Web Pulse Monitor

PUSHKAL KUMAR SHUKLA
(Assistant Professor at Inderprastha Engineering College, Ghaziabad)

1 st Kamlesh Singh
kamleshs1807@gmail.com

2 nd Anjali Singh
singh07anjali24@gmail.com

3 rd Aman Kumar
amankumar84256@gmail.com

INFORMATION TECHNOLOGY DEPARTMENT

INDERPRASTHA ENGINEERING COLLEGE

(2022-2023)

Abstract. We will calculate a person's heart rate using a basic webcam or even their phone's camera, making it a non-contact method. It will be easier to take care of your health as we move closer to a time when everyone is concerned about their wellbeing. Having a machine that measures your heartbeat is rather expensive, so our aim is to use our current resources to measure one's heartrate without incurring any additional costs. Elevated or abnormal heart rates can be fatal or serious for a normal or a middle aged person. Each time the heart beats, blood is expelled and travels through the body. This blood flow can be detected in the face using a standard webcam that is able to pick up subtle changes in color that cannot be seen by the naked eye. Due to the light absorption spectrum of blood, we are able to detect differences in the amount of light absorbed by the blood traveling just below the skin (i.e., photoplethysmography).Emotions can be detected by using haar cascades, feature extraction from ROI and then processing it.OpenCV and Python has been used for human computer interaction.For the detection of heartbeat, we use remote photoplethmysography.This paper discusses the reimplementation of one such approach that uses independent component analysis on mean pixel color values within a region of interest (ROI) about the face.

Keywords: ***PPG (Photoplethysmogram), HR(Heart rate), BPM(Beats per minute), HB(Heart beat)***

1. Introduction

As a result of sitting down jobs, a lack of exercise, and unhealthy eating habits, there are more and more cardiovascular illnesses in the population today, which increases the number of fatalities each year.

To live a proper healthy life, we should be able to understand our bodies. However, many pieces of equipment that are used to measure heartbeat are expensive and not readily available in many locations or even available 24/7, so this project will use a webcam to take live snapshots of your face and by calculating the variation in colours that your forehead undergoes it will determine the HB.

The health, fitness, level of activity, stress, and many other factors of an individual may be determined by their heart rate. Using the electrocardiogram (ECG), which requires patients to wear chest straps with adhesive gel patches that can be scratchy and painful for the user, cardiac pulse is often recorded in clinical settings. Pulse oximetry devices that may be worn on the fingertip or earlobe can also be used to measure heart rate. The pressure from these sensors might get unpleasant with time and they are not comfortable to wear all day.

The process begins with your heart pumping blood through your veins, which alters the intensity of the colours on your body. A simple camera or webcam will take a live video of you while focusing on your forehead, changing the image or dividing it into RGB, but only the green image is used to calculate your heart rate. Webcams will detect minute changes in green colour that are invisible to the naked eye. Our major objective is to develop a system that uses a consumer-grade camera to assess your heart rate. We will display blood pressure and heart rate after computing them in real time on the camera app.

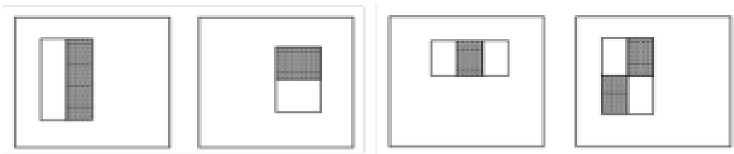


Figure 1. Example Haar-like features used in the boosted cascade face classifier from the original paper by Viola and Jones [11]. Each feature is calculated as the sum of pixels in the grey rectangles less the sum of pixels in the white rectangles.

A online pulse monitor is a self-monitoring device that enables real-time heart rate measurement and display. It is primarily used to collect heart rate information while engaging in various forms of physical activity. Hospitals often utilise many sensors and connected medical heart rate monitoring machines. Consumer-grade heart rate monitors are wire-free and intended for regular usage. We first recognise the patient's face, and then, utilising physiological data such as a photoplethysmogram (PPG), we compute the patient's heart rate utilising predetermined formulae and algorithms. Photo-plethysmograpy (PPG), which monitors fluctuations in blood volume by detecting changes in light reflectance or transmission during the cardiovascular pulse cycle, allows for the non-contact detection of heart rate.

PPG is typically carried out with specialised red or infrared light sources, just like a pulse oximetry sensor.

2. Literature Survey

- The plethysmographic signal may be seen in video taken with a typical colour camera, as demonstrated by Verkruysse et al. [10]. The researchers discovered that the signal could be seen in the red, green, and blue channels of colour video of exposed skin, but that it was highest in the green channel. This finding is consistent with the fact that haemoglobin has maxima in the absorption of green and yellow light wavelengths. In addition, they discovered that although the signal could be picked up all over the body, it was greatest on the face, particularly on the forehead.

- In order to properly extract the signal, Poh et al. observed that independent component analysis (ICA) should be used to distinguish distinct source signals from mixed colour data.

- Using Haar cascade classifiers, which were developed by Lienhart et al. [6] and suggested by Viola and Jones [11], face identification and tracking is carried out. We employ the OpenCV Cascade Classifier, which has been specifically trained on both positive and negative frontal face pictures [5]. Each classifier in the cascade of classifiers used to create the face detector employs one or more Haar-like characteristics.

- According to studies, colour changes in the face caused by pulse can be amplified by tiny variations between video frames [12], and vertical head motion can also be used to detect heart rate in addition to colour changes [1]. These are intriguing new advancements in the field, but they are less applicable to everyday life or medical applications since the former focuses more on visualisation than quantification and the later needs the patient to be perfectly motionless for measurements to be correct.

- E. Shih, J. Guttag, F. Durand, H.-Y. Wu, M. Rubinstein, W. Freeman [12] Eulerian Video Magnification employs spatial decomposition, followed by temporal filtering, to the frames of a typical video sequence as input. In order to disclose buried information, the resultant signal is then amplified. We can see the blood flow as it fills the face using our technology, as well as enhance and highlight minute movements. Our method may display events happening at the user-selected temporal frequency in real time.

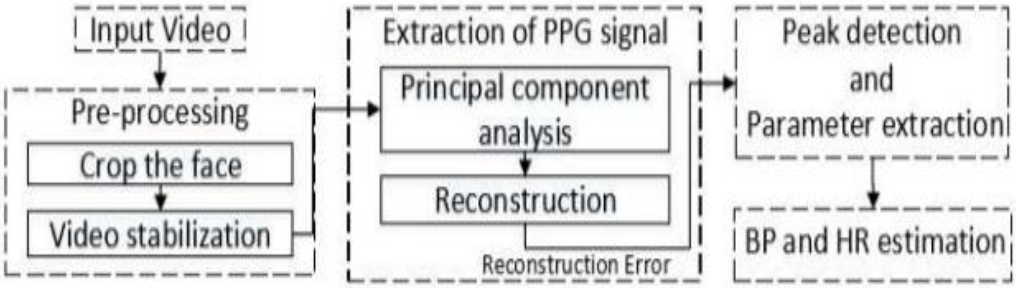
- D. McDuff, R. Picard, and M.-Z. Poh. automatic cardiac pulse readings without touch [8] Without the use of electrodes, pleasant physiological evaluation may be achieved by remote measurements of the heart pulse. Nevertheless, current methods are inefficient, prone to motion artefacts, and non-automated. In this study, we present a novel approach to solving these issues.

- M.-P. Jolly and Y. Boykov[3]. For the best border and region segmentation of objects in n-dimensional pictures, use interactive graph cuts. To give segmentation tight limits, the user labels certain pixels as "object" or "background." Information about boundaries and regions is incorporated into additional soft restrictions. The segmentation of the N-dimensional picture that is globally optimum is discovered using graph cuts.

- A. Gramfort, V. Michel, B. Thirion[7], F. Pedregosa, G. Varoquaux, A. Gramfort, This package focuses on utilising a general-purpose high-level language to make machine learning accessible to non-specialists. Usability, speed, documentation, and API consistency are prioritised. It is offered under the streamlined BSD licence, has few dependencies, and is useful in both professional and academic environments.

- Empirical analysis of detection cascades of boosted classifiers for quick object identification, R. Leinhart, A. Kuranov, and V. Pisarevsky[6]. It introduces a fresh set of rotational haar-like properties. The simple qualities of [6] are considerably enhanced by these unique features, which are likewise easily calculable. Our sample face detector displays an average 10% reduced false alarm rate at a given hit rate using these new rotating characteristics.

3.Methdology



3.1 TECHNICAL APPROACH

Heart rate detection primarily consists of three sections, the first of which concentrates on the face in the video frame since it is the part where this approach may successfully identify heart rate. Second, a region of interest (ROI), such as the forehead, must be selected for the bouncing box. The PPG signal must then be recovered from the ROI where the colours changed during that period, and it must then be examined to find the frequency within the heart rate range, which is between 60 and 100 beats per minute.

3.2 FACE DETECTION AND TRACKING

A machine learning-based method for face identification called Haar cascades involves training a cascade function using a collection of input data. For the face, eyes, and forehead, there are several pre-trained classifiers in our OpenCV library. Face recognition only functions on grayscale photos. So, it is crucial to convert the colour image to grayscale.

Here, we do so using the Face and Eye cascades, and after that, we discover a list of coordinates for the rectangular areas where faces were discovered. Similar to that, we use the coordinates of the eyeballs to draw rectangles in our video frame. The video frame is subjected to this face recognition algorithm, which generates a bounding box for each face it recognises.

According to figure 1, the features are made up of two, three, or four rectangular pixels. The total of pixels in the grey rectangles is subtracted from the amount of pixels in the white rectangles to compute each feature.

These characteristics are capable of identifying straightforward blobs, edges, and diagonals.

If there is no face identified in the frame, the face from the previous frame is utilised; if there are numerous faces detected in the frame, the face that is closest to that frame is used. This ensures consistency between frames. Each frame in the movie is subjected to this face identification algorithm, which generates a bounding box for each face it recognises.

3.3 REGION OF INTEREST SELECTION

Given that the face bounding box generated by the face identification method includes both background and facial pixels, a ROI (Region of Interest) must be selected from the area included inside the box. Since the bounding box is essentially outside the face region when measured in terms of face width but inside when measured in terms of face height. In essence, this approach is used to exclude background and exterior pixels, although some facial features, such as hairs at the corners of the bounding box, will still be visible.

The facedetection's bounding box has to be changed. A bounding box that was about 80% (width) of the original box's width and 20% (height) of the original box's height was used to guarantee that the bounding box contains all face pixels but excludes certain hair pixels.

We look at what happens when we remove the eye area, which has non-skin pixels that might change from frame to frame as a result of blinking or eye movement. The eyes could be eliminated by removing pixels from 25% to 50% of the bounding box height. As the forehead contains the highest plethysmographic signal, we also consider solely keeping the pixels in the area above the eyes.

The graph now has two more nodes that are joined to every other node. These nodes, which stand in for the foreground and background of the picture, will be arranged by the min-cut algorithm into opposing sets. The likelihood that a pixel belongs to either the foreground or the background is indicated by the weights of the terminal edges connecting it to those nodes. Gaussian Mixture Models (GMMs) are used to calculate these probabilities for the foreground and background pixel colour distributions.

Last but not least, we consider separating the face pixels from the backdrop pixels. GrabCut divides photos into segments for facial segmentation by repeatedly reducing an energy cost function. The foreground and background pixels of the picture may be represented by two sets of nodes, which can be obtained by constructing a graph model to describe the image and figuring out the minimal cut for the graph.

3.4 HEART RATE DETECTION

We may start extracting the rate from the colour picture data, such as a person's face, once we have achieved ROI for the video frame. In ambient lighting, a face video of the subject is taken while maintaining a distance of around half a metre between the camera and the face.

The individual is requested to sit motionless without closing their eyes for one minute while the video is being filmed.The face region of the video is used simply to extract the PPG signal, leaving out any further video frame content. By using a photoelectric approach, it measures the quantity of light that is absorbed in live tissue and finds changes in blood volume.

The amount of light that is absorbed by your skin varies in measurable proportions, just like when your heart pumps blood through your arteries and veins. Consumer cameras capture pictures in RGB (red, green, and blue) values, with the green channel supplying data that enables heart rate measurement.

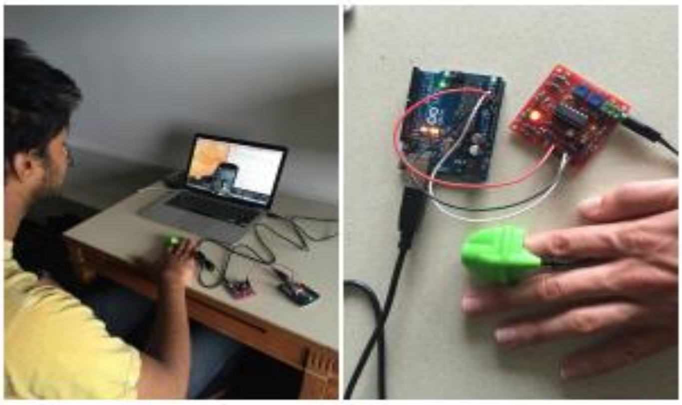


Figure 2. Experimental setup. One-minute videos of a subject's face were captured with a Samsung camera, and a reference PPG signal was recorded using the Easy Pulse fingertip PPG sensor and an Arduino Uno.

We can start to extract the heart rate from the colour picture data after we have a ROI for each frame. To acquire three signals, xR(t), xG(t), and xB(t), which correspond to the average red, green, and blue face pixels at time t, the ROI pixels are first averaged across each colour channel. The heart rate is then re-estimated every second as we normalise these signals over a 30-second sliding frame with a 1-second stride.

When the heart pumps blood into arteries close to the skin, more green light is absorbed and less is reflected because haemoglobin in the blood has an absorption peak for green light. After obtaining the source signals, we may use a Fourier transform to analyse the data and identify the dominant signal frequencies.

During the range of 0.75 to 4 Hz, which equates to natural heart rate ranges of 45 to 240 bpm, we are able to extract frequency peaks in the power spectrum. The frequency within the allowed range that corresponds to the peak with the greatest amplitude will be the heart rate that is measured.

While this assumption might not be accurate since variations in blood volume and the intensity of reflected light in skin tissue over distance might not be linear, it should be a good estimate for the 30-second time window.X

3. EXPERIMENTAL SETUP

Ten persons with different skin tones and lighting settings, including both natural and artificial light, were the topic of one-minute recordings. Videos were recorded using the front-facing camera on the phone and saved as mp4 files. The size of video frames was 480x640 pixels which were captured at 13.9 fps. Each subject was recorded in at least two videos: one in which the subject was as still as possible, and the other in which the subject moved a little.

Although the facial bounding box was typically well-centered on the face in each frame, we can imagine that in a noisier environment with more movement of the camera or the subject, poorer lighting, facial occlusions, and background clutter, there may be more error in the location of the facial bounding box.

4. **RESULT**

The experimental findings are described in the sections that follow. Some qualitative findings from each algorithmic stage are shown in Section 4.1. In Section 4.2, the algorithm accuracy for video of both stationary and moving subjects is displayed in comparison to the reference signal. The accuracy of the heart rate when different levels of noise are applied to the face bounding box are discussed in Section 4.3

3.1 The Algorithm Results

Figure 4 shows the outcomes of other ROI selection alternatives, such as a smaller box, a box without eyes, a box around the forehead, or a segmented face, but there was occasionally hair or background showing at the corners. Figure 6 illustrates how more background pixels entered the ROI as the participant tilted or rotated their head.

Figure 5 displays the segmentation outcomes for a straight face using the reimplementation of GrabCut, the original bounding box, and the updated bounding box. As can be observed, after a few cycles, the segmentation eliminates the majority of the hair and background pixels. This looks especially helpful when the subject bends or spins, as in picture 6, as it eliminates the background pixel fluctuation that was apparent when using the standard narrow bounding box. ROI.

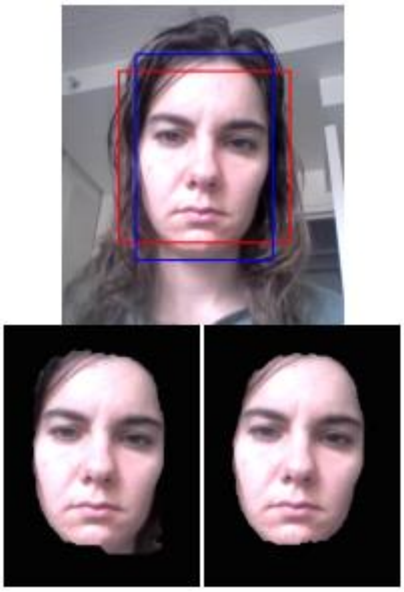


Figure 5. The original bounding box in red and adjusted bounding box that was input into the GrabCut segmentation algorithm in blue (top) as well as the first two iterations of the GrabCut implementation (bottom).

Figure 7 displays the three source signals identified by ICA for a 10-second window's worth of data, along with the normalised mean pixel intensities for the three colour channels.Figure 5 shows the initial bounding box in red, the modified bounding box that was entered into the GrabCut segmentation method in blue (top), and the first two implementations of the GrabCut algorithm (bottom).The power spectra for the three source signals within the physiological heart rate range are displayed in Figure 8. The heart rate value for this time step is correlated with the strongest frequency in this range.

3.2 HEART RATE ACCURACY

The estimated heart rate inaccuracy for movies of stationary faces was 3.4 +/- 0.6 bpm. The computed heart rate error for movies of participants tilting, rotating, and changing their faces was 2.0 1.6 bpm. Only inlier measurements—measurements within 10% of the reference heart rate—were used to determine these inaccuracies.

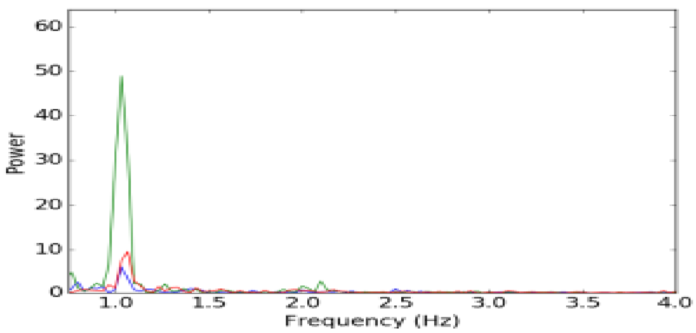


Figure 8. Example power spectra for three source signals over the physiological heart-rate range. The prominent frequency (0.97 Hz) corresponds to the heart-rate measurement of 58 bpm for this time-step.

Nonetheless, the quantity or percentage of outlier data does offer a decent indication of the algorithm's robustness as any outlier shows that the algorithm was unable to determine the heart rate inside that window. Thirty heart rate measurements were generated for every minute of video, with the ROI taking into account 30-second frames with a 1-second stride.

We can observe that increasing the forehead alone results in more outliers. This is probably due to the fact that if only the forehead area is used, there are less face pixels to offset the impact of any potential hair pixels included in the ROI.

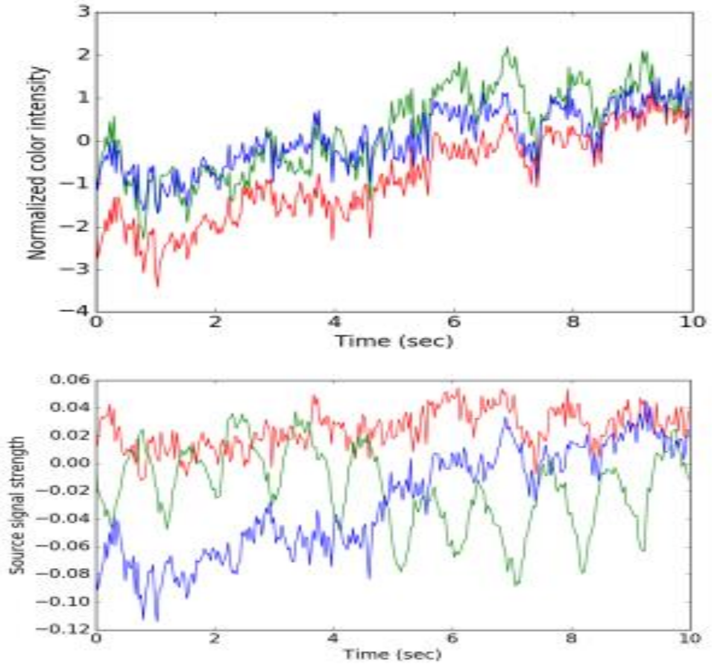
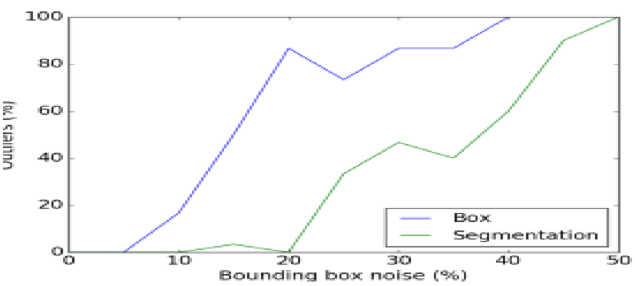


Figure 7. Example mean RGB color channel pixel values within the ROI (top) and associated source signals found through ICA (bottom). The source signal shown in green appears to oscillate at approximately 1 Hz and corresponds to the subject's pulse.

3.3 VALIDITY OF OUTSIDE NOISE

By intentionally adding random noise of different intensities to the face bounding box discovered in the algorithm's initial phase, we can additionally assess the algorithm's resistance to bounding box noise. As a function of the highest amplitude of injected noise, Figure 11 displays the percentage of measurements that were outliers.



gure 11. Percentage of measurements that are outliers as a function of the maximum bounding box noise introduced in each frame a percentage of the true bounding box size) for a video of a ll face. As expected, the number of outliers increases with the ount of noise, but a segmented ROI is more robust to handling s noise than a narrow box ROI.

4. CONCLUSION

We have shown that a person's heart rate can be determined by watching a standard colour video of their face. For videos of stationary faces and videos with movement, we found heart rate errors of 3.4 ± 0.6 bpm and 2.0 ± 1.6 bpm, respectively. With a low standard deviation and a constant difference between the computed heart rate and the reference rate throughout all movies, it is probable that the frame rate or sampling rate of the finger pulse sensor was calculated incorrectly as the cause of the base mistake.

If it is found that the offset seen in this study between the reference and computed heart rate is indeed constant throughout, the bias might be eliminated using a calibration step. A system like this could be able to detect heart rate in still images to within 0.6 bpm.

To guarantee that the reference heart rate is as precise as feasible, future research might make use of a medical-grade pulse oximeter. If it is found that the offset seen in this study between the reference and computed heart rate is indeed constant throughout, the bias might be eliminated using a calibration step.

5. Author Contributions

The journal to which the paper will be submitted has been decided upon by all authors, who have also given final consent to the published version and agreed to be responsible for all parts of the work.

5.1 Disclosure

In this paper, the authors disclose no conflicts of interest.

5.2 Data Availability

Data sharing is not relevant to this article because no new data were generated or examined in it.

8.REFERENCES

1. [1] J. Guttag, F. Durand, and G. Balakrishnan. identifying the pulse from video head movements. CVPR, 2013
2. R. Bhatt [2]. Integrated lab.
3. [3] M.-P. Jolly and Y. Boykov. For the best border and region segmentation of objects in n-dimensional pictures, use interactive graph cuts.2001 World Conference on Computer Vision.
4. [4] V. Kolmogorov and Y. Boykov. a comparison ofmin-cut/max-flow algorithms with the purpose of reducing visual energy. In 2004, IEEE Transactions on PAMI, 26(9):1124–7.G.

5. [5]. The library for OpenCV. 2000 edition of Dr. Dobb's Journal of Software Tools.
6. [6] V. Pisarevsky, A. Kuranov, and R. Leinhart. Empirical analysis of enhanced classifier detection cascades for quick object identification. MRL Technical Report, 2002.
7. [7]A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, and M. Passos. Python's Scikit-learn library for machine learning. 11:2825–2830, Journal of Machine Learning Research.
8. [8] D. McDuff, R. Picard, and M.-Z. Poh. automated, non-contact heart pulse measurements utilising blind source separation and video imaging. American Optical Society, 2010.
9. [9] "grabcut" - interactive foreground extraction via iterated graph cuts, C. Rother, V. Kolmogorov, and A. Blake. 2004 SIGGRAPH.
10. [10] L. Svaasand, J. S. Nelson, and W. Verkruysse. employing natural light for remote plethysmography. 2008's Opt Express.
11. P. Viola and M. Jones [11]. Quick object detection utilising a cascade of improved basic characteristics. Submitted to the 2001 Conference on Computer Vision and Pattern Recognition.
12. H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman are among the [12] authors. Eulerian video magnification to highlight minute alterations in the environment. 2012 SIGGRAPH.
13. [13] "Facial Recognition using OpenCV" by Shervin EMAMI and Valentin Petruţ SUCIU is available at https://www.researchgate.net/publication/26742 6877 Facial Recognition using OpenCV.
14. [14] H. Rahman, P. Funk, M.U. Ahmed, S. Begum Real-time heart rate monitoring using a webcam and facial RGB colour video.
15. [15] Face video based touch less blood pressure and heart rate estimation by Monika Jain, Sujay Deb, and A. V. Subramanyam (https://www.researchgate.net/publication/31248 7122 Face video based touchless blood pressure and heart rate estimation)
16. [16] "Emotion and heart rate detection using image processing," by Kaushal Kanakia, Saurabh Patil, Sujay Sabnis, and Vedant Shah (https://www.ijser.org/researchpaper/EmotionHeartbeat-Detection-using-Image-Processing.pdf).
17. You may get the source code for the algorithms covered in this paper at https://github.com/ibush/231A Project.
18. [18]Multiscale and Multisubgraph-Based Segmentation Algorithm for Ocean Remote Sensing Imagery, IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–20, 2023. Qianna Cui, Haiwei Pan, Kejia Zhang, Xiaokun Li, and Haoyu Sun
19. [19]. Interactive RGB Image Segmentation using Depth-modified Click Encoding and Estimated Depth, 2022 IEEE International Symposium on Multimedia (ISM), pages 10-17. Adrian Michl, Eckehard Steinbach, and Furkan Kaynar.
20. [20]. Statistical Learning of Multi-View Face Identification by Li, S.Z., Zhu, L., Zhang, Z., Blake, A., Zhang, H., and Shum, H. 7th European Conference on Computer Vision Proceedings, Copenhagen, Denmark (May 2002)
21. An Expanded Set of Haar-like Characteristics for Fast Item Detection. Lienhart, R., Maydt, J. 900–903 in IEEE ICIP 2002, September 2002, vol (2002)
22. Example-based object recognition in photos using components by Mohan, A., Papageorgiou, and T. 23(4), 349-361 IEEE Transactions on Pattern Analysis and Machine Intelligence (2001)
23. Analysis of fMRI data by blind separation into separate spatial components: M. J. McKeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, Hum. Brain Mapp. 6(3), 160–188 (1998). [CrossRef] [PubMed]
24. A Short Study to Assess the Potential of Independent Component Analysis for Motion Artifact Separation in Wearable Pulse Oximeter Signals," by Y. Jianchu and S. Warren, appeared in the Proceedings of the IEEE Conference of the Engineering in Medicine and Biology Society (IEEE, 2005), pp. 3585–3588.

25. 2Reducing motion artefact in photoplethysmography using independent component analysis," IEEE Trans. Biomed. Eng. 53(3), 566-568 (2006). [CrossRef]

26. Fast object recognition using a boosted cascade of basic characteristics was described by P. Viola and M. Jones in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (IEEE, 2001), p. 511.

27. An enlarged collection of Haar-like characteristics for quick object recognition by R. Lienhart and J. Maydt, Proceedings of IEEE Conference on Image Processing, 25 (IEEE, 2002).