A Benchmark Study of Machine Learning for Analysis of Signal Feature Extraction Techniques for Blood Pressure Estimation Using Photoplethysmography (PPG)

SUMBAL MAQSOOD[ID], **SHUXIANG XU**[ID], **MATTHEW SPRINGER**[ID], **AND RAMI MOHAWESH**[ID]

School of Information and Communications Technology, University of Tasmania, Hobart, TAS 7005, Australia

Corresponding author: Sumbal Maqsood (sumbal.maqsood@utas.edu.au)

ABSTRACT Cardiovascular related diseases are the most significant health concern around the globe. The most crucial health indicator is blood pressure because it gives essential information about the health of a patient's heart. Cardiovascular diseases can be detected early and prevented if blood pressure is monitored continuously and regularly. Blood pressure cuffs, which are widely used to control blood flow in the arm or wrist when measuring blood pressure, are not practical for continuous blood pressure measurement. However, biosignals can be used for blood pressure estimation; but it is still critical and challenging. In this paper, we conducted a comprehensive analysis of feature extraction techniques for blood pressure estimation by using PPG signals. The feature extraction techniques were further divided into three subgroups to analyse the significance of each group. Group A includes time-based features; group B presents statistical feature extraction, and group C presents frequency domain-based features. The analysis employed several machine learning algorithms and compared their performance from many perspectives. The experimental results from two publicly available datasets demonstrated that the set of features belonging to group A were more reliable than other techniques for blood pressure estimation. We found that deep learning models achieved better performance than all traditional machine learning methods. We also found that the GRU model and Bi-LSTM achieved the best performance for time-domain features for blood pressure estimation. We believe the findings of this benchmark study will help researchers choose the most appropriate method for feature extraction and machine learning algorithms.

INDEX TERMS PPG, datasets, deep learning, machine learning, blood pressure, features extraction.

I. INTRODUCTION

According to the World Health Organization (WHO), cardiovascular disease (CVD) is the leading cause of mortality, outnumbering all other diseases [1]. The predominant reasons for these diseases are high blood pressure, atrial fibrillation, and high cholesterol. There are a number of devices available to measure blood pressure (BP) manually or digitally, invasively or non-invasively [2]. Invasive BP estimation is the most accurate method that entails inserting a catheter into a peripheral artery's lumen [3]. This method can only be performed during high-risk surgeries or under intensive care [4]. It is a time-consuming method and needs a trained operator to

The associate editor coordinating the review of this manuscript and approving it for publication was Tony Thomas.

perform it [5]. For this reason, the standard approach to measuring blood pressure is non-invasive, allowing intermittent or continuous measurement [4].

The continuous non-invasive BP estimation is the process of volume clamping (or vascular unloading) [6]. In this method, A finger cuff is applied to the patient's finger to record arterial waveforms [7]. An integrated photoplethysmograph assesses the diameter of the finger artery, which ensures that the flow of blood in the finger artery is kept constant [8]. It continuously monitors blood pressure in real-time like an invasive arterial catheter device but is non-invasive like a regular upper arm sphygmomanometer. The accuracy of all continuous non-invasive instruments is susceptible to the patient's movement. In addition to this, continuous BP tracking is relatively expensive compared to
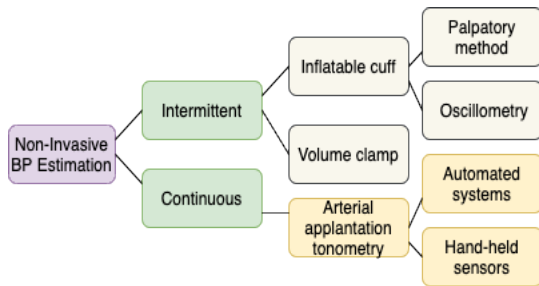
FIGURE 1. A hierarchical view of non-invasive blood pressure measurement depicts two types of non-invasive blood pressure estimation: intermittent and continuous, which are further divided.

traditional intermittent BP measurement systems [4]. Photo-plethysmogram (PPG) is a simple and low-cost technique that can measure the blood volume and heartbeat of the patients. It uses a semiconductor component to illuminate the skin tissue and measure light intensity and volume variations. However, recent developments in this method have demonstrated promising results concerning accuracy, convenience, and clinical acceptance. With technological advancement, it is imperative to have intelligent devices to get health checks, even sitting in the comfort of your home. Many researchers are working worldwide to get a non-invasive and cuffless blood pressure estimation device to achieve this. Continuous blood pressure estimation has attracted many researchers because it can detect early blood pressure problems. Electrocardiogram (ECG) and PPG are the two most commonly studied parameters for non-invasive BP estimation and have shown promising results.

The primary motivation of this study is that PPG is a non-invasive, simple, and low-cost technique that can be used to measure blood volume and other significant parameters related to CVD. Hypertension is an indicator for heart diseases risk, and it can be measured by screening blood pressure. There are conventional ways to monitor blood pressure, but a professional like a doctor or a nurse needs to measure it. Moreover, these methods cannot perform continuous measurements and are unsuitable for integrating with the latest technology. Furthermore, many researchers and companies have explored different methods for continuous BP measurement; development is required for the clinical benefits of PPG for mobile and wearable health technologies [9]. These devices can measure different physiological parameters like body temperature, physical activity, heart rate, including estimated systolic blood pressure (SBP) and diastolic blood pressure (DBP) [10]. The significance of PPG biosignals inspired us to do a detailed analysis of existing feature extraction techniques to assess their respective performance. To the best of our knowledge, this is the first detailed study to perform comparison analysis for feature extraction techniques used for PPG biosignals for blood pressure estimation. This paper proposes a continuous cuffless blood pressure estimation using PPG signals and compares feature extraction techniques by statistical and deep machine learning. For this study, we are particularly interested in these

follow-up research questions: RQ1: which traditional feature extraction techniques achieve high performance? RQ2: How accurate are conventional machine learning vs deep learning models for estimating blood pressure?

A. CONTRIBUTION
The key contributions of this paper can be summarised as follows:

- This paper aims to demonstrate how to use PPG signals to facilitate continuous monitoring of blood pressure of older or vulnerable patients. To be more specific, we analyse feature extraction techniques to estimate blood pressure more accurately.
- We analyse the performance of different types of feature extraction techniques. We also investigate the importance of these features individually to provide the best feature performance in this domain.
- We investigate the performance of machine learning approaches, including traditional machine learning (LR, RF, AdaBoost, SVM) and deep learning (vanilla LSTM, BI-LSTM, GRU).
- The experimental results on two datasets show that the performance of Time-domain features achieved better results with deep learning algorithms. To the best of our knowledge, this paper provides the first extensive comparison of different machine learning methods with different feature extraction techniques for blood pressure estimation using PPG signals.

The paper is arranged as follows: Section 2 assesses the proposed research taxonomy in the light of existing literature. Section 3 discusses publicly available databases used in this study, feature extraction techniques and machine learning algorithms. Section 4 details the experiment setup and results with discussion. Section 5 presents the conclusion and future work.

II. BACKGROUND AND RELATED WORK
This section sheds light on the background of the PPG signal and provides an overview of current research work to estimate blood pressure using PPG. Numerous methods for estimating the cuffless BP have been presented as alternatives to conventional blood pressure measurement techniques over the last decade. Pulse transit (PTT) time and pulse arrival time (PAT) have been widely used for this purpose [11]. PTT is defined as the amount of time a pressure wave takes to travel concerning two body parts. It is determined as the time difference concerning the proximal and distal locations of the arterial pulse [12]–[14]. Another cuffless approach is the Pulse Arrival Time (PAT) method which uses the total of PTT and the period of pre-ejection [15]. PTT can be extracted by simultaneously using PPG and ECG signals and is a valuable element in assessing BP values [13]. In several studies, researchers used vascular transit time (VTT) to estimate BP values. It is defined as the time interval between PPG and heart sounds [16]. Furthermore, the heart rate (HR) has

been used for blood pressure estimation [17]. However, PTT and PAT approaches are confronted with several problems, despite their advantages, which have limited their usage to research. Specific measurements are required at two separate locations on the body, and therefore, for evaluating those parameters, two sensors (ECG and PPG) are needed for measurement. Some patients may find this uncomfortable and challenging. More sensors also demand more work on signal pre-processing, which takes time and increases the complexity of the computation. Furthermore, these methods are dependent on complex models of artery wave propagation and need individual calibration because of their dependence on patient physiology. These are all reasons why PAT and PTT may not be trustworthy approaches and a substitute for regular blood pressure measurements [18], [19].

The reasons, as mentioned earlier, inclined the researcher's attention towards PPG-based devices to estimate BP as it requires one pulse sensor instead of two different sensors and is a non-invasive, cuffless approach. A photoplethysmogram (PPG) is a non-invasive and low cost optically taken plethysmograph technique, first explored in the 1930s. It can find changes in blood volume in the microvascular bed of tissue. A pulse oximeter can obtain the signals to measure the changes in light absorption by illuminating the skin. [20]. Pulse pressure is created from the heart pumping blood to the periphery in each cardiac cycle, and this can be measured by illuminating the human skin surface using an LED (Light Emitting Diode) to a photodiode. PPG can be used to monitor the cardiac cycle, heart rate, respiration, depth of anaesthesia, blood pressure, and conventional imaging as remote photoplethysmography. These signals are composed of non-pulsatile (DC) and pulsatile components (AC). DC has narrated the light absorption in the skin tissue, which correlates with the AC component synchronised with arterial pulsation and heart [21]. According to the Beer-Lambert law, blood volume changes represent optical path length (Δd) to generate AC as heartbeats [22]. It has traditionally been used in the healthcare domain to measure blood oxygen saturation [23] and heart rate [24] with a pulse oximeter. However, PPG signals can be analysed to benefit many areas, like the automotive field, to get information about driver behaviour [25], monitoring blood circulatory condition, breathing, and subjective analysis [26].

Over the last decade, machine learning with data analysis has been an attractive area for research. Many machine learning techniques and algorithms have been applied to calculate blood pressure from the study of PPG signals. In 2011, SBP was measured by a random forest model using a set of characteristics of PPG signal by Monte-Moreno [27]. Then, a similar methodology was performed on intensive care patients' data [28] for continuous measurement of blood pressure using a deep belief network model using PPG signal analysis [29]. Teng and Zhang [30] applied linear regression (LR) algorithm to investigate extracted features from the PPG relationship with blood pressure. It was observed that the diastolic time showed a stronger relationship with

systolic blood pressure than diastolic. It was determined that the outcomes were consistent with the American National Standards of the Association for the Advancement of Medical Instrumentation's (AAMI) [31] requirement for blood pressure estimation non-invasively. Nevertheless, they used a small dataset for testing consisting of data from individuals with normal blood pressure. The AdaBoost machine learning algorithm was employed to estimate blood pressure [32]. A threshold was used to classify SBP values, and then a machine learning model was used to estimate SBP using a nonlinear classification process. This method was only capable of assessing SBP and did not meet the criteria set by the AAMI standards.

In 2013, Kurylyak, *et al.* [33] extracted distinguishing features using PPG waveforms and used a feed forwards neural network to estimate SBP and DBP. The results were promising in terms of developing an accurate cuffless blood pressure monitoring system. Liu, *et al.* [34] extracted 21 time-scale PPG features proposed by Kurylyak, *et al.* [33] and 14 more features from PPG's second derivative (SDPPG). SDPPG has previously been shown to detect aortic stiffness and compliance, closely linked to blood pressure [35], [36]. They applied support vector regression (SVR) to estimate blood pressure. This study found that combining features improved the accuracy. In another research study, Khalid, *et al.* [37] applied three machine learning techniques to predict blood pressure categories and compared their accuracy. They extracted three PPG features: pulse area, 25% width, and pulse rising time. However, the results for blood pressure categories were unsatisfactory. Dey, *et al.* [38] presented a set of PPG features based on frequency and time domains. This new BP model used physiological and demographic partitioning. They proposed a smartphone app called InstaBP. However, this mobile application is no longer available. In addition to this, the findings were not AAMI standards-compliant. Yi, *et al.* [39] extracted nine morphological features from each PPG signal and applied linear regression, Elastic network, LASSO, KNN, and CART. The results showed that KNN outperformed all other algorithms, and they also stated that fewer PPG features could achieve the same accuracy as more PPG features.

Liang, *et al.* [40] used the GoogLeNet convolutional neural network for hypertension stratification with traditional features extractions using PPG signals. In the recent past, the cuffless blood pressure measurement studies by LSTM have received considerable attention [41]–[44]. Radha, *et al.* [45] applied many feature extraction techniques like time-domain, entropy-based, and frequency domain on PPG signals to retrieve more information and used machine learning neural nets. Further, numerous models of machine learning were applied to predict the relative SBP and DBP in a free-living context with 103 participants. It was found that the LSTM model worked better for the prediction of SBP [46]. The literature indicates that these models can measure blood pressure with a single PPG sensor, and it is the most straightforward and cost-effective technique. However, the link between the
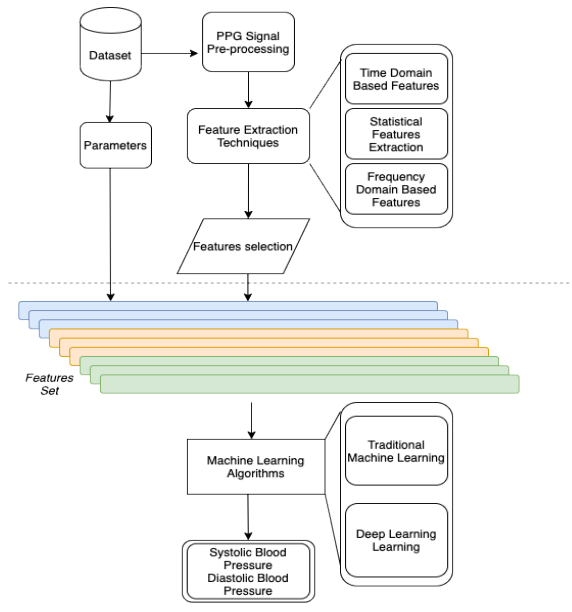
FIGURE 2. Structure of proposed methodology for a detailed comparison of feature extraction techniques and machine learning algorithms for systolic and diastolic blood pressure estimation.

PPG features and blood pressure were not necessarily linear. While evaluating a huge dataset from diversified groups, linear models might have difficulty accurately modelling the link between BP and PPG. Other machine learning models, like random forest and SVM, are more precise. Since these models require one for each objective, the two independent measures, SBP and DBP, can be computed separately. Though, DBP has a strong correlation with SBP and can be used to improve its estimation. Therefore, they should be modelled simultaneously using a single model architecture.

III. METHODS
The proposed approach has four stages:
- Datasets
- Pre-processing
- Feature extraction
- Machine learning algorithms

A. DATASETS
In this study, the PPG-BP dataset [47] and Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC-II) dataset [48] are used for the analysis of feature extraction techniques and machine learning models.

The MIMIC-II database [48] incorporates health-related information associated with a huge number of patients who remained at the Beth Israel Deaconess Medical Center critical care units from 2001 to 2008. This dataset is open-source and used extensively for blood pressure estimation analysis because it offers an array of signals containing PPG and references BP signals from a wide range of different demographics. It comprises physiological signals, including BP, PPG,

ECG and breathing. These signals are sampled at 125 Hz simultaneously. BP and PPG signals have been divided into 30-second windows, which display stable BP values.

The second dataset is PPG-BP, published recently and recorded at Guilin People's Hospital in Guilin, China. It consists of many patient parameters like age, height, weight, heart rate, BMI, blood pressure readings, PPG signals, etc. These records were collected from 219 patients aged between 21 to 86 years. It also contained other diseases records like hypertension and diabetes [47]. The MIMIC dataset has been used extensively in previous studies. This comprises much more information such as age, heart rate, ECG signals, PPG signals, arterial blood pressure, etc. Table 1 can be seen for statistical analysis of these two datasets.

B. RE-PROCESSING
A critical component to process a signal is accurately identifying and extracting PPG characteristics from the original signal data, as this is required to construct a reliable and widely applicable model. In the signal databases, there are a significant number of BP and PPG patterns that have distorted and irregular segments. We performed several steps in order to extract PPG features effectively. Firstly, we removed irregular signal segments and aligned PPG and BP signals. The patient's movement during the test could create a shift in the base of these signal acquisitions. These movement artefacts and noise can affect the signals subject to the sensor's properties, such as patient measurement point, movement, and ambient light. To denoise the signals and remove these movement artefacts, we applied the Butterworth filter, a signal processing filter. It is intended to provide frequency response in the passband that is as flat as mathematically possible. It is also known as a flat maximum magnitude filter [49]. Moreover, the min-max normalisation method was used to normalise the PPG signals to a range of [0,1] using the PPG signals. The signals were further segmented into the same time frames, and the resulting datasets were later used as input datasets for feature extraction and machine learning algorithms.

1) FEATURE EXTRACTION
In order to establish a trustworthy BP estimate model, it is necessary to identify integrated features pertinent to the task, especially the characteristics that describes best how blood pressure fluctuates. Due to the fact that PPG waveforms vary from person to person, there is no precise set of features that establish a direct relationship with blood pressure. Some features may not be extracted from all PPG signals; for instance, in the arterial pulse, there is a tiny downward deflection known as "aortic notch" or "dicrotic notch" [50] that can be calculated or approximated from some signals only [51]. This notch can be found in PPG waveforms recorded from healthy and young individuals [52]. However, the signals without a clear dicrotic notch may deteriorate the accuracy of BP estimation. Thus, building a dataset for an automatic blood pressure prediction model entails accurately extracting

TABLE 1. Datasets statistical summary of systolic and diastolic blood pressure.

Dataset	Statistics	Mean (mmHg)	STD (mmHg)	Min (mmHg)	Max (mmHg)
MIMIC II	SBP	137.15	22.43	58.18	199.69
	DBP	67.37	10.06	50.05	126.93
PPG-BP	SBP	127.94	20.33	80.00	182.00
	DBP	71.85	11.08	42.00	107.00

features from PPG signals, choosing the most influential parameters for blood pressure estimation and generalisation enhancement, reducing model overfitting. Despite this, the PPG-based strategy has gained popularity due to its simplicity and feasibility over the last decade. Numerous features that characterise the PPG waveform have been explored in the literature.

2) TIME DOMAIN-BASED FEATURES (GROUP A)

The PPG waveform is defined by the amplitude and duration of specific cardiac cycle components. Due to the movement artefacts present during the recording of PPG signals, the height of the pulse might change, considering that it is not suitable for blood pressure calculation. By contrast, the collection of frequencies by oximeters and other sensor devices varies according to the duration of each individual heartbeat for the PPG signal acquisition [53], [54]. This means that each heartbeat is signified by a unique set of values of samples, which cannot be used as the ANN input vector since input neurones numbers are fixed and cannot be modified during the ANN architecture. As a result, a different solution must be sought. Numerous factors can be used to define the pulsatile component of PPG. Apart from the Systolic upstroke Time (ST), Diastolic Time (DT), 2/3 and 1/2 pulse amplitude widths [30], the pulse height, cardiac period, and peak width at 10% of the pulse height are also employed [55].

Kurylyak, et al. [33] investigated additional data to extract more features and scrutinise the optimal parameter combination afterwards. In particular, they proposed that the width be calculated at 25%, 33%, and 75% of the pulse height and that different values be extracted for the systolic part (i.e., the interval between the minimal and maximal points) and the diastolic part (the interval between the maximal and the next minimal point). In this study, 21 parameters were extracted and used, including the timings and ratios of systole and diastole. These parameters can be seen in Figure 3.

Using the MIMIC-II dataset, the ABP signal provided the blood pressure values; what was required were the systolic and diastolic blood pressure single values that matched waveforms throughout the given period. Systolic blood pressure denotes the peak of the ABP pulse waveform's systole, while DBP denotes the peak of the ABP pulse waveform's end diastole. As a result, the ABP signal's peak and end at diastole points serve as the SBP, and DBP ground truth values, respectively.

- CP, ST, DT, pulse area
- 10 %: (DW10; DW10 + SW10; DW10/SW10)
- 25 %: (DW25; DW25 + SW25; DW25/SW25)
- 33 %: (DW33; DW33 + SW33; DW33/SW33)
- 50 %: (DW50; DW50 + SW50; DW50/SW50)
- 66 %: (DW66; DW66 + SW66; DW66/SW66)
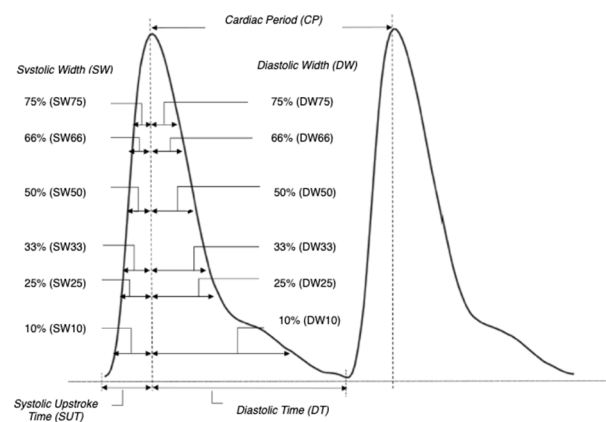- 75 %: (DW75; DW75 + SW75; DW75/SW75)

FIGURE 3. Time domain-based features that were extracted from the PPG waveform from each cycle. These features were calculated from the signal waveform width at a different signal level, such as at 10% for systolic and diastolic.

3) FEATURE IMPORTANCE

Dimensionality reduction is a pre-processing step in machine learning that effectively removes irrelevant and redundant information, increases learning accuracy, and improves the understandability of results [56]. In this study, we used several machine learning algorithms to reduce dataset dimensions to reduce complexity further and speed up calculations. It can help to exclude dataset redundancies and choose merely the efficient PPG features to estimate BP. Redundant features may negatively affect the output target, increasing the estimation error. Therefore, we have applied feature importance that refers to a group of techniques for assigning scores to the input features of a predictive model, assigning the relative importance of each feature when making a prediction. Scores of feature importance can be used to get insight into a dataset or model and can also be used to improve a predictive model. The relative scores can indicate which features are most important to the target and which are least important. Furthermore, it can help simplify the modelled problem, accelerate

the modelling process (removing features is referred to as dimensionality reduction), and in some cases, improve the model's performance. The findings of the analysis of each of the following tools were studied:

- **CART REGRESSION FEATURE IMPORTANCE:** Binary recursive partitioning is used for the Classification and Regression Trees (CART) methodology [57]. The process is binary because parent nodes are always divided into two children's nodes exactly and recursive since each child node is treated as a parent. CART analysis includes a number of regulations to split each node in a tree, decide when the tree is complete, and assign each terminal node a class result. The Gini rule is used for splitting, which is essentially a measure of the way in which the splitting rule separates classes in the parent node [57].
- **RANDOM FOREST FEATURE IMPORTANCE:** A practical algorithm for selecting features should assess subsets of features instead of individuals. The algorithm Random Forest (RF) constructs a large number of simple classifiers with the randomly selected features so that possible feature subsets are explored 58]. Moreover, since Bagging is a component of the algorithm, it does not include all the training data in constructing each of the basic assumptions. This data allows the evaluation of the feature's subsets without an independent test set. Therefore, random forest is well suited for feature selection.
- **XGBOOST FEATURE IMPORTANCE:** This is an improved gradient boosting algorithm based on the decision tree, which can efficiently build and operate boosting trees in parallel [59]. The core of the algorithm is the optimisation of the objective function. In comparison with the use of functional vectors for calculating the similarity between forecast and history, a gradient boost is built to give the boosted trees a smart indication of the importance of each feature's training model [60]. The algorithm considers "gain," "frequency", and "cover" as importance [61].

4) STATISTICAL FEATURE EXTRACTION (GROUP B)
In this section, statistical features were extracted. These features are further discussed below to provide more detail.

a: SKEWNESS
This can be calculated as given below, where σ and $\hat{\mu}_x$ present the empirical estimate of the standard deviation and average mean of x_i, respectively, and N shows the number of samples in the PPG signal.

$$S = 1/N \sum_{i=1}^{N} [x_i - \hat{\mu}_x/\sigma]^3 \qquad (1)$$

b: KURTOSIS
This is a statistical measure that can be used to describe the distribution of observed data around the mean. Moreover, it represents a heavy tail and peakedness or a light tail

and flat distribution relative to the normal distribution. Selvaraj, *et al.* [62] suggested that kurtosis is a reliable indication of the quality of PPG signals. It can be calculated below where N is the number of samples in the PPG signal, and σ and $\hat{\mu}_x$ present the average mean and standard deviation empirical estimate of x_i, respectively.

$$K = 1/N \sum_{i=1}^{N} [x_i - \hat{\mu}_x/\sigma]^4 \qquad (2)$$

c: PERFUSION
This is the gold benchmark for determining the quality of PPG signals. It is the ratio of pulsatile to non-pulsatile or static blood flow in tissue. To put it another way, this shows the difference between the quantity of light absorption by the pulse and the amount of light passing through the skin tissue; this can be expressed as given below where x signal (raw PPG signal) has statistical mean $\bar{x}$; moreover, filtered PPG signal is y.

$$P = [(y_{max} - y_{min})/|\bar{x}|] \times 100 \qquad (3)$$

d: MEAN ABSOLUTE DEVIATION, MAXIMUM AND MINIMUM
We included additional statistical features like mean absolute deviation of the given x PPG signal, maximum and minimum of x signal as given below:

$$maximum\,(x) = max\,(x_i)\,. \qquad (4)$$
$$minimum(x) = min\,(x_i) \qquad (5)$$
$$MeanAbs(x) = \frac{1}{N} \sum_{i=1}^{N} |x_i - \bar{x}| \qquad (6)$$

5) FREQUENCY-BASED FEATURES (GROUP C)
In this section, frequency-based features were extracted from PPG signals. The multitaper method has been used, an extended version of spectral representation and used to overcome classical spectral limitations. Combining high-frequency resolution and low variance gives a more robust spectral estimation than the classical and Welch's periodograms. This can be seen in equation 7. Further, for underlying stationary processes, x_t may include several recurring components given in equation 8. These processes are called conditional or central stationary processes with mixed spectra [63]. The discrete orthogonal increment process expected value dZ(f) is not zero for these processes and can be calculated by equation 9, where δ shows the Dirac delta function. The dZ(f) second central moment can be calculated by equation 10.

$$x_t = \int_{-1/2}^{1/2} e^{-i\omega t}\, dz(t) \qquad (7)$$
$$x_t = \sum_j C_j cos\,(\omega_j + \phi_j) + \xi_t = \sum_j \mu_j e^{-i\omega t} + \mu_j^* e^{-i\omega t} + \xi_t \quad (8)$$
$$E\{dZ\,(f)\} = \sum_j \mu_j \delta\,(f - f_j)\, df, \qquad (9)$$
$$E\left\{|dZ\,(f) - E\{dZ(f)\}|^2\right\} = S(f) df \qquad (10)$$

The multiwindow technique provides a practical yet straightforward likelihood ratio test for the relevance of periodic components in MTM spectral estimation. This method employs numerous data windows known as "discrete prolate spheroidal sequences" and "Slepian sequences" and can be defined as follows:

$$\lambda_k v_n^{(k)}(N, W) = \sum_{m=0}^{N-1} \frac{\sin\{2\pi W (n-m)\}}{\pi (n-m)} v_m^{(k)}(N, W),$$
(11)

where W represent spectral bandwidth, N shows the number of sample points of the PPG signal, and λ_k shows the values linked with the Slepian sequences $v_n^{(k)}(N, W)$ [64]. The Slepian functions can be calculated as in the following Fourier transformation:

$$v_k(f) = \sum_{n=0}^{N-1} v_n^{(k)}(N, W) e^{-i2\pi fn}.$$
(12)

The energy concentration of the above Slepian functions is most significant in the interval $(f - W, f + W)$. Moreover, the bias from all frequencies is far from the frequency of the window width multiplied by the number of observations, so using these sequences to eliminate window leakage is quite successful [65]. As an initial stage, the MTM computes the expansion or eigen coefficients of input Xt as shown below:

$$y_k(f) = \sum_{t=0}^{N-1} x_t v_i^{(k)}(N, W) e^{-i2\pi ft}.$$
(13)

The predicted value of $y_k(f)$ can be calculated by combining the previous formulae.

$$Ey_k(f)\} = \mu V_k(f - f_0) + \mu^* V_k(f + f_0)$$
(14)

$$Ey_k(f_0)\} = \mu V_k(0) + \mu^* V_k(2f_0) \approx \mu V_k(0); \quad where \, f = f_0.$$
(15)

This can be calculated by reducing the residual local squared error, that is, when $f = f_0$. The squared error is defined as follows by equation (16). The results can be provided as equation (17). The following formula (18) can be used to compute the continuous section of the spectrum:

$$e^2(\mu, f) = \sum_{k=0}^{N-1} |y_k(f) - \mu(f) V_k(0)|^2.$$
(16)

$$\hat{\mu}(f) = \frac{\sum_{k=0}^{K-1} V_k^*(0) y_k(f)}{\sum_{k=0}^{K-1} |V_k(0)|^2}.$$
(17)

$$\hat{S}(f) = \frac{1}{K} \sum_{k=0}^{K-1} |y_k(f)|^2.$$
(18)

However, a large line component with frequency F0 necessitates the reconstruction of the spectrum in accordance with equation (10).

$$\hat{S}_r(f) = \frac{1}{K} \sum_{k=0}^{K-1} |y_k(f) - \hat{\mu}(f_0) V_k(f - f_0)|^2.$$
(19)

6) MACHINE LEARNING ALGORITHMS

One of the fastest-growing technology areas today, machine learning lies at the foundation of artificial intelligence and data science, connecting computer science and statistics. Data-intensive learning methods can be applied in science, technology, and commerce, resulting in more evidence-based decision-making in many areas such as health, production, education, financial modelling, law enforcement, and marketing [66]–[68]. Machine learning and biosensor technology have piqued many researchers' interest in this area of research [69]. In this work, we tested a variety of classical and deep learning models. We detail all the machine learning models that we examined in this section.

C. TRADITIONAL MACHINE LEARNING MODELS

We constructed our first traditional machine learning models using Linear Regression (LR), Random Forest (RF), AdaBoost, and Support vector regression (SVR), each with a unique set of characteristics. We employed the linear SVM kernel, one of four significant variations. We provided the results comparison table for all the above-listed algorithms. Algorithm 1 presents the statistical machine learning process for blood pressure estimation given below:

Algorithm 1 SBP and DBP Estimation by Training Traditional Machine Learning
---
Input: PPG biosignal
Output: SBP, DBP
Blood pressure detection process:
1: begin
2:   for signal in signals
3:     Scale PPG signal
4:     Remove outliers
5:     Denoise signal (Butter low pass filtering)
6:     Find Min, Max (Signal)
7:     Scale signal size fixed length
8:       Extract feature vector (Apply different feature extraction techniques)
9:   end for
10: Train LR, RF, Adaboost, SVR

11: return evaluation metrics

---

1) LINEAR REGRESSION (LR)

Linear regression models are trained to assess the problem's linearity. These models are appropriately regularised through the use of a K-fold cross-validation procedure. It is well known that when there is a substantial nonlinear relationship between the feature vector and the target, the final trained models are inapplicable. These models are straightforward, simple to train, and less prone to over-fitting compared to other alternatives. They also entail fewer samples for training than other alternatives, making them more efficient and thus more efficient in terms of implementation [70].

2) RANDOM FOREST

When using ensemble learning methods, the final prediction is created by merging predictions from a number of weak learners, referred to as random forests (e.g., decision trees) [70]. Each tree is trained on a randomly selected subset

of the training data to achieve a low bias and a reasonable prediction variance. The ultimate prediction of a random forest model in a regression problem is the average of the predictions provided by each regression tree. The maximum depth of any tree in this area is unrestricted. Cross-validation was employed to decide the number of trees that will be included in the final regression model [70].

3) ADAPTIVE BOOSTING ALGORITHM (ADABOOST)

The AdaBoost algorithm is the most common and extensively used ensemble learning algorithm, more precisely, of the boosting family of ensemble learning algorithms [71]. AdaBoost's unique feature is that it uses the initial training data to build a weak learner and then adjusts the training data distribution based on the weak learner's prediction performance in the subsequent round of weak learner training. Note that the previous stage's training samples with low predictive accuracy will receive additional attention in the subsequent step. Finally, the weak learners are combined with a strong learner using varied weights.

4) SUPPORT VECTOR MACHINE (SVM)

In its most basic form, the SVM proposed by Vapnik [72] entails the construction of an optimal hyperplane that maximises the division of two distinct classes. Classification models with excellent generalisability are typically constructed using this approach, enhancing their capabilities for a wide range of applications [73]. We used SVM with a linear kernel in our implementation.

a: NEURAL MODELS

We compared three deep learning models for BP estimation using PPG biosignals: LSTM, Bi-LSTM and GRU. The models, as well as their experimental setups and algorithm process, are described further below.

Algorithm 2 SBP and DBP Estimation by Training Deep Learning

Input: PPG biosignal
Output: SBP, DBP
Blood pressure detection process:
1: begin
2:   for signal in signals
3:     Scale PPG signal
4:     Remove outliers
5:     Denoise signal (Butter low pass filtering)
6:     Find Min, Max (Signal)
7:     Scale signal size fixed length
8:     Extract feature vector (Apply different feature extraction techniques)
9:   end for
10: Train LSTM, Bi-LSTM, GRU

11: return evaluation metrics

---

5) LONG SHORT-TERM MEMORY (LSTM)

LSTM, a special RNN structure, has proven stable and robust for long-range modelling dependencies [74]–[76]. The LSTM's primary innovation is its memory cell c_t, which functions essentially as an accumulator of state information.
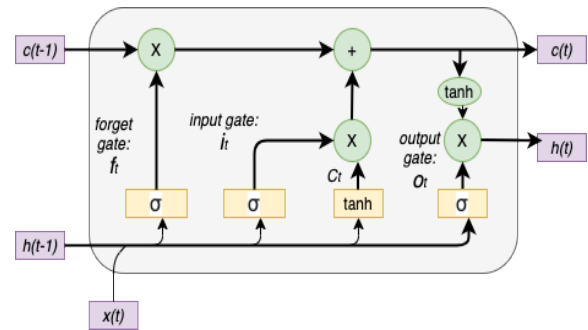

FIGURE 4. Structure of LSTM cell. Where σ presents the sigmoid function, and *i*, *o*, and *f* present the input gate, output gate, and forget gate. Furthermore, *c* and *h* present cell and hidden vectors. The subscripts for the weight matrix are as their name suggests. For instance, W_{xo} shows the input-output gate matrix, W_{hi} shows the hidden-input gate matrix, etc.

Several self-parameterised controlling gates access, write and clear the cell [77]. If the input gate is activated, each time a new input arrives, its information accumulates in the cell. Additionally, if the forget gate f_t is enabled, the previous cell c_{t-1} status may be "forgotten" during this process. The output gate o_t also controls whether the latest cell output is propagated to the final state h_t. A benefit of using a memory cell and gates to control information flow is that the gradient will be trapped within the cell and be kept from vanishing too rapidly, which is a crucial issue for the vanilla RNN model [75], [78]. The LSTM hidden states can be calculated from the equations below, further seen in Figure 4.

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (20)$$

$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f\right). \quad (21)$$

$$c_t = f_tc_{t-1} + i_t tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (22)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (23)$$

$$h_t = o_t tanh(c_t) \quad (24)$$

6) BIDIRECTIONAL LSTM

Bidirectional RNN developed by Schuster and Paliwal [79] is another variation of the RNN to train the network in the past and future using input data sequences. For the processing of input data, two connected layers are used. In the reversed time step direction, each layer performs operations. The results can be combined with various merging techniques. Likewise, Bidirectional LSTM has two layers so that one layer operates in the same data sequence direction and the second layer operating in the reverse sequence direction [80].

In some applications, such as phoneme classification, BLSTM was more effective than unidirectional LSTM [81]. For Bi-LSTM calculation, this can combine forward and backwards hidden layers, as shown in Figure. This connection can achieve the temporal flow of information in both directions with better network learning [82].

7) GATED RECURRENT UNITS (GRU)

GRU is a variation of RNN that has been developed to deal with the issue of long term dependency [83], shown
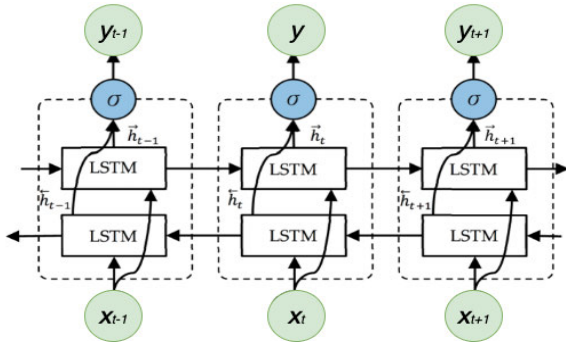
FIGURE 5. Structure of Bidirectional LSTM. The forward LSTM layer output $\vec{h}$ is calculated similarly as the unidirectional one, while the backward LSTM layer output $\overleftarrow{h}$ is computed utilising reversed inputs from time $t-1$ to $t-n$. Then the outputs are fed to σ function to achieve output vector y_t.
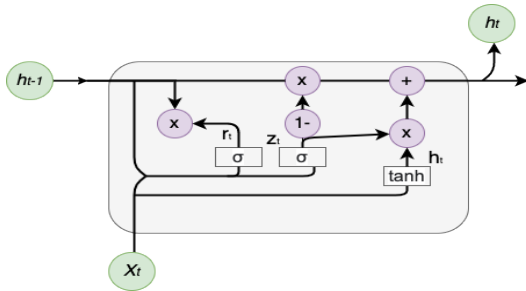

FIGURE 6. Structure of GRU cell.

in Figure. GRU is similar to LSTM. The GRU controls the information flow using several gates integrated into its cell and outperforms standard RNNs. However, the GRU offers a few benefits over the LSTM. First, it has two gates (rest gate and updated gate). Second, the GRU does not have a memory cell which makes it computationally faster than LSTM. It is simple to implement and compute. Like LSTM, GRU at each time step has two inputs: the output of the previous hidden state and the current input. Reset gate z and an update gate r combines input and forget gates. These both gates can be computed as follows:

$$r_t = \delta \left(W_r h_{t-1} + U_r x_t + b_r \right). \tag{25}$$

$$z_t = \delta \left(W_z h_{t-1} + U_z x_t + b_z \right). \tag{26}$$

where δ presents the logistic sigmoid function. The update gate chooses the volume of information that needs to be passed to the GRU current state. Moreover, the reset gate determines that when the previous hidden state needs to be ignored, that can be computed as given below:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \tag{27}$$

$$\tilde{h}_t = tanh \left(W_{\tilde{h}_t} \left(h_{t-1} \odot r_t \right) + U_{\tilde{h}_t} x_t \right) \tag{28}$$

IV. EXPERIMENTAL SETUP

This section discusses the findings of numerous machine learning algorithms tested on two different datasets that

TABLE 2. Traditional and deep machine learning algorithm best parameter setting.

Feature extraction technique & Models	Time-domain	Statistical features	Frequency domain
Best Parameters			
Random Forest Estimator Depth	(60, 1) (10, 10)	(110, 10) (120, 10)	(100, 10) (120, 10)
AdaBoost Estimator Learning rate	(40, 0.1) (20, 0.1)	(40, 1.0) (20, 0.1)	(20, 0.6) (20, 0.1)
SVR C, Kernal	(1.0, linear)	(1.0, linear)	(1.0, linear)
Number of hidden layers & learning rate			
Vanilla LSTM	(256,256) (0.001)	(64,256,256) (0.001)	(64,256,256) (0.001)
BiLSTM	(128,128) (0.001)	(512,512,512) (0.001)	(128,128,128) (0.001)
GRU	(64,512) (0.001)	(218,512) (0.001)	(64,256,512) (0.001)

contained multiple feature extraction strategies for each PPG signal. The mean absolute error (MAE), root mean square error (RMSE) and standard deviation of the error (SD) are the estimation accuracy metrics that can be used to evaluate the accuracy of the estimation. It provided the BP model's performance to facilitate comparisons across feature extraction approaches and machine learning algorithms. University of Tasmania[1] (UTAS) cluster and Google colaboratory[2] (Colab pro) were used to generate the experimental results on the MIMIC II and PPG-BP datasets, respectively. Further, Python 3.7[3] was the programming language that was used for the experimentation and testing. In order to better understand performance, the results tables have been presented in this section.

A. EVALUATION METRICS

In this study, blood pressure estimation was conducted by extracting features from PPG signals without a cuff. Estimation performance was evaluated using MAE, RMSE and SD performance criteria. It is the average of the absolute difference between the data's actual and anticipated values and is calculated as follows: A residual average is calculated by averaging all the residuals in a dataset where $\hat{y}$ is the predicted value of y and $\bar{y}$ is the mean of y value:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| y_i - \hat{y} \right| \tag{29}$$

$$RMSE = \sqrt[2]{\frac{1}{n} * \sum_{i=1}^{n} \left(y_i - \hat{y} \right)^2} \tag{30}$$

The standard deviation (SD) is a measure of how much a group of values varies or disperses. Standard deviation

[1] https://www.utas.edu.au/
[2] https://colab.research.google.com/
[3] https://www.python.org/

TABLE 3. Time domain-based (21 features) and machine learning algorithms error for SBP and DBP estimation.

Machine Learning Models	PPG-BP dataset				MIMIC dataset			
	SBP		DBP		SBP		DBP	
	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE
LR	12.77±3.42 mmHg	16.65	8.35±2.12 mmHg	10.41	12.90±10.35 mmHg	16.84	6.60±3.88 mmHg	9.35
Random Forest	11.59±3.24 mmHg	15.35	8.46±2.35 mmHg	10.53	8.85±17.76 mmHg	13.03	**4.35+7.10 mmHg**	7.24
SVM	15.14±3.93 mmHg	17.08	8.06±2.17 mmHg	10.37	15.71±11.07 mmHg	18.16	7.58±4.30 mmHg	9.51
AdaBoost	12.24±5.64 mmHg	16.41	7.94±3.49 mmHg	9.88	12.47±10.91 mmHg	15.86	6.69±4.84 mmHg	9.16
Vanilla LSTM	5.62±5.00 mmHg	7.53	5.20±5.43 mmHg	8.16	4.96±5.60 mmHg	7.18	7.35±5.89 mmHg	9.22
Bi-LSTM	4.93±4.40 mmHg	7.38	**4.18+4.43 mmHg**	8.23	**3.87+4.79 mmHg**	6.25	6.82±6.29 mmHg	8.04
GRU	**3.68+4.28 mmHg**	6.11	5.34±5.24 mmHg	8.22	4.49±5.58 mmHg	7.56	7.10±5.85 mmHg	9.00

TABLE 4. Time domain-based (11 features) and machine learning algorithms error for SBP and DBP estimation.

Machine Learning Models	PPG-BP dataset				MIMIC dataset			
	SBP		DBP		SBP		DBP	
	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE
LR	12.96±3.08 mmHg	16.80	8.31±1.57 mmHg	9.98	11.40±9.38 mmHg	14.63	6.67±3.67 mmHg	8.51
Random Forest	10.20±9.7 mmHg	12.77	8.36±1.22 mmHg	9.65	9.40±7.90 mmHg	12.56	4.63±6.07 mmHg	7.62
SVM	15.48±3.52 mmHg	17.01	8.14±1.54 mmHg	9.33	12.14±6.67 mmHg	13.88	6.65±1.16 mmHg	8.22
AdaBoost	12.00±5.43 mmHg	15.99	8.26±2.47 mmHg	8.15	12.99±10.71mmHg	13.76	**6.61+4.75 mmHg**	8.20
Vanilla LSTM	5.68±4.72 mmHg	7.35	5.85±5.27 mmHg	7.88	5.34±5.62 mmHg	8.48	8.01±5.80 mmHg	10.72
Bi-LSTM	**5.32+4.26 mmHg**	7.50	**5.84+5.35 mmHg**	7.92	5.59±5.92 mmHg	8.86	8.06±5.88 mmHg	10.80
GRU	5.42±4.32 mmHg	7.33	6.25±5.26 mmHg	8.17	**5.25+5.63 mmHg**	8.42	6.63±5.73 mmHg	10.37

measures how close a set's values tend to be to its mean (also known as anticipated value). On the other hand, a high standard deviation implies that the values are dispersed over a broader range of values. Expected values' standard deviation (SD) aids in understanding the dispersion of predicted values in a variety of models. This can be calculated as given below:

$$SD = \sqrt{\frac{\sum (x - \bar{x})^2}{N - 1}} \qquad (31)$$

B. PARAMETERS SETTING

In order to achieve high-performance accuracy, we used parameters setting and listed the best parameters listed in the tables given below for traditional machine learning and deep learning models. We used a grid search algorithm for traditional machine learning algorithms and selected the best parameters for model tunning. Similarly, each dataset was divided into train, validation, and test sets for deep learning to prevent overfitting and parameter optimisation. The accuracy of the test dataset was used to determine which model parameters were the most accurately optimised. Further, they were compared against machine-learning algorithms to see which was the best. The table provided below depicts the best parameters were found for each feature extraction technique.

C. RESULTS AND DISCUSSION

This section focuses on the performance analyses of traditional and deep learning models. We divided each dataset into three parts: training, validation, and testing to carry out these experiments.

We used parameter tuning for better performance for each machine learning model. This section is further divided into feature extraction groups to analyse the performance on two public datasets. We used mean absolute error (MAE), root mean square error (RMSE) and standard deviation (SD) to evaluate the models and compare the performance differences.

1) TIME DOMAIN-BASED FEATURES (GROUP A) PERFORMANCE

We extracted 21 time domain-based features from PPG signals and applied machine learning algorithms in this group. In addition to this, we applied feature importance algorithms and, following a thorough analysis of each feature importance technique's results, only 11 features were selected. Further, these datasets were given as input to machine learning algorithms to estimate blood pressure. Therefore, we constructed two datasets with the reference values of SBP and DBP, one with a set of 21 features and one with only a set of 11 features for group A features. Table 3, 4 and Figure 7 have been provided for the results for each algorithm.

The results showed that LSTM, Bi-LSTM and GRU achieved better accuracy as compared to traditional machine learning algorithms. The GRU performance showed that the average mean error was 3.68±4.28 mmHg and 5.34±5.25 mmHg for systolic BP and diastolic BP measurement using 21 features. However, after feature importance, eleven selected features were applied to all these

21- Time- domain based features
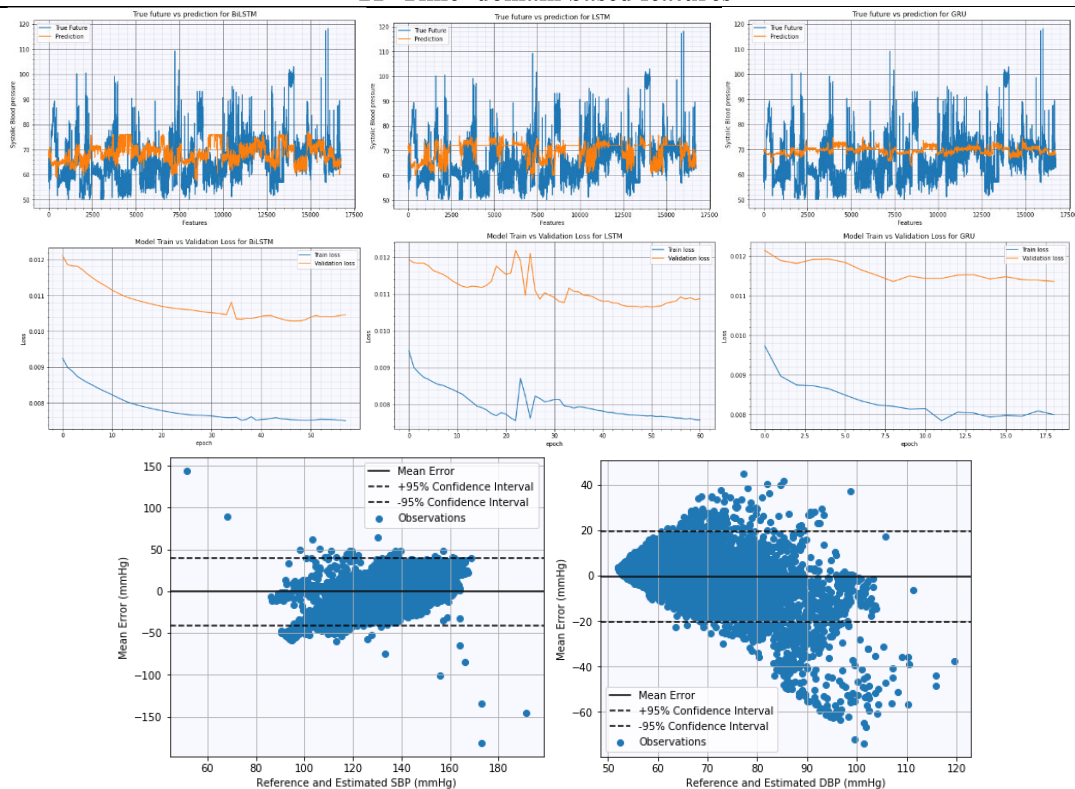
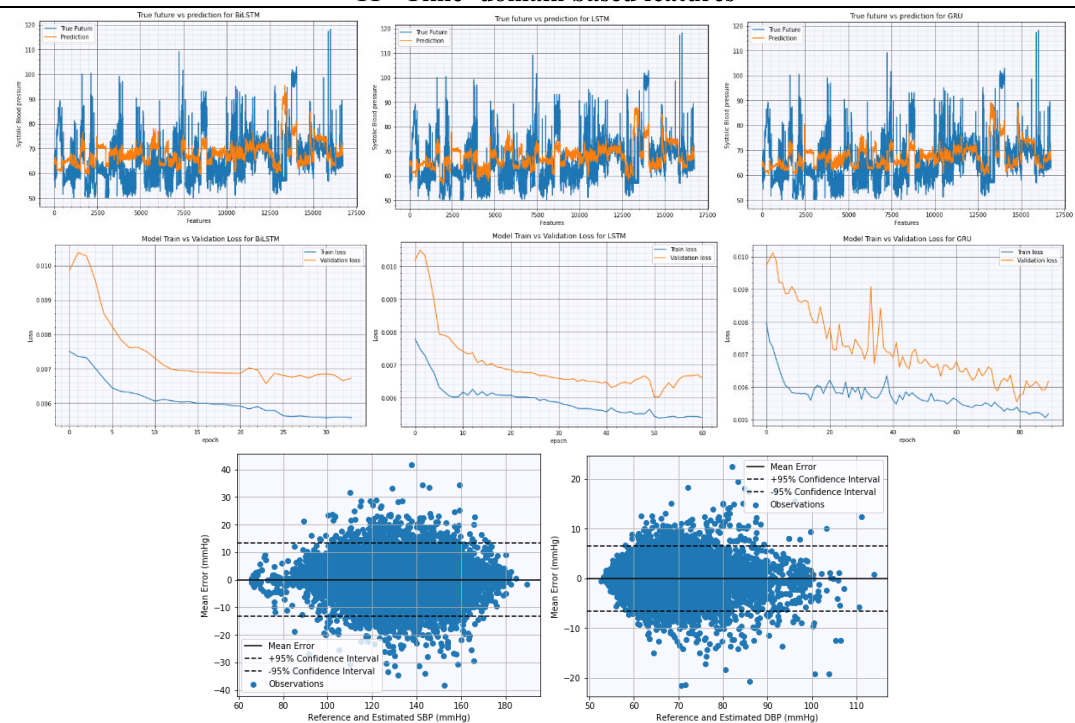## 11 - Time- domain based features

FIGURE 7. Time domain-based features before and after feature importance for Bi-LSTM, LSTM, and GRU with actual future and prediction values. There are also model training and validation results for all three deep learning models. Regression results analysis for SBP and DBP are provided by using Bland-Altman plots. The average estimated SBP and DBP values have been demonstrated with average mean error. The dotted line depicts the mean error difference between true and predicted values.
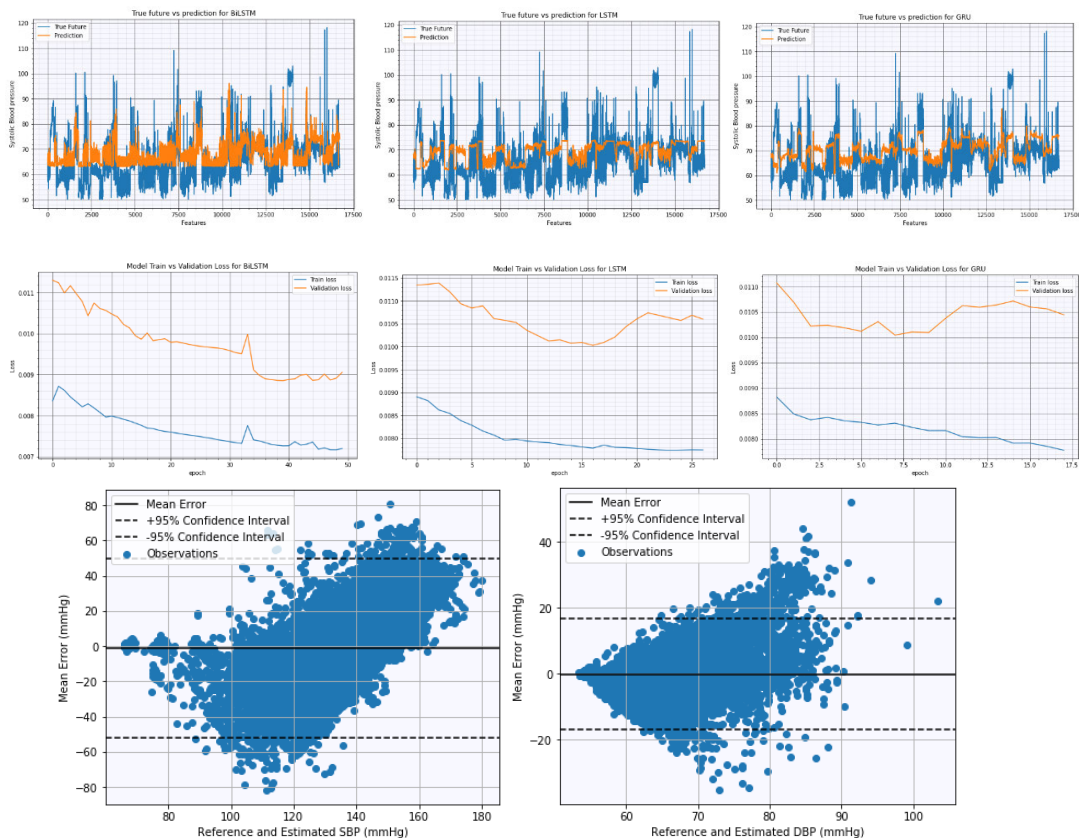
Statistical feature extraction

FIGURE 8. Statistical characteristics for Bi-LSTM, LSTM, and GRU. All three deep learning models have model train and validation outcomes. Bland-Altman charts are used to analyse SBP and DBP regression findings. The dotted line shows the mean error difference between the true and anticipated SBP and DBP values.

TABLE 5. Statistical feature extraction and machine learning algorithms error for SBP and DBP estimation.

Machine Learning Models	PPG-BP dataset				MIMIC dataset			
	SBP		DBP		SBP		DBP	
	MAE+ SD (mmHg)	RMSE	MAE+ SD (mmHg)	RMSE	MAE+ SD (mmHg)	RMSE	MAE+ SD (mmHg)	RMSE
LR	12.75+5.10 mmHg	16.06	7.33+1.36 mmHg	9.78	11.90+7.64 mmHg	15.88	6.94+2.87 mmHg	9.66
Random Forest	11.52+3.72 mmHg	15.88	7.24+1.38 mmHg	9.60	11.84+12.33 mmHg	15. 83	**6.17+4.46 mmHg**	9.27
SVM	15.56+1.60 mmHg	17.55	6.78+0.79 mmHg	8.78	10.98+7.52 mmHg	14.94	6.98+3.05 mmHg	8.34
AdaBoost	12.12+4.64 mmHg	15.99	7.23+2.06 mmHg	9.63	12.77+6.75 mmHg	16.11	7.03+3.08 mmHg	9.58
Vanilla LSTM	**4.70+5.21 mmHg**	6.73	**4.68+5.34 mmHg**	6.83	6.48+5.02 mmHg	8.01	6.32+5.50 mmHg	9.16
Bi-LSTM	4.83+4.28 mmHg	6.24	4.87+5.39 mmHg	6.94	**5.70+4.87 mmHg**	7.28	7.38+5.51 mmHg	9.21
GRU	5.20+4.37 mmHg	5.80	4.74+5.26 mmHg	6.79	7.02+4.84 mmHg	8.36	6.44+5.50 mmHg	9.25

algorithms that showed more average error than 21 based features findings. The best MAE and SD were achieved from Bi-LSTM for eleven features set using the PPG-BP dataset. The above findings conclude that 21 features can perform well as compared to the eleven features dataset.

2) STATISTICAL FEATURE EXTRACTION (GROUP B)
These features were extracted by statistical calculations for signal analysis. Figure 8 and Table 5 have been provided for

linear regression, random forest, AdaBoost, SVM, LSTM, Bi-LSTM and GRU to visualise the results.

The findings can be seen from Table 5 of the statistical features datasets achieved good accuracy using deep learning models for both datasets. For the PPG-BP dataset, vanilla LSTM performed better among other algorithms that portray mean absolute error + standard deviation of 4.70+5.21 mmHg and 4.68+5.34 mmHg for SBP and DBP, respectively. However, random forest performed better
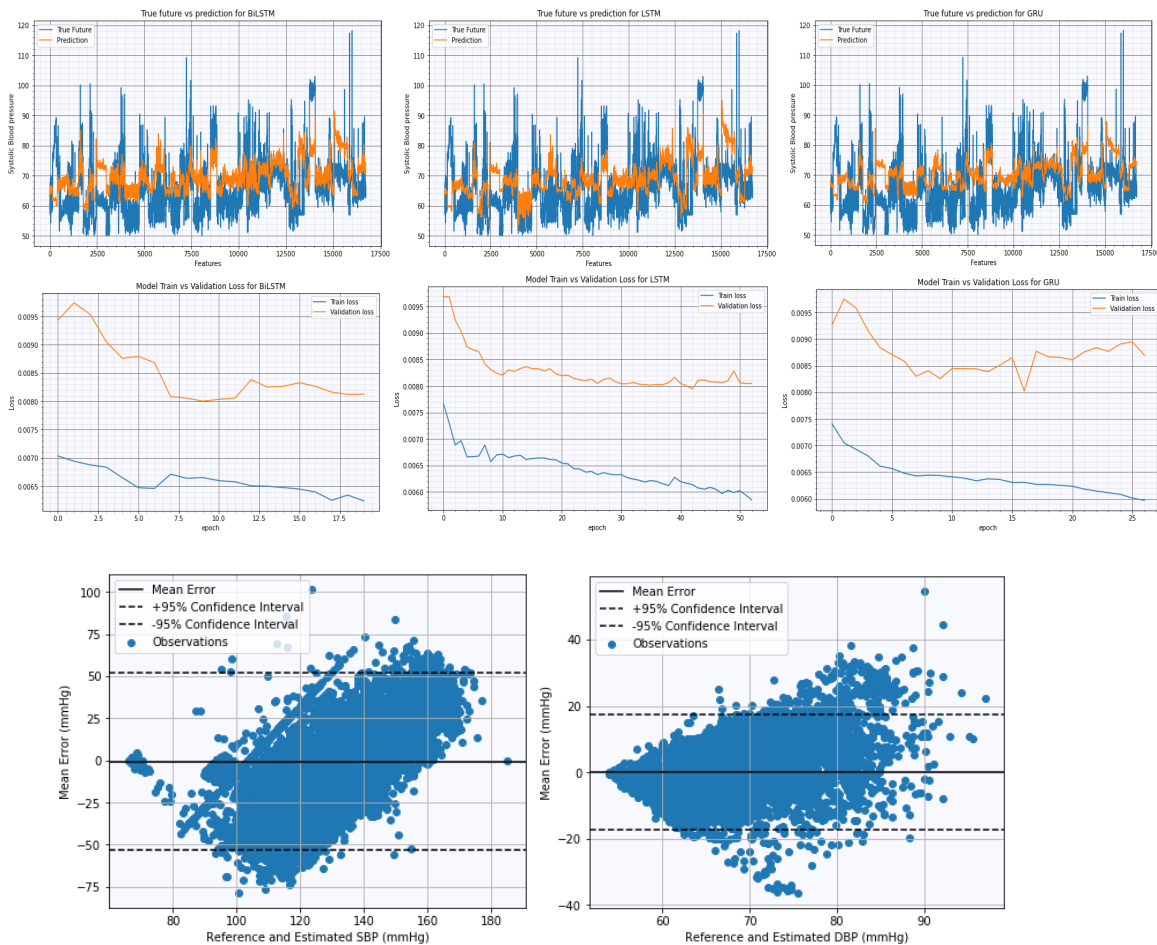
Frequency domain-based features

FIGURE 9. Bi-LSTM, LSTM, and GRU with real future and prediction values. Additionally, model training and validation results are provided for each of the three deep learning models. Bland-Altman plots are used to visualise the regression results for SBP and DBP. The average estimated SBP and DBP values have been plotted with the average mean error; the dotted line represents the mean error difference between the true and projected values.

than other machine learning algorithms for the MIMIC dataset for DBP and Bi-LSTM for SBP estimation.

3) FREQUENCY-DOMAIN BASED FEATURES (GROUP C)

In this group, frequency-based features have been discussed. We used the multitaper method to extract these features. Further, we calculated band power and relative band power.

The resultant features set showed a complex input dataset provided to the machine learning algorithms for systole and diastole estimation. Mean absolute error and standard deviation have been calculated to analyse the results given in the table below. Figure 9 and Table 6 have been provided for further analysis of these results. Frequency domain features depict better accuracy than statistical features extracted from PPG signals, But not better than time domain-based features. For the PPG-BP dataset, vanilla LSTM performed better than

all others and portrayed an average error of 4.60+5.17 mmHg and 4.98+5.52 mmHg for systolic and diastolic blood pressure predictions. On the other side, for the MIMIC dataset, Bi-LSTM gave an error of 5.42+5.21 mmHg and 6.17+5.89 mmHg for SBP and DBP, respectively.

D. COMPARISON WITH RELATED WORK

Over the last decade, numerous methods for determining cuffless blood pressure have been proposed as alternatives to conventional blood pressure measurement techniques. Most of the literature studies adopted PPG and ECG signals for BP estimation using some methods such as pulse transit time (PTT) and pulse arrival time (PAT). However, these techniques cannot be used in clinical practice because they depend on individual physiological parameters that require calibration. Furthermore, these methods require two sensors for measurement, which can increase the computational

TABLE 6. Frequency domain features and machine learning algorithms error for SBP and DBP estimation.

Machine Learning Models	PPG-BP dataset				MIMIC dataset			
	SBP		DBP		SBP		DBP	
	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE	MAE± SD (mmHg)	RMSE
LR	12.35±0.04 mmHg	15.55	8.65±0.02 mmHg	10.68	11.68±8.99 mmHg	15.82	7.09±2.14 mmHg	10.09
Random Forest	11.37±0.11 mmHg	14.60	8.65±0.16 mmHg	10.72	14.71±12.35 mmHg	17.01	6.46±4.41 mmHg	9.25
SVM	15.98±0.00 mmHg	17.47	8.27±0.00 mmHg	10.35	12.61±8.36 mmHg	16.10	7.04±2.43 mmHg	9.98
AdaBoost	12.77±0.47 mmHg	15.98	8.75±0.32 mmHg	10.86	11.09±8.59 mmHg	14.88	7.05±3.41 mmHg	10.01
Vanilla LSTM	**4.60±5.17 mmHg**	7.62	**4.98±5.52 mmHg**	7.14	6.05±5.43 mmHg	8.90	7.37±5.74 mmHg	9.26
Bi-LSTM	5.84±4.31 mmHg	7.26	5.24±5.50 mmHg	7.32	**5.42±5.21 mmHg**	7.07	**6.17±5.89 mmHg**	8.77
GRU	5.33±4.61 mmHg	7.05	6.02±5.05 mmHg	8.17	7.80±5.04 mmHg	9.29	7.92±5.59 mmHg	9.93

TABLE 7. Comparison of related work for blood pressure estimation using the MIMIC dataset.

Algorithm	Feature Extraction	Error (MAE + SD)	
		SBP	DBP
AdaBoost [70]	• Pulse Arrival Time	8.21±5.45 mmHg	4.31±3.52 mmHg
SVM [34]	• SDPPG • Time Scale	8.54±10.9 mmHg	4.34±5.8 mmHg
ResNet [84]	• SDPPG	15.41 mmHg	12.38 mmHg
MIMIC-II dataset (Ours)			
Bi-LSTM	• Time-domain features	3.87±4.79 mmHg	6.82±6.29 mmHg
	• Statistical Features	5.70±4.87 mmHg	7.38±5.51 mmHg
	• Frequency-domain features	5.42±5.21 mmHg	6.17±5.89 mmHg
GRU	• Time-domain features	4.49±5.58 mmHg	7.10±5.85 mmHg
	• Statistical Features	7.02±4.84 mmHg	6.44±5.50 mmHg
	• Frequency-domain features	7.80±5.04 mmHg	7.92±5.59 mmHg

complexity. This inclined researchers towards PPG-based blood pressure estimation, which looks promising.

Table 7 presents a performance assessment with literature work in this research area. A number of critical factors such as calibration, dataset size, different evaluation metrics, or distinct techniques make it challenging to compare our work directly to other relevant studies. Furthermore, the MIMIC II dataset have been extensively used in literature; however, PPG-BP is a new dataset, and very few studies have been conducted using this dataset so far. For this, we provide a comparison of our work with some calibration-free studies for BP estimation, for which they used the MIMIC public dataset and mean error (MAE or ME) as evaluation metrics. Kachuee, et al. [70] used the PAT method to estimate calibration-free SBP, DBP and MAP using several machine learning algorithms. The proposed study complied with AAMI standards for MAP and DBP estimation. Nevertheless, SBP estimation did not meet AAMI standards. Moreover, this method requires two signals for calculations from two different sensors. Liu, et al. [34] extracted second derivative features from PPG signals and also used

time-domain features for BP estimation. The results error was recorded 8.54±10.9 mmHg for SBP and 4.34±5.8 mmHg for DBP using 35 features. Obtaining the correct SDPPG features values, which are solely dependent on the visibility of the five peak points, is the most challenging aspect of this technique.

Slapničar, et al. [84] suggested a ResNet-based approach for BP estimating using first and second derivatives of PPG signals as an input. They tested the model on 510 individuals' data from the MIMIC III database. The input signals were analysed by a spectro-temporal block, five ResNet blocks, and the GRU layer in the sequence. The proposed technique complexity level was high. Moreover, it achieved low performance as compared to the state-of-art methods. For evaluation metrics, the MAE was 15.41 mmHg, 12.38 mmHg for SBP and DBP, respectively. They also did not provide standard deviation values. We can infer from the literature that traditional features extraction has been a significant factor in this research area, affecting the performance of the BP estimation models. Inspired by this, we presented a study of feature extraction techniques comparison for PPG signals for BP estimation using two benchmark datasets. Overall, given the high number of patient data and only one PPG signal usage, our proposed comparison study suggested that LSTM and GRU models performed well with time-domain features. Moreover, this demonstrates that a single PPG signal can estimate BP with reasonable accuracy.

V. CONCLUSION AND FUTURE WORK

This paper investigates the relationship between features extracted from PPG signals and their capacity to detect systolic and diastolic blood pressure levels. Studies have demonstrated that using ECG and PPG combined can help determine blood pressure directly, although synchronising both signals at the same time can be difficult. In addition, the computation of blood pressure needs the placement of several sensors on the body, which is a complex process to undertake. Therefore, the more sensors we have, the more difficult it is to compute blood pressure. As a result, our goal in this research is to detect blood pressure solely by using PPG signals. To achieve this, we conducted a comprehensive study to find the best feature extraction technique and machine learning algorithm for blood pressure estimation.

The findings of this study demonstrated that time domain-based features achieved better accuracy among other feature extraction techniques. More importantly, deep learning algorithms performed better to achieve high performance with less error for SBP and DBP estimation. Specifically, GRU and Bi-LSTM with time-domain features achieved the best performance for SBP and DBP estimation, respectively, than all other models for all datasets. The results and conclusions drawn from our comparative analysis can aid future research in this area and also assist researchers in selecting the best appropriate model for blood pressure estimation. Our future direction will focus on designing models that can minimise the error for more robust results. Further, we will investigate the attention process and weights to determine which signal's waveform sequence significantly impacts the desired result.

REFERENCES

[1] K. Yeates, L. Lohfeld, J. Sleeth, F. Morales, Y. Rajkotia, and O. Ogedegbe, "A global perspective on cardiovascular disease in vulnerable populations," *Can. J. Cardiol.*, vol. 31, no. 9, pp. 1081–1093, Sep. 2015.

[2] C. Dalvi, M. M. Lamego, M. J. E. Kiani, J. Poeze, and H. Vo, "Blood pressure monitor with valve-chamber assembly," Google Patent 10 555 678, Feb. 11, 2020.

[3] A. C. Rodriguez, Y. R. van Zeeland, N. J. Schoemaker, and J. C. de Grauw, "Agreement between invasive and oscillometric arterial blood pressure measurement using a high-definition oscillometric device in normotensive New Zealand white rabbits using two different anaesthetic protocols," *Veterinary Anaesthesia Analgesia*, vol. 48, no. 5, pp. 679–687, Sep. 2021.

[4] A. Meidert and S. Bernd, "Techniques for non-invasive monitoring of arterial blood pressure," *Frontiers Med.*, vol. 4, p. 231, Jan. 2018.

[5] B. V. Scheer, A. Perel, and U. J. Pfeiffer, "Clinical review: Complications and risk factors of peripheral arterial catheters used for haemodynamic monitoring in anaesthesia and intensive care medicine," *Crit. Care*, vol. 6, no. 3, p. 199, 2002.

[6] J. Penaz, A. Voigt, and W. Teichmann, "Contribution to the continuous indirect blood pressure measurement," *Zeitschrift Gesamte Innere Medizin Ihre Grenzgebiete*, vol. 31, no. 24, pp. 1030–1033, 1976.

[7] B. Saugel, R. Dueck, and J. Wagner, "Measurement of blood pressure," *Best Pract. Res. Clin. Anaesthesiol.*, vol. 28, no. 4, pp. 309–322, 2014.

[8] J. Y. Wagner, J. Grond, J. Fortin, I. Negulescu, M. Schöfthaler, and B. Saugel, "Continuous noninvasive cardiac output determination using the CNAP system: Evaluation of a cardiac output algorithm for the analysis of volume clamp method-derived pulse contour," *J. Clin. Monit. Comput.*, vol. 30, no. 4, pp. 487–493, 2016.

[9] T. B. Plante, A. C. O'Kelly, B. Urrea, Z. T. MacFarlane, R. S. Blumenthal, J. Charleston, E. R. Miller, L. J. Appel, and S. S. Martin, "User experience of instant blood pressure: Exploring reasons for the popularity of an inaccurate mobile health app," *npj Digit. Med.*, vol. 1, no. 1, p. 31, Dec. 2018.

[10] M. M. Baig and H. Gholamhosseini, "Smart health monitoring systems: An overview of design and modeling," *J. Med. Syst.*, vol. 37, no. 2, p. 9898, 2013.

[11] S. Rajala, T. Ahmaniemi, H. Lindholm, and T. Taipalus, "Pulse arrival time (PAT) measurement based on arm ECG and finger PPG signals–comparison of PPG feature detection methods for PAT calculation," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 250–253.

[12] J. Y. Kim, B. H. Cho, S. M. Im, M. J. Jeon, I. Y. Kim, and S. I. Kim, "Comparative study on artificial neural network with multiple regressions for continuous estimation of blood pressure," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 6942–6945.

[13] X. Ding, B. P. Yan, Y.-T. Zhang, J. Liu, N. Zhao, and H. K. Tsang, "Pulse transit time based continuous cuffless blood pressure estimation: A new extension and a comprehensive evaluation," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Dec. 2017.

[14] H. Gesche, D. Grosskurth, G. Küchler, and A. Patzak, "Continuous blood pressure measurement by using the pulse transit time: Comparison to a cuff-based method," *Eur. J. Appl. Physiol.*, vol. 112, no. 1, pp. 309–315, 2012.

[15] L. A. Geddes, M. H. Voelz, C. F. Babbs, J. D. Bourland, and W. A. Tacker, "Pulse transit time as an indicator of arterial blood pressure," *Psychophysiology*, vol. 18, no. 1, pp. 71–74, 1981.

[16] S. N. Shukla, K. Kakwani, A. Patra, B. K. Lahkar, V. K. Gupta, A. Jayakrishna, P. Vashisht, and I. Sreekanth, "Noninvasive cuffless blood pressure measurement by vascular transit time," in *Proc. 28th Int. Conf. VLSI Design*, Jan. 2015, pp. 535–540.

[17] K. Matsumura, P. Rolfe, S. Toda, and T. Yamakoshi, "Cuffless blood pressure estimation using only a smartphone," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, Dec. 2018.

[18] G. Zhang, M. Gao, D. Xu, N. B. Olivier, and R. Mukkamala, "Pulse arrival time is not an adequate surrogate for pulse transit time as a marker of blood pressure," *J. Appl. Physiol.*, vol. 111, no. 6, pp. 1681–1686, 2011.

[19] R. A. Payne, C. N. Symeonides, D. J. Webb, and S. R. J. Maxwell, "Pulse transit time measured from the ECG: An unreliable marker of beat-to-beat blood pressure," *J. Appl. Physiol.*, vol. 100, no. 1, pp. 136–141, 2006.

[20] K. Shelley, S. Shelley, and C. Lake, "Pulse oximeter waveform: Photoelectric plethysmography," *Clin. Monitor.*, pp. 420–428, 2001.

[21] K. H. Shelley, D. G. Silverman, and A. J. Shelley, "Volume status monitor: Peripheral venous pressure, hypervolemia and coherence analysis," Google Patent 8 727 997, May 20, 2014.

[22] T. K. Aldrich, M. Moosikasuwan, S. D. Shah, and K. S. Deshpande, "Length-normalized pulse photoplethysmography: A noninvasive method to measure blood hemoglobin, bronx, NY," *Ann. Biomed. Eng.*, vol. 30, no. 10, pp. 1291–1298, Nov. 2002.

[23] G. Yoon, J. Y. Lee, K. J. Jeon, K.-K. Park, H. S. Yeo, H. T. Hwang, H. S. Kim, and I.-D. Hwang, "Multiple diagnosis based on photoplethysmography: Hematocrit, SpO_2, pulse, and respiration," *Proc. SPIE*, vol. 4916, pp. 185–188, Sep. 2002.

[24] I. Reyes, H. Nazeran, M. Franco, and E. Haltiwanger, "Wireless photoplethysmographic device for heart rate variability signal acquisition and analysis," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 2092–2095.

[25] S. Koh, B. R. Cho, J.-I. Lee, S.-O. Kwon, S. Lee, J. B. Lim, S. B. Lee, and H.-D. Kweon, "Driver drowsiness detection via PPG biosignals by using multimodal head support," in *Proc. 4th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2017, pp. 0383–0388.

[26] F. Rundo, S. Conoci, A. Ortis, and S. Battiato, "An advanced bio-inspired photoplethysmography (PPG) and ECG pattern recognition system for medical assessment," *Sensors*, vol. 18, no. 2, p. 405, Jan. 2018.

[27] E. Monte-Moreno, "Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques," *Artif. Intell. Med.*, vol. 53, no. 2, pp. 127–138, Oct. 2011.

[28] J. C. Ruiz-Rodríguez, A. Ruiz-Sanmartín, V. Ribas, J. Caballero, A. García-Roche, J. Riera, X. Nuvials, M. de Nadal, O. de Sola-Morales, J. Serra, and J. Rello, "Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology," *Intensive Care Med.*, vol. 39, no. 9, pp. 1618–1625, Sep. 2013.

[29] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[30] X. F. Teng and Y. T. Zhang, "Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 4, Sep. 2003, pp. 3153–3156.

[31] *American National Standards for Electronic or Automated Sphygmomanometers*, document ANSI/AAMI SP 10-1987, A. F. T. A. O. M. Instrumentation, 1987.

[32] S. Suzuki and K. Oguri, "Cuffless blood pressure estimation by error-correcting output coding method based on an aggregation of AdaBoost with a photoplethysmograph sensor," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2009, pp. 6765–6768.

[33] Y. Kurylyak, F. Lamonaca, and D. Grimaldi, "A neural network-based method for continuous blood pressure estimation from a PPG signal," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (IMTC)*, May 2013, pp. 280–283.

[34] M. Liu, L.-M. Po, and H. Fu, "Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative," *Int. J. Comput. Theory Eng.*, vol. 9, no. 3, p. 202, 2017.

[35] L. A. Bortolotto, J. Blacher, T. Kondo, K. Takazawa, and M. E. Safar, "Assessment of vascular aging and atherosclerosis in hypertensive subjects: Second derivative of photoplethysmogram versus pulse wave velocity," *Amer. J. Hypertension*, vol. 13, no. 2, pp. 165–171, Feb. 2000.

[36] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, "Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform," *Hypertension*, vol. 32, no. 2, pp. 365–370, 1998.

[37] S. G. Khalid, J. Zhang, F. Chen, and D. Zheng, "Blood pressure estimation using photoplethysmography only: Comparison between different machine learning approaches," *J. Healthcare Eng.*, vol. 2018, pp. 1–13, Oct. 2018.

[38] J. Dey, A. Gaurav, and V. N. Tiwari, "InstaBP: Cuff-less blood pressure monitoring on smartphone using single PPG sensor," in *Proc. 40th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2018, pp. 5002–5005.

[39] C. Yi, C. Jian, and J. Wenqiang, "Continuous blood pressure measurement based on photoplethysmography," in *Proc. 14th IEEE Int. Conf. Electron. Meas. Instrum. (ICEMI)*, Nov. 2019, pp. 1656–1663.

[40] Y. Liang, Z. Chen, R. Ward, and M. Elgendi, "Photoplethysmography and deep learning: Enhancing hypertension risk stratification," *Biosensors*, vol. 8, no. 4, p. 101, Oct. 2018.

[41] U. Senturk, K. Polat, and I. Yucedag, "A non-invasive continuous cuffless blood pressure estimation using dynamic recurrent neural networks," *Appl. Acoust.*, vol. 170, Dec. 2020, Art. no. 107534.

[42] Ü. Şentürk, K. Polat, and İ. Yücedağ, "Towards wearable blood pressure measurement systems from biosignals: A review," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 27, no. 5, pp. 3259–3281, Sep. 2019.

[43] U. Senturk, I. Yucedag, and K. Polat, "Repetitive neural network (RNN) based blood pressure estimation using PPG and ECG signals," in *Proc. 2nd Int. Symp. Multidisciplinary Stud. Innov. Technol. (ISMSIT)*, Oct. 2018, pp. 1–4.

[44] C.-H. Goh, L. K. Tan, N. H. Lovell, S.-C. Ng, M. P. Tan, and E. Lim, "Robust PPG motion artifact detection using a 1-D convolution neural network," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105596.

[45] M. Radha, K. de Groot, N. Rajani, C. C. P. Wong, N. Kobold, V. Vos, P. Fonseca, N. Mastellos, P. A Wark, N. Velthoven, R. Haakma, and R. M Aarts, "Estimating blood pressure trends and the nocturnal dip from photoplethysmograph," 2018, *arXiv:1805.09121*. [Online]. Available: https://arxiv.org/abs/1805.09121

[46] M. Radha, K. de Groot, N. Rajani, C. C. P. Wong, N. Kobold, V. Vos, P. Fonseca, N. Mastellos, P. A. Wark, N. Velthoven, R. Haakma, and R. M. Aarts, "Estimating blood pressure trends and the nocturnal dip from photoplethysmography," *Physiol. Meas.*, vol. 40, no. 2, Feb. 2019, Art. no. 025006.

[47] Y. Liang, Z. Chen, G. Liu, and M. Elgendi, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," *Sci. Data*, vol. 5, no. 1, Dec. 2018, Art. no. 180020.

[48] M. Saeed, M. Villarroel, A. T. Reisner, G. Clifford, L.-W. Lehman, G. Moody, T. Heldt, T. H. Kyaw, B. Moody, and R. G. Mark, "Multiparameter intelligent monitoring in intensive care II (MIMIC-II): A public-access intensive care unit database," *Crit. Care Med.*, vol. 39, no. 5, p. 952, 2011.

[49] *Study of Signal Denoising Using Kaiser Window and Butterworth Filter*, I. B. PTU, India, 2012.

[50] S. Nara, M. Kaur, and K. L. Verma, "Novel notch detection algorithm for detection of dicrotic notch in PPG signals," *Int. J. Comput. Appl.*, vol. 86, no. 17, pp. 36–39, Jan. 2014.

[51] C. El-Hajj and P. Kyriacou, "Deep learning models for cuffless blood pressure monitoring from PPG signals using attention mechanism," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102301.

[52] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current Cardiol. Rev.*, vol. 8, no. 1, pp. 14–25, Feb. 2012.

[53] F. Lamonaca, Y. Kurylyak, D. Grimaldi, and V. Spagnuolo, "Reliable pulse rate evaluation by smartphone," in *Proc. IEEE Int. Symp. Med. Meas. Appl. Proc.*, May 2012, pp. 1–4.

[54] Y. Kurylyak, F. Lamonaca, D. Grimaldi, and F. Duro, "Smartphone based photoplethysmogram measurement," in *Digital Image and Signal Processing for Measurement Systems*. Denmark: River Publishers, 2012, pp. 135–164.

[55] S. P. Linder, S. M. Wendelken, E. Wei, and S. P. McGrath, "Using the morphology of photoplethysmogram peaks to detect changes in posture," *J. Clin. Monitor. Comput.*, vol. 20, no. 3, pp. 151–158, Jul. 2006.

[56] S. Khalid, T. Khalil, and S. Nasreen, "A survey of feature selection and feature extraction techniques in machine learning," in *Proc. Sci. Inf. Conf.*, Aug. 2014, pp. 372–378.

[57] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.

[58] J. Rogers and S. Gunn, "Identifying feature relevance using a random forest," in *Proc. Int. Stat. Optim. Perspect. Workshop Subspace, Latent Struct. Feature Selection*. Slovenia: Springer, 2005, pp. 173–184.

[59] J. H. Friedman, "Stochastic gradient boosting," *Comput. Statist. Data Anal.*, vol. 38, no. 4, pp. 367–378, 2002.

[60] H. Zheng, J. Yuan, and L. Chen, "Short-term load forecasting using EMD-LSTM neural networks with a xgboost algorithm for feature importance evaluation," *Energies*, vol. 10, no. 8, p. 1168, Aug. 2017.

[61] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Series in Statistics). USA: Springer, 2001.

[62] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in photoplethysmogram," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 4972–4975.

[63] D. J. Thomson, "Time series analysis of Holocene climate data," *Phil. Trans. Roy. Soc. London. Ser. A, Math. Phys. Sci.*, vol. 330, no. 1615, pp. 601–616, 1990.

[64] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty-V: The discrete case," *Bell Syst. Tech. J.*, vol. 57, no. 5, pp. 1371–1430, May 1978.

[65] A. R. Rao and K. Hamed, "Multi-taper method of analysis of periodicities in hydrologic data," *J. Hydrol.*, vol. 279, nos. 1–4, pp. 125–143, Aug. 2003.

[66] M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.

[67] M. Al-Hawawreh and E. Sitnikova, "Industrial Internet of Things based ransomware detection using stacked variational neural network," in *Proc. 3rd Int. Conf. Big Data Internet Things (BDIOT)*, 2019, pp. 126–130.

[68] M. Al-Hawawreh, E. Sitnikova, and F. den Hartog, "An efficient intrusion detection model for edge system in brownfield industrial Internet of Things," in *Proc. 3rd Int. Conf. Big Data Internet Things (BDIOT)*, 2019, pp. 83–87.

[69] S. S. Panicker and P. Gayathri, "A survey of machine learning techniques in physiology based mental stress detection systems," *Biocybern. Biomed. Eng.*, vol. 39, no. 2, pp. 444–469, Apr. 2019.

[70] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 4, pp. 859–869, Apr. 2017.

[71] D.-C. Feng, Z.-T. Liu, X.-D. Wang, Y. Chen, J.-Q. Chang, D.-F. Wei, and Z.-M. Jiang, "Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach," *Construct. Building Mater.*, vol. 230, Jan. 2020, Art. no. 117000.

[72] V. Vapnik, *The Nature of Statistical Learning Theory*. Germany: Springer, 2013.

[73] R. Mohawesh, S. Tran, R. Ollington, and S. Xu, "Analysis of concept drift in fake reviews detection," *Expert Syst. Appl.*, vol. 169, May 2021, Art. no. 114318.

[74] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*. [Online]. Available: https://arxiv.org/abs/1308.0850

[75] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[76] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3104–3112.

[77] R. Mohawesh, S. Xu, S. N. Tran, R. Ollington, M. Springer, Y. Jararweh, and S. Maqsood, "Fake reviews detection: A survey," *IEEE Access*, vol. 9, pp. 65771–65802, 2021.

[78] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2013, pp. 1310–1318.

[79] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997.

[80] K. A. Althelaya, E.-S.-M. El-Alfy, and S. Mohammed, "Evaluation of bidirectional LSTM for short- and long-term stock market prediction," in *Proc. 9th Int. Conf. Inf. Commun. Syst. (ICICS)*, Apr. 2018, pp. 151–156.

[81] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," *Neural Netw.*, vol. 18, no. 5, pp. 602–610, 2005.

[82] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Gener. Comput. Syst.*, vol. 115, pp. 279–294, Feb. 2021.

[83] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: https://arxiv.org/abs/1406.1078

[84] G. Slapničar, N. Mlakar, and M. Luštrek, "Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network," *Sensors*, vol. 19, no. 15, p. 3420, Aug. 2019.
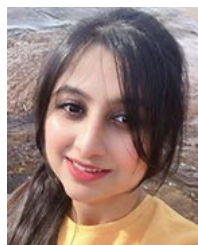
SUMBAL MAQSOOD received the B.S. degree (Hons.) in computer science from Punjab University College of Information Technology (PUCIT) and the M.S. degree in computer science from GC University, Lahore, Pakistan. She is currently pursuing the Ph.D. degree with the University of Tasmania, Hobart, TAS, Australia. Her research interests include machine learning, natural language cybernetics, biotechnologies, data science, and software engineering. She is also working on biosignals analysis using deep learning.
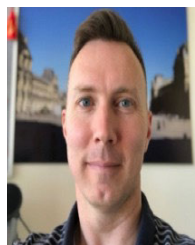
MATTHEW SPRINGER received the Ph.D. degree in information systems from the University of Tasmania, in 2010. He is currently a Lecturer with the School of Technology, Environments and Design, University of Tasmania. His major focus has been on improving teaching with the Discipline of Information and Communication Technology, but is also an Active Member of the Industry Transformation, and Games and Creative Technologies research groups.

SHUXIANG XU received the Bachelor of Applied Mathematics degree from the University of Electronic Science and Technology of China, China, the Master of Applied Mathematics degree from Sichuan Normal University, China, and the Ph.D. degree in computing from the University of Western Sydney, Australia. He is currently a Lecturer and a Ph.D. Student Supervisor with the School of Information and Communication Technology, University of Tasmania, Hobart, TAS, Australia. His research interests include artificial intelligence, machine learning, and data mining. Much of his work is focused on developing new machine learning algorithms and using them to solve problems in various application fields.

RAMI MOHAWESH received the B.E. and M.E. degrees in computer science. He is currently pursuing the Ph.D. degree with the University of Tasmania, Hobart, TAS, Australia.

In his Ph.D. research, he is the first Researcher who investigated the concept drift in fake review detection. His research interests include software engineering, cloud computing, natural language processing, cybersecurity, and machine learning. His current work is on Fake Review. He is a Reviewer of high impact factor journals, such as the *Information Processing and Management* journal, *Artificial Intelligence Review*, and *Secure Computing*.

● ● ●