Swin Transformer-based Cuffless Blood Pressure Estimation using Cardiovascular Signals

Vinit Kumar[1], Manu Mishra[2], Maulesh Gadani[3], Jayesh Jayarajan[3], Payal Sharma[3], Priya Ranjan Muduli[1]
[1]Department of Electronics Engineering, Indian Institute of Technology (BHU) Varanasi, India
[2]Department of Computer Science Engineering, Vellore Institute of Technology, Chennai, India
[3]Space Applications Centre (SAC), Indian Space Research Organisation, India
vinitkumar.rs.ece21@iitbhu.ac.in, manu.mishra2021@vitstudent.ac.in, drmaulesh@sac.isro.gov.in,
jayeshj@sac.isro.gov.in, payal@sac.isro.gov.in, prmuduli.ece@iitbhu.ac.in

2024 IEEE 21st India Council International Conference (INDICON) | 979-8-3503-9128-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/INDICON63790.2024.10958421

***Abstract*—A conventional sphygmomanometer used for blood pressure (BP) measurement is considered as a golden standard. However, this bulky device requires a pump, valve, battery, and relevant components to inflate and deflate the cuff. Prolonged BP monitoring using a sphygmomanometer may create discomfort for the patients. Therefore, cuffless blood pressure estimation has recently emerged as an excellent choice. Cuffless BP estimation can be performed using electrocardiogram (ECG) and photoplethysmogram (PPG) signals or both ECG and PPG measurements. Various machine learning-based techniques are proposed for cuffless BP estimation utilizing different time-domain and frequency-domain features of ECG and PPG signals. However, the accuracy of the measurements is one of the major concerns. BP estimation using the higher-order spectral features, including bispectrum, has yet to be explored in the literature. This paper proposes an effective approach that utilizes the bispectrum features and a Swin transformer-based network. Capturing the spectral variations in non-stationary cardiovascular signals is crucial. This task is accomplished by signal-to-image conversion of the cardiovascular measurements and utilizing the bispectrum features as input to the Swin transformer. The transformer architecture incorporates a layer normalization, two-level multi-layer perception, and a shifted window-based multi-head self-attention (SW-MSA) mechanism. The proposed method can handle long-range dependency and learn global structures effectively. The proposed technique is validated by considering various international standards for cuffless BP estimation tasks. A comparative study is performed on different evaluation metrics concerning contemporary techniques.**

***Index Terms*—Cuffless BP estimation, Electrocardiogram, Photoplethysmogram, Bispectrum, SWIN Transformer.**

I. Introduction

Case-control investigations have shown that hypertension is the leading risk factor for cardiovascular disease (CVD) worldwide. It is roughly calculated to be the reason behind 7.5 million, i.e., around 12.8% of all deaths globally. Additionally, it amounts to 57 million disability-adjusted life years (DALYs), i.e., 3.7% of all DALYs [1]. It is considered hypertension when the BP value surpasses 140/90 mmHg. Systolic blood pressure (SBP) and diastolic blood pressure (DBP) are the two primary components of a blood pressure measurement. When the human body is at idleness, hypertension is indicated if the DBP surpasses 90 mmHg or the SBP surpasses 140 mmHg [2]. The measurement of BP through a sphygmomanometer has been the gold standard for noninvasive blood pressure estimation. However, this method has some drawbacks, including the large, uncomfortable cuff that must be worn constantly and the bulky electronics that use pneumatic systems (such as a pump, valve, battery, etc.) to inflate and deflate the cuff. Therefore, developing a reliable cuffless BP estimation method is essential for monitoring the patient's continuous or prolonged blood pressure. The fluid wave propagation theory is the basis for the cuffless blood pressure measuring technique. It depends on the inherent relationship between wave propagation velocity and fluid pressure. This theory contends that BP can be computed using the cardiac wave's velocity. The ECG and PPG signals are widely used in cardiovascular surveillance, which includes blood oxygen saturation, blood pressure, and heart rate monitoring [3]. ECG signals, PPG signals, or both ECG and PPG signals can be used for cuffless BP Estimation. Various researchers have used either ECG or PPG signals for this task, whereas some researchers utilize ECG and PPG signals to obtain more comprehensive features for cuffless BP estimation tasks. The literature suggests several machine/deep learning-based approach combinations for cuffless BP estimation. A machine learning technique based on the ECG signal's frequency domain features is proposed for cuffless BP estimation tasks [4]. The authors use the first and second derivatives as the features of the PPG signal, along with multiple linear regression (MLR) techniques [5]. An adaptive boosting (AdaBoost) regression-based technique with morphological features is developed for cuffless BP estimation [6]. In [7], authors have utilized the wavelet-based features of PPG signals along with the AdaBoost regression technique for BP estimation. The demographic features and the Gaussian process regression (GPR) are utilized for the BP estimation using PPG signals in [8]. A spectro-temporal deep learning approach is presented for cuffless BP estimation in [9]. An attention mechanism-based deep learning method is proposed in [10]. Apart from only ECG or PPG signals, various researchers have combined both ECG and PPG signals to obtain more comprehensive features for BP estimation. The authors incorporate the Womersley number-based features and different regression techniques for BP estimation in [11]. A group of researchers combines the complexity features with

the SVM regressor for BP estimation using both ECG and PPG signals [12]. Authors combine different ECG features and PPG signals with the random forest regression in [13]. In [14], authors have tried to use phonocardiogram (PCG) along with PPG signal to estimate blood pressure. Most existing techniques utilize either machine learning-based regressors with time or frequency domain features or deep learning techniques where the ECG/PPG signal is directly fed to the model for BP estimation. Since cardiovascular signals show non-stationary properties, these physiological signals have varying temporal and spectrum information dynamically. The efficiency of real-time monitoring and decision-making may be impacted by delayed or out-of-date information resulting from non-stationarity. These limitations can be overcome by combining innovative approaches like signal-to-image conversion with deep learning models, which can automatically learn robust features from 2D representations of raw physiological data [15]. The higher-order spectrum is a robust method for non-linear signal analysis [16]. It provides the spectrum of third or higher-order statistics (i.e., moments and cumulants). The proposed technique employs the third-order statistics, i.e., bispectrum, of the PPG and ECG signals. This paper proposes a vision transformer with a shifted windows (Swin transformer) based deep learning model that utilizes bispectrum-based higher-order spectral features for cuffless BP estimation. The following are this paper's major contributions.

- We propose an efficient cuffless BP estimation method. Our proposed model introduces novel higher-order spectral features, including the bispectrum of ECG and PPG measurements, to a Swin transformer-based deep learning model for cuffless BP estimation. The transformer architecture can efficiently handle the long-range dependency and global structures, improving the accuracy of BP estimation.
- The proposed method's effectiveness is assessed concerning several international standards.
- The method is validated by comparing its estimation performance with state-of-the-art methods using several assessment metrics.

The overall format for the rest of the paper is as follows. Section II contains the details of the proposed method. In section III, the experimental details and the numerical results are included. The concluding remarks are presented in section IV.

II. Methodology

This section presents a detailed explanation of the proposed cuffless BP estimation technique and a detailed description of the dataset and pre-processing. The graphical representation of the proposed BP estimation technique is presented in Fig.1.

A. Dataset Details and Pre-Processing

The experiments in this work utilize the cuffless BP dataset provided by the University of California, Irvine machine learning repository [13]. This dataset is derived from the multi-parameter intelligent monitoring in intensive care units (MIMIC-II version 3) [17] database. The ECG, PPG, and arterial blood pressure (ABP) signals are obtained from the MIMIC-II online waveform database. SBP and DBP values are computed using the ABP signal. The data pre-processing is performed according to [13]. The sampling rate of both ECG and PPG measurements is 125 Hz. The cardiovascular measurements are highly prone to the noise and artifacts during signal acquisition process [18], [19]. Hence, in order to suppress these artifacts, the signals are decomposed using discrete wavelet transform (DWT) [20]. The level of decomposition is 10, and the mother wavelet function used is Daubechies 8 (db8). The low-frequency components between $0 - 0.25$ Hz are suppressed to resolve the baseline wander issues. Similarly, the high-frequency elements in the 250–500 Hz range are also removed to suppress the muscular activity artefacts. The wavelet-based denoising is implemented over the left-over wavelet coefficients using the soft Rigrsure (SURE) technique [21]. The denoised signal is recovered using the inverse DWT-based reconstruction approach. Train, test, and validation data are separated from the dataset in the ratios of 0.8, 0.1, and 0.1, respectively.

B. Proposed BP Estimation Technique

The proposed technique computes the bispectrum of both ECG and PPG signals. The Fourier transform of third-order cumulants is called the bispectrum of a signal [16]. In this work, the FFT-based technique is used for bispectum computation. Let the input signal be $S(n)$, and the third-order cumulant is computed by follows [22].

$$C_s(\delta_1, \delta_2) = E[S(n)S(n+\delta_1)S(n+\delta_2)] \tag{1}$$

Here, δ_1 and δ_2 denote the time shift, and $E[.]$ is the expectation operator. Further, the bispectrum from this cumulant can be computed by performing the Fourier transform as given below.

$$B_s(f_1, f_2) = \mathcal{F}(C_s) \tag{2}$$

Here, $\mathcal{F}$ denotes the fast Fourier transform operation. In frequency domain the bispectrum $B_s(f_1, f_2)$ is as follows [16].

$$B_s(f_1, f_2) = E[S(f_1).S(f_2).S^*(f_1+f_2)] \tag{3}$$

Here, $S(f_1)$ and $S(f_2)$ are the Fourier coefficients at frequencies f_1 and f_2, whereas $S^*(f_1+f_2)$ is the complex conjugate of the Fourier coefficient at the frequency f_1+f_2. From Eqn.3, we can say that the bispectrum is defined as the expectation of the product of the Fourier transform at two frequencies f_1 and f_2, and the complex conjugate of the Fourier transform at their sum f_1+f_2. The bispectrum has both magnitude and phase, which makes it a complex quantity.

$$|B_s(f_1, f_2)| = \sqrt{Re(B_s(f_1, f_2))^2 + Im(B_s(f_1, f_2))^2}$$
$$Phase(B_s(f_1, f_2)) = tan^{-1}\big(\frac{Im(B_s(f_1, f_2))}{Re(B_s(f_1, f_2))}\big) \tag{4}$$

The magnitude of the bispectrum gives information about the strength of the interactions, and the phase provides the phase relationship between the interacting components. The bispectrum can detect the phase coupling (non-linear interactions)

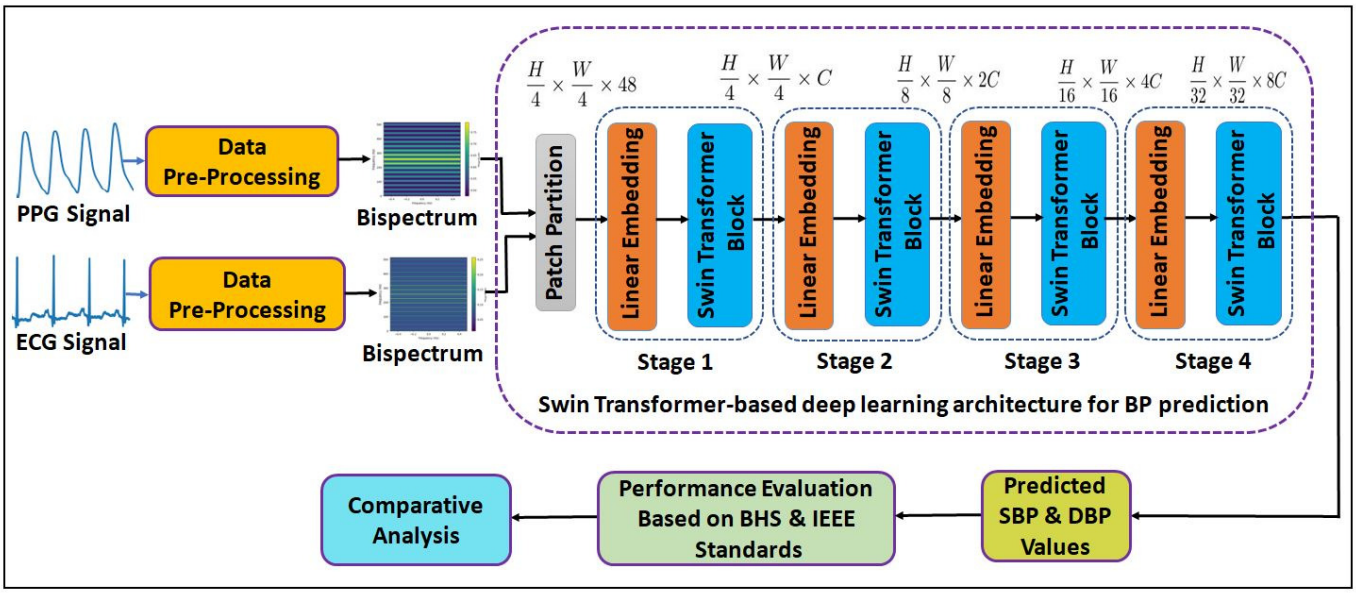

Fig. 1. Architecture of the proposed BP prediction model.

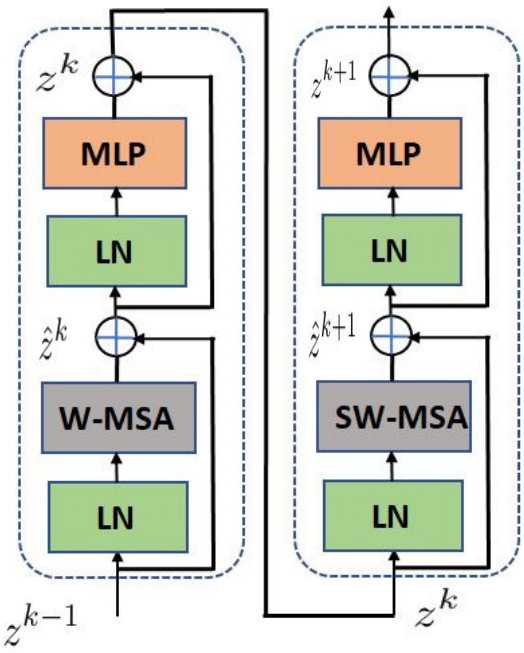

Fig. 2. Architecture of Swin transformer block.

between different frequency components. After computing the bispectrum of the input cardiovascular signals, the obtained bispectrum image is provided as the input to the Swin transformer [23] based deep neural network. The Swin transformer is a hierarchical vision transformer that increases the scalability and efficiency of vision transformers, enabling them to be used for various computer vision applications. It allows for more effective and adaptable processing of high-resolution images by introducing a hierarchical framework with shifting windows. In the Swin transformer, the first step is patch partitioning, the input bispectrum image $B_s \in \mathbb{R}^{H \times W \times C}$. Here, H, W, and C represent the height, width, and number of channels, respectively. It is divided into non-overlapping patches of size $P \times P$. In this work, we consider the patch with size 4×4, which makes the feature dimension of each patch as $(4 \times 4 \times C = 48)$, with the initial value of $C = 48$. This raw-valued feature is projected onto an arbitrary dimension (denoted by C) using the linear embedding layer. These patch tokens are given to the Swin transformer unit. This whole process is referred to as stage 1. The feature map's resolution decreases in each stage, doubling the number of channels. This is achieved by merging patches. Here, the feature vectors from neighbouring patches are concatenated and linearly transformed to produce a new sequence of embeddings by down-sampling with an output dimension set to $2C$. It is repeated in further stages, producing a hierarchical representation similar to a convolutional neural network. This procedure is repeated up to stage 4, finally having the output resolutions $\frac{H}{32} \times \frac{W}{32}$ and output dimension set to $8C$. The two major components of a Swin transformer block are the window-based multi-head self-attention (W-MSA) and shifted window-based multi-head self-attention (SW-MSA). As depicted in Fig.2, a shifted window-based MSA module is followed by a multi-layer perceptron (MLP) in the Swin transformer block. The W-MSA module applies the self-attention to the non-overlapping window of size $M \times M$ and is computed as follows.

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d}} + G\right)V \quad (5)$$

Here, $M = 4$. Q, K, and V denote the query, key, and value matrices obtained, and d is the dimension of the keys. $G \in \mathbb{R}^{M^2 \times M^2}$ is the relative position bias, here M^2 is the

number of patch in each window. The SW-MSA module is used to enable cross-window connections and capture long-range dependencies. Here, the windows are shifted by $\frac{M}{2}$ pixels from their original positions. The outputs for different units in a Swin transformer block are computed as follows.

$$\hat{z}^k = \text{W-MSA}(LN(z^{k-1})) + z^{k-1}$$

$$z^k = \text{MLP}(LN(\hat{z}^k)) + \hat{z}^k$$

$$\hat{z}^{k+1} = \text{SW-MSA}(LN(z^k)) + z^k$$

$$z^{k+1} = \text{MLP}(LN(\hat{z}^{k+1})) + \hat{z}^{k+1} \tag{6}$$

Here, z^{k-1} is the output from the previous block as an input to the current Swin transformer unit. The MLP block consists of a 2-layer MLP consisting of Gaussian Error Linear Unit (GELU) [24] non-linear activation function. The GELU activation function is designed to combine properties of both ReLU (Rectified Linear Unit) and sigmoid functions. The GELU activation is mathematically represented as follows.

$$GELU(x) \approx \frac{x}{2}\Big(1 + erf(\frac{x}{\sqrt{2}})\Big) \tag{7}$$

Here, $erf(x)$ represents an error function related to Gaussian distribution. Unlike ReLU, which zeroes out negative values, GELU scales them according to the Gaussian cumulative distribution, allowing smoother, non-linear transitions. Further, the output of two-layer MLP for an input Z is computed as follows.

$$MLP(Z) = GELU(Zw_1 + b_1)w_2 + b_2 \tag{8}$$

Here, w_1 and w_2 are the weights, whereas b_1 and b_2 are the biases for both the layers of the MLP. These are learnable parameters. The layer normalization (LN) unit normalizes the inputs across the features in a layer, helping to stabilize the learning process in deep neural networks. For any input $Z = [z_1, z_2, \cdots, z_n]$, the normalized ouput is computed as follows.

$$LN(Z) = \frac{z_i - z_\mu}{\sigma_z^2 + \epsilon} \tag{9}$$

Here, z_μ and σ_z^2 are the mean and variance of the input vector Z. The constant ϵ is added to prevent the division by zero. It provides numerical stability. The Adam optimizer is used in the proposed technique, and its initial learning rate is set to 0.0001 empirically. L_1 loss is the loss function used during training. The batch size and number of epochs are 32 and 300, respectively.

III. Experimental Assessments

This section includes the experimental details and results substantiating the proposed technique's effectiveness.

A. Evaluation Parameters

The evaluation parameters used to verify the effectiveness of the proposed approach are mean absolute error (MAE), standard deviation (SD), and cumulative percentage (CP). The MAE is computed as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |q_i - p_i| \ . \tag{10}$$

Here, q_i and p_i are predicted and actual BP values, respectively. N denotes the total size of the test set. The standard deviation is computed as follows:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(e_i - \bar{e_i})^2}{N}} \ . \tag{11}$$

Where, $e_i = q_i - p_i$ and $\bar{e_i} = mean(e_i)$. The cumulative percentage (CP) provides information regarding the percentage of predictions that lie below a particular error value. It can be mathematically given by

$$CP(in\%) = \frac{Cumulative\ Frequency}{N} \times 100 \ . \tag{12}$$

Here, $Cumulative\ Frequency$ denotes the number of predictions having MAE equal to or below a particular value, and N denotes the total number of predictions.

B. Numerical Results

We obtained promising results using the proposed technique. The MAE between predicted and actual values of the test-set for both SBP and DBP are 5.72 mmHg and 2.64 mmHg, respectively. The SD for SBP and DBP estimation are 5.22 mmHg and 3.45 mmHg, respectively. The comparative analysis with recent techniques is presented in Table I. It is observable in Table I that the proposed approach outperforms several state-of-the-art techniques. Furthermore, the proposed method also satisfies the criteria for international standards, including the Institute of Electrical and Electronics Engineers (IEEE) [25], and British Hypertension Society (BHS) standard [26]. The detailed information regarding the standards is presented in section III-C.

To observe the effect of using dual cardiovascular signals, i.e., ECG and PPG, we performed an ablation study where we tried to estimate the BP values using only ECG or PPG signals. Here, all the methodological processes and model parameters are kept the same. While using only ECG signals for BP estimation, we obtained the MAE of 5.95 and 2.82, and the standard deviation of errors was 5.57 and 3.39 for SBP and DBP, respectively. While using only the PPG signal, we obtained the MAE 7.30 and 3.76 and standard deviation of errors 6.45 and 4.02 for SBP and DBP, respectively. When both ECG and PPG signals were used for BP estimation, the obtained MAE were 5.72 and 2.64, and the standard deviation of errors was 5.22 and 3.45 for SBP and DBP, respectively. This study shows that using both cardiovascular signals improves performance significantly.

TABLE I
COMPARATIVE EVALUATION WITH STATE-OF-THE-ART TECHNIQUES.

Method	Signal type	SBP		DBP	
		MAE	SD	MAE	SD
Frequency domain features + RFR [4]	ECG	12.75	12.15	6.04	6.42
DNN with Attention [10]	PPG	12.08	15.67	5.56	7.32
Spectro-Temporal Deep neural network [9]	PPG	9.43	—	6.88	—
Morphological features + AdaBoost [6]	PPG	8.22	10.38	4.17	4.22
Wavelet-based features + AdaBoost [7]	PPG	6.43	11.35	3.07	5.69
Multilinear regression [5]	PPG	6.10	8.08	4.65	6.22
Womersley number based features + RFR [11]	ECG and PPG	9	—	5.48	—
Whole-based features + RFR [13]	ECG and PPG	11.80	9.87	5.83	5.70
Complexity features + SVM [12]	ECG and PPG	8.41	10.05	5.81	7.13
PTT feature + Gradient descent optimization algorithm [14]	PCG and PPG	6.22	9.44	3.97	5.15
Proposed Approach	ECG and PPG	5.72	5.22	2.64	3.45

TABLE II
REQUIREMENTS FOR PASSING IEEE STANDARD.

Standard	Grade	Metrics (MAD in mmHg)
IEEE [25]	A	≤ 5
	B	5 - 6
	C	6 - 7
	D	> 7

TABLE III
REQUIREMENTS FOR PASSING BHS STANDARD.

Standard	Grade	CP (5 mmHg)	CP (10 mmHg)	CP (15 mmHg)
BHS [26]	A	$\geq 60\%$	$\geq 85\%$	$\geq 95\%$
	B	$\geq 50\%$	$\geq 75\%$	$\geq 90\%$
	C	$\geq 40\%$	$\geq 65\%$	$\geq 85\%$

TABLE IV
CUMULATIVE FREQUENCY (%) FOR BHS STANDARD ASSESSMENT.

BP Type	CP (5 mmHg)	CP (10 mmHg)	CP (15 mmHg)
SBP	57 % (Grade B)	85 % (Grade A)	93.5 % (Grade B)
DBP	87 % (Grade A)	95 % (Grade A)	98.5 % (Grade A)

C. Evaluation on Various International Standards

We selected two standards to validate the proposed technique on international standards, i.e., IEEE and BHS. MAE and CP metrics are considered to verify the BHS standard. The mathematical formulation of MAE and CP is given in Eqn.10 and Eqn.12, respectively. For grading in IEEE standard, mean absolute distance (MAD) is an important metric. MAD is computed using the similar formulation Eqn.10. The requirement for grading in IEEE standard is presented in Table II. The MAD for SBP and DBP are 5.72 mmHg and 2.64 mmHg, respectively. Hence, according to the IEEE standards requirements, the proposed method falls in "$Grade\ A$" for DBP prediction, whereas in "$Grade\ B$" for SBP prediction. The necessary conditions for BHS standard grading are presented in Table III. According to BHS standards, the proposed method demonstrates good performance. The CP values for both SBP and DBP are presented in Table IV. For SBP prediction, the proposed method achieves "$Grade\ B$" in 5 mmHg and 15 mmHg category with 57% and 93.5% of the predictions having MAE values below 5 mmHg and 15 mmHg, respectively. In the 10 mmHg category, the proposed technique achieves "$Grade\ A$" with 85% predictions having MAE below 10 mmHg. The method achieves "$Grade\ A$" in all three categories in DBP prediction. It achieved 87%, 95%, and 98.5% prediction with MAE below 5 mmHg, 10 mmHg and 15 mmHg, respectively.

IV. CONCLUSION

This paper proposes an efficient deep-learning technique for cuffless BP estimation using ECG and PPG measurements. The proposed method incorporates a Swin transformer-based model for cuffless BP estimation. The bispectrum of ECG and PPG signals are the input features for the Swin transformer unit. The experimental findings show that the proposed method outperforms several cutting-edge methods, satisfying the IEEE, and BHS standard grading criteria. The improvement in estimation accuracy is due to the novel input features and the incorporation of LN, W-MSA, two-layer MLP, and SW-MSA in the transformer architecture. The proposed technique falls under "$Grade\ A$" of IEEE standards for DBP prediction and "$Grade\ B$" for SBP prediction. The proposed technique in the BHS standard assessment also falls under "$Grade\ A$" for most cases.

ACKNOWLEDGEMENT

This research was carried out as a part of the project "Development of Wearable Internet of Medical Things for Continuous Health Monitoring of Astronauts (project sanction number: ISRO/RAC-S/IIT(BHU)/2022-23)", funded by the Space Applications Center, ISRO, Department of Space, Govt. of India.

REFERENCES

[1] World Health Organization, "Hypertension," Available online on: https://www.who.int/news-room/fact-sheets/detail/hypertension.

[2] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," *Current cardiology reviews*, vol. 8, no. 1, pp. 14–25, 2012.

[3] V. Kumar and P. R. Muduli, "Infimal convolution and AM-GM majorized total variation-based integrated approach for biosignal denoising," *Signal, Image and Video Processing*, vol. 18, no. 2, pp. 1919–1927, 2024.

[4] S. S. Mousavi, M. Hemmati, M. Charmi, M. Moghadam, M. Firouzmand, and Y. Ghorbani, "Cuff-Less blood pressure estimation using only the ECG signal in frequency domain," in *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*. IEEE, 2018, pp. 147–152.

[5] S. Haddad, A. Boukhayma, and A. Caizzone, "Continuous PPG-based blood pressure monitoring using multi-linear regression," *IEEE journal of biomedical and health informatics*, vol. 26, no. 5, pp. 2096–2105, 2021.

[6] N. Hasanzadeh, M. M. Ahmadi, and H. Mohammadzade, "Blood pressure estimation using photoplethysmogram signal and its morphological features," *IEEE Sensors Journal*, vol. 20, no. 8, pp. 4300–4310, 2019.

[7] V. Kumar, G. V. S. S. Bharadwaj, J. Jayarajan, M. N. Gadani, P. Sharma, and P. R. Muduli, "Wavelet-based adaptive boosting method for cuffless blood pressure estimation on pynq-z2," in *2024 International Conference on Signal Processing and Communications (SPCOM)*, 2024, pp. 1–5.

[8] M. H. Chowdhury, M. N. I. Shuzan, M. E. Chowdhury, Z. B. Mahbub, M. M. Uddin, A. Khandakar, and M. B. I. Reaz, "Estimating blood pressure from the photoplethysmogram signal and demographic features using machine learning techniques," *Sensors*, vol. 20, no. 11, p. 3127, 2020.

[9] 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, 2019.

[10] N. Aguirre, E. Grall-Maës, L. J. Cymberknop, and R. L. Armentano, "Blood pressure morphology assessment from photoplethysmogram and demographic information using deep learning with attention mechanism," *Sensors*, vol. 21, no. 6, p. 2167, 2021.

[11] G. Thambiraj, U. Gandhi, U. Mangalanathan, V. J. M. Jose, and M. Anand, "Investigation on the effect of womersley number, ECG and PPG features for cuff less blood pressure estimation using machine learning," *Biomedical Signal Processing and Control*, vol. 60, p. 101942, 2020.

[12] S. Yang, W. S. W. Zaki, S. P. Morgan, S.-Y. Cho, R. Correia, and Y. Zhang, "Blood pressure estimation with complexity features from electrocardiogram and photoplethysmogram signals," *Optical and Quantum Electronics*, vol. 52, pp. 1–16, 2020.

[13] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany, "Cuffless blood pressure estimation algorithms for continuous health-care monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4, pp. 859–869, 2016.

[14] A. Esmaili, M. Kachuee, and M. Shabany, "Nonlinear cuffless blood pressure estimation of healthy subjects using pulse transit time and arrival time," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 12, pp. 3299–3308, 2017.

[15] K. C. Vidyasagar, K. R. Kumar, G. A. Sai, M. Ruchita, and M. J. Saikia, "Signal to image conversion and convolutional neural networks for physiological signal processing: A review," *IEEE Access*, 2024.

[16] A. Swami, J. M. Mendel, and C. L. Nikias, "Higher-order spectral analysis toolbox," *The Mathworks Inc*, vol. 3, pp. 22–26, 1998.

[17] 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," *Critical care medicine*, vol. 39, no. 5, p. 952, 2011.

[18] V. Kumar and P. R. Muduli, "Attentive Bi-LSTM based method for noise suppression in ambulatory ECG measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–9, 2023.

[19] P. R. Muduli and V. Kumar, "A proximity operator-based method for denoising biomedical measurements," *Circuits, Systems, and Signal Processing*, vol. 42, no. 10, pp. 6253–6277, 2023.

[20] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

[21] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[22] S. Shahnawazuddin, A. Kumar, V. Kumar, S. Kumar, and W. Ahmad, "Robust children's speech recognition in zero resource condition," *Applied Acoustics*, vol. 185, p. 108382, 2022.

[23] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 9992–10 002.

[24] D. Hendrycks and K. Gimpel, "Gaussian error linear units (gelus)," *arXiv preprint arXiv:1606.08415*, 2016.

[25] "IEEE standard for wearable cuffless blood pressure measuring devices," *IEEE Std 1708-2014*, pp. 1–38, 2014.

[26] E. O'Brien, J. Petrie, W. Littler, M. de Swiet, P. L. Padfield, K. O'Malley, M. Jamieson, D. Altman, M. Bland, and N. Atkins, "The british hypertension society protocol for the evaluation of automated and semi-automated blood pressure measuring devices with special reference to ambulatory systems," *Journal of hypertension*, vol. 8, no. 7, pp. 607–619, 1990.