**Literature Survey**

**Literature Survey -01**

**Title: automatic region-based heart rate measurement using remote photoplethysmography.**

This paper was published by IEEE Xplore. The authors are Benjamin, Kossack, eric, anna Hilsman,

Preterist. This paper presents a model-based approach to measuring the vital signs from RGB video files focusing on the heart rate. We use the plane-orthogonal-to-skin (POS) remote photoplethysmography (rPPG) transformation performed individually at five well-defined regions of interest (ROI) in the face. We extract the heart rate information by a correlation of the different rPPGNet signals in these ROIs and a magnitude-based reliability calculation. This increases the robustness of the heart rate extraction from videos. our model-based method is entirely automatic and does not require large amounts of data for training or time-consuming training sessions; our approach can be applied immediately. The ubfc-rppg dataset is used in this research model. our model-based method is entirely automatic and does not require large amounts of data for training or time-consuming training sessions; our approach can be applied immediately

**Literature Survey -02**

**Title: Remote Photoplethysmography signal measurement from facial videos using Spatio-temporal networks.**

This paper is published by the centre for machine vision and signal the analysis university of Oulu. The author is zitong Yu, Xiaobo. In this paper, they estimate heart rate using the techniques like network architecture,3dcnn based phys net, rnn based phys net. MAHNOB-HCL dataset is used. they implemented end to end framework with Spatio-temporal networks which can recover rppg signals from raw facial videos fast and efficiently. Phys net can recover rppg signals with accurate time location of each pulse weak, which allows measuring not only the average hrv level features that enable potential applications in remote a detection and emotion recognition. There was one drawback in implementing these techniques that facial expression analysis for multimode emotional recognition has not been found.

**Literature Survey -03**

**Title: Assessment of deep learning-based heart rate estimation using rppg under different illuminations**

This paper is published by IEEE Xplore. The authors are ze yanghoafei hang, Feng leu.

Remote photoplethysmography (rPPG) monitors heart rate without requiring physical contact, which allows for a wide variety of applications. Deep learning-based rPPG has demonstrated superior performance over the traditional approaches in a controlled context. However, the lighting situation in indoor spaces is typically complex, with uneven light distribution and frequent variations in illumination. It lacks a fair comparison of different methods under different illuminations using the same dataset. In this paper, they present a public dataset, namely the BH-rPPG dataset, which contains data from twelve subjects under three illuminations: low, medium, and high illumination. We also provide the ground truth heart rate measured by an oximeter. We evaluate the performance of three deep learning-based methods to that of four traditional methods using two public datasets: the UBFC-rPPG dataset and the BH-rPPG dataset. The experimental results demonstrate that traditional methods are generally more resistant to fluctuating illuminations. We found that the rPPGNet achieves the lowest MAE among deep learning-based methods under medium illumination, whereas the CHROM achieves 1.5 beats per minute (BPM), outperforming the rPPGNet by 60%. These findings suggest that while developing deep learning-based heart rate estimation algorithms, illumination variation should be taken into account. This work serves as a benchmark for rPPG performance evaluation, and it opens a pathway for future investigation into deep learning-based rPPG under illumination variations.

**Literature Survey -04**

**Title: Multihirerchial conventional network for efficient rppg signal and heart rate from video clips**

This paper is published by the school of information science and technology, northwest university. The authors have panned Zhang, bin li, jin ye Peng, Wei Jiang. UBFC-RPPG AND COFACE datasets are used. The proposed method aims to reconstruct rPPG from RGB facial videos by designing a 3D spatiotemporal convolutional network with multi-hierarchical fusion. the proposed network includes four modules: low-level face feature generation (LFFG), 3D Spatio-temporal stack convolution (STSC), MHFF, and signal predictor (SP). Dataset Our dataset contained a total of 300 VIS videos with a frame rate of 30 fps from 300 objects at the age of 18-26 years. The length of each video was 1 min with a pixel resolution of 1920×1080. These videos were collected by an Honor v30 mobile phone in a well-lit environment. Physiological signals were collected by BIOPAC MP160, including the average HR, respiratory rate, SpO2, ECG signal, and blood volume pulse (BVP) wave of each subject. The physiological signal sampling rate was 1000 Hz. The BF -RPPG database contained 42 videos from 42 subjects. The videos were recorded using a simple low-cost webcam (Logitech C920 HD Pro) at 30 fps with a resolution of 640×480 pixels in an uncompressed 8-bit RGB format. A CMS5OE transmissive pulse oximeter was used to obtain the ground truth PPG waveform and PPG HRs. During the recording, the subject sat in front of the camera with his/her face visible. All experiments were conducted indoors with varying amounts of sunlight and indoor illumination. The OHFA dataset [26] contained 160 videos with high compression rates from 40 subjects (12 women and 28 men); each of the subjects contributed four one-minute videos: two videos in 20 well-lit conditions, and the other two captured under natural light. The videos were recorded using a Logitech HD C525 with a resolution of 640×480 pixels and a frame rate of 20 fps. Each subject wore a contact PPG sensor to obtain the BVP data.

**Literature Survey -05**

**Title: A Deep learning approach for remote heart rate estimation**

This paper was published by agh university of science and technology. the author of this paper was Jaromir Przybyla. The mr-nirp dataset is used. Three data sets, characterized by great diversity, were used to evaluate the proposed algorithm. Two different cameras, eight participants with ages ranging from 22 to 70 years, seven locations with various lighting conditions, and various activities performed by the participants. Ethical review and approval are not applicable for this study, because the article presents a non-contact and non-invasive method of measuring pulse rate. This is only preliminary research and the results have not been used to assess human health. All devices used for collecting the ground truth pulse rate are battery powered and are commercially available products for personal use. Participants were not exposed to any stress - they performed only daily activities, as they did every day. Informed consent was received from all human subjects. The first set of data was recorded using the following configuration. RGB and infrared video sequences were captured using the Intel® RealSense™ camera (model D425). The video acquisition parameters were the following: a resolution of 640 × 480 pixels and a frame rate of 60 FPS. The camera was located 0.5 to 0.6 m from the volunteers. Video duration ranges approx. from 2 to 5 min. Details are provided. Three different locations with various illumination levels were selected. Additional signals were also recorded using a Simple Link sensor Tag CC2650. It is a low-energy Bluetooth device that includes 10 low-power MEMS sensors. The Sensor Tag was placed on the chest of the subject near the neck and face. To measure the ground truth HR and PPG signal, two devices connected via Bluetooth were used. The ECG-based H10 Heart Rate Sensor measured the reference HR. The optical heart rate sensor OH1 captured the PPG signal. Recorded data were used for both: training the LSTM network and testing. The ground truth heart rate (HR) varies from 48 bpm to 128 bpm.

**Literature Survey -06**

**Title: Fusion Method to Estimate Heart Rate from Facial Videos Based on RPPG**

Remote sensing of vital signs has been developed to improve the measurement environment by using a camera without a skin-contact sensor. The camera-based method is based on two concepts, namely color, and motion. The color-based method, remote photoplethysmography (RPPG), measures the color variation of the face generated by the reflectance of blood. In this article, the proposed method is the fusion method for estimation of heart rate using rppg.

First, the face was detected and tracked from the consequence frame of the facial video. Then, photoplethysmography signals were extracted from the face by RPPG. These signals were used to minimize noise and maximize cardiac components. Finally, the heart rate was estimated from the combined signal in the frequency domain. Thus, this study employed the single shot detector] with ResNet trained by the WIDER FACE dataset as mentioned in the article that they conducted several experiments.

This paper mainly focuses to overcome the noise of illumination variance and motion artifacts in RPP

Title: super-high resolution for video-based heart rate estimation with a semi-blind source separation method.

This article states that selecting an appropriate resolution based on a given shooting distance also plays a crucial role to improve the quality of rPPG measurements. Remote photoplethysmography (rPPG), a non-contact technique to estimate heart rates (HR) from video recordings, has attracted much attention from researchers in recent years. It is well-known that rPPG signals can be extracted from low-resolution videos.

**Literature Survey -07**

**Title: Rhythm Net: End-to-End Heart Rate Estimation from Face via Spatial-Temporal Representation**

This article was published by IEEE. The authors of this article are Xuesong Nau,Shiguang Shan, Hu Han. In this paper, they propose an end-to-end Rhythm Net for remote HR estimation from the face. In Rhythmed, they used spatial-temporal representation encoding the HR signals from multiple ROI volumes as to their input. Then the spatial-temporal representations are fed into a convolutional network for HR estimation. they also take into account the relationship of adjacent HR measurements from a video sequence via the Gated Recurrent Unit (GRU) and achieve efficient HR measurement. In addition, they build a large-scale multi-modal HR database (named VIPL-HR 1), which contains 2,378 visible light videos (VIS) and 752 near-infrared (NIR) videos of 107 subjects. The VIPL-HR database contains various variations such as head movements, illumination variations, and acquisition device changes, replicating a less-constrained scenario for HR estimation. The proposed approach outperforms the state-of-the-art methods on both the public-domain and VIPL-HR databases.

**Literature Survey -08**

**Title: Heart Rate Estimation from Facial Videos Using a Spatiotemporal Representation with Convolutional Neural Networks**

This paper is published by IEEE on 30 March 2020. Remote photoplethysmography (rPPG) is a kind of noncontact technique to measure heart rate (HR) from facial videos. As the demand for long-term health monitoring grows, rPPG attracts much attention from researchers. However, the performance of conventional rPPG methods is easily degenerated due to noise interference. Recently, some deep learning-based rPPG methods have been introduced and they revealed good performance against noise. In this article, we propose a new rPPG method with convolutional neural networks (CNNs) to build a mapping between a spatiotemporal HR feature image to its corresponding HR value. The feature map is constructed in a time-delayed way with noise-contaminated pulse signals extracted from existing rPPG methods. The CNN model is trained using transfer learning where images built from synthetic rPPG signals are taken to train the model first in order to generate initials for the practical one. The synthetic rPPG signals are interpolated from blood volume pulses or electrocardiograms through a modified Akima cubic Hermite interpolation. The proposed method is tested in both within-database and cross-database configurations on public databases. The results demonstrate that our method achieves overall the best performance compared to some other typical rPPG methods. The mean absolute error reaches 5.98 beats per minute and the mean error rate percentage is 7.97% in the cross-database testing on the MAHNOB-HCI data set. Besides, some key factors affect the performance of this method.

**Literature Survey -09**

**Title: Remote Heart Rate Measurement from Highly Compressed Facial Videos:**

**An End-to-end Deep Learning Solution with Video Enhancement.**

The authors of this paper are Zitong Yu1, Wei Peng1, Xiaobai Li1, Xiaopeng Hong, and GuoyingZhao. In this paper, the heart rate was estimated by highly compressed facial videos. They proposed an end-to-end deep learning-based propose a two-stage, end-to-end method using hidden rPPG information enhancement and attention networks, which is the first attempt to counter video compression loss and recover rPPG signals from highly compressed videos. The method includes two parts:

1) a Spatio-Temporal Video Enhancement Network (STEVEN) for video enhancement, and

2) an rPPG network (rPPGNet) for rPPG signal recovery. The rPPG Net can work on its own for robust rPPG measurement, and the STVEN network can be added and jointly trained to further boost the performance, especially on highly compressed videos. Comprehensive experiments are performed on two benchmark datasets to show that,

1) the proposed method not only achieves superior performance on compressed videos with high-quality videos pair,

2) it also generalizes well on novel data with only compressed videos available, which implies the promising potential for real-world applications. Two datasets are used – OBF and MAHNOB-HCI.

**Literature Survey -10**

**Title: A deep learning framework for heart rate estimation from facial videos**

This paper was published on 5 December 2020 by Elsevier. The authors are Gee-Sern JisonHsu and ArulMurugan Ambika path. In this paper, their approach is one of the pioneering works that propose a deep learning framework with TFRs as input for solving the heart rate estimation from facial videos. In addition, they have developed a heart rate database, named the Pulse from Face (PFF), and used it along with the existing PURE database to train the CNN. The PFF database is released for research purposes with this paper. They have evaluated the proposed framework on the MAHNOB-HCI database and the VIPL-HR database and compared its performance with that of other contemporary approaches to demonstrate its efficacy. develop a novel deep learning framework for real-time estimation of heart rates by using an RGB camera. This approach consists of the following four steps. We begin Step 1 by detecting the face and facial landmarks in the video to identify the required facial Region of Interest (ROIs). In Step 2, extract the sequence of the mean of the green-channeled video from the facial ROIs, and explore three-stage sequential filtering, including illumination rectification, trend removal, and signal amplification. In Step 3, the Short-Time Fourier Transform (STFT) is employed to convert the 1D filtered signal into the corresponding 2D Time-Frequency Representation (TFR) for characterizing the frequencies over short time intervals. The 2D TFR allows the formulation of the heart rate estimation as a video-based supervised learning problem, which can be solved by exploring a deep Convolutional Neural Network (CNN), as is carried out in Step 4. The approach is one of the pioneering works that propose a deep learning framework with TFRs as input for solving the heart rate estimation from facial videos.