Real-Time Heart Rate Estimation Using a Standard Camera

Ali Sameer Salim[1] and Abdul Sattar Mohammed Khidhir[2]
[1]*Department of Computer Techniques Engineering, Technical Engineering College, Northern Technical University, 41001 Mosul, Iraq*
[2]*Department of Networks and Computer Software Techniques, Mosul Technical Institute, Northern Technical University, 41001 Mosul, Iraq*
alisameer90@ntu.edu.iq, abdulsattarmk@ntu.edu.iq

Keywords: Remote Photoplethysmography (rPPG), Non-Contact Heart Rate Monitoring, Remote Health Monitoring, Computer Vision, Deep Learning.

Abstract: Health indicators are among the most important ways to assess health status and predict sudden health relapses. Non-contact health monitoring is critical for overcoming limitations of traditional vital sign measurement methods, such as infection risks and patient discomfort during continuous monitoring. This study will present a method for real-time heart rate measuring without contact with people by using a standard webcam and applying processing on the extracted signal from frames of video, applying a Butterworth filter, Chrominance-based method algorithm to obtain a signal similar to the heart rate signal, then calculate the heart rate using two methods, the first peaks detection and second method by Fast Fourier transform (FFT) and compare between them. These Experiments were conducted indoors with a lighting source in front of participants, and the number of participants was 15. The results obtained from the proposed method were compared with those of the Pulse Oximeter, which was medically approved, and the final results of median absolute error (MAE) were 1.28 for the Peak detection method and 2.47 for the FFT method. These results demonstrate the potential of webcam-based heart rate monitoring as a reliable, non-invasive alternative for health assessment.

1 INTRODUCTION

Remote Photoplethysmography (rPPG) is a non-contact method for calculating health indications, including blood pressure, heart rate estimation, and breathing rate, using standard cameras. This paper will discuss remote heart rate estimation. Remote tracking of health indicators has been widely popular because it offers a more comfortable and unobtrusive method when calculating the health indicators, especially for people with sensitive or burnt skin or who require continuous vital sign measurement, as in an intensive care unit (ICU) also in some cases, contact with a Patient carrying an infectious disease must be avoided when taking healthy measurements [1]. The principle of operation of the rPPG involves more than one step. After recording the video of the face with a traditional camera and then dividing the recorded video into frames to apply processing signals on each frame, the first step is to select the region of interest (ROI) after detecting the face using pre-trained network algorithms. The selected ROI has thin skin and veins close to the skin surface [2]. The second step is to count each frame's average colour to obtain the raw signal. Changed in colour value calculated It depends on the amount of light absorbed and reflected from the blood vessels under the surface of the skin, as the heartbeats with each beat cause an increase in the amount of blood inside the veins, which is the leading cause of the difference in the amount of absorbed and reflected light as shown in Figure 1 [3]. Then apply filtering on the raw signal to remove the noise after that, apply the algorithms such as Independent Component Analysis (ICA) [4], Principal Component Analysis (PCA) [5], Laplacian Eigenmap (LE) [6], Eulerian Video Magnification (EVM) [7], CHROM [8], Plane Orthogonal-to-Skin (POS) [9], and convolutional neural network (CNN) [10], [11]. All previous algorithms used to extract the final signal were similar to the heart rate signal. The results of the rPPG were acceptable compared with the availability of the technologies used and comfortable for the patient during the measurement [12]. There are some challenges in the rPPG approach, and needs to take

into account several different matters during the design stage of the rPPG, and among these matters is the physical distance from the camera that records the video and the person, controlling the number of frames during processing, the various environments and lighting conditions, the diversity of Samples that the systems are tested on it in terms of age, gender and skin colour; The movement of people during the test must also be taken into account; the methods that use deep learning must consider the previous matters in the database used to make the results more accurate [13], [14]. This paper proposes a method to estimate the heart rate remotely based on a web camera. The suggested method focuses on removing the noise from the main signal and then applying the Chrome method to obtain the heart rate signal. Afterwards, two methods are applied to determine the heart rate: first, counting the peaks of the obtained signal, and second, applying the FFT to the signal. For further assessment, the results of the two methods are compared with those obtained by the pulse oximeter.

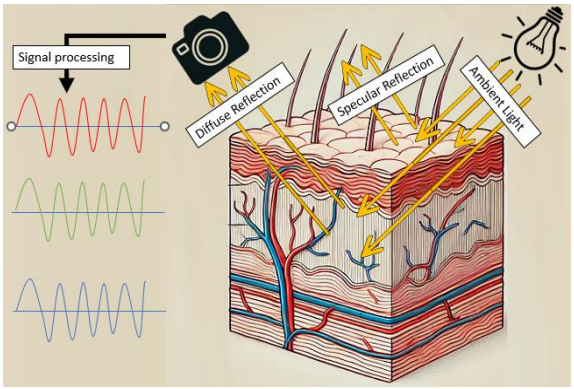

Figure 1: Principles of absorption and reflection of light on the skin.

The rest of the paper is organized as follows: Section Two provides a Theoretical background, Section Three describes the Experiment setup, Section Four describes the experiment's methodology, Section Five presents the results and discussion, and Section Six gives the conclusions.

2 THEORY

This section will explain the background theory of the parts used in the proposed method.

2.1 Face Detection

YuNet is a model developed by Shiqi Yu et al., an efficient, lightweight face detector created for edge devices that need low power consumption and high-speed processing. It is a single-stage face detector based on a fully convolutional neural network (CNN), which can accurately predict the precise location of faces in an image. YuNet utilizes depth-wise separable convolution and channel attention mechanisms to reduce computational complexity while maintaining accuracy. It is lightweight and smaller than other models, such as MTCNN and YOLO, for face detection while giving a strong performance, achieving 81.1% accuracy on the WIDER FACE dataset [15], [16]. It is optimized for real-time performance on low-power devices like mobile phones or embedded systems. As a result, it can achieve at least 100 frames per second (FPS) efficiently while maintaining high-performance accuracy [17].

2.2 Facial Landmark

Dlib’s 68-point facial landmark detector is widely used in facial recognition and analysis applications. Developed with machine learning techniques, it can detect 68 distinct facial landmarks on a human face. These landmarks include key points around the eyes, nose, mouth, jawline, and brows, allowing for a detailed mapping of facial features. The model works by detecting the face from the image using a histogram of oriented gradients (HOG) or other more efficient methods based on CNN. Then, the face is aligned, and the points are found in the eyes, nose, and mouth. Then, combine them to obtain an array of 68 points in the form of (x, y) coordinates that map to the facial structures of the face, as shown in Figure 2. [18]. The time it takes for Dlib's 68-point facial landmark detector to process and detect points depends on factors such as the CPU’s specifications and the image resolution. Typically, on mid-range processors (e.g., Intel Core i5 or i7) and using standard image resolutions (such as 640x480 pixels), the detector operates at around 15 to 30 milliseconds per image, which translates to approximately 33 to 66 FPS. This type of algorithm currently has many applications, including those in augmented reality and emotion detection [19], [20]. The proposed method will use the facial landmarks to determine the ROI.

2.3 Signal Filtering

Noise is one of the most critical obstacles we face when processing signals. The proposed method uses several filters to overcome these obstacles, removing noise according to its properties. Then, the required data can be extracted from the given signal [21].

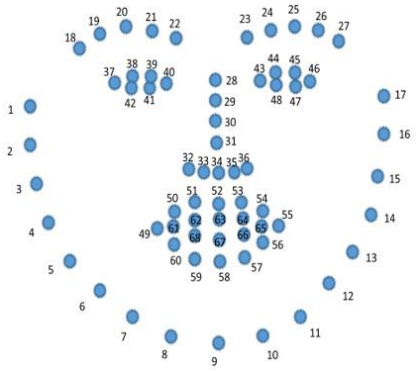

Figure 2: 68 Facial landmarks [19].

2.3.1 Butterworth Filtering

It is a filter used to let a specific range of frequencies pass through while blocking frequencies outside that range. The Butterworth design gives it a uniquely smooth and flat response within the passband, meaning it doesn't have any ripples or unwanted bumps, so the signal stays as clean as possible within the allowed frequency range. Butterworth filters are known for their simplicity and reliability. They're widely used for a clean, balanced signal, particularly in systems where phase accuracy [22].

2.3.2 Detrending

It removes the "drift" from a signal- those gradual ups and downs unrelated to the data of interest. Detrending eliminates these slow changes during heart rate monitoring, allowing us to focus on the rapid, beat-to-beat variations that indicate blood flow [23].

2.3.3 The Savitzky-Golay

Filter is a smart way to smooth out data without affecting essential details. This filter smooths the signal but keeps the peaks and troughs representing meaningful changes. Here, we use the Savitzky-Golay filter to reduce random noise in the heart rate signal, ensuring that the critical variations of heartbeat cycles are preserved and accessible to analyze [24].

2.4 Chrominance

The Chrominance Algorithm is integral to image and video processing. It handles colour information by separating it from luminance (brightness), enhancing colour correction, segmentation, and compression tasks. This approach originates from the YCbCr and YUV colour spaces or uses custom transformations to emphasize specific colours or tones by linearly combining RGB channels. This approach is useful in colour-based object detection, where chrominance channels represent colour without brightness, optimizing data processing and compression by reducing the data needed for colour while preserving visual quality. This method isolated the light from the obtained signal to get pure color data from video frames [25].

2.5 Peak Detection Algorithm

The peak detection algorithm is an important algorithm that detects the peaks of the signal through which the frequency that represents this signal is calculated accurately. This algorithm is based on determining the threshold or derivatives of the signal [26]. This algorithm was used in the proposed method to detect the peaks of the final signal. These peaks represent the highest recorded value of the rate of change in colour, which corresponds to the heartbeats. Following peak detection, the intervals between successive peaks are measured to calculate the final heart rate.

3 EXPERIMENT SETUP

The signal processing was implemented in Python version 3 using a computer with an Intel Core i7 2.5 GHz CPU and 8 GB RAM. The computer's webcam was used at 640 × 480 pixels with 30fps. The experiments were conducted on 15 participants between 18 and 60 years of age. The experiments were conducted indoors with an LED light source at a distance and height of 3 meters from the person sitting in front of the camera at a distance from 0.5 to 0.75 meters from the camera, as in Figure 3. Participants were asked not to move excessively and to sit generally in front of the camera. The results obtained from the method will be compared with those of a pulse oximeter.

Figure 3: Experimental setup.

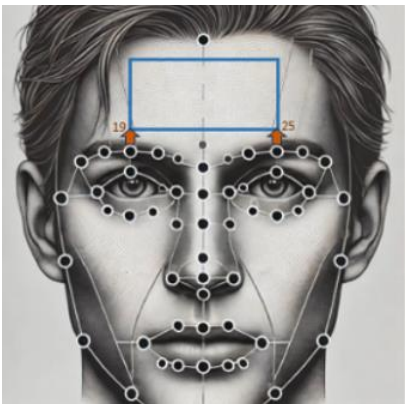

Figure 4: Selecting ROI.

4 METHODOLOGY

This study employed real-time heart rate monitoring based on video analysis using a combination of Python libraries and pre-trained models. Essential libraries include NumPy for numerical computations, cv2 (OpenCV) for video processing, YuNet for face detection, dlib for facial landmark detection, and matplotlib for data visualization. The Scipy signal library implements advanced signal filtering and peak detection functions.

4.1 ROI Extraction and Tracking

Non-skin areas (including eyes, hair, and eyebrows) are not useful for obtaining the heartbeat signal. In most cases, the movement of these areas causes noise that affects the initial signal that needs to be extracted. To increase the signal-to-noise ratio (SNR), the initial signal is extracted from the ROI in the face, specifically from those with thin skin and where the blood vessels are close to the surface of the skin, which contain the initial signal closest to the heartbeat and not from the entire face. After opening the webcam, the recorded video is processed for each frame by Model YuNet, which is pre-trained to detect the entire face. This method is characterized by its high speed in identifying the face, and this is useful for us because the proposed method focuses on presenting HR in a real-time model. After that, the 68-point facial landmark is used to identify the facial features. Points 19 and 24, located above the eyebrows, are chosen to draw a rectangle in the middle of the forehead and consider ROI to take the initial heartbeat signal from it, as shown in Figure 4. The initial signal containing the heart rate data from the ROI will be extracted from each frame.

4.2 Count Frame Per Second

One of the most important things that must be calculated accurately is measuring the frame per second (FPS). The period between processing one frame and the next was calculated, and the value of FPS was calculated through this period. During the experiments noticed that these values change between one frame and another, and thus, the sampling frequency is variable. If it depends on one value for the FPS, the obtained results will be in error. In the proposed method, the FPS value is measured continuously throughout the heart rate measurement period, and to make the measurements more accurate, every 5 seconds, the average rate of FPS is calculated and not relying on one of the FPS values measured during the five seconds, FPS value is variable during run the code according to the availability of processing resource.

4.3 Signal Processing

The rate of change in colours is extracted by calculating the pixel values for red, green, and blue for the ROI and dividing them by the number of pixels in the ROI to obtain the average of these three colours for each processed frame. This is the initial signal which will apply the process below.

4.3.1 Detrending

The focus is on the periodic or subtle variations of the skin-reflected RGB signal, which represents the physiological signal. Therefore, the trend removal principle was used to improve the data analysis. Making the signal centered around zero reduces the complexity of operations.

4.3.2 Band-Pass Butterworth Filter

The initial signal obtained from the ROI suffers from unwanted frequencies resulting from noise, lighting changes, or movement. To remove these frequencies, a Butterworth filter with a 5th-order and frequency range between (0.8 - 3) HZ will be applied to the RGB signal so that the potential heartbeats will be between (48- 180). By focusing only on this range, the filter removes frequencies outside the range of the heart rate frequency, helping us get a cleaner signal directly linked to the heartbeat signal.

4.3.3 Savitzky-Golay Filter

Applying the Savitzky-Golay Filter with window length=11 and polyorder=2 to an RGB signal smooths out the data obtained while preserving essential details like the main peaks of the signal. This filter works to polish the received signal from the Butterworth filter to remove the oscillations centred on this signal, thus improving the results of the following operations.

4.4 Chrominance

This method has been used previously and has proven its effectiveness in dealing with the problems of changes in lighting and its effect on the results. The basic principle of this method is that the original signals (RGB) contain useful information but are mixed or disturbed by noise or external influences. This means it Amplifies colour channels that contain physiological signals and subtracts them from others. By Linear Mathematical Operations of the channels and analyzing their relationship, we get a signal close to the heartbeat signal. From the experiments, we found that the red channel contains a more significant physiological signal than the others. so will amplify the red channel due to its unique interaction with human skin and blood. Specifically, red light has a longer wavelength, which allows it to penetrate deeper into the skin layers. This deeper penetration enables it to reach blood vessels beneath the surface, making it less susceptible to interference from minor skin movements and surface changes. Consequently, red light can provide a more stable signal in varying environmental conditions, including different lighting settings. After applying this method, the final signal is obtained, which will be relied upon in the next step, which is the last, to extract the heart rate. Figure 5. shows the signals obtained after each step for the red channel, specifically because it has strong physiological signals.

4.5 Heart Rate Estimation

The proposed method focused on two methods to measure heartbeats from the final signal. The first was using a Fast Fourier transform to find the frequency representing this signal, then multiplying the result by 60 seconds to obtain the final result for Beats Per Minute (BPM). The second was to detect the peaks of the final signal and calculate the time taken between the peaks, then divide the 60 seconds by the average time between peaks to obtain the final BPM. The results of the two methods are compared in the results section. After executing the code, the previous steps will be executed to obtain the final signal, and then the heartbeats will be calculated. The pulse reading will be taken every second and stored in a stack without displaying it on the real-time screen, with the removal of values that carry a severe increase or decrease within a short time, because this data carries high noise and thus affects the final results. After that, the average of these readings is calculated every 5 seconds to be displayed on the display screen.

5 RESULTS AND DISCUSSION

The proposed method calculated the final heart rate result from the final signal using the FFT and peak detection. The recorded results were for a period of 50 seconds, showing the minimum and maximum HR recorded by the proposed method, the average of recorded values, Standard Deviation (SD), absolute error (AE), and the average of the Pulse Oximeter. Table 1 shows the results of the FFT method. It is clear from this table that the error range extends from (0.2-5.06), whereas the standard deviation values fluctuate between (2.14-8.29). The results of the detection peaks method are shown in Table 2. It can be seen that lower error values and standard deviation were obtained using this method. Table 3 shows the comparison between the results for the two methods. The peak detection method gives HR a lower error rate than the FFT method. However, both methods estimate the HR with an acceptable error range. Figures 6 and 7 illustrate the differences between each method and the pulse oximeter. The FFT method showed differences ranging from -5 to 3 bpm, while Peak detection differences were tighter, ranging from -2 to 2 bpm, with most values clustered around zero. Figure 8 highlights the distribution of absolute errors: FFT errors reached up to 5 bpm, whereas Peak detection errors peaked at approximately 2.5 bpm. These findings indicate that Peak detection is more accurate and consistent, making it preferable for applications requiring precise heart rate monitoring.

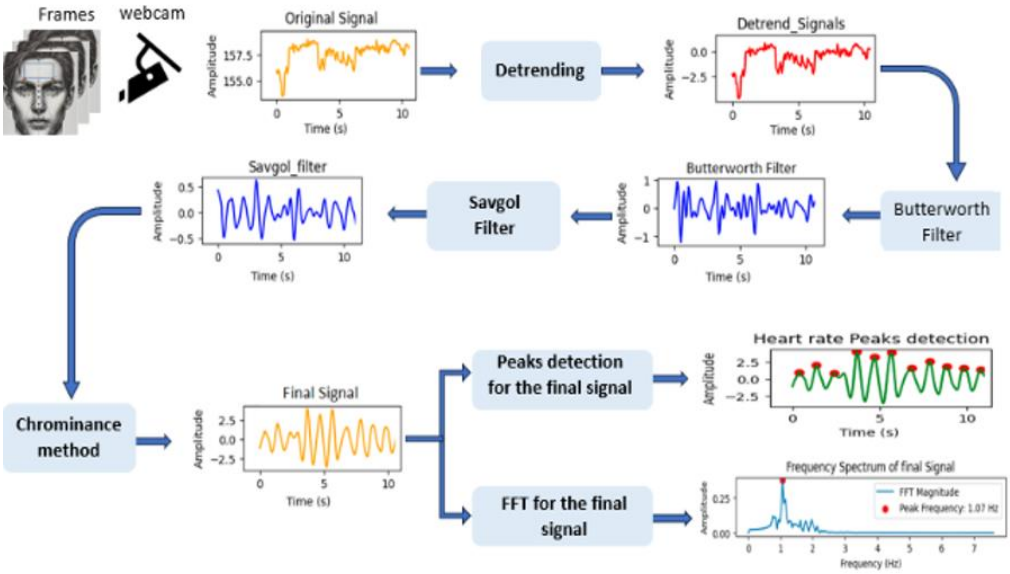

Figure 5: Shows the steps of the proposed method.

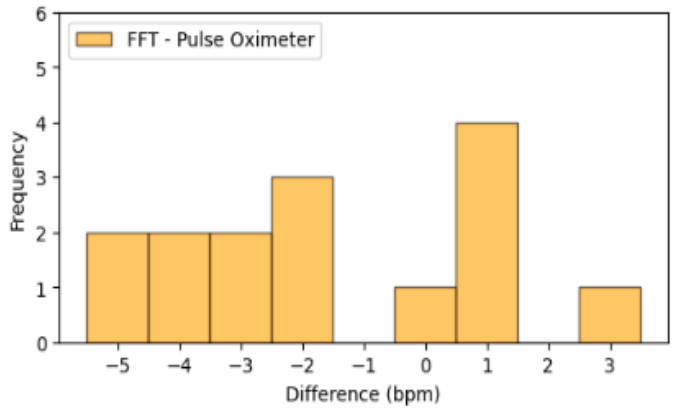

Figure 6: Histogram of Differences between FFT and Pulse Oximeter.

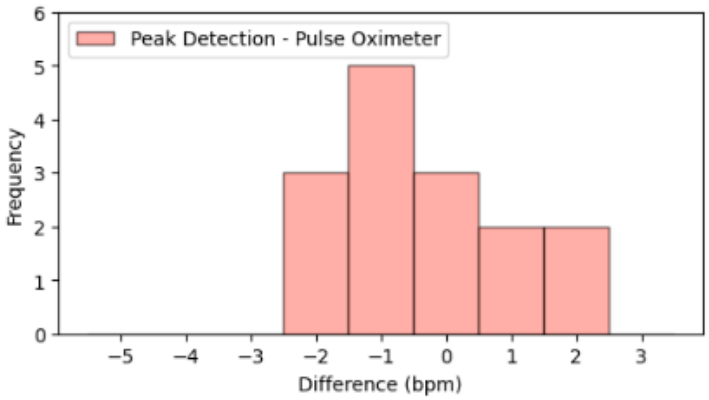

Figure 7: Histogram of differences between peak detection and pulse oximeter.

Table 1: Results of HR using the FFT method.

Objects	Min. HR	Max. HR	Av. HR	S.D.	A.E.	Av. Pulse Oximeter
Object 1	78.3	92.02	87.99	6.64	3.91	91.9
Object 2	63.88	67.05	66.7	2.48	3.6	70.3
Object 3	50.96	61.19	57.2	4.08	2.8	60
Object 4	73.2	90.2	80.14	7	5.06	85.2
Object 5	70.5	85.57	80.16	6.02	2.78	82.94
Object 6	50.27	54.51	52.97	1.72	1.53	54.5
Object 7	60.68	65.55	63.03	2.14	1.97	65
Object 8	49.31	63.7	57.8	4.87	0.2	58
Object 9	78.18	89.92	83.64	4.07	1.17	82.47
Object 10	60.37	66.99	64.5	2.49	1.09	63.41
Object 11	66.66	75.95	71.74	3.45	3.16	68.58
Object 12	56	62.02	58.64	2.87	1.14	57.5
Object 13	73	94.08	81.96	8.29	2.34	84.3
Object 14	69.4	78.63	76.03	3.84	5.02	81.05
Object 15	54.55	62.86	59.47	2.91	1.29	58.18

Table 2: Results of HR using the Peak detection method.

Objects	Min. HR	Max. HR	Av. HR	S.D.	A.E.	Av. Pulse Oximeter
Object 1	86.51	95.62	90.55	4.08	1.35	91.9
Object 2	64.3	75.41	69.1	4.38	1.2	70.3
Object 3	55.97	60.55	58.71	2.07	1.29	60
Object 4	80.16	86.17	82.78	2.16	2.42	85.2
Object 5	73.03	85.14	80.69	4.14	2.25	82.94
Object 6	55.25	58.07	56.75	1	2.25	54.5
Object 7	61.28	67	63.2	2.04	1.8	65
Object 8	55.5	62.46	58.2	2.64	0.2	58
Object 9	77.82	87.63	82.43	3.74	0.04	82.47
Object 10	56.76	66.49	62.49	3.84	0.92	63.41
Object 11	75.72	64.39	69.3	4.27	0.72	68.58
Object 12	54.76	60.82	58.39	2.41	0.89	57.5
Object 13	79.29	94.72	86.54	6.22	2.24	84.3
Object 14	76.44	81.79	79.72	2.54	1.33	81.05
Object 15	55.11	60.93	58.54	2.18	0.36	58.18

Table 3: Comparison between HR results for FFT and the Peak detection method.

Objects	FFT		Peak detection		Av. Pulse Oximeter
	Av. HR	A.E.	Av. HR	A.E.	
Object 1	87.99	3.91	90.55	1.35	91.9
Object 2	66.7	3.6	69.1	1.2	70.3
Object 3	57.2	2.8	58.71	1.29	60
Object 4	80.14	5.06	82.78	2.42	85.2
Object 5	80.16	2.78	80.69	2.25	82.94
Object 6	52.97	1.53	56.75	2.25	54.5
Object 7	63.03	1.97	63.2	1.8	65
Object 8	57.8	0.2	58.2	0.2	58
Object 9	83.64	1.17	82.43	0.04	82.47
Object 10	64.5	1.09	62.49	0.92	63.41
Object 11	71.74	3.16	69.3	0.72	68.58
Object 12	58.64	1.14	58.39	0.89	57.5
Object 13	81.96	2.34	86.54	2.24	84.3
Object 14	76.03	5.02	79.72	1.33	81.05
Object 15	59.47	1.29	58.54	0.36	58.18

The superior performance of Peak detection can be attributed to its focus on identifying distinct pulse peaks, which aligns well with the periodic but variable nature of heartbeat signals. Heart rates often fluctuate slightly due to factors like breathing or movement, and Peak detection effectively captures these peaks to calculate beats per minute (bpm).

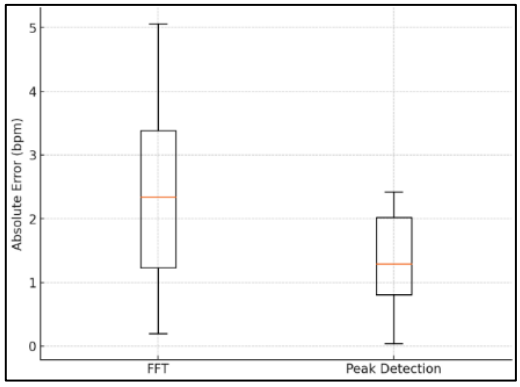

Figure 8: Absolute error distribution for FFT and peak detection.

Conversely, FFT, while powerful for analyzing complex signals, is more sensitive to noise, such as lighting changes or motion in non-controlled environments. This sensitivity can distort frequency analysis, leading to less accurate heart rate estimates. Therefore, for non-contact heart rate monitoring using a webcam, Peak detection appears to be the more reliable method

6 CONCLUSIONS

This paper presents a method for measuring heart rate in real time using a Webcam. Experiments conducted indoors with a light source in front of the person have proven high accuracy in heart rate estimation. In addition, the results were compared between the peak detection and FFT methods, where the first method outperformed the FFT regarding the absolute error rate. These results highlight the potential of webcam-based monitoring as a reliable, non-invasive solution for healthcare applications. The findings revealed that peak detection consistently outperformed FFT in terms of absolute error rate, providing more precise and stable results. These outcomes highlight the promise of webcam-based monitoring as a reliable, non-invasive solution for healthcare applications, particularly for remote patient monitoring and telemedicine services. The future work could explore enhancing robustness under varying lighting conditions, incorporating advanced signal processing or machine learning techniques, and extending the system to measure additional vital signs, ultimately paving the way for more comprehensive and accessible digital health solutions.

REFERENCES

[1] A. Al-Naji, K. Gibson, S. H. Lee, and J. Chahl, "Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods," IEEE Access, vol. 5, pp. 15776–15790, Aug. 2017, doi: 10.1109/ACCESS.2017.2735419.

[2] X. Chen, J. Cheng, R. Song, Y. Liu, R. Ward, and Z. J. Wang, "Video-based heart rate measurement: Recent advances and future prospects," IEEE Trans. Instrum. Meas., vol. 68, no. 10, pp. 3600–3615, Oct. 2019, doi: 10.1109/TIM.2018.2879706.

[3] K. M. van der Kooij and M. Naber, "An open-source remote heart rate imaging method with practical apparatus and algorithms," Behav. Res. Methods, vol. 51, no. 5, pp. 2106–2119, Oct. 2019, doi: 10.3758/s13428-019-01256-8.

[4] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," in Proc. Conf., 2010.

[5] M. Lewandowska and J. Nowak, "Measuring pulse rate with a webcam," J. Med. Imaging Health Inform., vol. 2, no. 1, pp. 87–92, Mar. 2012, doi: 10.1166/jmihi.2012.1064.

[6] L. Wei, Y. Tian, Y. Wang, T. Ebrahimi, and T. Huang, "Automatic webcam-based human heart rate measurements using Laplacian eigenmap," in Lecture Notes in Computer Science, vol. 7725, 2012.

[7] H. Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," ACM Trans. Graph., vol. 31, no. 4, Jul. 2012, doi: 10.1145/2185520.2185561.

[8] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," IEEE Trans. Biomed. Eng., vol. 60, no. 10, pp. 2878–2886, 2013, doi: 10.1109/TBME.2013.2266196.

[9] W. Wang, A. C. Den Brinker, S. Stuijk, and G. De Haan, "Robust heart rate from fitness videos," Physiol. Meas., vol. 38, no. 6, pp. 1023–1044, May 2017, doi: 10.1088/1361-6579/aa6d02.

[10] B. Huang, C. L. Lin, W. Chen, C. F. Juang, and X. Wu, "A novel one-stage framework for visual pulse rate estimation using deep neural networks," Biomed. Signal Process. Control, vol. 66, Apr. 2021, doi: 10.1016/j.bspc.2020.102387.

[11] R. A. Firmansyah, Y. A. Prabowo, T. Suheta, and S. Muharom, "Implementation of 1D convolutional neural network for improvement remote photoplethysmography measurement," Indonesian J. Elect. Eng. Comput. Sci., vol. 29, no. 3, pp. 1326–1335, Mar. 2023, doi: 10.11591/ijeecs.v29.i3.pp1326-1335.

[12] P. V. Rouast, M. T. P. Adam, R. Chiong, D. Cornforth, and E. Lux, "Remote heart rate measurement using low-cost RGB face video: A technical literature review," Front. Comput. Sci., vol. 12, no. 5, pp. 858–872, Oct. 2018, doi: 10.1007/s11704-016-6243-6.

[13] C. H. Cheng, K. L. Wong, J. W. Chin, T. T. Chan, and R. H. Y. So, "Deep learning methods for remote heart rate measurement: A review and future research agenda," Sensors, vol. 21, no. 18, Sep. 2021, doi: 10.3390/s21186296.

[14] R. J. Lee, S. Sivakumar, and K. H. Lim, "Review on remote heart rate measurements using photoplethysmography," Multimed. Tools Appl., vol. 83, no. 15, pp. 44699–44728, May 2024, doi: 10.1007/s11042-023-16794-9.

[15] W. Chen, H. Huang, S. Peng, C. Zhou, and C. Zhang, "YOLO-face: A real-time face detector," Vis. Comput., vol. 37, no. 4, pp. 805–813, Apr. 2021, doi: 10.1007/s00371-020-01831-7.

[16] O. A. Naser, S. M. S. Ahmad, K. Samsudin, M. Hanafi, S. M. B. Shafie, and N. Z. Zamri, "Facial recognition for partially occluded faces," Indonesian J. Elect. Eng. Comput. Sci., vol. 30, no. 3, pp. 1846–1855, Jun. 2023, doi: 10.11591/ijeecs.v30.i3.pp1846-1855.

[17] W. Wu, H. Peng, and S. Yu, "YuNet: A tiny millisecond-level face detector," Mach. Intell. Res., vol. 20, no. 5, pp. 656–665, Oct. 2023, doi: 10.1007/s11633-023-1423-y.

[18] G. Amato, F. Falchi, C. Gennaro, and C. Vairo, "A comparison of face verification with facial landmarks and deep features," in Proc. Int. Conf. Advances Multimedia (MMEDIA), Pisa, Italy, 2018.

[19] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), pp. 1867–1874, 2014.

[20] K. Ramyasree and C. S. Kumar, "Acoustic and visual geometry descriptor for multi-modal emotion recognition from videos," Indonesian J. Elect. Eng. Comput. Sci., vol. 33, no. 2, pp. 960–970, Feb. 2024, doi: 10.11591/ijeecs.v33.i2.pp960-970.

[21] I. Berezhnyi and A. Nakonechnyi, "Analysis of methods and algorithms for remote photoplethysmography signal diagnostic and filtering," Adv. Cyber-Phys. Syst., vol. 9, no. 1, pp. 82–88, May 2024, doi: 10.23939/acps2024.01.082.

[22] S. F. Hussin, G. Birasamy, and Z. Hamid, "Design of Butterworth band-pass filter," Politeknik & Kolej Komuniti J. Eng. Technol., vol. 1, pp. 32–46, 2016.

[23] S. Akdemir Akar, S. Kara, F. Latifoglu, and V. Bilgiç, "Spectral analysis of photoplethysmographic signals: The importance of preprocessing," Biomed. Signal Process. Control, vol. 8, no. 1, pp. 16–22, Jan. 2013, doi: 10.1016/j.bspc.2012.04.002.

[24] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," Anal. Chem., vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[25] F. Pierre, J. F. Aujol, A. Bugeau, N. Papadakis, and V. T. Ta, "Luminance-chrominance model for image colorization," SIAM J. Imaging Sci., vol. 8, no. 1, pp. 536–563, Mar. 2015, doi: 10.1137/140979368.

[26] Q. Qin, J. Li, Y. Yue, and C. Liu, "An adaptive and time-efficient ECG R-peak detection algorithm," J. Healthc. Eng., vol. 2017, 2017, doi: 10.1155/2017/5980541.