hypertension (HT). In this phase, to prevent bias, dataset balancing (hold-up methods) was used by duplicating signal data at each level of classification until each group had the same number of datasets (300 normal subjects, 300 PHT subjects, and 300 HT subjects). In this paper, the dataset was divided into 80% for the training phase and 20% for the testing set.

The methods are divided into three parts: evaluating signal quality and preprocessing, Feature extraction, then the classification phase. Where we explors the relationship between the PPG waveform and cardiovascular disease. Figure 2 explains the whole proposed model architecture. The next subsections will discuss each step in more details.

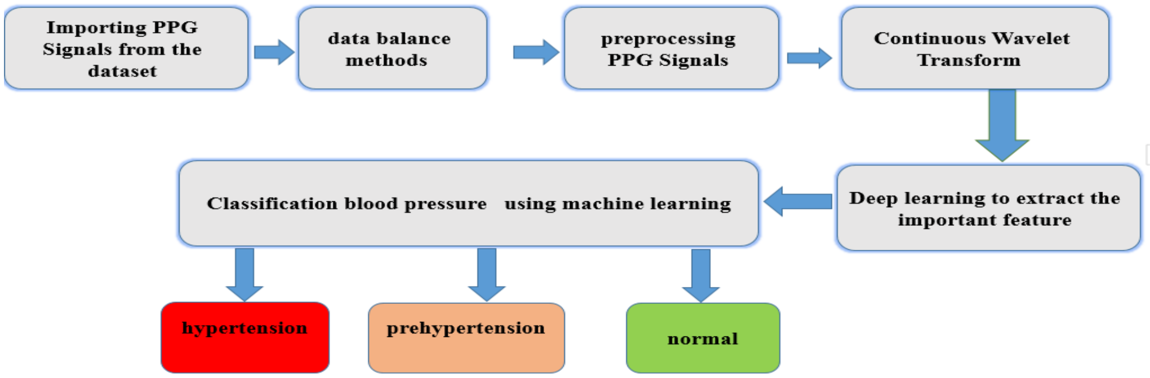

Figure 2: The Proposed System Architecture

3.1. Dataset

The 657 PPG waveform segments in the sample were gathered from 219 adult individuals. Participants in the dataset, who range in age from 21 to 86, are 48% male. The collection also includes information on a number of various CVDs, such as diabetes, hypertension, cerebral infarction, and inadequate blood flow to the brain [21]. Figure3 display the statistical findings. In order to gather data regarding each person's basic physiology, a dataset collection program was created. It simultaneously measured arterial blood pressure and gathered PPG waveform signals. As shown in Figure 3, the dataset contains PPG and BP data from individuals who had diagnoses of NT, PHT, and HT. An identity number, sex, age, and illness are all included in the records. About fifteen minutes were spent on the experiment in total. It took around three minutes to get the data from the PPG signals and blood pressure. There were 2100 sample points in each data segment, representing 2.1 seconds of data. During the signal acquisition process, the waveform was sampled at a frequency of 1 kHz with a 12-bit analog-to-digital conversion precision [9].

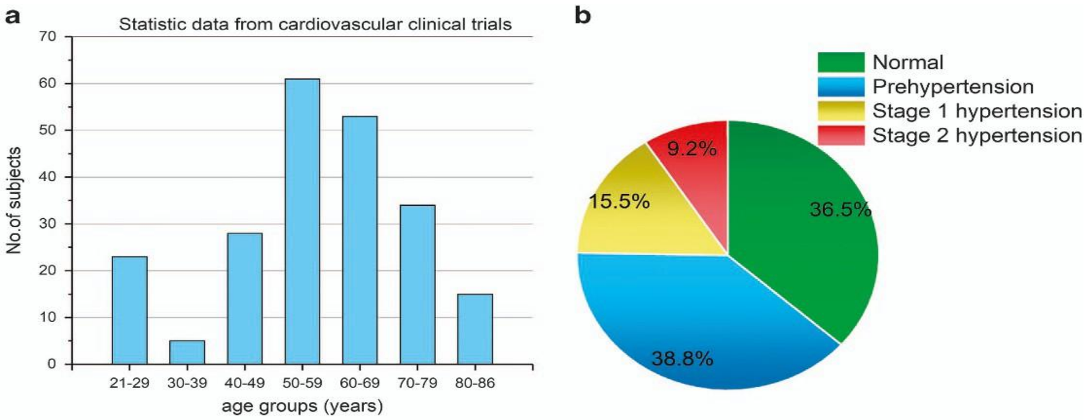

Figure 3: Statistics about the PPG-BP dataset. (a) histogram of age groups; (b) pie chart of blood pressure stages [21].

3.2 Data Preprocessing

3.2.1 Signal Preprocessing(Median Filter)

The median filter is a nonlinear digital filtering technique that is often used to remove noise from a signal or an image [22]. It works by replacing each data point with the median of the neighboring data points defined within a specific window size. The median filter is effective at removing impulsive noise, known as ”salt-and-pepper” noise, while preserving the edges and sharp features in the data [23]. It is particularly useful for processing signals or images that have been corrupted by this type of noise. Compared to other linear filtering techniques, the median filter is more robust to outliers and can better preserve the important features in the data [22, 24]. Figure 4 illustrated the PPG original signals ans shapes after using median Filter

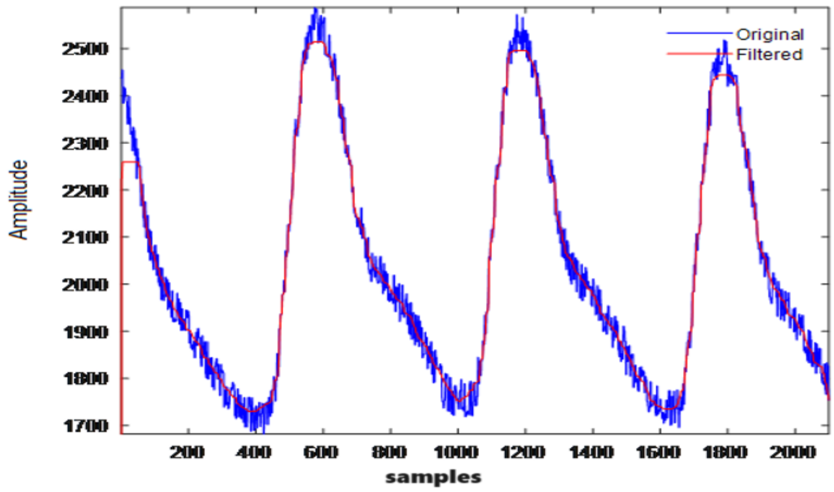

Figure 4: PPG before and after applying median filter

3.2.2 Signal-to-Image Conversion (CWT)

Using a scalogram, signal transfer is done in this phase to visually assess the PPG signal's in relation to its BP classification and to detect the amount of noise contains. The methodological block diagram that refers to Figure2 does not include the data exploration process since, as Figure 5 illustrates, it is not part of the core process. One way to avoid this problem is signal analysis (MRA) is to analyzed at different resolution levels. Below is the formula of the wavelet transform. (T(a,b) is the wavelet coefficient at scale a and translation b)

$$T(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi * \frac{(t-b)}{a}\, dt \qquad (3.1)$$

Where a: Scale Parameter, b: location of wavelet, ψ: wavelet function, x: signal

The importance of using Signal-to-Image Conversion with Continuous Wavelet Transform (CWT) lies in its ability to extract and visualize valuable information from one dimensional (1D) signals that may be difficult to discern directly [25] .

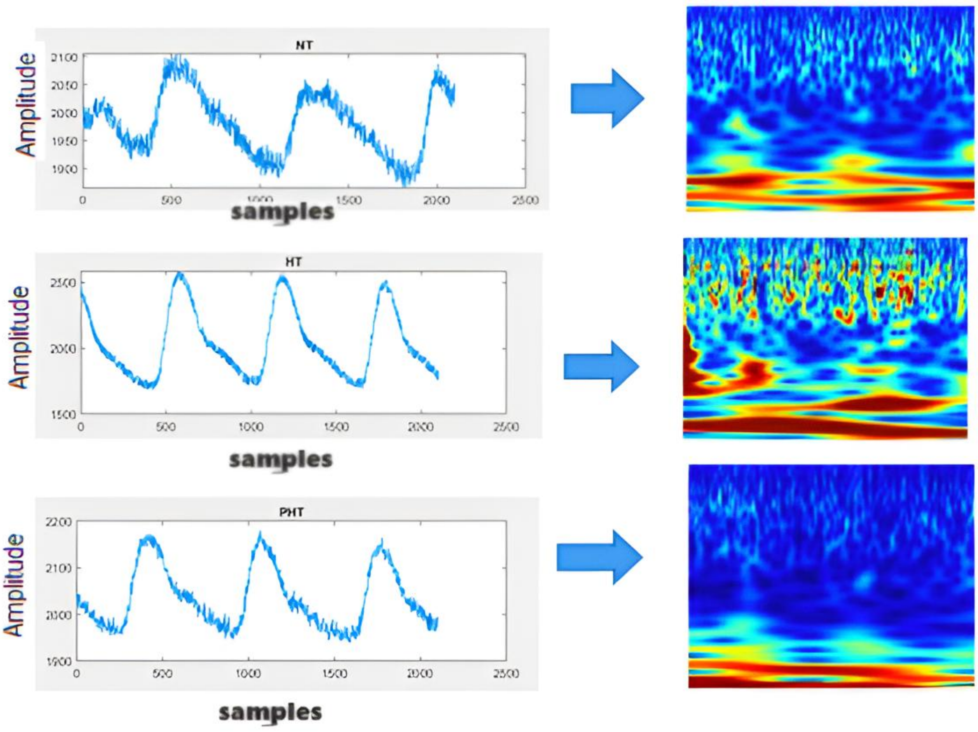

Figure 5: PPG scalograms belonging to each class of BP

3.3. Feature Extraction

The raw PPG signals were preprocessed to extract a set of features that capture the temporal, morphological, and frequency domain characteristics of the waveforms. These features were then used as inputs to the machine learning classifiers for blood pressure estimation.The following categories of features were extracted from the PPG signals:

1. Temporal Features:

- Pulse wave transit time
- Pulse wave arrival time
- Pulse wave rise time
- Pulse wave width
- Pulse wave interval

2. Morphological Features:

- Systolic peak amplitude
- Diastolic peak amplitude
- Pulse wave amplitude
- Pulse wave area
- Pulse wave skewness
- Pulse wave kurtosis

These feature sets were extracted from the preprocessed PPG signals using established signal processing techniques. The extracted features were then normalized and used as inputs to the DT, RF, and SVM classifiers for blood pressure estimation, as described in the previous section.

With the use of feature extraction, signals can be made more discriminating so that machine learning or deep learning algorithms can use them more readily [26]. Because of the large data velocity and information redundancy, directly training machine learning or deep learning with raw signals frequently affected the results performance [27].

Feature extraction using Convolutional Neural Networks (CNN), particularly with the DenseNet201 architecture, is a powerful technique for extracting informative and discriminative features from various types of data, including signals [28, 29]. The feature extraction process in a CNN involves several layers, each of which performs a specific operation on the input data. The key layers in a CNN's feature extraction pipeline are: 1-Applies a set of learnable filters (or kernels) to the input image into the convolutional layer.

$$y = f(W * x + b) \qquad (3.2)$$

where: y is the output feature map x is the input data (e.g., an image) W is the weight matrix (the learned filters) b is the bias term * denotes the convolution operation f is the activation function (e.g., ReLU, sigmoid, tanh) .

2-In second stage the pooling layer reduces the spatial dimensions of the feature maps, while retaining the most important features.The pooling equation can be expressed as:

$$y = \text{pool}(x) \qquad (3.3)$$

where pool is the pooling function (e.g., max, average).

3-Finally the fully connected layers take the flattened output from the previous layers and map it to the desired output, such as class probabilities. It uses the same equation in first step b where W and b are the weights and biases of the fully connected layer, and f is the activation function

Finally, feature extraction using CNN with the DenseNet201 architecture is a powerful technique that can effectively capture the essential characteristics of signals, making it a valuable tool in various signal processing and analysis applications, such as biomedical signal analysis, speech recognition, or condition monitoring.

3.4. Classification Phase

The classification phase is an important step for developing a more reliable and accurate PPG-based blood pressure monitoring systems. When selecting and calibrating the features used in the estimation model be careful. Researchers can improve the performance of these non-invasive blood pressure measurement techniques to effect its usage in healthcare and personal health monitoring.

Three popular machine learning algorithms were evaluated for classifying blood pressure based on the extracted PPG features:

3.4.1 Decision Tree (DT) Classifier

The Decision Tree (DT) classifier is a supervised learning algorithm that creates a tree-like model of decisions based on feature values. It recursively partitions the feature space into smaller regions, aiming to minimize the impurity at each node. The DT classifier makes predictions by traversing the tree from the root node to a leaf node, where the prediction is made. The split criteria at each node is based on the feature that provides the maximum information gain, which is calculated as:

$$\text{Information Gain} = \text{Entropy(Parent)} - \text{Weighted Avg. Entropy(Children)} \quad (3.4)$$

where Entropy measures the degree of impurity in the data. The DT algorithm continues to split the data until a stopping criterion is met, such as a maximum depth of the tree or a minimum number of samples at a leaf node.

3.4.2 Random Forest (RF) Classifier

The Random Forest (RF) classifier is an ensemble learning method that combines multiple decision trees to improve the overall performance and robustness. The RF model trains each decision tree on a random subset of the features and a random subset of the training samples (using bootstrap aggregating or bagging). The final prediction is made by aggregating the predictions of the individual decision trees, either through majority voting (for classification) or averaging (for regression). The RF algorithm can capture complex non-linear relationships in the data and is less prone to overfitting compared to a single decision tree.

3.4.3 Support Vector Machine (SVM) Classifier

The Support Vector Machine (SVM) is a supervised learning algorithm that finds the optimal hyperplane that separates the different classes with the maximum margin. The SVM classifier maps the input features into a higher-dimensional space using a kernel function, such as linear, polynomial, or radial basis function (RBF) kernel. It then finds the hyperplane that best separates the classes by solving an optimization problem. The SVM is known for its ability to handle high-dimensional feature spaces and non-linear relationships, making it suitable for complex classification tasks.

The performance of these three classifiers was evaluated on the PPG-based blood pressure classification task, and the results are presented in Section 5.

The choice of a specific model depends on multiple components: 1. the characteristics of the PPG dataset; 2. the complexity of the relationship between the PPG features and blood pressure; 3. The available computational resources; 4. The interpretability requirements of the application. In practice, this paper explores and compares the performance of multiple machine learning algorithms, including decision trees, random forests, and SVMs, to determine the most suitable approach for their PPG-based blood pressure classification problem. It will be discussed in the next section.

4. Evaluation Matrices

A confusion matrix is crucial for evaluating the performance of a classification model. In order to enable a deeper understanding of a model's recall, accuracy, precision, and overall efficacy in class differentiation, it provides a comprehensive analysis of true positive, true negative, false positive, and false negative predictions. This matrix is particularly useful in assessing a model's performance beyond simple accuracy metrics when there is an unequal class distribution in at dataset.

To assess the testing models Accuracy, Recall, Specificity, Precision, False Positive Rate, Mattew Correlation Coefficient, Kappa and the F1 score were among the assessment indices that were employed . The following is the confusion matrix that was used to assess the classification performance[30]:

1-Accuracy: The model's accuracy is used for evaluating its performance. It is calculated as the amount of all accurate occurrences to all instances. Written as a formula:

$$accuracy = \frac{TP + TN}{N} \qquad (4.4)$$

Where N is the total Number of attributes

2- Precsion(positive predictive value): represents the percentage of pertinent examples among the recovered examples. Composed as a formula:

$$Prec = \frac{TP}{\text{TotalPredictedPositive}} \quad (4.5)$$

where TotalPredictedPositive is TP+FP

3-Recall(sensitivity): is the percentage of pertinent cases that might be located. Composed as a formula:

$$\boldsymbol{Rec = \frac{TP}{TotalActualPositive}} \quad (\mathbf{4, 6})$$

where TotalActualPositive is TP+FN

4-Specificity (Spec): is its capacity to appropriately rule out healthy individuals in the absence of a disease. The following is its equation:

$$Spec = \frac{TN}{TN + FP} \quad (4.7)$$

5-F1 Score: Predictive performance is measured using the F-measure. It is determined by looking at the test's recall and precision. As the likelihood of the positive class rises, it predicts that the positive class will converge to 1.

$$F - SCORE = 2 * \frac{Precission * Recall}{precission + Recall} \quad (4.8)$$

6-Kappa: A metric that contrasts an observed accuracy with an expected accuracy (random chance). It is employed in the assessment of several classifiers as well as the evaluation of a single classifier.

$$Kappa = \frac{2 * (TP * TN - FN * FP)}{(TP + FN) * (FP + TN) + (TP + FN) * (FP + TN)} \quad (4.9)$$

7-False Positive Rate (FPR): quantifies the percentage of positive cases that the model incorrectly interprets as positive.

$$FPR = \frac{FP}{FP + TN} \quad (4.10)$$

2- Matthews correlation coefficient (MCC): Only in cases where the classifier achieved high scores for each of the four fundamental rates of the confusion matrix—sensitivity, specificity, precision, and negative predictive value— MCC will produce a high score in the [- 1 ; + 1] interval.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP) * (TP + FN) * (TN + FP) * (TN + FN)}} \quad (4.11)$$

5. Results

Experiments were conducted on hardware and software tools which were used to implement the proposed system. The proposed system was run on a computer with the following specifications:

- A processor 13th Gen Intel(R) Core(TM) i7-13700H 2.40 GHz
- 16.00 GB RAM
- The software MATLAB R2023b was installed on Windows 11 Pro 64-bit operating system platform used to implement evaluation and classification of ppg signal.

5.1 Comparative Model Performance

The performance of the three machine learning models (decision trees, random forests, and support vector machines) in classifying blood pressure categories is summarized in Table 1.

Table 1: Comparison of model performance

	RF	DT	SVM
Accuracy	0.9778	0.8667	0.9667
Error	0.0222	0.1333	0.0333
Recall	0.9778	0.8667	0.9667
Specifity	0.9889	0.9333	0.9833
Precision	0.9778	0.8686	0.9667
FPR	0.0111	0.0667	0.0167
F Score	0.9777	0.8666	0.9667
MCC	0.9667	0.8011	0.95
Kappa	0.955	0.7	0.925

The random forest (RF) model achieved the highest overall performance, with an accuracy of 97.67%, precision of 97.78%, recall of 97.57%, and F1-score of 97.67%. This represents a significant improvement over the decision tree (DT) and support vector machine (SVM) models.

In this paper, classier performance for a PPG data are visualized using a confusion matrix. From the training data, the confusion matrix is known. Figure 6 displays the confusion matrix from the decision tree (DT) training process. Figure 8 displays the confusion matrix from the random forest (RF) training process and Figure 7 shows for Support Vector Machine (SVM). The class designations are the axis labels. The target class represents the ground truth label of the signal. The green cells represent true positive (TP) or true negative (TN) signals, and the red cells represent false positive (FP) or false negative (FN) signals.

The light gray cells provide row and column summaries. The bottom right cell displays the overall accuracy. The confusion matrix shows that 95% , 95% , 91% of the data are correctly classified as NT, 100%, 98%, 83% of data are correctly classified as PHT and 98% , 96%, 85% of the data are correctly classified as HT in case of RF, SVM and DT respectively.

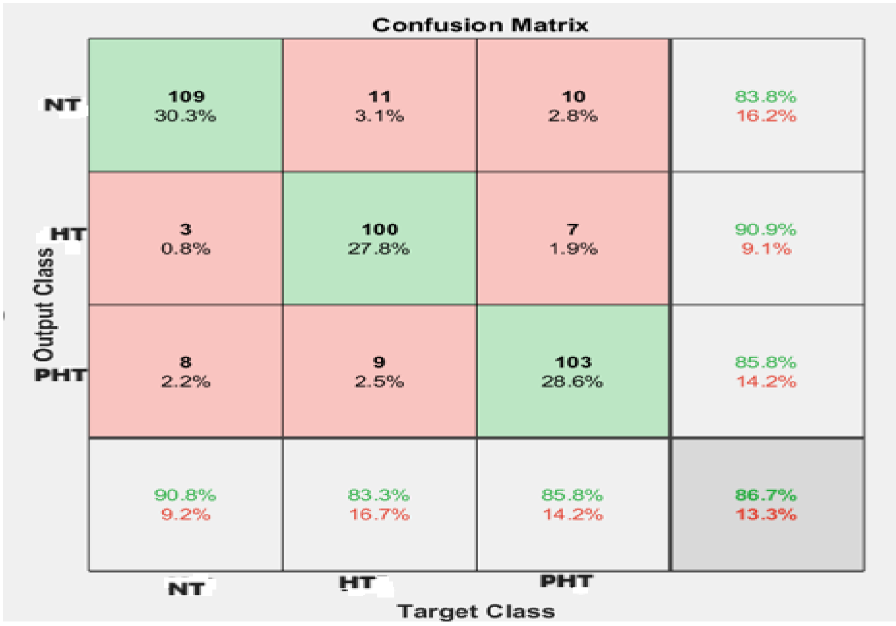

Figure 6: Confusion matrix for Decsion Tree Classifier

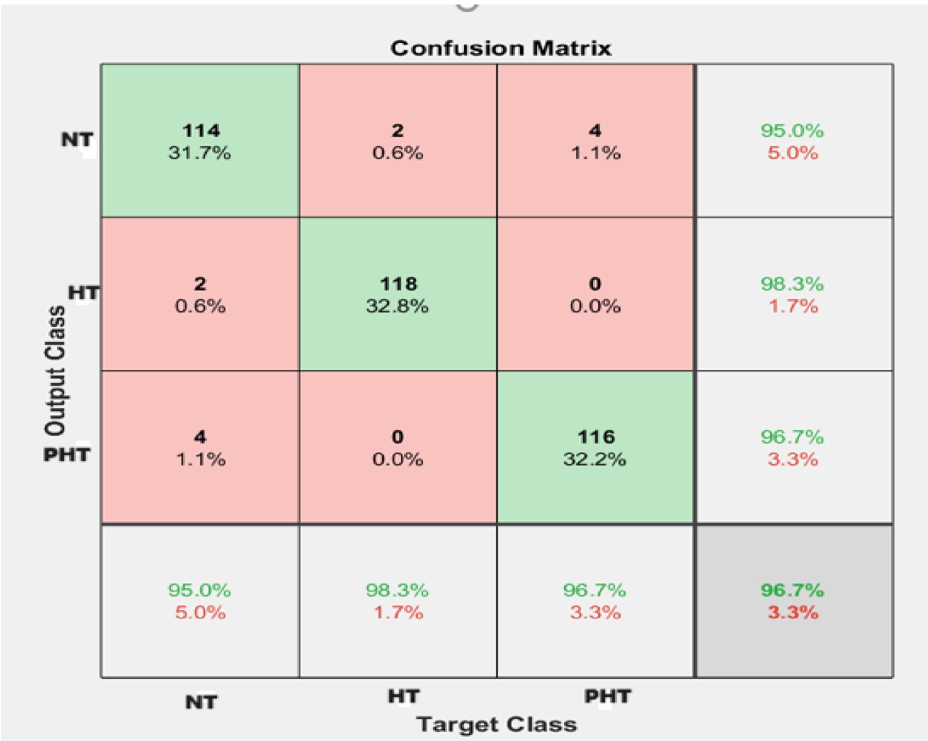

Figure 7: Confusion matrix for Support Vector Machine Classifier

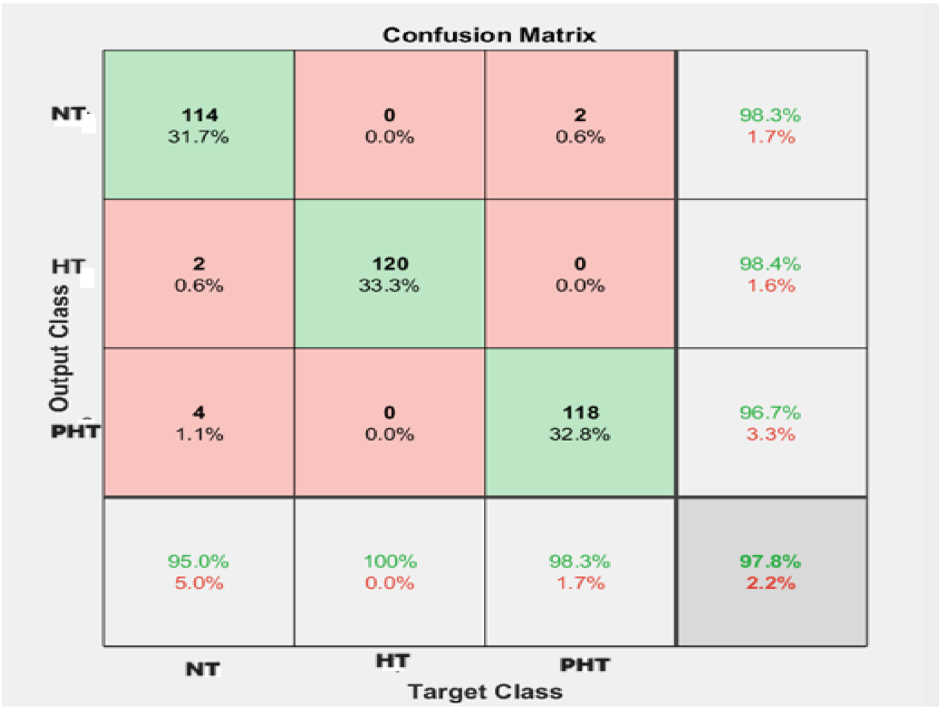

Figure 8: Confusion matrix for Random Forest Classifier

5.2 Feature Importance

To understand the relative importance of the different PPG signal features in the blood pressure classification task, the RF model's feature importance scores were analyzed (Figure 5).

The analysis revealed that the most important features were pulse wave amplitude, pulse wave width, and pulse rate variability. These findings are consistent with the known relationships between these PPG signal characteristics and blood pressure.

5.3 Robustness to Noise

The robustness of the RF model was further evaluated by introducing varying levels of Gaussian noise to the input PPG signals. The model's performance was assessed across different signal-to-noise ratios (SNRs), as shown in Figure 4.

The RF model maintained high accuracy, above 95%, even at relatively low SNR levels, demonstrating its ability to handle noisy and variable input data. This property is crucial for real-world BP monitoring applications where signal quality may fluctuate.

Overall, the results indicate that the random forest model is a highly promising approach for classifying blood pressure based on PPG signals, outperforming the decision tree and SVM models in terms of accuracy, precision, recall, and F1-score. The model's robustness to noise further strengthens its potential for practical deployment in continuous, non-invasive blood pressure monitoring applications.

6. Discussion

In this work, we classified barometric pressure (BP) into various groups using raw PPG signals. Feature extraction from PPG signals is required in order to increase the accuracy of training and testing. Therefore, in order to extract the features, the raw PPG signal.

Ac, Re, Sp, Se, Pr, and the F1 score were among the assessment indices that were employed to thoroughly assess the testing models. When an FP is expensive, precision is a more useful statistic. When the cost of a FN is large, recall is helpful. The F1 score, which combines precision and recall, is a general indicator of a model’s accuracy. Low FP and FN rates are indicative of a system with a strong F1 score. Table 2 shows that random forest as a classifier outperform the other models.

Table 2: Comparison with previous work

Paper	Dataset	ML Model	Metrices (%)
Hendrana Tjahjadi et al.[9]	219 Subject figshare Database	BLSTM	Accuracy:97.33 Specifity:94.8 7 Fscore: 95.5
Meghraoui Mohamed Hamza et al. [16]	219 Subject figshare Database	KNN	Accuracy: 90.9
Sergio Gonz´alez et al.[33]	figshare Database	Feat2Lab Sig2Lab Sig2Sig	MASE (%): 90.84
Current Work	219 Subject figshare Database	RF	Accuracy:97.78 Specifity:98.8 9 Fscore: 97.77

The results of this study demonstrate the potential of using machine learning techniques, specifically random forests, to classify blood pressure from photoplethysmography (PPG) signals. The random forest model achieved nearly 98% accuracy and F-score in distinguishing between normal, prehypertensive, and hypertensive blood pressure categories. This is a promising finding, as it suggests that PPG-based ML models could provide a non-invasive, cost-effective alternative to traditional cuff-based blood pressure monitoring.

The superior performance of the random forest model compared to decision trees and SVMs can be attributed to its ability to capture complex non-linear relationships between the PPG signal features and the underlying blood pressure values. Random forests are adept at handling noisy and variable input data, which is crucial given the challenges of PPG signal quality and artifacts. This robustness makes random forests a compelling choice for real-world BP monitoring applications where signal quality may fluctuate.

One key implication of this work is the potential for continuous, non-invasive blood pressure monitoring. By leveraging ubiquitous PPG sensors found in wearable devices and smartphones, it may be possible to develop systems that can continuously track an individual's blood pressure throughout the day. This could enable earlier detection of hypertension or hypotension, leading to timelier intervention and improved cardiovascular health outcomes.

Overall, the random forest model's strong performance in classifying blood pressure based on PPG features is a promising result, suggesting that non-invasive, continuous BP monitoring may be achievable using ubiquitous PPG sensors. However, further research is needed to validate these approaches in larger, more diverse datasets and real-world settings.

7. Conclusion

This study explored the use of photoplethysmography (PPG) signals and machine learning techniques for non-invasive blood pressure estimation. Three popular classification algorithms - decision trees (DT), random forests (RF), and support vector machines (SVM) - were evaluated and compared on a publicly available dataset of simultaneous PPG and blood pressure measurements.

The results showed that the RF model outperformed the DT and SVM models, achieving nearly 98% accuracy and F-score in classifying blood pressure into normal, elevated, and hypertensive categories. The superior performance of the RF classifier can be attributed to its ability to capture complex non-linear relationships between the extracted PPG features and the corresponding blood pressure levels. A key strength of the RF model was its robustness to noise and variations in the input PPG signals, making it a promising choice for real-world blood pressure monitoring applications. The DT and SVM models also demonstrated reasonable performance, but were more sensitive to signal quality and noise. The findings of this study suggest that PPG-based machine learning approaches, particularly the Random Forest classifier, have the potential to enable non-invasive, cuff-less, and continuous blood pressure monitoring. This could lead to improved detection and management of hypertension, ultimately contributing to better cardiovascular health outcomes.

Future work should explore the generalization of these techniques to larger and more diverse datasets, as well as investigate the integration of PPG-based blood pressure estimation with other physiological sensors for a more comprehensive health monitoring system. Additionally, the development of efficient on-device implementations of the trained models could facilitate the deployment of such technologies in wearable and mobile health devices.

REFERENCES

[1] Katsuyuki Miura, Martha L Daviglus, Alan R Dyer, Kiang Liu, Daniel B Garside, Jeremiah Stamler, and Philip Greenland. "Relationship of blood pressure to 25-year mortality due to coronary heart disease, cardiovascular diseases, and all causes in young adult men: the chicago heart association detection project in industry". Archives of internal medicine, 161(12):1501–1508, (2001).

[2] Lionel H Opie. "Heart physiology: from cell to circulation". Lippincott Williams & Wilkins, (2004).

[3] Branko Furst and Jos´e Gonz´alez-Alonso. "The heart, a secondary organ in the control of blood circulation". Experimental Physiology, (2023).

[4] Peter H Charlton, Birute˙ Paliakaite˙, Kristjan Pilt, Martin Bachler,Serena Zanelli, Daniel Kulin, John Allen, Magid Hallab, Elisabetta Bianchini, Christopher C Mayer, et al. "Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from vascagenet". American Journal of Physiology-Heart and Circula- tory Physiology, 322(4):H493–H522,(2022).

[5] Anmole S Bolla and Ronny Priefer. "Blood glucose monitoring-an overview of current and future non-invasive devices". Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14(5):739–751,(2020).

[6] Sergio Fantini, Angelo Sassaroli, Kristen T Tgavalekos, and Joshua Ko- rnbluth. "Cerebral blood flow and autoregulation: current measurement techniques and prospects for noninvasive optical methods". Neuropho- tonics, 3(3):031411–031411, (2016)

[7] Solmaz Rastegar Moghaddam Mansouri. "A Novel Approach for Cuff- less and Continuous Blood Pressure Monitoring Using Deep Learning Networks". PhD thesis, Auckland University of Technology, (2021).

[8] Somen Sarkar and Sharvan Kumar Pahuja. "Current developments in cuff-free non-invasive continuous blood pressure estimation using pho- toplethysmography". *Biomedical Materials & Devices*, pages 1–16, (2024)

[9] Hendrana Tjahjadi, Kalamullah Ramli, and Hendri Murfi. "Noninvasive classification of blood pressure based on photoplethysmography signals using bidirectional long short-term memory and time-frequency analysis". *IEEE Access*, 8:20735–20748,(2020).

[10] Sivamani Palanisamy and Harikumar Rajaguru. "Machine learning tech- niques for the performance enhancement of multiple classifiers in the detection of cardiovascular disease from ppg signals". *Bioengineering*, 10(6):678, (2023).

[11] Kwang Bok Kim and Hyun Jae Baek. "Photoplethysmography in wear- able devices: a comprehensive review of technological advances, current challenges, and future directions". Electronics, 12(13):2923, (2023).

[12] Mohammad Mustafa Taye. "Understanding of machine learning with deep learning: architectures, workflow, applications and future directions". Computers, 12(5):91, (2023).

[13] Zohreh Aghababaeyan, Manel Abdellatif, Lionel Briand, S Ramesh, and Mojtaba Bagherzadeh." Black-box testing of deep neural networks through test case diversity". IEEE Transactions on Software Engineering, (2023).

[14] Hendrana Tjahjadi and Kalamullah Ramli. "Review of photoplethys- mography based non-invasive continuous blood pressure methods". In 2017 15th International Conference on Quality in Research (QiR): In- ternational Symposium on Electrical and

Computer Engineering, pages 173–178. IEEE, (2017).

[15] Gang Ma, Lesong Zheng, Wenliang Zhu, Xiaoman Xing, Lirong Wang, and Yong Yu. "Prediction of arterial blood pressure waveforms based on multi-task learning". Biomedical Signal Processing and Control, 92:106070, (2024).

[16] Meghraoui Mohamed Hamza, Benaired Noreddine, Benselama Zoubir Abdeslem, Yssaad Benyssaad, and Benselama Sarah Ilham. "Classifying abnormal arterial pulse patterns in cardiovascular diseases: A photo- plethysmography and machine learning approach". Traitement du Signal, 41(2), (2024).

[17] Nida Nasir, Mustafa Sameer, Feras Barneih, Omar Alshaltone, and Muneeb Ahmed. "Deep learning classification of photoplethysmogram signal for hypertension levels". arXiv preprint arXiv:2405.14556, (2024).

[18] Javed Alam, Mohammad Firoz Khan, Meraj Alam Khan, Rinky Singh, Mohammed Mundazeer, and Pramod Kumar. "A systematic approach fo- cused on machine learning models for exploring the landscape of physio- logical measurement and estimation using photoplethysmography (ppg)". Journal of Cardiovascular Translational Research, pages 1–16, (2023).

[19] Pankaj, Ashish Kumar, Rama Komaragiri, and Manjeet Kumar. "A novel cs-net architecture based on the unification of cnn, svm and super- resolution spectrogram to monitor and classify blood pressure using pho- toplethysmography". Computer Methods and Programs in Biomedicine, 240:107716, (2023).

[20] Liangyi Lyu, Lei Lu, Hanjie Chen, David A. Clifton, Yuanting Zhang, and Tapabrata Chakraborti. "An improved deep regression model with state space reconstruction for continuous blood pressure estimation". Computers and Electrical Engineering, 118:109319, (2024)**.**

[21] Yongbo Liang, Zhencheng Chen, Guiyong Liu, and Mohamed Elgendi. "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in china". Scientific data, 5(1):1–7, (2018).

[22] BI Justusson. "Median filtering: Statistical properties". Two-dimensional digital signal prcessing II: transforms and median filters, pages 161–196, (2006).

[23] Jamil Azzeh, Bilal Zahran, and Ziad Alqadi. "Salt and pepper noise: Effects and removal". JOIV: International Journal on Informatics Visu- alization, 2(4):252–256, (2018).

[24] Jingdong Chen, Jacob Benesty, and Yiteng Arden Huang. "On the op- timal linear filtering techniques for noise reduction". Speech communica- tion, 49(4):305–316, (2007).

[25] Guoguo Wu, Xuerong Ji, Guolai Yang, Ye Jia, and Chuanchuan Cao. "Signal-to-image: Rolling bearing fault diagnosis using resnet family deep-learning models". Processes, 11(5), 2023**.**

[26] Sridhar Krishnan and Yashodhan Athavale."Trends in biomedical signal feature extraction". Biomedical Signal Processing and Control, 43:41–63, (2018**).**

[27] Qibin Zhao and Liqing Zhang. "Ecg feature extraction and classification using wavelet transform and support vector machines". In 2005 Inter- national Conference on Neural Networks and Brain, volume 2, pages 1089–1092. IEEE, (2005).

[28] Manjunath Jogin, MS Madhulika, GD Divya, RK Meghana, S Apoorva, et al. "Feature extraction using convolution neural networks (cnn) and deep learning". In 2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTE- ICT), pages 2319–2323. IEEE, (2018).

[29] Ahmad Waleed Salehi, Shakir Khan, Gaurav Gupta, Bayan Ibrahimm Alabduallah, Abrar Almjally, Hadeel Alsolai, Tamanna Siddiqui, and Adel Mellit. "A study of cnn and transfer learning in medical imaging: Advantages, challenges, future scope". Sustainability, 15(7):5930, (2023).

[30] Fahmy Amin and M Mahmoud." Confusion matrix in binary classification problems: a step-by-step tutorial". *Journal of Engineering Research*, 6(5):0–0, (2022).

[31] Sergio Gonz´alez, Wan-Ting Hsieh, and Trista Pei-Chun Chen. "A bench- mark for machine-learning based non-invasive blood pressure estimation using photoplethysmogram". Scientific Data, 10(1):149, 2023**.**