[1] **Hemlata Biradar**

[2] **Jayanand Gawande**

# Facial Video based Heart Rate Estimation using Wavelet Packet Transform and Singular Value Decomposition.

Journal of Electrical Systems

***Abstract: -*** Traditional techniques of heart rate (HR) measurement rely upon optical or electronic sensors. This research proposes a HR measuring method in a noncontact way that enables a physiological examination of cardiac pulse without using electrodes. The proposed method is based on automated face tracking and semi blind source separation singular value decomposition (SVD). SVD is introduced in order to increase the probability of detection of vital signs. Wavelet packet transform (WPT) is also used where the signal is successively split into sub-bands at each decomposition level, similar to the wavelet transform. The experimentation is carried out on COHFACE and DROZY datasets. The HR predicted by the proposed method and by the finger blood volume pulse sensor is compared using the Bland-Altman and correlation analysis. The proposed method when compared to other methods utilizing comparable datasets outperforms in terms of mean absolute error (MSE), root mean square error (RMSE), standard deviation (SD), and correlation coefficient (CC).

***Keywords:*** Heart Rate, Scale Invariant Feature Transform, Photoplethysmography, Wavelet packet transform, Singular Value Decomposition, Power Spectrum.

## I. INTRODUCTION

Human vital indicators, such as the respiratory rate (RR) and heart rate (HR) are measured using photoplethysmography (PPG) using either contact or non-contact (remote) methods [1]. In order to collect pre-diagnosis indicators or to keep an eye on patients' vital signs in intensive care units, medical facilities commonly use fingertip pulse oxymeters, is an example of contact photoplethysmography. Hertzman et al. made the first reference of PPG in the 1930s [2]. PPG signals are derived from changes in light that are reflected off of or refracted through the skin. Since haemoglobin in the blood binds to oxygen to produce oxyhemoglobin, which absorbs light, especially wavelengths associated with green hue, a significant amount of light is absorbed by the skin's capillary layer in proportion to the amount of blood that is pierced. The capillary layer's volume of oxy-hemoglobinized blood fluctuates with each contraction-relaxation cycle of the heart. We can determine the heart rate by measuring the fluctuations. An essential physiological signal for tracking the state of the human body's health and emotions is the CARDIAC signal [3]. Abnormal heart rates are defined as follows: HR lower than 60 bpm is referred to as bradycardia, and a HR beyond 100 bpm is known as tachycardia [4].

Physiological and pathological variables that interfere with the normal electrical impulse that regulates the heart's pumping function are the cause of abnormalities in HR. High blood pressure, smoking, temperature, unexpected stress, drug side effects, and damage to the heart's tissues from heart disease are the usual causes of tachycardia. Bradycardia is typically the cause of rheumatic fever, obstructive sleep apnea, myocarditis, and aging-related damage to the heart's tissue. Either of these abnormalities would have a risk factor that could lead to death. Two popular methods for collecting heart signals are photoplethysmography (PPG) and electrocardiograms (ECG). Numerous gadgets, such as fitness smart watches and pulse oxymeters, use the photoplethysmography method (PPG). The primary purpose of this method of measuring blood volume variation is to ascertain heart rate. Furthermore, this metric can also be used to calculate breath rate or oxygen saturation. The PPG measurement is often carried out in close proximity to the skin. The tissue is illuminated by a light source, and minute differences in the reflected or transmitted light are measured.

The electrocardiogram is among the simplest and fastest tests to evaluate the heart (ECG). Electrodes are small plastic patches that stick to the skin in certain places on the arms, legs, and chest. To connect the electrodes to the ECG equipment, lead wires are utilized. After that, the heart's electrical activity is measured, deciphered, and printed. Both of them rely on particular sensors to make touch with subjects' skin, which could be uncomfortable

[1] hemabiradar@gmail.com

[2] jayanand.gawande@rait.ac.in

or inappropriate for those with delicate skin [5]. The development of non-contact heart rate monitors using computer vision or microwave Doppler methods has been popular in recent years. Using consumer-level cameras, remote photoplethysmography (rPPG) is a type of computer vision-based technology that records color changes in the skin on the face produced by associated heartbeats. In non-contact PPGs, the physiological information is gathered from the slight color variations on the face brought on via the heartbeat using cameras. The skin color signal can be viewed as a mixture of pulsatile, specular reflection, and time varying intensity signals [6], [7]. The basic idea behind PPG is the variations in blood vessel volume during a cardiac cycle alter ambient light route length that is incident, which sequentially results in variations in the amount of light that is reflected and displays the timing of cardiovascular events. The skin color changes can also be used, in newborn, for identifying the severity of jaundice [8]. A nontouch, non-intrusive technique called imaging photoplethysmography (iPPG) can measure or observe heart activity instead of requiring the user to touch a sensor on their skin. As it can monitor human vital indications with the help of video taken with a webcam or a regular user level digital camera. The first remote PPG (rPPG) imaging method was projected by in the presence of ambient light. Participants' PPG signals who were more than one meter away were measured by the researchers using a digital camera. Fast Fourier transform (FFT) and bandpass filters applied to the green channel of RGB videos of faces were found to produce top-quality HR measurement when related to the channel with red or blue. In this work contactless HR estimation method is implemented using singular value decomposition (SVD) to increase the probability of detection of vital signs and Wavelet packet transform (WPT). WPT is used to split the signal into sub-bands at each decomposition level. Organization of the paper is as follows: Section 2 describes the Literature survey, Dataset information is given in section 3, Section 4 contains the detailed explanation of proposed method. Heart rate estimation is enlightened in section 5, Results are discussed in section 6, and lastly section 7 consists of the conclusion and future scope of the work.

## II. LITERATURE SURVEY

We provide a quick overview of related work, The first remote PPG (rPPG) imaging method was projected by [1] in the presence of ambient light. Five distinct face segmentation and ROI selection techniques, including a fixed region and a deformable model fitting approach, were used by [9]. By modelling the variations in backdrop illumination and subtracting it from the variations in face color, [10] developed an illumination rectification technique in order to improve the PPG signal. In order to evaluate the PPG signal's quality distribution, spatiotemporal segmentation films were used in [11], and an adaptive-ROI was computed using mean-shift clustering and adaptive SNR map thresholding. Super pixels were produced in the photos by [12] using the adjacency and color uniformity features of the pixels. A model-based ROI segmentation method using a training generated spatial map was recently proposed in [13]. Managing changes in illumination Surface and subsurface reflectance models were employed by [14] in their research. To correct for errors brought on by variations in light, three color components are converted into two orthogonal bases using the CHROM [15] color space conversion method. In 2SR [16], a spatial color subspace is created for every frame, in the RGB color space depending on the distribution of the skin pixels.

By calculating the degree of rotation and scaling variations between the skin color subspaces caused by heartbeats, the PPG signal is retrieved. Signal decomposition approach [17] method was one of the first to demonstrate how the blind source separation (BSS) technique of independent component analysis (ICA) [18] can be applied to isolate the rPPG signal from other contributive dynamics, hence enhancing its quality. Later, in [19], this work was enhanced by including additional pretreatment and postprocessing procedures. Using the CHROM constraint, Reference [20] built a restricted ICA algorithm. In additional investigations, the principal component analysis (PCA) was also employed. In order to differentiate external illumination variations from the PPG signal, reference [21] applied the ensemble empirical mode decomposition (EEMD) approach to the G-channel signal. Reference [22] presented nonlinear mode decomposition (NMD) technique is more noise-resistant. The history-based consistency check (HBCC) is suggested to choose the best HR.

Based on differentiable local transformations, the Local Group Invariance approach (LGI) [23] generates features that are motion and action invariant. Pulse Blood Volume (PBV) [24] leverages the characteristic of blood volume changes in different wavelengths to clearly discriminate between pulse-induced color changes and motion noise in RGB measurements. Principal Component Analysis (PCA) is a statistical approach used to select a subset of uncorrelated components from temporal RGB recordings [25]. Plane- Orthogonal-to-Skin (POS) [26] is based on a plane orthogonal to the skin tone in the temporally normalized RGB space. SSR [27] uses a spatial subspace of skin

pixels and temporal rotation data to extract pulses. The benefit of superhigh resolution is examined in, and a new semi-BSS approach is introduced. SVD is utilized to increase the likelihood that vital signs will be detected. In a complicated environment, Singular value decomposition (SVD) performs well at detecting faint motion signals. The current work proposes a SVD aims to decompose a matrix into three components: a left singular matrix, a diagonal singular value matrix, and a right singular matrix. It is primarily used for matrix factorization and dimensionality reduction.

## III. DATASET INFORMATION

Here we have used existing datasets, named COHFACE and DROZY. The 160 films and physiological information gathered from 40 healthy participants make up the COHFACE dataset. The campaign to gather data resulted in multiple days. When compared to the carefully regulated environment used to capture the Manhob HCI-Tagging data, the data obtained in this new corpus includes more realistic settings. The mean age of the subjects is 35.6 years, with a 11.47-year standard deviation. There were 12 women (30%) and 28 men (70%), divided by gender [28]. This database, known as the "ULg Multimodality Drowsiness Database" or DROZY, is designed to support researchers in their experimentation, development, and evaluation of systems (i.e., algorithms) related to drowsiness monitoring. It contains a variety of drowsiness-related data, such as images and signals. The Laboratory for Signal and Image Exploitation (INTELSIG), a division of the Department of Electrical Engineering and Computer Science at the University of Liège (ULg), Lige, Belgium, gathered the (multimodality) data. This lab is identified by the notation "ULg-INTELSIG"[29].

## IV. PROPOSED METHOD

Estimating heart rate using SVD is a signal processing technique that can be applied to certain types of physiological data, such as photoplethysmography (PPG) signals obtained from a pulse oximeter or wearable devices. SVD is used to extract the underlying periodic components in the signal, which correspond to the heart rate. Here are the general steps involved in estimating heart rate using SVD and Wavelet packet transform (WPT) as in Fig. 1.

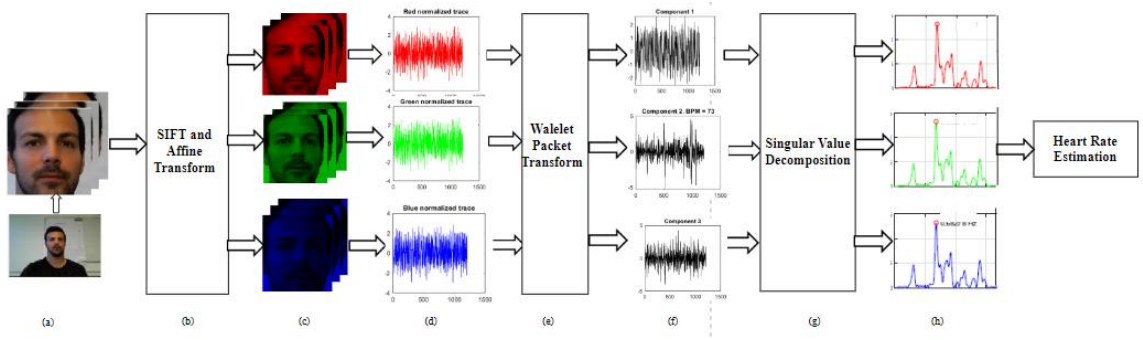


Fig. 1: Block diagram of proposed method

### *A. Face Detection and Tracking*

Extract the face region from the input video frames using face detection algorithms. First, the measuring ROI for each of the video frame was located and faces within the frames were detected using an automated face tracker Fig. 1(a). To acquire the coordinates of the face location, we used a free version of the Open Computer Vision (OpenCV) software that is compatible with MATLAB [30]. The work of Lienhart and Maydt [31] and Viola and Jones [32] is the foundation of the OpenCV face detection method. The algorithm returns the height, width, and x- and y-coordinates that construct a box around each face that is discovered. We used the center 60 percent width and full height of the box as the ROI for our further computations based on this information. To reduce noise caused by face motion in video frames, we use Scale Invariant Feature Transform (SIFT) features from the current and past frames to adjust the face of the current frame to the location of the face in the prior frame using the Affine Transform (AT). SIFT bestowed by Lowe, is one of the best techniques for getting local descriptors [33]. The image is changed into a set of local feature vectors via the SIFT technique as in Fig. 1(b). These feature vectors are intended to be mean and unit variance. After that, the ROI was divided into its three RGB channels and its pixels were spatially averaged

to produce the raw traces a1(t), a2(t), and a3(t) which are the red, green, and blue measurement points for each of the frame as shown in Fig. 1(c) and (d). The raw RGB lines are normalised as follow:

$$a_i'(t) = \frac{a_i(t) - \mu_i}{\sigma_i} \qquad (1)$$

the value of i is taken as 1, 2, and 3, $\mu_i$ is the mean and $\sigma_i$ is the standard deviation of $a_i(t)$. The standardization transform $a_i(t)$ to $a'_i(t)$ with zero mean and unit variance.

### *B. Wavelet Packet Transform (WPT):*

Apply Wavelet Packet Transform as in Fig. 1(e) to decompose the pre-processed face region into its frequency components. WPT is an extension of the wavelet transform (WT), which is a powerful tool in signal processing and data compression. The WT is a technique for representing time-scales that breaks down signals into time- and scale-based base functions. These basis functions are translated and dilated copies of a basis function known as the mother wavelet. Discrete wavelet transform (DWT) is a useful tool for examination at several resolutions [34]. With precise temporal resolution, the DWT breaks down a signal's high frequency components. The most recent method to get around the drawbacks of the Fourier transform (FT) is most likely the WT or wavelet analysis. Wavelet transform decomposes a signal into a set of wavelet coefficients at different scales and positions, wavelet packet transform further extends this decomposition by allowing both time and frequency localization, offering a richer representation of the signal. The Haar wavelet, which has the following wavelet function (φ) and scaling function (Π), is the mother wavelet used in this work.

$$\psi(x) = \begin{cases} +1 \ when\ 0 \le x < \frac{1}{2} \\ -1 \ when\ \frac{1}{2} < x \le 1 \\ 0 \quad otherwise \end{cases} \qquad (2)$$

$$\Phi(x) = \begin{cases} 1 \ when\ 0 \le x < 1 \\ 0 \quad otherwise \end{cases} \qquad (3)$$

The reconstruction and decomposition of signals are carried out by wavelet transform. After one coefficient is passed through low pass and the other through high pass filters, the two coefficients are obtained as approximate and detailed coefficients. Weighted by the wavelet coefficients, the detailed coefficients are derived from the sum of the translated and dilated wavelet mother ψ(t). The scaling function $\Phi$(t)'s translated and dilated versions make up the weighted total of the approximated coefficients. The detailed and approximated coefficients can be denoted as follows [35]:

$$D_j(t) = \sum_k d_x(j,k)\psi_{(j,k)}(t) \qquad (4)$$

$$A_j(t) = \sum_k a_x(j,k)\Phi_{(j,k)}(t) \qquad (5)$$

Compared to wavelet decomposition, WPT provides a more accurate frequency resolution. Based on the Haar wavelet filter, which is crucial to biomedical research, is the revised version of DWT. Here, the quality of a signal and the amount of time needed for analysis lead to the selection of the Haar wavelet, and the suggested method illustrated in Fig. 2 performs three-level decomposition. This algorithm's benefit is that it eliminates redundant samples from each level of decomposition, greatly enhancing execution speed and accuracy. In wavelet packet transform, the signal is successively split into sub-bands at each decomposition level, similar to the wavelet transform. However, unlike the wavelet transform, which only decomposes the signal into approximation and detail coefficients at each level, wavelet packet transform decomposes the signal into both approximation and detail coefficients, but also allows for further decomposition of each detail coefficient. This results in a binary tree structure where each node represents a sub-band of the signal, with the root node representing the original signal and the leaf nodes representing the finest level of decomposition.

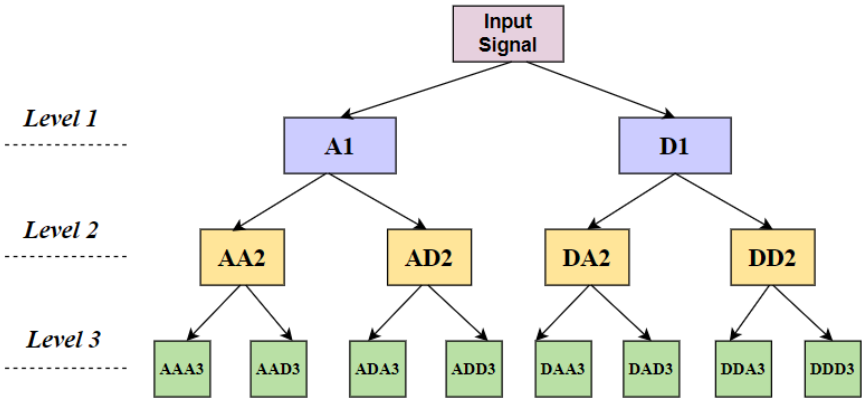


Fig. 2: Three level structure of WPT

The main advantage of wavelet packet transform over wavelet transform is its flexibility and adaptability in capturing both local and non-local signal features. By allowing more freedom in the decomposition process, wavelet packet transform can provide a more detailed and informative representation of the signal, which can be useful in various signal processing tasks such as denoising, feature extraction, and compression. Here's a basic algorithm for performing Wavelet Packet Transform (WPT) [36].

---

**Algorithm 1: Algorithm for Wavelet Packet Transform (WPT)**

---

1: **Input**: A discrete signal x[n], where n=0, 1, 2, . . ., N-1

2: **Choose a Wavelet Function**: Select a wavelet function $\psi(t)$ (e.g., Daubechies, Haar, etc.) and its corresponding scaling function $\phi(t)$.

3: **Decomposition**:

- Start with the original signal x[n] at level 0.
- Apply a high-pass filter (detail filter) h[n] and a low-pass filter (approximation filter) g[n] to obtain detail coefficients dapprox[n] and approximation coefficients aapprox[n] for the current level. Subdivide the detail coefficients dapprox[n] into further detail and approximation coefficients by applying the same filters.
- Repeat the decomposition process recursively for each subband until reaching the desired level of decomposition.

4: **Output**: A tree structure representing the decomposition of the signal into various subbands, each containing detail and approximation coefficients

---

### *C. Singular value decomposition*

Perform Singular Value Decomposition on the selected frequency components obtained from WPT. Extract the most significant singular values or components that capture the variation related to heart rate as seen in Fig. 1(g).

a. **Data Matrix**: Construct a data matrix where each row represents a time sample, and each column represents a data point in the PPG signal. The PPG signal is typically a timeseries with discrete samples.

b. **Perform SVD**: Apply Singular Value Decomposition to the data matrix. SVD decomposes the data matrix into three matrices: U, S, and V T (transpose of V). The Sigma matrix contains singular values, which represent the importance of each component. The columns of U and V represent the left and right singular vectors, respectively. U (Left Singular Vectors): U is a square or rectangular matrix whose columns are orthogonal unit vectors. These vectors are often referred to as the left singular vectors and capture the structure of the original data in a new coordinate system. S (Singular Values): S is a diagonal matrix containing singular values.

The singular values are non-negative real numbers and are listed in descending order. They represent the importance or variance of the information contained in the corresponding singular vectors. VT (Right Singular Vectors): VT is the transpose of a square or rectangular matrix, and its columns are orthogonal unit vectors as well. These vectors are known as the right singular vectors and capture the relationships between the original data points. Singular value decomposition (SVD) is extensively employed in various domains, including signal detection [37] and noise

reduction [38]. Additionally, components of the fetal and maternal electrocardiogram (ECG) signals are extracted using the SVD [39].

### *D. Feature extraction and Heart rate calculation*

Combine the information obtained from both WPT and SVD to extract features relevant to heart rate estimation. These features may include amplitude variations, frequency content, and temporal patterns. Utilize the extracted features to estimate the heart rate. This can be done using various methods such as peak detection, spectral analysis, or machine learning algorithms. As shown in Fig. 1(h) The power spectrum is obtained by applying the fast Fourier transform (FFT) on the chosen source signal. Within an operational frequency band, the frequency corresponding to the maximum power in the spectrum was identified as the pulse frequency. Once you have the dominant frequency, you can calculate the heart rate by converting it to beats per minute (BPM). Typically, heart rate is calculated as HR = (Frequency * 60), where Frequency is in Hertz (Hz).

**Algorithm II: Algorithm for Singular Value Decomposition (SVD)**

---

**Input**: Matrix A of size m × n.

**Step 1**: **Compute $A^T A$ and $AA^T$**

- Compute the matrices $A^T A$ and $AA^T$ These are symmetric matrices of size n × n and m × m, respectively.

**Step 2**: **Compute Eigenvectors and Eigenvalues of $A^T A$ and $AA^T$**

- Compute eigenvectors and eigenvalues of the matrices $A^T A$
- Let $v_1, v_2, v_3 \ldots . v_n$ be the eigenvectors $A^T A$ and $u_1, u_2, u_3 \ldots u_n$ be the eigenvectors $AA^T$
- Arrange the eigenvectors in descending order of their corresponding eigenvalues.

**Step 3**: **Compute Singular Values**:

- The singular values $\sigma_i$ are the square roots of the nonnegative eigenvalues of $A^T A$ or $AA^T$ .

Arrange them in descending order.

**Step 4**: **Compute Singular Values:**

- The left singular vectors U are the eigenvectors of $AA^T$ normalized by the corresponding singular values.
- The right singular vectors V are the eigenvectors of $A^T A$ normalized by the corresponding singular values.

**Step 5**: **Construct Singular Value Decomposition:**

- The SVD of matrix A is given by A = $U\Sigma V^T$, where U is an m×m orthogonal matrix, Σ is n×n diagonal matrix with singular values on the diagonal, and V is an n × n orthogonal matrix.

---

### *E. Feature extraction and Heart rate calculation*

Combine the information obtained from both WPT and SVD to extract features relevant to heart rate estimation. These features may include amplitude variations, frequency content, and temporal patterns. Utilize the extracted features to estimate the heart rate. This can be done using various methods such as peak detection, spectral analysis, or machine learning algorithms. As shown in Fig. 1(h) The power spectrum is obtained by applying the fast Fourier transform (FFT) on the chosen source signal. Within an operational frequency band, the frequency corresponding to the maximum power in the spectrum was identified as the pulse frequency. Once you have the dominant frequency, you can calculate the heart rate by converting it to beats per minute (BPM). Typically, heart rate is calculated as HR = (Frequency * 60), where Frequency is in Hertz (Hz).

## V. RESULTS AND DISCUSSION

To validate the heart rate estimation obtained from the combined approach against ground truth data or reference measurements, Using the same procedures, we were able to derive the reference HR readings from the recorded finger BVP signal. The two measurement approaches were merged into a graphical and statistical interpretation using Bland Altman plots [40]. Plotting the discrepancies between the Flexcomp finger BVP sensor and the estimates from the suggested method versus the averages of both systems was done. Calculations were performed to determine the mean and standard deviation (SD) of the differences, the mean of the absolute differences, and the 95% bounds of agreement (± 1.96 SD). The parameter like mean absolute error (MAE), root mean squared error (RMSE), Pearson's correlation coefficients (CC) and the standard deviation were calculated for the estimated HR

from proposed method and the finger BVP. These parameter values are calculated for face video sequences using proposed method and reference blood volume pulse (BVP) signals using following equations:

$$MAE = \frac{1}{N}\sum_{1}^{N} HR_{GT} - HR_{\text{PM}} \qquad (6)$$

$$RMSE = \sqrt{\frac{\frac{1}{N}\sum_{1}^{N}(HR_{GT} - HR_{\text{PM}})^2}{N}} \qquad (7)$$

where, $HR_{GT}$ = HR obtained using ground truth, $HR_{\text{PM}}$ =HR estimated using proposed method.

$$SD = \sigma = \sqrt{\frac{\sum(a_i - \mu_i)^2}{N}} \qquad (8)$$

where, σ = Standard Deviation of number of subjects, N = Number of subjects, $a_i$ = Each value from the no. subjects, μ= Mean of no. of subjects.

$$CC = r = \frac{\sum(a_i - \bar{a})(bi - \bar{b})}{\sqrt{\sum(a_i - \vec{a})^2 \sum(bi - \vec{b})^2}} \qquad (9)$$

where, r = correlation coefficient, $a_i$= values of the a-variable in a sample, $\bar{a}$ = mean of the values of the a-variable, $b_i$= values of the b-variable in a sample, $\bar{b}$= mean of the values of the b-variable.

Power spectrums of normalized RGB components are shown in Fig. 3. The cardiac pulse wave is clearly discernible in the second component among the three separate sources obtained after SVD, and it resembles the referenced BVP signal, it is observed that the peak pulse frequency for the signal from the BVP matches the reference's power spectrum at 1.389 Hz. Peaks in the power spectra of the SVD components and raw RGB traces are visible at that frequency. However, the power spectrum density (PSD) of the second SVD component coupled with the most precise estimated pulse frequency (0.74139 Hz) and produce the highest signal to ratio. Using the Bland-Altman plot for analysis, two measurements of the same variable are compared.

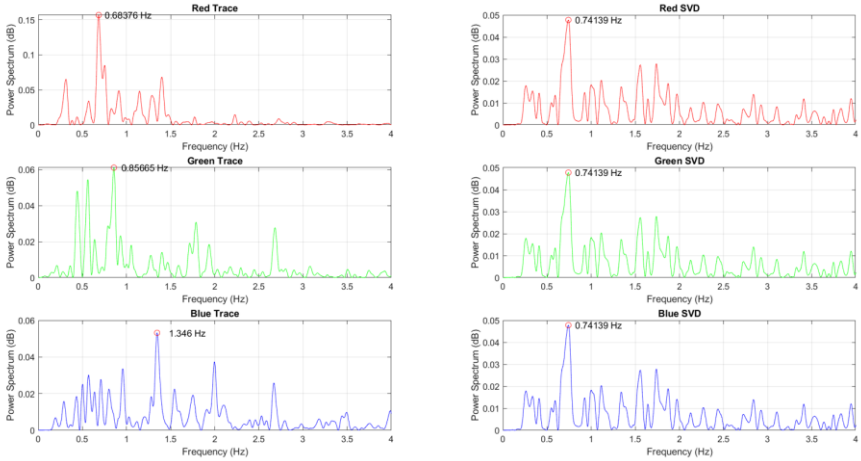


Fig. 3. Power spectrum of RGB components obtained using proposed method.

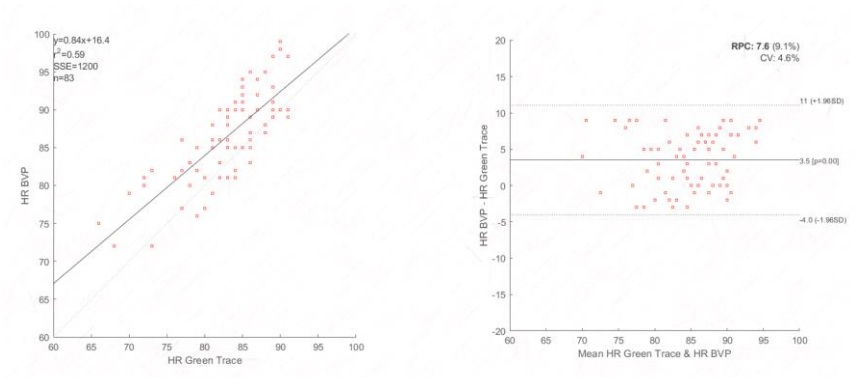


Fig. 4. Bland-Altman (BA) plots of HR with green traces and $HR_{BVP}$

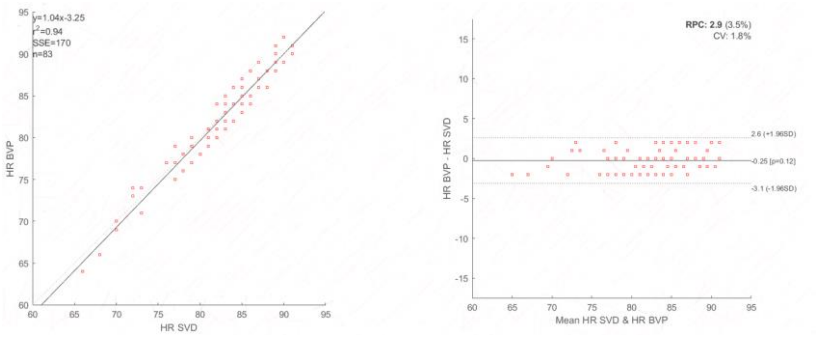


Fig. 5. Bland-Altman (BA) plots of HR with proposed method and $HR_{BVP.}$

It is an approach for evaluating different methods. For instance, a costly measuring system can be compared with a low-cost one, or an invasive measuring method might be compared with a non-invasive one. Here, we calculated the absolute differences in mean, standard deviation (SD), and 95 percentile boundaries of agreement (1.96 SD). From BA plots, the significant differences are seen in the distribution of points before and after proposed method. Before application of proposed method, as depicted in Fig. 4, the mean error is 3.5 beats per minute (BPM), with limits of agreement ranging from 11 to −4.0 BPM. After applying the proposed method error is reduced to −0.25 BPM, with limits of agreement ranging from 2.6 to -3.1 BPM as seen in Fig. 5. Estimation of HR is carried out with the suggested method as well as with the BVP sensor, resulted values were almost identical. It shows that the suggested method allows for superior HR estimation as compared to other existing methods. This non-contact HR measurement technique can be incorporated into fitness trackers, smartwatches, or sports monitoring devices. It enables users to monitor their heart rates during physical activity without the need for chest straps or other contact-based sensors. Also, the proposed face video based HR estimation method can be used in mental health research and clinical settings to monitor HR variability and assess stress levels, emotional arousal, or physiological responses. Table 1 compares the parameters (MAE, RMSE, CC and SD) after applying proposed method with the parameters of other approaches using the COHFACE database. Similarly, Table 2 shows the parameter comparison with different method using DROZY dataset.

Table 1: Performance comparison on the COHFACE database using different HR measurement methods.

| **Methods** | **MSE** | **RMSE** | **SD** | **CC** |
|---|---|---|---|---|
| GREEN [1] | 14.69 | -- | -- | .05 |
| PBV [24] | 13.27 | -- | -- | 0.05 |
| LGI [23] | 12.78 | -- | -- | 0.01 |
| CHROM [15] | 12.49 | -- | -- | 0.02 |
| POS [26] | 11.99 | -- | -- | 0.03 |
| SSR [27] | 11.89 | | | 0.03 |
| ICA [17] | 11.16 | -- | -- | 0.19 |
| PCA [25] | 10.47 | -- | -- | 0.27 |
| Fast ICA [41] | 8.06 | 8.45 | 14.02 | 0.9 |
| SIFT AND FASTICA [44] | 7.02 | 7.47 | 14.01 | 0.89 |
| **WPT and SVD (Proposed Method)** | **6.52** | **6.62** | **13.44** | **0.97** |

Table 2: Performance comparison on the DROZY database using different HR measurement methods.

| **Methods** | **MSE** | **RMSE** | **SD** | **CC** |
|---|---|---|---|---|
| SCICA [42] | 7.75 | 12.50 | 9.82 | 0.52 |
| SCF [1] | 7.57 | 12.45 | 9.90 | 0.39 |
| EEMD [43] | 8.39 | 12.81 | 9.7 | 0.32 |

| | | | | |
|---|---|---|---|---|
| Fast ICA [41] | 8.01 | 8.43 | 9.9 | 0.83 |
| **WPT and SVD (Proposed Method)** | 7.00 | 7.15 | 9.45 | 0.9 |

To compare the distributions of HR estimations across the COHFACE and DROZY datasets, we used box and whisker plots. This is useful for graphically representing the distribution of data through its quartiles. The "whiskers," or lines stretching parallel from the boxes, are used to denote variability outside the upper and lower quartiles. Outliers are sometimes displayed as individual dots that are in-line with whiskers. Both vertical and horizontal lines can be used to create box plots. Box plots have the benefit of taking up less room than a histogram or density Plot, which is helpful when comparing distributions across numerous groups or datasets. Despite the fact that they may appear simple in comparison. The Box and whisker plots of MAE, RMSE, SD and CC for COHFACE and DROZY dataset are shown in Fig. 6.

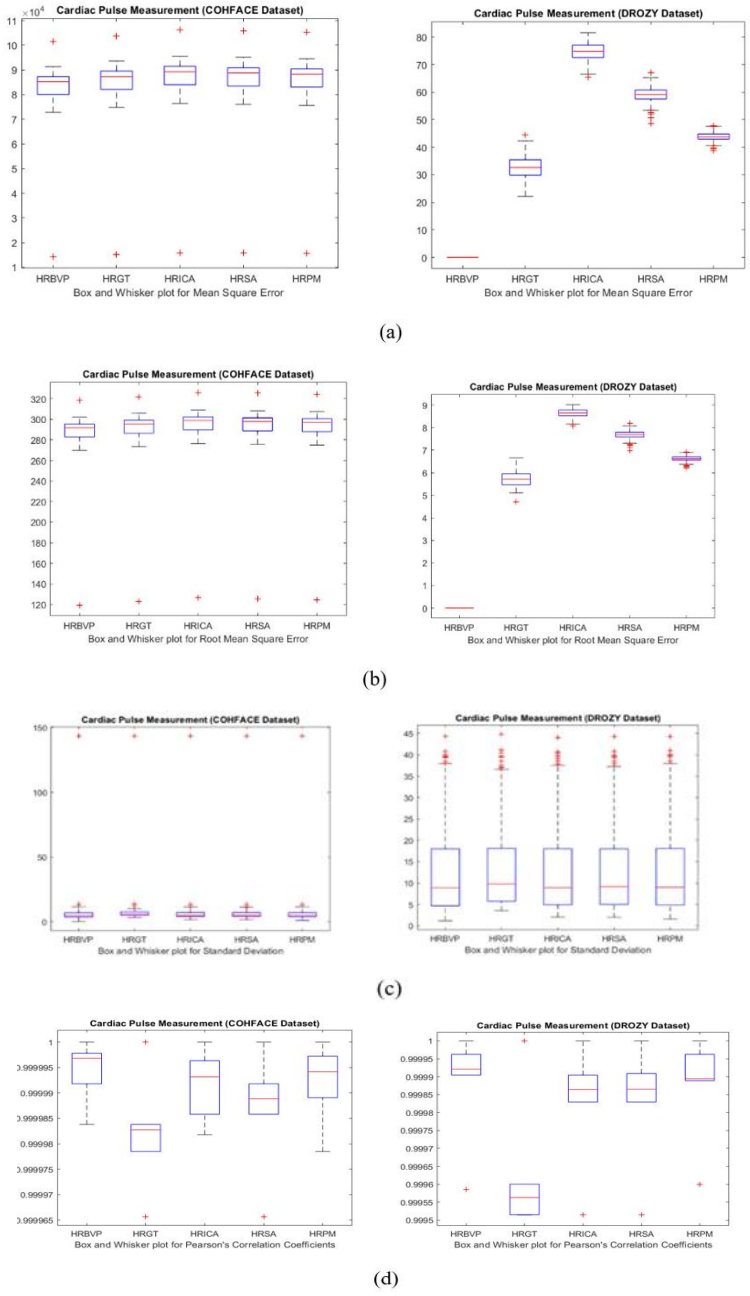


Fig. 6. Box and Whisker plots for MSE, RMSE, SD and CC of various methods and proposed method for COHFACE and DROZY datasets are shown in (a), (b), (c) and (d) respectively.

## VI. CONCLUSION

In this study, a non-contact method to recover the heart rate from video recordings of human faces called singular value decomposition (SVD) and wavelet packet transform (WPT) has been described, put into practice, and tested. The findings of proposed method are equated to the cardiac pulse measured by the finger blood volume pulse (BVP) sensor, and the Bland Altman analysis plots are used to determine the minimum error. The suggested method's mean absolute error, root mean square error, standard deviation, and correlation coefficient were minimal when compared to other HR estimate methods. The proposed method can be applied for fitness and sports tracking, healthcare monitoring, mental health assessment, etc. In future, the suggested method has the potential to estimate a wide range of other significant physiological characteristics, including respiratory rate, heart rate variability, and arterial blood oxygen saturation.

## References


[1] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," Opt. Exp., vol. 16, no. 26, pp. 21434–21 445, 2008.

[2] Hertzman, A. B. "The blood supply of various skin areas as estimated by the photoelectric plethysmography", American Journal Physiology, vol. 124(2), pp.329-340, 1938, https://doi.org/10.1152/ajplegacy.1938.124.2.328.

[3] Song, R., Chen, H., Cheng, J., Li, C., Liu, Y., and Chen, X. PulseGAN: Learning to generate realistic pulse waveforms in remote photoplethysmography. ArXiv. /abs/2006.02699

[4] Hassan, M.A., Malik, A.S., Fofi, D., Saad, N., Karasfi, B., Ali, Y.S. and Meriaudeau, F., "Heart rate estimation using facial video: A review. Biomedical Signal Processing and Control", 38, pp.346-360, 2017.

[5] Sun, Y., Hu, S., Azorin-Peris, V., Greenwald, S., Chambers J., Zhu, Y. "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise", Journal of Biomedical Optics, vol. 16, no. 7, 2011, DOI: 10.1117/1.3602852.

[6] Park, C. and Choi, H., "Effective methods to extract PPG signals from face using stochastic state space modeling approach", 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), pp. 172-173, 2015, DOI: 10.1109/URAI.2015.7358857.

[7] More, A.V., Wakankar, A., Gawande, J.P., "Automated Heart Rate Measurement Using Wavelet Analysis of Face Video Sequences", 6th International Conference on Innovation in Electronics and Communication Engineering (ICIECE 17), vol. 33, pp.113-120, 2019, Springer, Singapore, https://doi.org/10.1007/978-981-10-8204-7-11.

[8] Vignesh Kumar, K., Periyasamy, R., Senthil Kumar, K., Suresh Chelliah, D., Senguttuvan, D. , "Development of non-contact optical device for monitoring neonatal jaundice based on the skin color of the upper trunk using skin reflectometry", International Journal of Biomedical Engineering and Technology, vol. 42, No.2, pp.135-149, 2023, DOI: 10.1504/IJBET.2023.131707.

[9] Stricker, R., Müller, S. and Gross, H.M., "Non-contact video-based pulse rate measurement on a mobile service robot", In The 23rd IEEE International Symposium on Robot and Human Interactive Communication (pp. 1056-1062), 2014.

[10] Li, X., Chen, J., Zhao, G. and Pietikainen, M., "Remote heart rate measurement from face videos under realistic situations", In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4264-4271, 2014.

[11] Po, L. M., et al. "Block-based adaptiveROI for remote photoplethysmography", Multimed. Tools Applications **77**, 6503–6529, 2018.

[12] Bobbia, S., et al. "Real-time temporal superpixels for unsupervised remote photoplethysmography", In: IEEE CVPRW, pp. 1341–1348, 2018.

[13] Li, P., Benezeth, Y., Nakamura, K., Gomez, R. and Yang, F., "Model-based region of interest segmentation for remote photoplethysmography". In 14th International Conference on Computer Vision Theory and Applications SCITEPRESS-Science and Technology Publications, pp. 383-388, February 2019.

[14] Kumar, M., Veeraraghavan, A. and Sabharwal, A., "DistancePPG: Robust non-contact vital signs monitoring using a camera", Biomedical optics express, 6(5), pp.1565-1588, 2015.

[15] De Haan, G. and Jeanne, V., "Robust pulse rate from chrominance-based rPPG", IEEE transactions on biomedical engineering, 60(10), pp.2878-2886, 2013.

[16] Wang, W., Stuijk, S. and De Haan, G., "A novel algorithm for remote photoplethysmography", Spatial subspace rotation. IEEE transactions on biomedical engineering, 63(9), pp.1974-1984, 2015.

[17] Poh, Ming-Zher, Mc Duff, D.J., and Picard, R.W. "Non contact, automated cardiac pulse measurements using video imaging

and blind source separation", Optics Express, vol. 18, No. 10, pp.10762−10774, 2010, https://doi.org/10.1364/OE.18.010762.

[18] Song, R., Zhang, S., Cheng, J., Li, C. and Chen, X., "New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method", Computers in biology and medicine, 116, p.103535, 2020.

[19] Huang, P.W., Wu, B.J. and Wu, B.F., "A heart rate monitoring framework for real-world drivers using remote photoplethysmography", IEEE journal of biomedical and health informatics, 25(5), pp.1397-1408, 2020.

[20] Macwan, R., Benezeth, Y., Mansouri, A., Nakamura, K. and Gomez, R., "Remote photoplethysmography measurement using constrained ICA". In 2017 E-Health and Bioengineering Conference (EHB) (pp. 430-433), June 2017.

[21] Chen, D.Y., et al., "Image sensor-based heart rate evaluation from face reflectance using Hilbert huang transform", IEEE Sens. J. **15**, 618–627, 2015.

[22] Demirezen, H., Erdem, C.E. "Heart rate estimation from facial videos using nonlinear mode decomposition and improved consistency check", Signal, Image and Video Processing, vol. 15, No. 7, pp.1415-1423, 2021, https://doi.org/10.1007/s11760-021-01873-x.

[23] C. S. Pilz, S. Zaunseder, J. Krajewski, and V. Blazek, "Local group invariance for heart rate estimation from face videos in the wild," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1254–1262, 2018.

[24] G. de Haan and A. van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," Physiological Measurement, vol. 35, no. 9, pp. 1913–1926, aug 2014.

[25] Lewandowska, M., Rumiński, J., Kocejko, T., Nowak, J. "Measuring pulse rate with a webcam - a non-contact method for evaluating cardiac activity", Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 405-410, 2011.

[26] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic principles of remote ppg," IEEE Transactions on Biomedical Engineering, vol. 64, no. 7, pp. 1479–1491, 2016.

[27] W. Wang, S. Stuijk, and G. De Haan, "A novel algorithm for remote photoplethysmography: Spatial subspace rotation," IEEE transactions on biomedical engineering, vol. 63, no. 9, pp. 1974–1984, 2015.

[28] Heusch, G., Anjos, A. and Marcel, H. "A reproducible study on remote heart rate measurement", arXiv, 2016, https://doi.org/10.48550/arXiv.1709.00962.

[29] Massoz, Q., Langohr, T., François, C., Verly, J.G. "The ULg Multimodality Drowsiness Database (called DROZY) and Examples of Use", Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV 2016), pp.1-7, 2016, DOI: 10.1109/WACV.2016.7477715.

[30] A. Noulas, and B. Krıse, "EM detection of common origin of multi-modal cues," in Proceedings of ACM Conference on Multimodal Interfaces (ACM, 2006), pp. 201–208, 2006.

[31] R. Lienhart, and J. Maydt, "An extended set of Haar-like features for rapid object detection," in Proceedings of IEEE Conference on Image Processing (IEEE, 2002).

[32] P. Viola, and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2001), p. 511, 2001.

[33] Lowe, D.G. "Distinctive Image Features from Scale Invariant Keypoints", International Journal of Computer Vision 60, pp.91-110, 2004, https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[34] Fu, T.H., Liu, S.H. and Tang, K.T. "Heart rate extraction from photoplethysmogram waveform using multiresolution analysis", Journal of Medical and Biological Engineering, Vol. 33, pp.131–136, 2008.

[35] German-Sallo, Z. "Wavelet transform based HRV analysis", Procedia Technology, Vol. 12, pp.105–111, 2014.

[36] Jegan, R., Anusuya, K.V. and George, E.M., "Real-time ECG peak detection for heart rate measurement using wavelet packet transform", International Journal of Biomedical Engineering and Technology, 19(3), pp.244-254, 2015.

[37] Alonsoa F J, Castilloa J M, Pintadob P. "Application of singular spectrum analysis to the smoothing of raw kinematic signals", Journal of Biomechanics, 38(5): 1085–1092, 2005

[38] Salgado D R, Alonso F J. Tool wear detection in turning operations using singular spectrum analysis [J]. Journal of Materials Processing Technology, 171(3): 451–458, 2006.

[39] Hao J, Yang Y, Zhou Z, Wu S., "Fetal Electrocardiogram Signal Extraction Based on Fast Independent Component Analysis and Singular Value Decomposition", Sensors (Basel), 12;22(10):3705, 2022 May, doi: 10.3390/s22103705.

[40] Bland, J.M. and Altman, D.G., "Bland-Altman Plot and Analysis", NCSS Statistical Software, 1999.

[41] Biradar, H.G. and Gawande, J.P., "Heart Rate Estimation from Facial Video Sequences using Fast Independent Component Analysis", National Conference on Communications (NCC), pp. 88-93, 2022, DOI: 10.1109/NCC55593.2022.9806810.

[42] Zhao, F., Li, M., Qian, Y., and Tsien, J.Z. , "Remote measurements of heart and respiration rates for telemedicine", PLoS ONE, vol. 8, no. 10, 2013, DOI: 10.1371/journal.pone.0071384.

[43] Zhang, Q., Zhou, Y., Song, S., Liang, G. and Ni, H., "Heart rate extraction based on near-infrared camera: Towards driver state monitoring", IEEE Access, vol. 6, pp. 33076-33087, 2018, DOI: 10.1109/ACCESS.2018.2845390.

[44] Hemlata Biradar and Jayanand Gawande, "Non-contact Heart Rate Measurement from Face Video Sequences using SIFT and FastICA", International Journal of Biomedical Engineering and Technology, in press.