A Review of Face Processing for Telehealth: Research Survey of Remote Visual Photoplethysmography (rvPPG)

Hilmi Demirhan

A Capstone Project Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Department of Computer Science

University of North Carolina Wilmington

2022

Approved by

Advisory Committee

Dr. Devon Simmonds

Dr. Ulku Clark

Dr. Karl Ricanek
Chair

Contents

ABSTRACT

This study will focus on an in-depth survey of commercial and open source resources for face processing for telehealth using Remote Visual Photoplethysmography (rvPPG). We can use Remote Visual Photoplethysmography (rvPPG) as a tool to monitor vital health signs and use its results to observe health conditions. There is a growing demand for non-invasive remote (tele-services) patient care information technology solutions. Advancement in the area of machine learning and deep learning especially as it relates to computer vision made it possible to predict various vital signs such as heart rate, blood pressure and oxygen saturation from pixel changes in recorded or live video feeds of human skin. Ability to extract vital signs via face processing could be a major contribution for telehealth solutions.

Recent studies have shown that rvPPG could be a sufficient substitute for traditional contact based methods for PPG, i.e. activity trackers, smart watches and pulse oximeter. In order to help the scientific community take advantage of the benefits of rvPPG methods, we would like to identify the rvPPG research and tools that currently exist in academia and the industry. There may be missed opportunities for tools developed for research purposes to be transferred to the industry. Our research will shed light on tools available therefore creating an opportunity for academia to discover what has been done in industry. The research will be unique in the sense that it will be an extensive survey of existing literature and tools.

In this thesis, we will perform a systematic literature review of the rvPPG research and tools that exists today, and identify the various directions that this research area can take. This work will provide detailed information about open source datasets and that can be used for developing a rvPPG framework. Furthermore, we will compare the selected commercial and open source rvPPG tools' specifications.

In this review, we have included thirteen novel methods that have significant contributions, twelve open source public datasets, fifteen open source tools and eleven commercial tools for literature review. We also discovered that there is an opportunity for the transferal of rvPPG methods from academia to industry to aid face processing for telehealth software development and maintenance.

ACKNOWLEDGMENTS

I would like to thank Dr. Karl Ricanek, my advisor, for introducing me to this area of study, for supporting and guiding me in my research. I would like to thank my committee members Dr.Ulku Clark and Dr.Devon Simmonds for their intellectual contributions to my capstone project.

List of Tables

List of Figures

LIST OF SYMBOLS

1 Introduction

There is a growing demand for non-invasive remote patient care technology to address the disparity gaps in healthcare. Recent advancements in computer science especially in computer vision field made it possible to predict vital signals from pixel changes in recorded or live video feeds of human skin. We can monitor various metrics such as heart rate, respiration, blood pressure, oxygen saturation and use the results as a measure to detect health conditions. There were known two famous heart rate measurement techniques which are electrocardiography (ECG) and photoplethysmography (PPG). The first actual electrocardiogram (ECG) [1] measurement of a human heart beat was recorded with a mercury capillary electrometer by British physiologist Augustus Desire Waller in May 1887 at St. Mary's Hospital, London. The PPG was described in late 1930's for the first time [2]. Both ECG and PPG techniques requires professional equipment that has to be attached to the skin, which might be inconvenient. However Remote Visual Photoplethysmography (rvPPG) allows contactless heart rate measurement [3]. Recent studies have shown that rvPPG could be a sufficient substitute for older methods of heart rate prediction.

Telehealth is used for delivering health and health related services via telecommunication and digital communication tools that allows remote connection between patient and the health care provider. Mobile apps, video conferencing, remote patient monitoring and remote robotic surgery are some of the examples of telehealth. Continuous innovation in technology and convenience of remote care will result in developing new telehealth tools and services. In the past some insurance providers were reluctant to cover the telehealth costs unless the patient is in a rural area. In most recent years private insurance and medicare coverage for the telehealth services have increased due to changes in health accessibility requirements. [4]. Telehealth can significantly contribute to increasing access to healthcare by decreasing time and costs. Employers are encouraging usage of telehealth services to their employees due to increase in employee productivity and lower medical expenditure[5].

1.1 Purpose

Many researchers have found that the ability to measure heart rate and various health metrics using rvPPG have great benefits. Especially in circumstances where visiting the health facility is not convenient, rvPPG can aid telehealth immensely. Having vital signs information with the help of rvPPG can be critical during emergency remote patient care especially when there is time and location limitation.

About 60 million people, nearly twenty percent of US population live in rural America [6]. The rural areas are sparsely populated and not close to urban centers such as hospital and clinics that have most of the necessary medical equipment for patient care. For that reason, many people who are living in rural areas have limited access to important health services and health professionals.

The socioeconomic, racial and gender disparities are important factors that affect access to quality healthcare services and life expectancy negatively in the United States. Even though many factors can potentially contribute to racial and socioeconomic disparities, some of them are more significant such as level of income, education and access to foods that support healthy eating habits [7]. Telehealth solutions can be used to decrease barriers and increase access to quality healthcare among diverse demographics.

This paper aims to survey a variety of known rvPPG tools that exist, either academically or commercially, and categorize and compare them. Another goal is to identify the features of these tools that are useful for developers and researchers in order to perform measurement mechanisms. The outcome of these findings should be able to provide a foundation for future development of rvPPG tools that researchers/developers can use to measure health metrics via face processing using telehealth.

There is a recognized need for distant contactless measurement of health metrics in order to ensure the social, economical and psychological impacts of remote medicine using telehealth. The goal for effective rvPPG usage is to give the developer and researcher the tools to monitor vital signs, and allow the developer, researcher, user to assess various health conditions. By cataloging and analyzing the capabilities that existing rvPPG tools offer, we can establish what exists today, and identify any potential opportunities for future rvPPG tools.

1.2 Scope

There are some pioneering work on using the rvPPG for heart rate prediction and other vital signs detection. This study focuses on an exhaustive literature survey on research and tools that use rvPPG. For a systematic literature review, it is important to have research questions identified to drive the methodology. The following research questions have been developed:

1.3 Research Questions

- What capabilities and features are offered by rvPPG tools and research methods? Identify possible areas of usage that could benefit from rvPPG?
- What are the available datasets and tools in the commercial and open source space? Classify the face based rvPPG datasets according to the number of subjects and application domain.

1.4 Organization

This thesis is organized into five sections: Introduction, Background, Related Work and Methods, Datasets, Tools, and Conclusion. In the Introduction, we will present an initial background and the purpose for carrying out this work. In the Background section, we will discuss the details of the techniques used in rvPPG. In the Related Work and Methods section we will give details about the novel contributions and developed methods in the field. In the Datasets section, we will present the available datasets details and provide an comparison of the datasets. In the Tools section we will list available tools both commercial tools and open source tools. Finally, in the Conclusion section, we will provide some final thoughts on the results of the study and identify opportunities for future work.

2 Background

The need for the telehealth solutions arise from the demand for remote patient care. First known usage of telehealth started with technological development of distance communication with the invention of telegraph and telephone in the 1800s [8]. During the civil war telegraph used for medical purposes. In a Lancet article from 1879 telephone usage suggested to reduce unnecessary doctor's office visits. NASA developed physiologic monitoring from distance during Mercury space program to understand if spaceflight has any effects on astronauts in late 1950s. Chronic diseases which require frequent doctor visits and close monitoring are main cause for high healthcare costs. There can be a significant drop in healthcare expenditure for individuals and governments if some of the office visits could be changed to telehealth visits [9].

The rvPPG can monitor the heart rate and other vital signs with via face processing without skin contact. This feature allows us to do monitoring patients closely at a home environment or at a clinical environment. There are commercially available applications are used in medical field such as telehealth applications. Since heart conditions can cause fatal results, fast diagnosis and treatment is very important. Remote heart rate monitoring can aid detection of a heart rate disorder easily without necessarily being in hospital settings. The telehealth will be more critical to provide long distance assessments by monitoring the vital signs. It is also used for measuring heart rate of multiple persons [10], which may be used to assess crowds of people. This feature can be used in hospitals and nursing homes.

There are important telehealth solutions such as Open Telehealth which is a remote patient monitoring healthcare software solution that is funded by the Danish government with the explicit purpose to demonstrate the efficacy of large scale Remote Healthcare Management. After successfully completing the initial software platform OpenTele in 2015, the Open Telehealth was founded as a commercial alternative to bring affordable large scale platform to international market. The OTH platform allows healthcare professionals, health clinics and hospitals to interact with patients in their homes. It has patient web portal and patient app that allows collecting important patient data through real-time video, messages and texts. The app simulates real doctor and patient consultation remotely [11].

Face processing is the area of study and development of systems and devices for face recognition, face modeling, facial analytics, facial attributes, facial affect and affective computing. Face processing for telehealth is used for improving delivering health and health related services via telecommunication and digital communication tools that allows remote connection between patient and the health care provider. Face processing technology using Remote Visual Photoplethysmography (rvPPG) can help analyzing vital and physiological signals from human skin mostly on facial regions.

There is a growing demand for non-invasive remote patient care technology to address the disparity gaps in healthcare. Almost half of the US population have hypertension defined as a systolic blood pressure greater than 130, a diastolic blood pressure greater than 80 or are taking medication for hypertension [12]. More than half of the people who die suddenly from heart disease show no symptoms [13]. Using the rvPPG methods and tools can be very helpful to develop various applications that can be used for prevention or diagnosis of heart related issues. Chronic diseases which require frequent doctor visits and close monitoring are main cause for high healthcare costs [14].

First known work related to the PPG method was developed in late 1930s [2] by Alrick B. Hertzman. The PPG method has established foundational bases for contactless version which is called the Remote Visual Photoplethysmography (rvPPG). There is strong interest in using remote heart rate monitoring for various purposes including medical vital sign monitoring. Remote heart rate monitoring is based on an optical technique called Photoplethysmography (PPG) that can detect changes in blood flow.

The rvPPG allows us measuring heart rate utilizing video by analyzing the color variations from the human skin surface, most widely used techniques analyzes the facial skin parts [15]. How the light reflects from the face can give us information about various metrics such as volume of blood, movement of the wall of blood vessels [16, 17]. The changes in blood volume creates a plethysmographic signal [18]. A plethysmographic signal shows the changes in blood volume or blood flow during a heartbeat. Plethysmograph is composed of two ancient Greek words 'plethysmos' which means increase and 'graph' [19]. With rvPPG it is possible to read a plethysmographic signal and use this signal to infer data such as heart rate variability, respiration rate, blood pressure and oxygenation [20, 21], quality of sleep, heart rhythm disturbances [22], and also mental stress [23] and drowsiness [24, 18].

2.1 Working Principles of Photoplethysmography (PPG)

The cardiovascular system plays a key role moving the blood, oxygen, nutrients between organs through arteries, veins and capillaries. The changes in blood volume creates a plethysmographic signal. A plethysmographic signal shows the changes in blood volume or blood flow during a heartbeat. The oxygenated blood is pumped from heart to other organs and deoxygenated blood returns back to the heart during continuous cardiovascular circulation. The oxygen enters to human body through respiratory system. The physical and physiological changes occurs in human body due to cardiovascular and respiratory activities.

It is possible to extract cardio-respiratory signal from skin color changes. There will be differences between oxygenated and deoxygenated blood cells because of the light spectrum absorbed by the hemoglobin. Oxygenated blood cells will have lighter color than deoxygenated blood cells. The photoplethysmographic (PPG) can be described as optical properties skin color changes due to cardiorespiratory activity [25, 26, 27].

As shown on the Figure 1, working principles of PPG, [28] on the left: lower pressure preceding the pulse wave means narrower arteries and less absorption (higher reflectivity) of the green light source and on the right: a higher blood pressure pulse causes wider arteries and more light absorption (lower reflectivity).

The cardiac cycle of human human heart activity represents from one heartbeat to

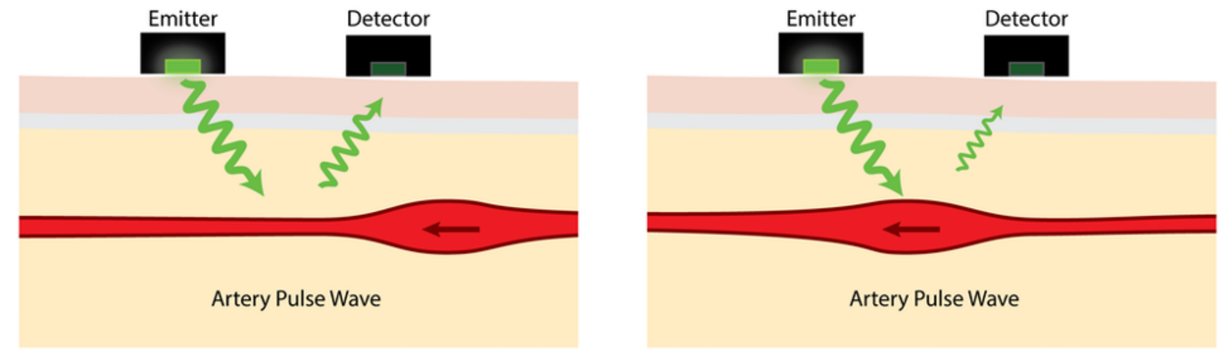

Figure 1: Working Principle of rvPPG

another. There are two activity types of the heart: one during which the heart muscle relaxes and refills with blood, called diastole, following a period of heart contraction and pumping of blood, called systole. The Photoplethysmogram (PPG) represents the variations in blood volume or blood flow in the body which goes from the heart to the fingertips and toes through the blood vessels within a cardiac cycle. As shown on Figure 2 the typical PPG signal [29], on the right shows the time interval between beginning and end of the PPG signal, known as the pulse interval. The distance between two consecutive systolic peaks is referred to as peak-to-peak interval. The time interval between the systolic and diastolic peak is related to the time taken for the pulse wave to propagate from the heart to the periphery and back. Heart rate is calculated by counting the number of systolic peaks per minute.

The PPG method is able to work with a light source and photo detector. The PPG method requires physical contact since detection sensor needs to be placed on the skin.

As shown on the Figure 3, the wearable PPG has two modes which are transmission and reflectance modes [28]. In transmission mode, the light transmitted through the medium is detected by a PD opposite the LED source, while in reflectance mode, the PD detects light that is back-scattered or reflected from tissue, bone and/or blood vessels. The basic form of PPG technology requires only a few opto-electronic components: a light source to illuminate the tissue (e.g. skin), and a photodetector to measure the small variations in light intensity

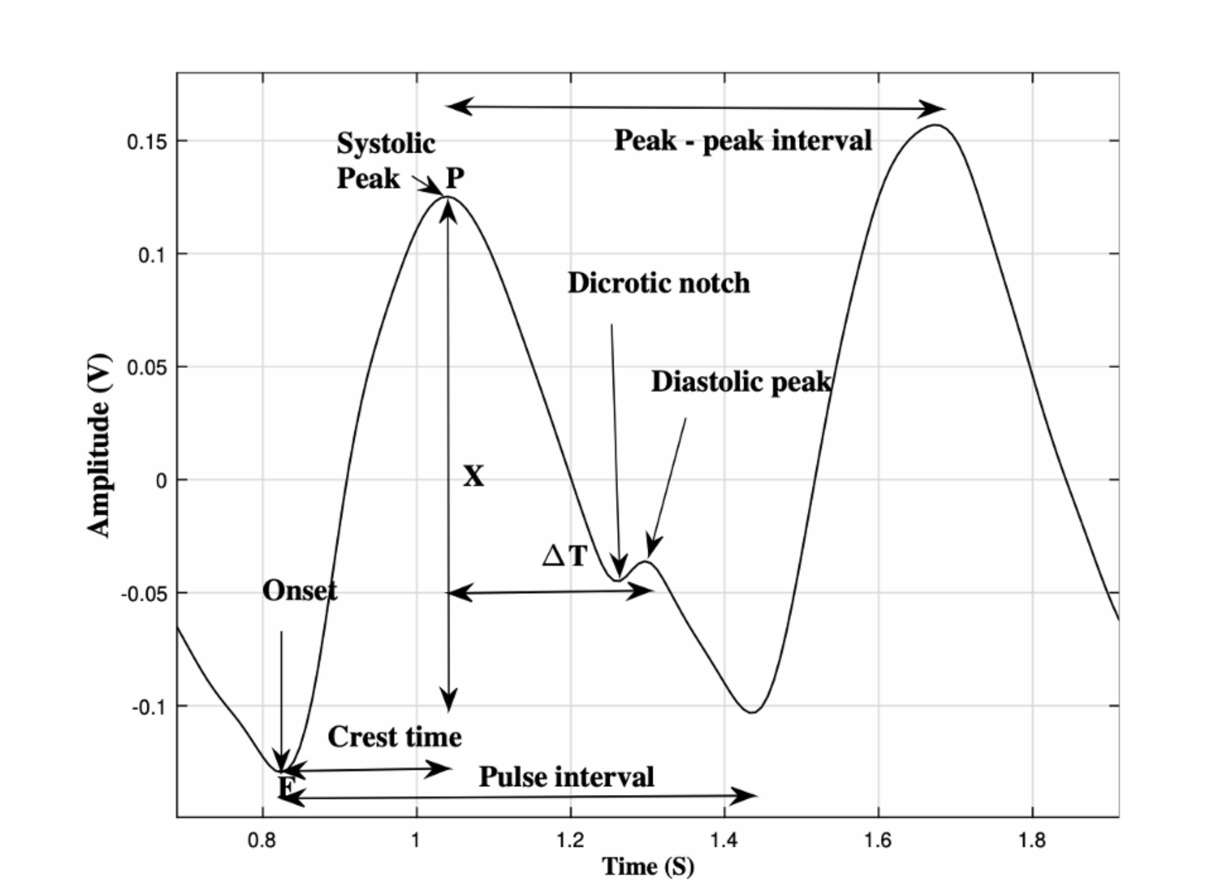

Figure 2: A typical PPG signal

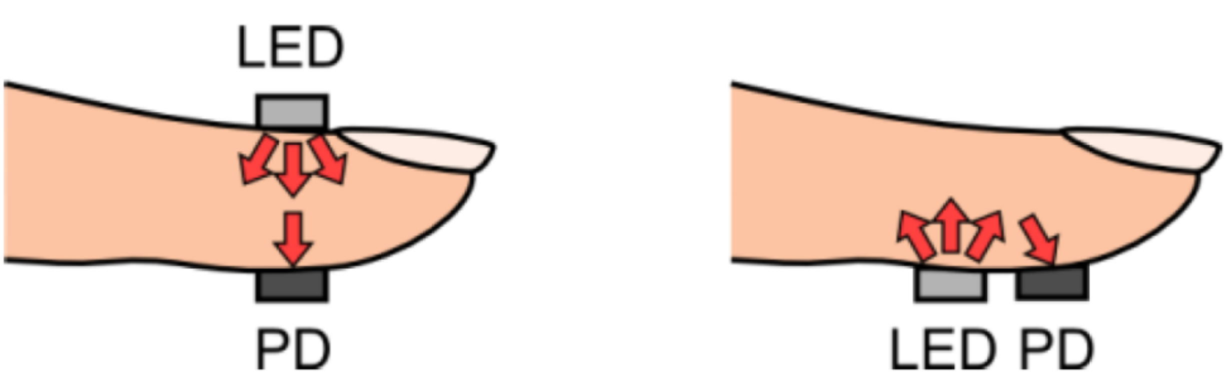

Figure 3: PPG Modes

associated with changes. The PPG based devices such as pulse oximeter, smart watches and activity trackers can be used for diagnosing heart conditions and irregular heartbeat. Even though PPG method was introduced in 1930s it started to become popular after usage of pulse oximiter in clinical settings in 1980s.

Figure 4: PPG Device Examples

As shown on the Figure 4, the example of a wearable PPG devices, such as a heartrate-sensing watch, light emits to the skin from one sensor and a second sensor detects how much light is returned to the device. Most of the light that penetrates the skin is absorbed by body tissues, but some is reflected. The amount of light that is reflected depends on several factors, one of which is the volume of arteries near the skin's surface. This forms the contrast between emitted light and reflected light. The amount of reflected light changes according to the blood volume.

2.2 Working Principles Remote Visual Photoplethysmography (rvPPG)

The PPG method has established foundational bases for contactless version which is called the Remote Visual Photoplethysmography (rvPPG). The rvPPG has proven to be highly accurate[30] and, in many cases, more accurate than trusted wearable devices prediction methods. The rvPPG allows us measuring heart rate utilizing video by analyzing the color variations from the human skin surface, most widely used techniques analyzes the facial skin parts [15]. The rvPPG methods chooses facial area as region of interest, around the chicks on facial area has most of the blood vessels. How the light reflects from the face can give us information about various metrics such as volume of blood, movement of the wall of blood vessels [16, 17]. The changes in blood volume creates a plethysmographic signal [18].

A plethysmographic signal shows the changes in blood volume or blood flow during a heartbeat. Plethysmograph is composed of two ancient Greek words 'plethysmos' which means increase and 'graph' [19]. With rvPPG it is possible to read a plethysmographic signal and use this signal to infer data such as heart rate variability, respiration rate, blood pressure and oxygenation [20, 21], quality of sleep, heart rhythm disturbances [22], and also mental stress [23] and drowsiness [24, 18].

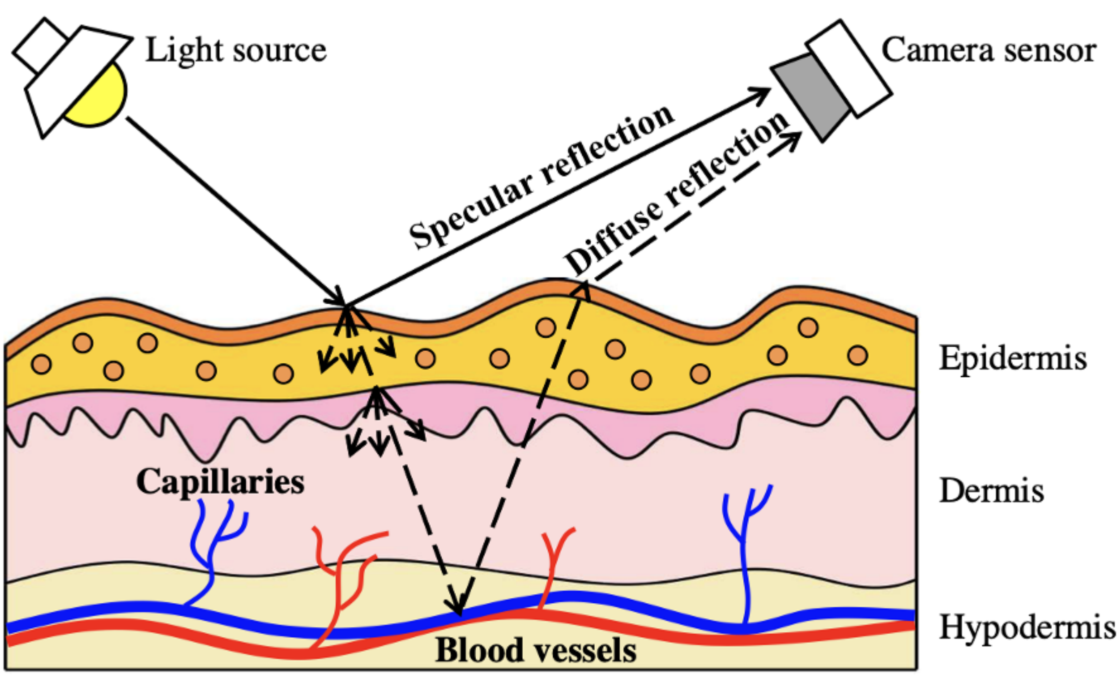

Figure 5: Skin Reflection Model of rvPPG

The rvPPG has the same principle with PPG but it is a contactless measurement from a distance, usually half to 1.5 meter. Even though we cannot see with our eyes our skin will have different reflection to light. Depending on the blood vessels flow we can measure changes in light reflection based on the absorption. It measures the variance of red, green, and blue light reflection changes from the skin, as the contrast between specular reflection and diffused reflection. As shown on the Figure 5 specular reflection is the pure light reflection from the skin. Diffused reflection is the reflection that remains from the absorption and scattering in skin tissue, which varies by blood volume changes.

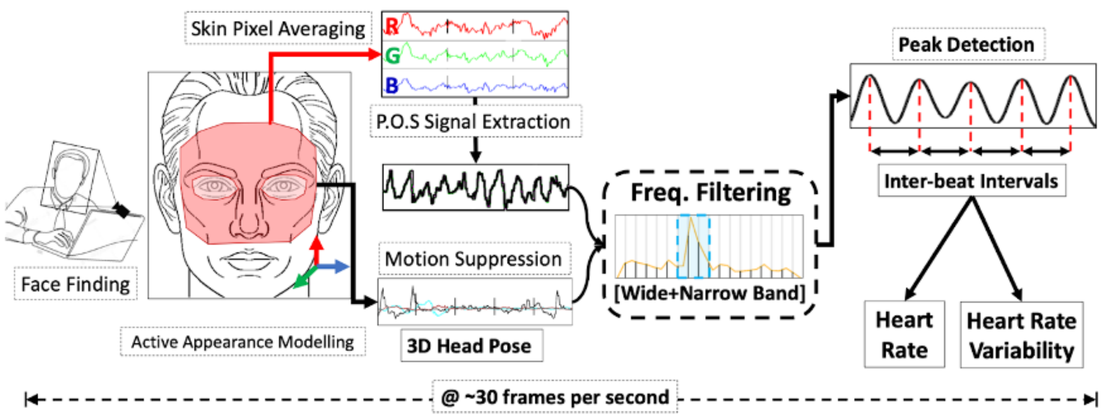

Figure 6: Heart Rate and Heart Rate Variability Estimation Pipeline rvPPG

As shown on the the heart rate and heart rate variability estimation pipeline Figure 6, the skin pixel selection part is for face capturing via webcam and it is detected and modeled in order to determine facial landmarks and head orientation. The selected area is approximately the top two-thirds of the face, where most of the blood vessels are concentrated, is selected as the region of interest. The signal extraction part is average of each pixel colors (red, green, blue) of the region is measured over time both specular and diffuse reflections. The Signal filtering is the noise from the head motions is detected by fitting the facial model and then noise-free heart rate is produced. The output calculations part is calculated by detecting peaks, inter-beat intervals are measured and then the heart rates and heart rate variability are estimated [31].

2.3 Areas of Usage of Remote Visual Photoplethysmography (rvPPG)

As stated above, rvPPG can be used to detect many vital health signals besides heart rate by the aid of a camera without need of any other equipment. It is an inexpensive compared to other methods and do not require a clinical setting. Heart rate is a predictor and biomarker of various situations such as emotional state that are linked to physiological health, psychological health and mental health [32]. Thus it can be used in many different areas. Some of these areas are listed below.

2.3.1 Heart Rate and Health Monitoring

The rPPG can monitor the heart rate by camera without skin contact. This feature allows us to do heart rate monitoring at a home environment or at a clinical environment. However these applications needs to get approval from FDA in order to be used in medical field such as telehealth applications. Since heart conditions can cause fatal results, fast diagnosis and treatment is very important.

Remote heart rate monitoring can aid detection of a heart rate disorder easily without necessarily being in hospital settings. In our recent circumstances with COVID19 telehealth will be more critical to provide long distance assessments by monitoring the vital signs. The rvPPG method can be used for measuring heart rate of multiple persons [10], which may be used to assess crowds of people. This feature can be used in hospitals and nursing homes.

2.3.2 Physical Exercise Monitoring

We can use rvPPG to monitor physical exercises. When exercising people want to track their pulse-rate for cardiac training, fat burning and health safe movements. ECG sensors in the chest or PPG sensors in the wrist bands could be used [33]. However wearing sensors can be quite uncomfortable during trainings. The rvPPG can be used as an alternative. Cameras can be attached to the fitness machines to monitor pulse-rate of the exercising person. Moreover with Machine Learning methods, using the data coming from the rvPPG method predictions can be made about the person such as when the subject will get tired based on the historical data collected by the rvPPG camera.

It is possible to give personalized suggestions or training advice to the person based on his/her past data. These systems could replace or assist personal health trainers which are costly and inaccessible for many people. When exercising people want to track their pulse-rate for cardiac training, fat burning and health safe movements. ECG sensors in the chest or PPG sensors in the wrist bands could be used [33]. However wearing sensors can be quite uncomfortable during trainings. The rvPPG can be used as an alternative. Cameras can be attached to the fitness machines to monitor pulse-rate of the exercising person.

The smart home gym and personal trainer devices like Tonal and The Mirror gives personalized programs and automatic weight suggestions using artificial intelligence. These smart home gym devices can benefit using rvPPG vital sign monitoring capabilities to improve their personalized suggestions. Moreover with Machine Learning methods, using the data coming from the rvPPG method predictions can be made about the person such as when the subject will get tired based on the historical data collected by the rvPPG camera. Also personalized suggestions or training advice can be given to the person based on his/her past data. These systems could replace or assist personal health trainers which are costly and inaccessible for many people.

2.3.3 Vehicle Operator Monitoring

Automotive, Aircraft and Spacecraft Cameras in vehicles have been used for various purposes such as eye movement detection[34], head movements[35] to prevent road accidents. Heart rate of the driver of a car, watercraft, an airplane, spacecraft, railed vehicle or any kind of vehicle can be detected using rPPG by using the cameras in these vehicles. Emotional states are closely connected to heart rate and emotions can effect driving [36]. For instance angry drivers are more likely to drive faster.

Accidents can be prevented by detecting the state of the driver and giving an early alarm. The rvPPG can help determining the wake status of the driver or changes in emotions and stress levels. The vital signs feed from face processing device inside vehicles be used to reduce road rage incidents if the increase in heart rate and stress could be captured, then possible intervention could include playing calming music or not allowing abrupt speed increases.

2.3.4 Deep Fake Video Detection and Prevention

The rvPPG can be used to detect Deepfake videos. DeepFake is a deep learning based method to create fake videos by changing the face of a person by the face of another person. Research has been done on using the rvPPG features in the heart rate estimation process for deepfake detection. Basically the heart rate information coming from videos can aid in deciding whether the visual is a deepfake or not. When the video is detected as a deepfake, its distribution can be prevented by online platforms such as YouTube, Facebook or TikTok.

Recent advances in computer vision and deep learning fields made it possible to create more realistic deepfake videos by replacing the faces of people in the videos with different faces. Deepfakes have been center of attention not only in academia or industry but in social media as well [37]. Many methods have been proposed to detect Deepfakes as no matter how real the Deepfake videos look, there are certain aspects of them such as eye blinking, natural micro-expressions that makes them detectable [38]. Research has shown that heart rate signal estimated from a real face is very different that heart rate signal measured from deepfake videos [39].

2.3.5 Public Facilities, Workplace Health and Safety Monitoring

The commercial tools like NervoHealth [40] and Binah.ai [41] can help workplace health and safety monitoring. The NervoHealth by Nervotech can be integrated to any business platform to check employees well being for example before starting a shift in construction site or factory where employees need to have overall good health conditions being able to do physically challenging activities. The BinahAi solution provides employee stress level measurement which is a important factor that needs to be assessed carefully especially for customer facing roles. Study shows that one of the main factor is stress for employee absences or decreases in productivity levels [42]. The Vastmindz's Employee Wellbeing [43] solution can be integrated to existing workplace wellness apps to track how employees respond to new and existing initiatives in real time, allowing them to tailor programs for individual needs.

Health and safety monitoring is important for the public facilities such as airport terminals, hospitals, care facilities, malls and office buildings. It is possible to integrate rvPPG capabilities to existing surveillance infrastructure or implement a new rvPPG based health and safety monitoring systems. For example in aviation industry individual's physiological signals monitoring can be done for travelers, employees, pilots and cabin crew to improve safety for everyone.

2.3.6 Customer and Retail Advertisement Monitoring

Many companies are already using cameras to track and analyze customer behaviors in retail stores. Emotional tracking systems can be integrated with rvPPG techniques to make better prediction using vital signals for better shopping experience at different location in a retail store. For example VitalSigns Camera Technology [44] provided self-tracking of the moods feature which can be used to analyze customer reactions to certain product or services while they are in physical retail stores or in online platforms. Companies can decide what products to bring market and they can personalized their products for various customer segments.

2.3.7 Insurance Policy Underwriting, Health Assessment

Insurance policy providers are invested healthiness of their customers. A good health assessment for life and health insurances can help determining premiums. Integrated rvPPG solutions to insurance policy apps can help both assessing and predicting customers health. The AI insurance underwriting tool can decide on fair insurance premiums and it can suggests incentives, discounts when customers improve their health and lifestyles such as exercising more and maintaining healthy weights. Having ability to check heart rate and blood pressure often can contribute to better health outcomes via preventive care suggestions such as changing diet and avoiding from unhealthy food options.

2.4 Challenges of Remote Visual Photoplethysmography (rvPPG)

Although rPPG is used in many different areas, it has some challenges. There exists a problem with the current rPPG methods; the accuracy of heart rate prediction drops sharply other than facial skin parts of the body. It is also less accurate different parts of the facial skin based on the angle and environment lighting variance. Different factors leads to inaccurate prediction. It makes optimally measuring heart rate from videos difficult. With this in mind, many research have utilized deep learning based on Convolutional Neural Network (CNN) to make heart rate predictions more accurate [15].

Most of the Remote Visual Photoplethysmography (rPPG) techniques aim to be designed for high performance smooth region of interest detection, prevention of background intervention while tracking and processing the facial vital signals. Visual object detection and tracking have always been a important challenge for researchers and organizations.

There are various challenges such as PASCAL Visual Object Classes (VOC) challenge which is considered a benchmark in visual object category recognition and detection, providing the vision and machine learning communities with a standard dataset of images and annotation, and standard evaluation procedures [45], the Visual Object Tracking (VOT) challenge which is a novel performance evaluation methodology for single-target trackers with fully annotated dataset [46], and ImageNet Large Scale Visual Recognition Challenge (ILSVRC) which is a large-scale hierarchical image database with 3.2 million diverse images [47].

The rvPPG methods try to increase consistency through the rvPPG pipeline especially consistency of the tracked area between old and new frames. Otherwise, the measurement results after spatial averaging will introduce noise. Because the detection and tracking algorithms are not designed to solve these problems it is important to consider the impact of visual object detection and tracking algorithms on the rvPPG measurement and the final calculation of the accuracy [48].

While most of the novel rvPPG methods achieves high accuracy on measuring rPPG signals in stationary case, they poorly perform when there is a motion or changes in lighting conditions. Improving the accuracy for real world scenarios such as driving a vehicle or during a physical exercise. It is not realistic to expect from the driver to stay steady while driving a vehicle or working out in a gym. Accurate reading of the heart rate is considered a important measure and is a frequently used indicator to assess the medical well being status or the physiological state of the person [49].

3 Related Work and Methods

3.1 Related Work

There has been some work done regarding heart rate prediction with rvPPG in the last decade[50, 51], however the foundations of rvPPG goes back to the PPG method which first introduced in late 1930s which requires attaching a device to capture blood flow via direct contact with the skin. In 1937 Alrick B. Hertzman published a first description of a photoelectric plethysmograph (PPG) on his article "Observations on the finger volume pulse recorded photoelectrically" and he became the first to introduce the term PPG [2].

In 2008 National Institute of Health (NIH) funded work and its publication "Remote plethysmographic imaging using ambient light" is the first known method of rvPPG [52]. This work was considered major milestone for capability of remote vital signal processing for both commercial and open source tools. During this work, simple, inexpensive less than 200 dollars digital cameras were used. With the camera on a tripod, movies were recorded of the facial area. Duration varied from 30 s to several minutes. This method was called Green Channel method because green channel featuring the strongest plethysmographic signal during the experiments.

After proving the possibility that PPG signals can be obtained remotely without need of physical attachments, various novel study and research contributions have been increased with new improvements for rvPPG in both by academic researchers and commercial organizations. The Independent Component Analysis (ICA) Method is a statistical and computational technique [15], and it is low-cost method for measuring multiple physiological parameters using a basic webcam. The Principal Component Analysis (PCA) Method is a simple and robust method of measuring the pulse rate.

Developments in artificial intelligence field increased popularity of using new machine learning and deep learning methods for rvPPG heart rate estimation in most recent years. One of the most frequently used deep learning technique is Convolutional Neural Networks[53, 54, 55, 56]. Zhan et. al. investigate efficiency of CNN-based rvPPG methods [30]. They concluded that black box nature of CNN methods can be a challange in tuning the parameters. They also found that the choice and parameters (phase, spectral content, etc.) can be very critical in developing a good model.

In another deep learning based work researchers propose a new framework, which uses deep spatiotemporal networks with data augmentation for contactless heart rate variability (HRV) measurements from raw facial videos. The databases they have used include face videos with synchronized physiological signals[57].

Kopeliovich et. al. looked into possible architectural modification to enhance CNN methods [58]. One advancement they suggested is adding convolutional-based filter for postprocessing of network outputs for better accuracy of heart rate prediction. They also proposed combined loss function where the first component is a cross entropy and the second one is a mean squared error between the network output and smoothed one-hot vector.

Chen and McDuff developed a convolutional attention network extracting the PPG signal from a video using two parallel models [59]. They propose two models. The first model is a classical "appearance model" which learns to find the skin region-of-interest (RoI) using a finger oximeter-derived signal as a reference. Their second model is a motion model which uses skin motion rather than blood absorption.

Authors propose a general-to-specific transfer learning method called SynRhythm [56]. They convert the spatial-temporal features into heart rate based on the pre-trained network[60]. In PhysNet [53] authors developed a 3 dimensional CNN. They look into time based and space based features of face images to extract the rvPPG signal. They measure the heart rate variability from the rvPPG signal.

One of the most studied area is effects of motion in rvPPG measurement. Some of the early papers focused on

3.2 Methods

In this section we have identified and compared the important methods that have significantly contributed rvPPG.

3.2.1 Plethysmography (PPG) Method

Photoelectric Plethysmography Method [2] was the initial method which require an attaching a device physically to the patient. Detection of changes in the blood volume using dedicated light source and photo detector. Photoelectric plethysmographs for the fingers and toes are described which use electrocardiographs for the recording and which have definite advantages in routine clinical observations on the circulation.

The validity of the technique is established by comparison of the photoelectric records with simultaneous records obtained with transmission plethysmographs, and by comparison of the photoelectric records in instances of circulatory disturbances with independent directional confirmation by other methods in the literature

Foundational related work regarding the PPG methods goes back to 1930s when it was originally introduced for the first time. As shown on the Figure 7 this PPG technique requires attaching a device to capture blood flow via direct contact with the skin. In 1937 Alrick B. Hertzman from the Department of Physiology at St. Louis University, published a first description of a photoelectric plethysmograph on his article ”Observations on the finger volume pulse recorded photoelectrically” and he became the first to introduce the term PPG [61].

3.2.2 Green Channel Method

The Green Channael method is measuring PPG signals from human face remotely using only simple digital camera and ambient light using green channel. It is called green channel because green channel has the strongest signal, because hemoglobin absorbs green light

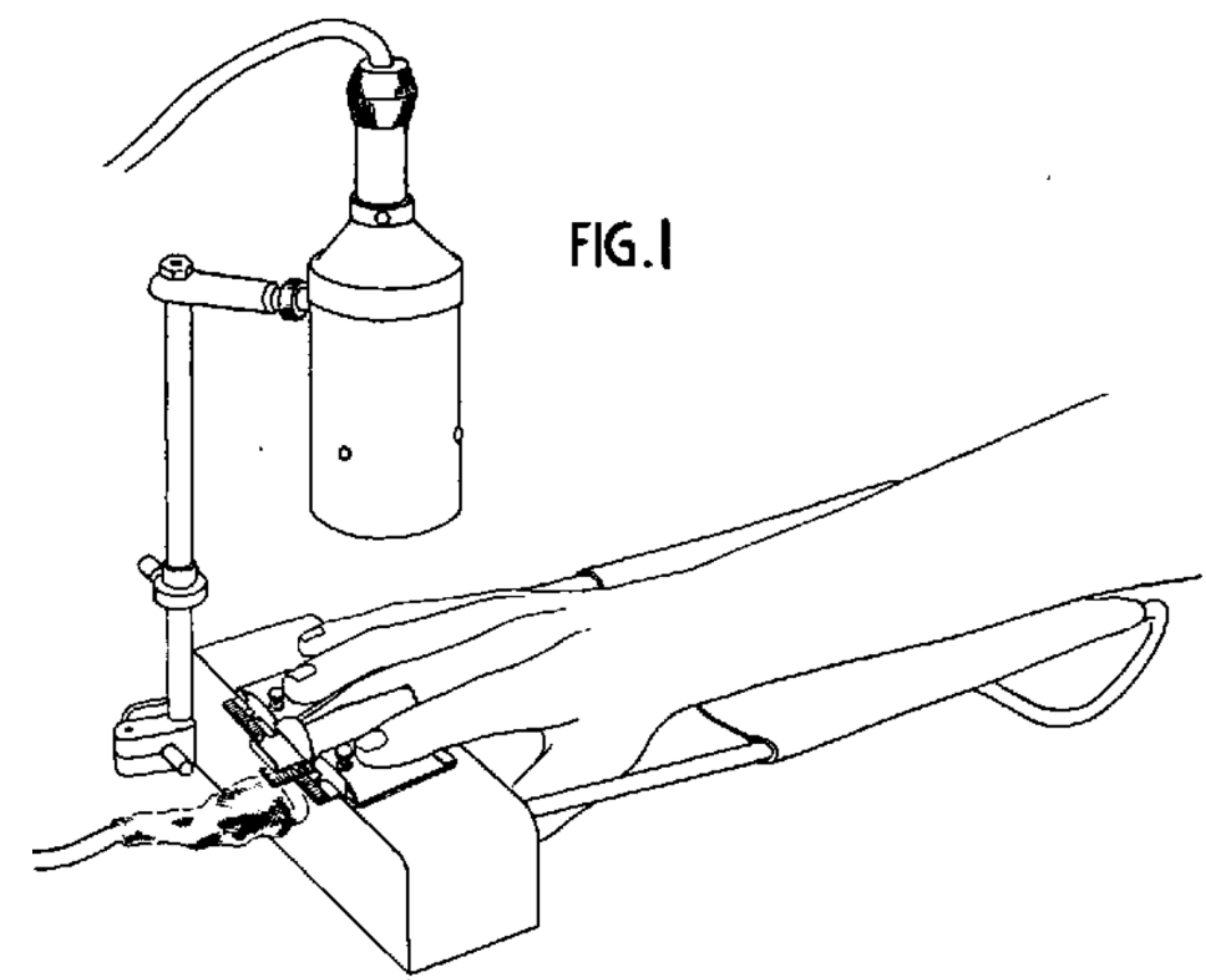

Figure 7: Finger Photoelectric Plethysmograph

better than it does red and blue. It consists of 30 seconds to several minutes videos of the volunteers.

It is a simple but effective approach in estimating pulse rate via rvPPG. It consist of three steps. First step is identifying suitable ROIs within the subject's face, second step is calculating the average color intensity for the green channel and, the third step by spatial averaging over the ROI, extracting the spectral content to look for highest frequency component [52].

As shown on the Figure 8 In 2008 National Institutes of Health (NIH) funded work show that signal can be remotely (about several m) measured on the human face with normal

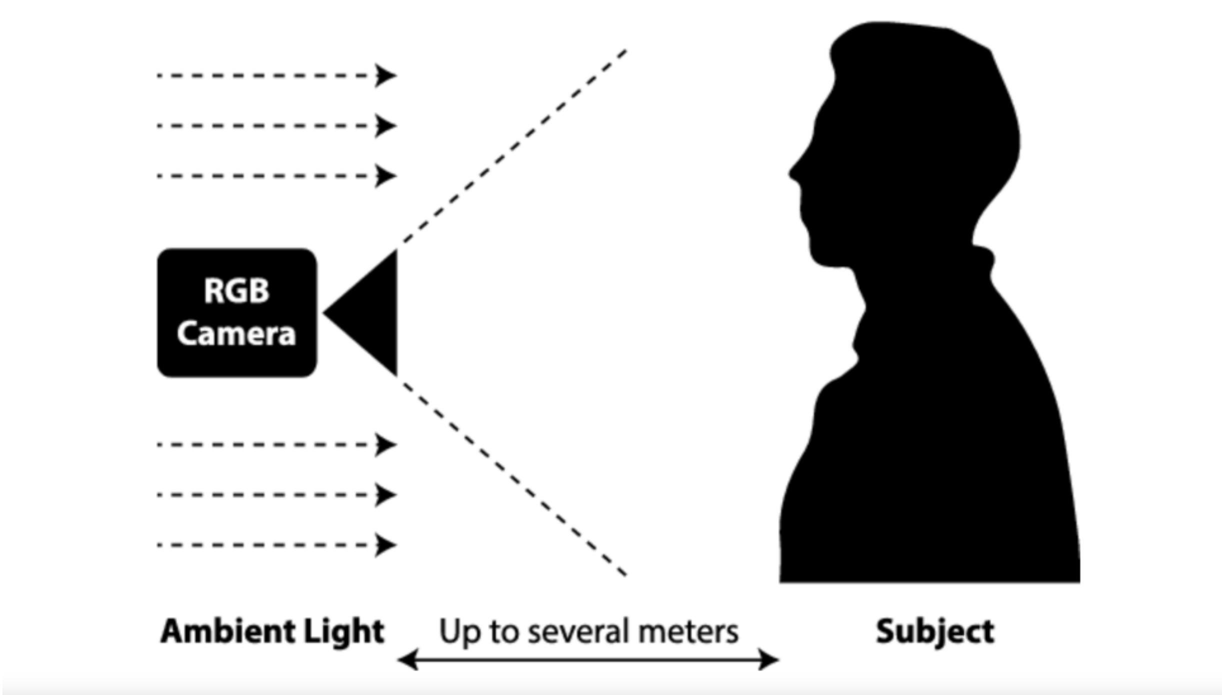

Figure 8: Remote Plethysmographic Imaging Using Ambient Light

ambient light as the source and a simple digital, consumer level photo camera in movie mode.

3.2.3 Independent Component Analysis (ICA) Method

Independent component analysis (ICA) [15] is a technique for uncovering independent signals from a set of observations that are composed of linear mixtures of the underlying sources. It measures multiple physiological parameters using a basic webcam by applying independent component analysis on the color channels. It has 12 participants of both genders four females and 8 males, different ages between 18– 31 years old and with different skin colors.

The underlying source signal of interest in this study is the BVP that propagates throughout the body. During the cardiac cycle, volumetric changes in the facial blood vessels modify the path length of the incident ambient light such that the subsequent changes in amount of reflected light indicate the timing of cardiovascular events. By recording a video of the facial region with a webcam, the red, green, and blue (RGB) color sensors pick up a mixture of the reflected plethysmographic signal along with other sources of fluctuations in light due to artifacts.

3.2.4 Principal Component Analysis (PCA) Method

The Principal Component Analysis, (PCA) [62] is sometimes called the Karhunen-Loeve Transformation, which is a technique commonly used for data reduction in statistical pattern recognition and signal processing. The PCA is a transformation that identifies patterns in data, and expresses the data in such way that it highlights the similarities and differences. It measures pulse rate with a webcam non-contact method for evaluating cardiac activity. It consist of video recordings of 10 volunteers who are 2 females and 8 males, of different ages between 20 and 64 years old.

A simple processing of image data and then applying PCA allows extracting the changeable component containing information of the heart rate. The results obtained from independent and principal component analysis' show that those two methods extract the "pulse" component with similar accuracy. However, comparison of time of calculation and other studies allows for the conclusion that PCA is less computationally complex so it is reasonable to choose that method to reduce time and complexity of analysis.

3.2.5 Chrominance (CHROM) Based Method

The Chrominance (Chrom) [63] based method which is robust pulse rate from Chrominance-Based rPPG. We present an analysis of the motion problem, from which far superior chrominance-based methods emerge. It consists of subjects and 1 minute video recording of each subject. The method tested in a fitness setting using a simple spectral peak detector, the obtained pulse-rate for modest motion (bike) improves from 79 percent to 98 percent correct, and for vigorous motion (stepping) from less than 11 percent to more than 48 percent correct.

It analyzes how motion enters the pulse signal. As a result of this analysis, it provides the assessment details, the results of which are shown for 117 stationary subjects over a broad range of skin-types, and for some subjects exercising in a gym to test motion robustness. It assesses the motion robustness of various methods by comparing pulse rates obtained from exercising subjects in a fitness.

3.2.6 Pulse Blood Volume (PBV) Method

The Pulse Blood Volume (PBV) method explores the fact that PPG signals have known ratios in the red, green and blue channels. When these are the coefficients of a pulse-blood-volume (PBV) "signature" vector, the least-mean-squared criteria can be suitably applied to separate distortions from actual signals. The method show that the different absorption spectra of arterial blood and bloodless skin cause the variations to occur along a very specific vector in a normalized RGB-space. The exact vector can be determined for a given light spectrum and for given transfer characteristics of the optical filters in the camera.

This method shows that this 'signature' can be used to design an rvPPG algorithm with a much better motion robustness than the recent methods based on blind source separation, and even better than the chrominance-based methods which is previoulsy published. The PBV method uses six video recording in a gym, with four subjects exercising on a range of fitness devices, and the new improved rvPPG method provides superior motion robustness [64].

3.2.7 Plane Orthogonal to the Skin (POS) Method

The Plane Orthogonal to the Skin (POS) Method [65] investigate the algorithmic principles of rvPPG in a mathematical context with optical and physiological reasoning. Its exploration based on the skin reflection model shows that different characteristic properties of rvPPG can be used to design algorithmic solutions for pulse extraction. This method not only gives an integral view on and insight into the core rvPPG methods, but also leads to a new alternative that demonstrates tractable algorithm development. The new method defines a plane orthogonal to the skin tone in the temporally normalized RGB space for pulse extraction and is, therefore, referred to as the "plane-orthogonal-to-skin" (POS).

This method proposed a mathematical model for rvPPG measurement which is based on the optical and physiological considerations and assumption of a single light source with a constant spectrum. This model can be used to understand the commonalities and differences between existing rvPPG methods in pulse extraction.

3.2.8 Local Group Invariance (LGI) Method

The Local Group Invariance (LGI) Method is examined the invariance for the task of heart rate estimation from face videos of 25 participants in presence of disturbing factors like rigid head motion, talking, facial expressions and natural illumination conditions under different scenarios. This method shows functional approach for the task of heart rate estimation from face videos under the load of nuisance factors. It performed an evaluation on the data collected under everyday natural facial motions and environmental conditions [66].

3.2.9 DeepPhys Method

The DeepPhys method introduces video-based physiological end-to-end system for heart and breathing rate using a deep convolutional network. The DeepPhys method features a new motion representation based on a skin reflection model and a new attention mechanism using appearance information to guide motion estimation, both of which enable robust measurement under heterogeneous lighting and various motions. This method significantly outperforms all previous state-of-the-art methods on both RGB and infrared video datasets.

The performance improvements were especially good for the tasks with increasing range and angular velocities of head rotation. This improvement achieved by the end-to-end nature of the model which is able to learn an improved mapping between the video color and motion information. The method tested on four different datasets, each featuring participants of both genders, different ages, a wide range of skin tones (Asians, Africans and Caucasians) and some had thick facial hair and/or glasses to improve accuracy on diverse users [67].

3.2.10 Bounded Kalman Filter (BKF) Method

The Bounded Kalman Filter (BKF) method [68] is a real-time measurement of heart rate across different lighting and motion conditions. The method's algorithm aims to minimize motion artifacts such as blurring and noise due to head movements by employing a blur identification and denoising algorithm for each frame and a bounded Kalman filter technique for motion estimation and feature tracking. They have created a data set containing 200 video sequences (each video lasting 60 seconds) using the front HD camera which is a CMOS type sensor of a Surface 4 tablet. All videos are recorded in a 24-bit RGB color format having a resolution of 1920 x 1080, recorded at 29.97 frames per second and it was stored as uncompressed data.

The proposed method has a consistent performance with the 95 percent agreements ranging from 52-81 percent, surpassing the benchmarked methods. The proposed method was applied in an environment featuring subjects walking towards the camera from a distance of 4ft., attaining a mean error of $\pm$ 3 B.P.M., thus opening up a new paradigm in rvPPG research domain. This advancement will improve the detection of heart rate in real world scenarios such as while a person performs an exercise in the gym or perform a task in a workplace.

3.2.11 PhysNet Method

The PhysNet method is end-to-end spatio-temporal network for rPPG signal measurement from raw facial videos. It takes temporal context into account which was ignored in previous works. It multiples commonly used spatio-temporal modeling methods are explored and compared, which can serve as a foundation for future network optimization specially for rPPG measurement task. It compared with state-of-the-art methods, the proposed PhysNet method achieved good performance for measuring not only the average heart rate, but also the heart rate variability features, which were further demonstrated to be effective for atrial fibrillation issue detection and emotion recognition. Comprehensive experiments are conducted on two benchmark datasets which are OBF and MANHOB-HCI.

The OBF dataset has large number of facial videos and it is used for both training and testing the results. The MANHOB-HCI is used for cross testing the generalization of the model. Three previous methods Green, CHROM, POS were replicated and the PhysNet's accuracy outperformed all of them with results of 5.96 HRmae, 7.88 HRrmse and 0.76 HRr [69].

3.2.12 Signal to Noise Ratio Bandpass Filtering (CWT-SNR) Mehtod

The Signal to Noise Ratio Bandpass Filtering" (CWT-SNR) mehtod uses the SNR, instead of the more common fixed frequency range provided by the Band-Pass Filtering (BPF). The proposed CWT-SNR method can be used as a post-processing step in general remote visual plethysmography (rvPPG) algorithms to improve their robustness. The UBFC-rPPG dataset is used for performance comparision and validation of the developed algorithm.

The ground truth results were computed from the ground truth pulse signal data included in the UBFC-rPPG dataset. Face video of the participant were recorded using a webcam at 30fps and a resolution of 640x480. The experiment result shows a regression result of 0.88 correlation in the results from the face video compared to the ground truth data in SDRR and 0.63 in RMSSD [70].

3.2.13 Singular Spectrum Analysis (SSA) Method

The Singular Spectrum Analysis (SSA) method investigate the effect of ROI selection for rvPPG using the singular spectrum analysis. It demonstrates the effectiveness of the proposed method using 154 videos from the public COHFACE dataset excluding 6 videos that were too challenging to process due to the effects of the skin color of the subject and camera shake. The SSA method defines patches from a full ROI covering skins, and the patches are ranked and selected through quality metrics. The impact of ROI on the performance of SSA method is then evaluated through using different numbers of optimal patch ROIs.

It is experimentally given that the selection of the optimal patch ROIs can most effectively eliminate illumination noise and achieve reliable HR measurements. The performance of SSA-7 with the HRrmse of 4.76, the HRmae of 2.05, the HRsd of 4.29 and the r of 0.91 is the most outstanding among all methods. It has less quantization noise compared to less ROIs selection, and less facial illumination noise compared to larger ROIs selection [71].

Lead Author	Methods	Description	Dataset	Accuracy
Hertzman et al. 1937	Photo-electric Plethys-mography (PPG)	Detection of changes in the blood volume using dedicated light source and photo detector. Foundation of rvPPG.	Measured blood volume changes in the fingers induced by the Valsalva manoeuvre, exercise and with exposure to cold.	Potential sources of error with the technique have been identified, good contact with skin was needed.
Verkruysse et al. 2008	Green Verkruysse (Green Channel)	Measuring PPG signals from human face remotely using simple digital camera and ambient light using green channel.	Recorded 30 second to several minutes videos of volunteers. Some recording was from port wine stain (PWS) patient.	Strongest plethysmographic signal on Green Channel. Power spectrum for the Green channel amplitude modulation of RR (0.12 Hz) and HR (1.12 Hz). Impovement on SNR.
Poh et al. 2010	Independent Component Analysis (ICA)	Measuring multiple physiological parameters using a basic webcam by applying independent component analysis on the color channels in video recordings.	12 participants of both genders (four females), different ages (18– 31 years) and skin color. Collected their own data.	The root mean square error was reduced nearly threefold from 6.00 (obtained from raw green channel trace before ICA) to 2.29 bpm and the correlation coefficient r increased from 0.89 to 0.98 ($p < 0.001$ for both).
Lewandowska et al. 2011	Principal Component Analysis (PCA)	Measuring pulse rate with a webcam non contact method for evaluating cardiac activity.	Recorded a video of 10 white volunteers, 2 women and 8 men, of different age (20 - 64 years) they were sitting still and distant 1 m in front of the camera.	Signal to noise ratio and comparison of mean heart rate calculated by PCA and ECG. Exact SNR s and differences in mean heart rates are not given.
De Han et al. 2013	Chrom	Robust pulse rate via chrominance based method with a dataset collected with a camera.	117 subjects, 1 minute video recordings.	Method performs 92% good agreement (±1.96) with contact PPG, with RMSE and standard deviation both a factor of 2 better than BSS-based methods.
De Han et al. 2014	PBV	Robust pulse-rate from chrominance based rPPG, improved motion robustness of remote-PPG by using the blood volume pulse signature.	Six videos recorded in a gym, with four subjects exercising on a range of fitness devices four different subjects exercising on a bike, a stepping device, a handbike, a treadmill and a synchro-device.	The signal-to-noise-ratio for long (512) overlap window: 7.6 and short (64) overlap-add window: 5.6 compared to CHROM method SNR of the pulse signal improves from -5 dB to -4 dB.
Wang et al. 2015	SSR	A novel algorithm for remote photoplethysmography, spatial subspace rotation that improves the robustness of remote photoplethysmography	A benchmark dataset containing 54 video sequences (with 108000 frames) has been built to evaluate the proposed rPPG algorithm.	Signal-to-noise ratio. 2SR improves on average the SNR of ICA by 2.22 dB, CHROM by 1.56 dB, and PBV by 1.95 dB
Wang et al. 2016	POS	Algorithmic principles of remote PPG, mathematical model that incorporates the pertinent optical, physiological properties of skin reflections.	Dataset containing 60 video sequences (with 147100 frames) the subject is illuminated by a frontal fluorescent lamp, and sits in front of the camera with the face visible.	The SNR of the pulse frequency derived by the ratio between energy around the first two harmonics and remaining parts in the frequency spectrum.

Table 1: Related Work and Methods with Significant Contributions

Lead Author	Methods	Description	Dataset	Accuracy
Pilz et al. 2018	LGI	Local group invariance for heart rate estimation from face videos. The energy of the blood volume signal is re-arranged in vector space.	Face video recordings of 25 subjects during head resting conditions, during head rotations, during an exercise on a bicycle ergometer in a gym and during an urban conversation.	POS archives a correlation of 0.35 with a RMSE of 21 BPM and LGI a correlation of 0.87 with a RMSE of 11 BPM.
Chen et al. 2018	DeepPhys	Video-based end-to-end system for heart and breathing rate using a deep convolutional network.	Tested on four datasets, participants of both genders, different ages, a wide range of skin tones (Asians, Africans and Caucasians).	Introduced an attention mechanism to improve signal extraction accuracy, 6.89 HRmae, 13.89 HRrmse and 0.34 HRr.
Prakash et al. 2018	BKF	Real-time measurement of heart rate across different lighting and motion conditions, aims to minimize motion artifacts.	Data set containing 200 video sequences (each video lasting 60 seconds) using the front HD camera.	The proposed method has a consistent performance with the 95% agreements ranging from 52%-81%, surpassing the benchmarked methods.
Yu et al. 2019	PhysNet	End-to-end spatio-temporal network for rvPPG signal measurement from raw facial videos.	Comprehensive experiments are conducted on two benchmark datasets, OBF and MANHOB-HCI.	Accuracy outperformed all Green, CHROM, POS methods with 5.96 HRmae, 7.88 HRrmse and 0.76 HRr.
He et al. 2021	CWT-SNR	Uses signal to noise ratio to assist in the selection adaptive range of the frequency that may contain HRV information.	UBFC-rPPG dataset is used for validation. Subject faces were recorded using a web-cam at 30fps and a resolution of 640x480.	Experiment result shows a regression result of 0.88 correlation in the results from the face video compared to the ground truth data in SDRR and 0.63 in RMSSD.
Wang et al. 2021	SSA	Investigate the effect of ROI selection for rPPG using the singular spectrum analysis.	Demonstrate the effectiveness of the proposed method on 154 videos from the public COHFACE dataset.	SSA with different optimal patch ROIs selected on dataset COHFACE. SSA-7 with the HRrmse of 4.76, the HRmae of 2.05, the HRsd of 4.29.

Table 2: Related Work and Methods with Significant Contributions Continued

4 Datasets

In this section we will look at the available datasets for the Remote Visual Photoplethysmography (rvPPG). There are various datasets that have been recorded with both with a remote camera and a attached device for ground truth data. The datasets have diverse application domains such as heart rate prediction, emotion, stress, psycho-physiological analysis and fatigue monitoring. While some datasets have limited participants some of them have large number of participants from various ethnic and racial backgrounds such as MMSE-HR dataset. Since head motion is a big challange for getting accurate predictions there are datasets with presence of motion under different scenarios such as LGI-PPGI dataset.

4.1 PURE Dataset

This PURE Pulse Rate Detection dataset [72] consists of 10 people performing different, controlled head motions in front of an Eco274CVGE camera by SVS-Vistek camera at a frame rate of 30 Hz with a cropped resolution of 640x480 pixels and a 4.8mm lens. During the experiment, the image sequences of the head, as well as reference pulse measurements were recorded. The recording was done in six different setups that include steady, talking, slow translation, fast translation, small rotation, and medium rotation resulting in a total number of 60 sequences of 1 minute each.

The participants were placed in front of the camera at an average distance of 1.1 meters with available daylight through the large window and ground truth data was recorded with a finger clip Pulox CMS50E pulse oximeter. It is built for pulse rate detection and can be used to evaluate the performance of various rvPPG methods.

4.2 UBFC-rPPG Dataset

This dataset is called UBFC-rPPG [73] and it stands for University Bourgogne Franche-Comte Remote PhotoPlethysmoGraphy. It consists of two datasets that are focused specifically on rvPPG analysis. It was created using a custom C++ application for video acquisition with a simple low-cost Logitech C920 HD Pro webcam at 30fps with a resolution of 640x480 in uncompressed 8-bit RGB format. A CMS50E transmissive pulse oximeter was used to obtain the ground truth PPG data comprising the PPG waveform as well as the PPG heart rates. The participants were placed in front of the camera at an average distance of 1 meter and their faces were visible.

In the first dataset (simple), participants were asked to sit still but some videos present significant movement (especially at the beginning of the sequence). The dataset is composed of 8 videos (about 16500 frames). In the second dataset (realistic), the participants sit in front of the camera and are required to play a time-sensitive mathematical game that is aimed at augmenting their heart rate while simultaneously emulating a normal human-computer interaction scenario.

4.3 UBFC-Phys Dataset

The UBFC-Phys [74] is a public multimodal dataset, in which 56 participants underwent an experiment with a rigorous protocol inspired by the well-known Trier Social Stress Test (TSST), and was conducted in three stages: a rest, a speech, and arithmetic tasks with different levels of difficulty. During the experiment, participants were filmed and wore a wristband that measured their blood volume pulse (BVP) and electrodermal activity (EDA) signals. Before the start of the experiment and after the experiment was completed, participants completed a form to calculate their self-reported anxiety scores. The dataset contains participants' video recordings and BVP and EDA signals measured during the three tasks, as well as their anxiety scores calculated before and after the experimental sessions.

The ground truth measurements were captured using the Empatica E4 wristband 1, which records BVP, skin temperature, and EDA responses. The E4 bracelet has also an accelerometer and computes the Inter-Beat Intervals (IBI) from the BVP signal, which constitute the PRV. The UBFC-Phys dataset can be used to investigate the link between stress and physiological responses.

4.4 VIPL-HR Dataset

The VIPL-HR [75] dataset contains various face variations, such as head movements and illumination change, and acquisition diversity, to evaluate methods designed for real-world heart rate estimation. It has more than 107 participants and their videos are recorded under various illumination conditions and different head movements.

All the videos are recorded using three different cameras which are Logitech 310, Realsense F200, Huawei P9 smartphone. The relative physical measurements, such as heart rate (HR), oxygen saturation (SpO2), and blood volume pulse (BVP) signal, are also simultaneously recorded. The ground truth measurements were captured using Contech cms60c pulse oximeter.

4.5 COHFACE Dataset

The COHFACE [76] is a remote photoplethysmography (rPPG) dataset that contains RGB video sequences of faces, synchronized with the heart rate and breathing rate of the recorded subjects. The dataset includes 160 one-minute-long video sequences of 40 subjects (12 females and 28 males). The video sequences have been recorded using Logitech HD C525 at a resolution of 640x480 pixels and a frame rate of 20Hz. Physiological recordings, namely blood volume pulse (BVP) and breathing rate have also been recorded using BVP SA9308M, SA9311M blood volume pulse sensors. Physiological signals have been acquired using devices from Tought Technologies and using the provided BioGraph Infiniti software suite, version 5.

4.6 MANHOB-HCI Dataset

The MAHNOB-HCI [77] is a multimodal dataset recorded in response to affective stimuli with the goal of emotion recognition and implicit tagging research. A multimodal setup was arranged for the synchronized recording of face videos, audio signals, eye gaze data, and peripheral/central nervous system physiological signals. The dataset provides baseline results for the emotion recognition and implicit tagging for researchers who are going to use the database. The baseline results set a target for the researchers to reach.

There were twenty-seven participants which included 11 males and 16 females from different cultural backgrounds and they participated in two experiments. In the first experiment, they watched 20 emotional videos and self-reported their felt emotions using arousal, valence, dominance, and predictability as well as emotional keywords. In the second experiment, short videos and images were shown once without any tag and then with correct or incorrect tags.

4.7 LGI-PPGI Dataset

The LGI-PPGI [78] dataset contains video recordings of 25 participants which include 20 males and 5 females. The videos are recorded during head resting conditions, during head rotations, during an exercise on a bicycle ergometer in a gym, and during an urban conversation. To estimate the heart rate, a low-cost Logitech HD C270 camera was used under rigid and non-rigid facial motions and varying illumination. The ground truth measurements were captured using a CMS-50E pulse oximeter.

The first session consists of a resting scenario where no head or facial motion is performed and the illumination is more or less static. In the second session, the users are asked to perform head as well as facial motions but the illumination remains static. The third session is performed during an exercise on a bicycle ergometer in a gym and the fourth session is recorded during an urban conversation including head and facial motions as well as natural varying illumination conditions. Every session is recorded over a one-minute time span, except the ergometer session which is recorded over a 5 minute period.

4.8 OSF Dataset

The OSF dataset [79] contains recordings to evaluate and train remote photoplethysmography (rPPG) methods to estimate heart rate remotely by processing video that contains facial area. There were 7 participants which include 6 males and 1 females. The color signals acquired from videos using Logitech C920 HD Pro at a frame rate of 15Hz and heart rate values measured using MD300C318 pulse oximeter. The distance from the face to the webcam was in the range of 0.5–0.7 meter. The pixel size of the facial area was from 350 $\times$ 350 pixels to 550 $\times$ 550 and each video was recorded in daylight. The face is detected by OpenCV implementation of Viola-Jones algorithm.

4.9 iPPG Dataset

The iPPG - Imaging Photoplethysmography dataset [80] contains video recordings for heart rate and heart rate variability(HRV) from imaging Photoplethysmography signals acquired using cameras. It consists of activities with complex facial movement, researchers can use it to develop iPPG algorithms that are robust to such high facial motion scenarios, and measuring heart rate variability during different activities such as deep breathing and engagement is of interest. The dataset consists of facial video recordings of 14 people (8 males, six females) under different engagement activities such as reading a webpage, talking about their day, watching a movie trailer, and deep breathing. It also contains videos when they were stationary to establish a baseline. The video for each activity is 2 minutes. Ground truth PPG data were collected simultaneously using a CMS50D pulse oximeter. It can be used to develop applications such as stress or fatigue monitoring at the workplace as well as early detection of autonomic cardiac neuropathy.

4.10 Toadstool Dataset

The Toadstool dataset [81] consists of video, sensor, and demographic data collected from ten participants which included 5 males, and 5 females playing Super Mario Bros. The Toadstool dataset aims to contribute to the field of reinforcement learning, multimodal data fusion, and the possibility of exploring emotionally aware machine learning algorithms. The videos were recorded using Samsung Series 9 Notebook NP900X4C at a frame rate of 30Hz. The sensor data is collected through an Empatica E4 wristband, which provides high-quality measurements and is graded as a medical device. This dataset provides the (i) video frames of a person's facial expressions, (ii) the sensory output of the person playing a game, and (iii) data from the video game synchronized with the facial expressions and sensor data. The dataset opens up for a wide range of new and interesting analyses, and a proper and fair comparison between different methods, both from a psychological and a multimedia perspective. It can also be useful to researchers interested in facial expressions, biometric sensors, sentiment analysis, and game studies.

4.11 MMSE-HR Dataset

The MMSE-HR Multimodal Spontaneous Expression-Heart Rate dataset [82] is a well-annotated, multimodal, multidimensional spontaneous emotion corpus of 140 participants from various ethnic/racial backgrounds which include 58 males, and 82 females. The videos were recorded using Biopac MP150 and ground truth data was collected using Biopac NIBP-100D. Data were acquired from a variety of sensors of the face that included high-resolution 3D dynamic imaging, high-resolution 2D video, and thermal (infrared) sensing, and contact physiological sensors that included electrical conductivity of the skin, respiration, blood pressure, and heart rate. The facial expression was annotated for both the occurrence and intensity of facial action units from 2D video by experts in the Facial Action Coding System (FACS). The corpus further includes derived features from 3D, 2D, and IR (infrared) sensors and baseline results for facial expression and action unit detection.

4.12 Oulu Bio-Face (OBF) Dataset

The Oulu Bio-Face (OBF) dataset [83] is a large face video database for remote physiological signal measurement and atrial fibrillation detection. Atrial Fibrillation (AF) patients were recruited from Oulu University Hospital by cardiologists and nurses. The OBF database includes large number of facial videos with simultaneously recorded reference physiological signals . The data were collected both from healthy subjects and from patients with atrial fibrillation (AF) issue, which is the most common sustained and widespread cardiac arrhythmia encountered in clinical practice. Accuracy of heart rate, heart rate variability and radio-frequency measured from OBF videos are provided as the baseline results for evaluation. The recording set includes a computer, a color (RGB) camera, a near infrared (NIR) camera, a biosignal data acquisition device with three sets of sensors, and two LED lights. The ground truth was recorded using ECG sensor and pulse oximeter.

Dataset	Description	Device & FPS	Ground Truth	Number of Subjects	Application Domain
PURE	Consists of performing different, controlled head motions in front of a camera. Image sequences of the head and reference pulse measurements were recorded.	Eco274CVGE camera by SVS-Vistek & 30Hz	Pulox CMS50E pulse oximeter	Videos of 10 people including 8 females, 2 males.	Built for pulse rate detection, evaluation the performance of various rvPPG methods
UBFC-RPPG	Comprising of two datasets which are focused specifically for rvPPG analysis.	Logitech C920 HD Pro & 30Hz	CMS-50E pulse oximeter	First Dataset has 8 videos, second dataset has 42 videos	Heart rate monitoring and rvPPG analysis.
UBFC-PHYS	Provides more than 500 minutes of videos with the corresponding PPG signals. Investigates the link between stress & physiological responses.	Logitech C920 HD Pro & 30Hz	Empatica E4 wristband	56 people about 1 minute each video.	Built for rvPPG analysis and psycho-physiological studies.
VIPL-HR	Multi-modal database recorded with various head movement, illumination variations and has base lines for HR, SpO2 and blood volume pulse.	Logitech 310, Realsense F200, Huawei P9 smartphone & 30Hz	Contech cms60c pulse oximeter	Videos of 107 subjects	Built to be a large scale open dataset for rvPPG techniques.
COHFACE	RGB videos sequences of faces, synchronized with heart-rate and breathing-rate of the recorded subjects.	Logitech HD C525 Webcam & 20Hz	BVP SA9308M, SA9311M	40 people indivudials, 12 females, 28 males.	Estimation of blood volume pulse, breathing rate and physiological measurements.
MANHOB-HCI	Multimodal database, vaious light conditions, recorded in response to affective stimuli with the goal of emotion recognition and implicit tagging research.	6 video cameras, microphone, eye gaze tracker & 60Hz x 6	ECG, EEG 32 Channels	30 people, 17 females, 13 males.	Understanding of human emotional experience, individual's response to media items.
LGI-PPGI	Databse in presence head motion, talking, facial expressions and natural illumination conditions under different scenarios.	Logitech HD C270	CMS-50E pulse oximeter	25 people, 20 males, 5 females	Heart rate estimation under the load of nuisance factors.

Table 3: Open Source Datasets

Dataset	Description	Device & FPS	Ground Truth	Number of Subjects	Application Domain
OSF	Contains 70 color signals where the faces are already extracted through Viola-Jones, and the baseline measurements.	Logitech C920 HD Pro & 15Hz	MD300-C318 pulse oximeter	7 people, 6 males, 1 female.	Evaluation and training of rvPPG methods to estimate heart rate remotely.
iPPG	Imaging Photo-plethysmography Dataset, measurement of Heart Rate Variability	Blackfly BFLY-U3-23S6C & 30Hz	CMS-50D pulse oximeter	14 people, 8 males, 6 females.	Stress or fatigue monitoring, early detection of autonomic cardiac neuropathy.
Toadstool	Multimodal dataset which focuses on the human component of intelligent machines, participants playing Super Mario Bros.	Samsung Series 9 Notebook NP900X4C & 30Hz	Empatica E4 wristband	10 people, 5 males, 5 females.	Predicting specific sensor values using a combination of data from the video game and facial expressions.
MMSE-HR	Contains 2D facial expression video sequences and the corresponding heart rate and blood pressure sequences.	Biopac MP150 & 30Hz	Biopac NIBP-100D	140 people with various ethnic/racial background, 58 males, 82 females	Predicting the heart rate while simultaneously discovering the best regions of the face.
OBF	Facial videos with simultaneously recorded reference physiological signals, from healthy subjects and patients with atrial fibrillation.	RGB camera, NIR camera, Biosignal device & 30Hz	ECG sensor, pulse oximiter	100 healthy participants, 6 AF patient participants.	Physiological signal, diagnosing atrial fibrillation, cardiac arrhythmias.

Table 4: Open Source Datasets Continued

4.13 Pros and Cons of Datasets

	PROS	CONS
PURE	Six different recording positions	Low subject number(10) No changing light conditions Not Augmented heart rate
UBFC-RPPG	Augmented heart rate with stress test Varying amount of sunlight	Low number of videos(8)
UBFc-PHYS	Augmented heart rate with stress test Additional data with wristband High participant number(56)	No changing light conditions No audio data
VIPL-HR	Changing illumination Very high participant number(107) Additional data with sensors	Not augmented heart rate
COHFACE	High number of participants(40) Additional data with sensors	Not augmented heart rate No changing light conditions
MANHOB-HCI	Medium participant number(27)	No changing light conditions
LGI-PPGI	Medium participant number(25) Physical exercise of participants Varying illumination	Low camera resolution
OSF	Collection of Heart rate data with pulse oximeter	No changing light conditions Ntot augmented heartrate Low number of participants (7)
iPPG	Complex facial movements	No changing light conditions
Toadstool	Augmenting heart rate with video game	Low number of participants (10)
MMSE-HR	Very high participant number(140) High demographic profile Annotated facial expressions	No changing light conditions Not augmented heart rate
OBF	Very high participant number(100) Includes unhealthy participants	Unnatural lighting conditions with LED lights

Table 5: Pros and Cons of Datasets

5 Tools

This section include information about open source tools and commercially available tools. The rvPPG can be support for the doctors and individual users who want to monitor their heart health. The main goals of commercial tools include helping to diagnose cardiovascular problems accurately and give recommendations to prevent future complications. Commercial implementations of rvPPG are using remote heart rate monitoring with some additional features such as emotional signs monitoring and their solutions can be used by healthcare professionals, insurance companies, individual users, fitness and wellness companies. It is also possible to integrate these rvPPG solutions to other existing products to improve capabilities of the applications.

Based on the recent data published in June 2021 about the physician's workforce demand by the AAMC (Association of American Medical Colleges) [84], the United States could see an estimated shortage of up to 124,000 physicians by 2034, including shortfalls in both primary and specialty care. The main contributor to the shortage trend will be population growth and aging among the U.S population. As a result disparities and vulnerabilities in the healthcare system will deepen more. There are no quick solutions for this issue because healthcare professionals need extensive training which usually takes a long time.

Due to worker shortages in the healthcare industry, the demand for supportive tools that have capabilities of vital signal processing will increase. The available commercial tools can be very beneficial especially when the patients need immediate or flexible access to devices that provide vital signal processing capabilities. There are various commercial solutions are using the Remote photoplethysmography (rvPPG) technique. In this research paper, the relevant information about commercial solutions and their software specifications have been identified. These tools offer remote heart rate monitoring with some additional features such as emotional signs monitoring and their solutions can be used by healthcare professionals, insurance companies, individual users, and fitness and wellness companies.

Big tech companies like Google, Apple, Amazon and Microsoft are all investing on healthcare related Artificial Intelligence based solutions such as The Check Up app by Google Health, Amazon Care and Microsoft Cardiolens.

5.1 Commercial Tools and Applications

5.1.1 VitalSigns Biosensing

Philips develops medical tools and devices for the healthcare industry. The Biosensing VitalSigns Camera Technology [44] measures heart and breathing rates using the remote photoplethysmography (rvPPG) method and it senses skin color changes and movements in the chest and abdomen. Philips Biosensing maintains high accuracy even when the user is moving, this technology could be a good fit for fitness monitoring, vehicle monitoring, gaming, outdoor advertisement, and security besides vital signal monitoring in clinical settings. The camera tracks the rise and fall of the chest and abdomen to determine respiration rate and the tiny changes in skin color caused by heart rate are detected. Philips was able to solve many challenges associated with this technique such as motion, light, and other environmental factors.

Using any standard video camera, it can simultaneously measure pulse and breathing rate for a single person or multiple people at the same time within a range from 1 to 5 meters distance. It can be used with any standard camera and it does not need any special tools or devices. Since it is motion robust it uses facial tracking to get an accurate reading during motion, for example when the user is moving in front of the camera or the whole environment is moving. This feature ensures always a pulse and breathing rate can be calculated from the subject. The technology works with any skin type. Philips Biosensing VitalSigns Camera Technology is available through licensing to 3rd party manufacturers.

5.1.2 Oxehealth Oxevision

The Oxford University spin-off Oxehealth [85] has founded in 2012 and it has been accredited as a Class II a medical device by European regulators in 2018. The development project was funded by Health Education England to support clinical care through the use of physical activity reports and non-contact vital signs measurements. The Oxehealth Vital Signs system can be installed in a patient's hospital room, where it can spot measure pulse, heart, respiratory, and breathing rates with no physical contact. The company uses algorithms to analyze video to identify pixel level changes that correlate with pulse and breathing rates.

The Oxevision has received an FDA De Novo clearance in 2021 for software that can estimate pulse rate, heart rate, respiratory rate, and breathing rate from a video camera signal. The Oxehealth Vital Signs device is for use on humans 18 years of age or older who do not require critical care or continuous vital signs monitoring. It can be used for patient activity monitoring and medical-grade cardio-respiratory vital signs monitoring. It needs an installation of an Oxehealth Vital Signs device where it allows professionals to check patients' safety.

5.1.3 Noldus FaceReader

Noldus FaceReader [86] is an information technology company that develops software solutions for research on behavioral analysis for both humans and animals. The FaceReader is able to measure heart rate and heart variability using the rvPPG method without any attachments. Their technology can be used only with standard cameras. The FaceReader can be used for detecting emotional states and stress levels.

It offers solutions for the healthcare and wellness with facial expressions by analyzing and classification of the emotions and stress levels.It can be helpful for market researches, consumer behavior analysis, human computer interaction by observing facial expressions. It offers integration to the websites and applications, runs online on a cloud server.

5.1.4 BinahAi

The Binah.ai is a health data platform that provides solution for various industries that in need of health data and wide range of health measurements. It can be used using personal electronic devices and it does not require a purchase of any devices for both companies or their users. The measurements include blood pressure, heart rate, heart rate variability, oxygen saturation, respiration rate, sympathetic stress, parasympathetic activity, and pulse-respiration quotient.

The Binah.ai's solutions can be used by healthcare, insurance, digital health, pharmaceutical, fitness and wellness companies. It provides SDK integration with iOS, Android and Windows operating systems. Companies can easily integrate the Binah.ai solutions to their applications without requiring any physical investment. When customers are using the solution all the data stays in their personal device, Binah.ai does not store any of the data from facial scanning [41].

5.1.5 ShenAI

The ShenAI tool analyzes facial skin texture and vital physiological signs in real-time via remote visual photoplethysmography (rPPG). The platform provides vital signals like heart rate, heart rate variability, respiration rate, oxygen saturation and stress levels. It is a contactless optical measurement technique of recording skin blood pulsations at different vascular depths. The user needs just a smartphone, tablet, or any other device equipped with a camera.

The Shen.AI API can be integrated in mobile apps and web pages as a white-label solution. The mobile app heart monitor visualizes blood flow in real-time, enabling to track cardiovascular health. It helps to recognize early symptoms of hypertension. Their aim to reduce gaps in early diagnostics of heart related issues.

5.1.6 Anura.ai

The Anura.ai AI powered solution is digital health app monitoring app developed by Nuralogix. Mobile app allows instant health and wellness vital signals data from your smartphone. It provides solutions for healthcare, wellness, digital health, neuroscience, physiological indexes and physiological analysis. It allows heart rate, heart rate monitoring, blood pressure, mental state, stress monitoring. It is available through integration to Anura Core SDK for organizations and personal users can download mobile applicatio on iOS and Android marketplaces [87].

5.1.7 Wellfie

The WellFie is an AI based wellness selfie startup platform based in India. It allows convenient remote real-time monitoring and vital signs assessment. It provides reading for heart rate, breathing rate, stress level, oxygen saturation and blood pressure. It can be used for chronic illness monitoring, elderly care and self monitoring. It is available through web portal and mobile app. It provides SDK integration for both iOS and Android platforms [88].

5.1.8 Nervotec

The Nervotec remotely and effectively measures a person's vital signs within a minute to provide an accurate wellness index through NervoHealth platform [40]. It is headquartered in Singapore and currently provides solutions only for the businesses through their read-to-deploy platform. The NervoHealth solution help companies minimize health risks of their workforce. It can be used for workplace health monitoring such as health check up through the smartphone selfie before starting the work.

5.1.9 Cardiolens

The Cardiolens application [89] run on Microsoft HoloLens mixed reality smartglasses device and related research was published ACM SIGGRAPH 2017 Emerging Technologies conference for the first time in 2017. The mixed reality application that enables measurement and visualization of blood flow and vital signs through the HoloLens smartglas camera. The remote visual photoplethysmography software allow users to view the physiological state of a person who are in front of them via front-facing webcam video capture of the facial area.

This mixed reality system gives medical professionals the means to capture the signals of these aforementioned newer methods and view the heart rate and its variability in real time via head-mounted display. It allows surgeons to see whether a transplanted organ and tissue has blood flowing to it. Physical trainers would be able to see if athletes are in their "zone" and achieving their goals. It also allows to understand physiological states of people.

5.1.10 Vastmindz

The Vastmindz is AI based health monitoring system uses proprietary technology and works on any device with an embedded camera. It is able to extract various physiological parameters by analyzing a video of an individual. It can give reading for the heart rate, blood pressure, respiration rate, stress and oxygen saturation. It provides solutions for the healthcare, wellness, telehealth, aviation health screening, deep fake prevention, pet health monitoring (Biopawz). It is built for fast easy integration through SDK [43].

5.1.11 Lifelight

The Lifelight is digital healthcare platform that measures a person's vital signs using remote visual photoplethysmography. It is clinically validated and regulatory approved. It helps removing the cost and complexity of hardware blood pressure cuffs and pulse oximeters. It allows to be embedded into existing healthcare platforms and pathways, allowing them to be more efficient, more effective and more scalable for healthcare solutions. It is supported by the NHS England and it is built based on medical grade ISO 13485 and IEC 62304 standards.

The Lifelight platform allows long-term condition monitoring and remote primary care consultations. It gets accurate vital signs in one 40 second reading, just by the patient or user looking into a smartphone or tablet device, with no additional device needed. The Lifelight has been completed over 8500 patient physiological patient clinical study at Portsmouth Hospitals Trust and recorded over 1 million heartbeats. It allows primary care and secondary care through virtual consultations. It can be integrated to healthcare platform or application using smartphones for both iOS and Android devices.

Name	Solutions	Monitoring	Channels
VitalSigns Camera Technology	Healthcare, wellness, vehicle driver, sport and fitness, gaming, broadcasting, outdoor advertisement, and security.	Heart rate, breathing rate, sensing changes in skin color and abdomen movement.	Available for licensing to 3rd party manufacturers from Philips.
Oxehealth Vital Signs	Healthcare, wellness patient activity based alerts and warnings.	Medical grade cardio-respiratory vital signs monitoring.	Require installation of the Oxevision device.
FaceReader	Healthcare, wellness, facial expressions, emotions and stress analysis, classification.	Heart rate, behavior, emotions and stress monitoring.	Integration to the websites, runs online on a cloud server.
Binah.ai	Healthcare, wellness, insurance, digital health, and pharma.	Heart rate, respiration rate, oxygen saturation, emotion and stress.	SDK integration with iOS, Android and Windows.
Shen.AI	Healthcare, wellness, digital health, stress analysis and health reports.	Heart rate, respiration rate, oxygen saturation, and stress.	Heart Monitor app, SDK integration with iOS and Android.
Anura	Healthcare, digital health, wellness, physiological indexes, analysis, neuroscience.	Heart rate, blood pressure, mental state, stress monitoring.	Anura Core SDK, mobile application is available for iOS and Android devices.

Table 6: Commercial Tools and Applications

Name	Solutions	Monitoring	Channels
Wellfie	Healthcare, wellness life sciences, insurance, elder care, remote care self-monitoring.	Heart rate, blood pressure, respiration rate, stress and oxygen saturation.	Mobile app, SDK integration with iOS and Android.
Nervotec	Healtcare, wellness, digital health, workplace safety.	Heart rate, respiration rate, stress and oxygen saturation.	Available through integration to NervoHealth platform.
Cardiolens	Healtcare, wellness, fitness, emotions and physiological analysis.	Blood flow, heart rate, vital signs, physiological signals.	Require Microsoft Hololens headset.
Vastmindz	Healthcare, wellness, telehealth, aviation health screening, deep fake prevention, pet health (Biopawz).	Heart rate, blood pressure, respiration rate, stress and oxygen saturation.	Web based solution, SDK integration with iOS, Android.
Lifelight	Healthcare, wellness, digital healthcare, virtual health consultations.	Heart rate, behavior, mental health monitoring, health screening.	Integration to healthcare platform using smartphones, iOS and Android.

Table 7: Commercial Tools and Applications Continued

5.2 Open Source Tools and Applications

The open source tools and frameworks are available for research and development purposes that are using different techniques such as pyVHR pulse rate estimation, webcam pulse detection using OpenCV and heart rate detection through laptop webcam using Independent Component Analysis.

5.2.1 PyVHR

Package pyVHR [90] is a comprehensive Python framework for studying methods of pulse rate estimation relying on video, also known as remote photoplethysmography (rvPPG). The methodological rationale behind the framework is that in order to study, develop and compare new rvPPG methods in a principled and reproducible way. It can extend the pool of elements to be evaluated with newly developed rvPPG methods and any kind of video datasets.

5.2.2 PPGI-Toolbox

The PPGI-Toolbox [91] is a MATLAB toolbox for photoplethysmography imaging. This project funded by European Union Regional Development Fund, Federal Ministry of Education and Research and supported by the CanControls, ixp, Uniklinik and Medit Rwthaachen University. It has links for the example data, tests and algorithms information. Evaluation was conducted on the UBFC-RPPG and LGI Multi Session databases.

5.2.3 Irissometry

The Irissometry [92] is a rvPPG implementation using rPPG toolbox for MATLAB. The code has been tested in MATLAB 2019b on a Microsoft Windows 10 operating system. It detects person's face, it detects and tracks unique face points to calculate face rotations over time and it puts a mask on top of the face to detect only the skin. Applies time frequency analysis to extract heart rate as a function of video time. Creates several plots to allow inspection of signal and HR detection results. It allows modification of parameters for the signal processing steps.

5.2.4 RemotePPG

The RemotePPG [93] is official implementation of The Way to my Heart is through Contrastive Learning: Remote Photoplethysmography from Unlabelled Video. Similar to current state-of-the-art methods for rvPPG, it applies neural networks to learn deep representations with invariance to nuisance image variation. In contrast to such methods, it employs a fully self-supervised training approach, which has no reliance on expensive ground truth physiological training data.

5.2.5 MTTS-CAN

The MTTS-CAN, Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement [94] leverages a novel multi-task temporal shift convolutional attention network (MTTS-CAN) and enables real-time cardiovascular and respiratory measurements on mobile platforms. They evaluate their system on an ARM CPU and achieve state-of-the-art accuracy while running at over 150 frames per second which enables real-time applications. It provides a link for the camera-based remote PPG (Pulse) sensing demo.

5.2.6 rPPG

The rPPG [95] is robust heart rate estimation from facial videos. It monitors real time cardiac activities of a person through remote photoplethysmography(rPPG) without any physical contact by detecing blood volume pulse induced subtle color changes from video stream through webcam sensor or a video file. After detection and tracking region of interest for signal extraction it computes the spatial red, green and blue channel mean of skin segmented pixels to minimise camera quantization error. It uses deep learning for semantic segmentation of skin and non skin pixels from frames and the segmentation requires cuda enabled device.

5.2.7 Inter-beat

The Inter-beat Interval Estimation [96] provides video-based IBI estiamation. It estimates a person's cardiac information such as inter-beat interval (IBI) and average heart rate (HR) from a facial video. It provides a demo and it allows instantly running the whole code and display a comparison between IBIs from contact PPG and imaging PPG visually using Tokyo Tech Dataset. It estimates candidate blood volume pulse signals with a random patch-based ICA method. In default, it uses the green channel signals averaged within two different patches on the face.

5.2.8 ALT

The Artificial Light Texture (ALT) [97] technology is for non-contact vital signs such as heartbeats and respiration monitoring. It uses a video camera and a special light source to obtaining heart rate and respiration rate information for a person in real time. It works for any pose of the person and even when the person is covered by a very thick blanket or wears loose-fitting clothes. It does not use depth data to obtain the vital signs information. It can use inexpensive computing and image capture devices such as Raspberry Pi single-board computer and Pi NoIR camera. It is compatible with light emitting elements of various consumer electronics devices such as light projectors of standalone or embedded depth sensing devices like Microsoft Kinect, Intel RealSense cameras and Occipital Structure Sensor.

5.2.9 Heart Rate

The Heart Rate [98] is non-contact based real-time application system using camera to measure heart rate. It extracts heart rate information from facial skin color variation caused by blood circulation. The application's purpose is for monitoring drivers' physiological state. The data is collected via specialized device and Compact 5 medical Econet is used for the ground truth. It can only detect heart rate for 1 person at a time and sudden changes in motion can cause incorrect heart rate calculation. In the most case, heart rate can be correctly detected after 10 seconds being stable in front of the camera.

5.2.10 Heartbeat

The Heartbeat [99] is for measuring heart rate using remote photoplethysmography (rvPPG). It is a simple implementation of rvPPG technique to measure heart rate without skin contact. It uses a video recording or live feed of the face to analyse subtle changes in skin color. The face is detected and continuously tracked, signal series is obtained by determining the facial color in every frame and heart rate is estimated using frequency analysis and filtering of the series. The OpenCV libraries are required to run Heartbeat.

5.2.11 PulserateV2

The PulserateV2 Video Pulse Monitor [100] features a python script which is able to measure your heartbeat using a webcam or some other camera attached to a computer. It is a python implementation of a pulse rate monitor using rvPPG technique from video with the chrominance method. It uses OpenCV for face detection, for skin selection two methods are available. Skin classification using a hsv color range and forehead estimation. Due to the state of the art chrominance method the system is fairly motion robust. The framework also features a GUI that depicts the measured raw rvPPG signal and the resulting fourier spectrum.

5.2.12 iPhys

The iPhys [101] is an open non-contact imaging-based physiological measurement toolbox using MATLAB. This toolbox contains implementations of a number of algorithms for non-contact physiological measurement. The toolbox includes implementations of many of the most commonly used baseline methods for imaging photplethysmography (iPPG) and image ballistocardiography (iBCG). Imaging PPG (iPPG) focuses on the measurement of volumetric changes in blood flow at distance from the body using imaging devices to capture changes in transmitted or reflected light. Imaging ballistocardiography (iBCG) typically leverages optical flow estimation to track the vertical motion of the head or body from a video sequence. Both iPPG and iBCG methods can be used to recover human vital signals.

5.2.13 Eulerian

The Eulerian [102] is an implementation of the Eulerian video magnification computer vision algorithm initially developed by MIT CSAIL. This program uses the method for the application of remotely detecting an individual's heart rate in beats per minute from a still face video. The Eulerian Video Magnification method takes a standard video sequence as input, and applies spatial decomposition, followed by temporal filtering to the frames. The resulting signal is then amplified to reveal hidden information. Using this method, it is possible to visualize the flow of blood as it fills the face and also to amplify and reveal small motions. This technique can run in real time to show phenomena occurring at temporal frequencies selected by the user.

5.2.14 Pulse-detector

The Pulse-detector [103] is a python framework that detects the heart-rate of an individual using a common webcam or network IP camera. It tested on OSX, Ubuntu, and Windows. The application uses OpenCV to find the location of the user's face, then isolate the forehead region. Data is collected from this location over time to estimate the user's heart rate. This is done by measuring average optical intensity in the forehead location. Physiological data can be estimated this way thanks to the optical absorption characteristics of haemoglobin. When there is good lighting and minimal noise due to motion, a stable heartbeat should be isolated in about 15 seconds.

5.2.15 Pulse

Pulse [104] is a browser-based, non-contact heartrate detection application using Python and Javascript. It can detect the heartrate in thirty seconds or less, requiring only a browser and a webcam. Facial recognition, pixel manipulation, and some frequency extraction are done in Javascript, and there is a Python backend for Independent Component Analysis (JADE algorithm) and Fast Fourier Transform. A region of interest (ROI) is selected from the forehead hased on the tracked head area. It takes pulse about 10 seconds to gather enough data to start and pulse is most accurate after 30 seconds.

Name	**Description**	**Language**	**Year Started**	**Last Updated**
PyVHR	Python framework for Virtual Heart Rate is a comprehensive framework for studying methods of pulse rate estimation.	Python	2020	2022
PPGI-Toolbox	A MATLAB Toolbox for PPG Imaging	MATLAB	2019	2022
Irissometry	Remote heart rate detection of rPPG implementation toolbox for MATLAB	MATLAB	2017	2022
RemotePPG	Measurement of blood volume changes from observations of a person's face or skin.	Python MATLAB	2021	2021
MTTS-CAN	Multi-task temporal shift attention networks for on-device contactless vitals measurement	Python	2020	2021
rPPG	Robust Heart rate estimation, monitoring real time cardiac activities of a person through remote photoplethysmography.	Python	2019	2021
Inter-beat	Video-based inter-beat interval estimation	MATLAB	2019	2021
ALT	Real-time heart monitoring and respiration rates using Artificial Light Texture (ALT)	Python	2018	2021
Heart Rate	Real time application to measure heart rate using facial landmarks with camera	Python	2017	2021
Heartbeat	Simple implementation of rPPG, face analysis subtle changes in skin color.	C++	2016	2021
PulserateV2	Python implementation of a pulse rate monitor using chrominance method.	Python	2020	2020
iPhys	MATLAB iPPG implementations of a number of algorithms for physiological measurement	MATLAB	2018	2020
Eulerian	Remote heart rate detection through Eulerian magnification of face videos	Python	2017	2020
Pulse-detector	Webcam pulse detector using OpenCV	Python	2013	2017
Pulse	Pulse is a browser-based, non-contact heartrate detection application.	Javascript Python	2013	2017

Table 8: Open Source Tools and Applications

5.2.16 Pros and Cons of Open Source Tools

	PROS	CONS
PyVHR	Python supports Deep Learning Informative user interface Detailed Github Documentation Active GitHub with recent discussion threads Can be used with different datasets	No Mobile app Not tested on darker skin tones Weak debugging interface
PPGI-Toolbox	Tested on diverse face video databases Evaluated on different skin tones Strong publication of authors on Nature Funded by European Union Good Github documentation Active Github with recent commits	MATLAB does not support Deep learning Not a user friendly interface No mobile app
Irissometry	Good Github documentation Active Github with recent commits	MATLAB does not support Deep learning Adaptive user interface No live user community Works only on Windows
RemotePPG	Works with self-supervised approach No need for ground truth data Good Github documentation Active Github with recent commits	No user friendly interface No active user community Not tested with different datasets
MTTS-CAN	Python supports Deep Learning Works on mobile platforms A web based live demo	No recent Github update to codebase Not tested with different datasets
rPPG	Python supports Deep Learning Cross platform support Available pre-trained model	No recent Github update to codebase No graphical interface Not tested with different datasets
Inter-beat	Works on related part of the face Different head motions	No recent Github update to codebase No graphical interface Not tested with different datasets
ALT	Python supports Deep Learning Can work in real time Can work with low quality video Recent commit to Github	No graphical interface Not tested with different datasets
Heart Rate	Python supports Deep Learning Works in real time Graphical user interface	No support for motion No recent Github update to codebase No graphical interface Not tested with different datasets
Heartbeat	Works in real time or pre-recorded Uses OpenCV Works on desktop and mobile	No recent Github update to codebase No graphical interface Not tested with different datasets Insufficient Github documentation
PulserateV2	Python supports Deep Learning Motion tracking	No recent Github update to codebase No graphical interface Not tested with different datasets Insufficient Github documentation
iPhys	Utilized different rPPG methods Works with different video parameters Dynamic Video parameters	MATLAB does not support AI Works on pre-recorded videos No recent commit to Github
Eulerian	Python supports Deep Learning	Works on pre-recorded videos No recent commit to Github No motion tracking
Pulse-detector	Python supports Deep Learning Ease of installation User friendly interface	No recent updates Consisten lighting Minimal head motion No recent commit to Github
Pulse	Browser-based, cross-platform Easy installation Supports different illumination	No recent Github updates No motion tracking

Table 9: Pros and Cons of Open Source Tools

6 Discussion

The first research question of this thesis is "What capabilities and features are offered by rvPPG tools and research methods? Identify possible areas of usage that could benefit from rvPPG?". Findings of this thesis have addressed this research question as follows. Most possible areas of usage for rPPG such as telemedicine, vehicle opetor monitoring, physical exercise monitoring are listed in the current rPPG publications and they are listed in section 2. As heart and cardiac activity is so essential and relevant to all human conditions, we can anticipate higher and more variant uses of rPPG technology in the future. Contactless nature of rPPG also opens doors to many uses as remote, online and distant applications are going to shape the future of health in post COVID era.

The second research question of this thesis is "What are the available datasets and tools in the commercial and open source space?". The available datasets and tools are listed in sections 4 and 5. Their properties and pros and cons are listed in Tables 4.12, 4.13, **??**, 5.2.15 and 5.1.11.

There is not a best dataset among the current rPPG datasets. Each one of them has cons and pros as listed in 4.13. The properties of an ideal dataset can be listed as follows. First of all, current rPPG datasets do not include subjects with a variability of skin tone, gender, race etc. Most of them has narrow demographic profile who are mostly students who are taking classes with the faculty who are conducting the research if the research is done by academics. The experiments were done with the partipants found as the researchers who carried out the experiments might have felt a pressure to find the enough number of volunteer participants. rPPG works different for people for different skintones especially for darker skintones [105]. A good dataset for rPPG should include high number of participants with different skin tone, racial features and genders. Secondly, rPPG datasets should be recorded under different illumination conditions. Because in real life if rPPG is going to be used for telemedicine, patients are going to connect to telehealth application under different light conditions based on time of the day, room lighting etc. Another criteria for a good rPPG dataset is creating different psychological conditions for participants in order to mimic real life varying heart rates. For instance in a real world setting, patient might be anxious, scared or upset which can affect heart rate. In the laboratory setting, stress tests can be given to induce these emotions in partipants. Moreover, participants should have different physical conditions such as talking, running, sitting, standing. They should also have varying head movements during the recordings. A recording which the participant is sitting still with his/her head towards the camera will fail to replicate the real world conditions. On the other hand, besides rPPG camera additional physiological data such as blood volume pulse, breathing rate should be recorded to assist the rPPG findings and to collect ground truth data.

Likely there is not a best tool in the current rPPG tools. Every tool has disadvantages and advantages. Properties that should be in a good tool can be listed as follows. None of the solutions have been tested these solutions on darker skin tones. The impact of skin type on rPPG methods is still an open issue, especially on darker skin tones. A good tool should address this issue and should be tested on darker skin tones and different demographics as well [105]. One other property of a good tool is being tested on multiple datasets. Some tools have been tested on various datasets which increases their credibility. Because it is always easy to get reliable results on the data you have created. Some tools have been tested on very low number of videos which decreases their reliability. One of the most important attribute of an optimal tool is being cross platform and supporting different operating systems such as Windows, Mac, Linux. Also browser based tools are good for platform independence. Mobile based tools are very convenient as users can easily have access to a mobile phone compared to a computer. For a tool to be convenient, it has to have a good user interface with changeable parameters. Most of the tools fail to develop a solid user interface, they mostly print out results out of the command window as the results have generally been used in research. The tools with Github pages with the most recent updates also attrack more users. As listed in **??** table, some tools have uptodate Github pages with live discussion forums such as PyVHR. In addition to all these characteristics of a good tool, one other important feature will be working with the unlabelled data via semisupervised approaches. Most of the rPPG methods use a traditional PPG device in addition to the rPPG cameras to use the measurements of the PPG device as a ground truth to compare the rPPG calculation. Tools that work with deep learning approach such as rPPG tool use elf-surpervised deep learning approach to find out the accuracy of the method without the need for a ground truth validation data. One other trait is a good Github documentation and ease of installation. If the user faces a lot of frustration while installing the tool, they are very likely to give up using the tool. Some tools such as MTTS-CAN, rPPG have not been updated for many years. The Python library versions they have used are not recent. So the users installing these tools have to find the older libraries which always causes a lot of heeadache for the users, resulting in withdrawal from the tool usage. To sum up, we cannot nominate any of the tools as the best tool, a tool with properties above will be the best.

7 Conclusion

In this research, survey of Remote Visual Photoplethysmography (rvPPG) literature, methods, datasets and tools was performed in addition to an informal usage areas inspection. The results of this review and analysis led the discovery of some opportunities to improve rvPPG tools for telehealth. Data privacy will remain a crucial part of the development process, and the necessity for usability is a critical goal for health professionals. With the discoveries being made in research tools, there is an opportunity for tool developers to migrate beneficial and cost-saving capabilities to the industry for development of efficient rvPPG tools.

When academic research give empirical studies showing the tangible benefits of using these tools, there should be a plan in place to migrate these technologies to a more lasting venue. By equipping rvPPG software developers with effective tools to understand impacts of potential changes, it is hoped that the rvPPG will make a great impact on how telehealth is done opening easier and more affordable health care access to everyone including disadvantaged people with poor health benefits and disabled people who cannot make a visit to a doctor's office physically.

References

[1] S. S. Barold, "Willem einthoven and the birth of clinical electrocardiography a hundred years ago," *Cardiac electrophysiology review*, vol. 7, no. 1, pp. 99–104, 2003.

[2] A. B. Hertzman, "Observations on the finger volume pulse recorded photoelectrically," *Am. J. Physiol.*, vol. 119, pp. 334–335, 1937.

[3] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light." *Opt. Express*, vol. 16, no. 26, pp. 21 434–21 445, Dec 2008. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-16-26-21434

[4] R. V. Tuckson, M. Edmunds, and M. L. Hodgkins, "Telehealth," *New England Journal of Medicine*, vol. 377, no. 16, pp. 1585–1592, 2017.

[5] N. Catalyst, "What is telehealth?" *NEJM Catalyst*, vol. 4, no. 1, 2018.

[6] J. P. Marcin, U. Shaikh, and R. H. Steinhorn, "Addressing health disparities in rural communities using telehealth," *Pediatric research*, vol. 79, no. 1, pp. 169–176, 2016.

[7] M. D. Wong, M. F. Shapiro, W. J. Boscardin, and S. L. Ettner, "Contribution of major diseases to disparities in mortality," *New England Journal of Medicine*, vol. 347, no. 20, pp. 1585–1592, 2002.

[8] C. A. Olson and J. F. Thomas, "Telehealth: no longer an idea for the future," *Advances in Pediatrics*, vol. 64, no. 1, pp. 347–370, 2017.

[9] T. S. Nesbitt *et al.*, "The evolution of telehealth: where have we been and where are we going," in *Board on Health Care Services; Institute of Medicine, eds. The Role of Telehealth in an Evolving Health Care Environment: Workshop Summary. Washington (DC): National Academies Press (US)*, 2012.

[10] H. Nisar, M. B. Khan, W. T. Yi, Y. V. Voon, and T. S. Khang, "Contactless heart rate monitor for multiple persons in a video," in *2016 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*. IEEE, 2016, pp. 1–2.

[11] "The Open Telehealth (OTH) Remote Patient Monitoring," https://www.oth.io.

[12] "Estimated Hypertension Prevalence, Treatment, and Control Among U.S. Adults," https://millionhearts.hhs.gov/data-reports/hypertension-prevalence.html.

[13] "Many Die Of Heart Attacks Without Prior History Or Symptoms: PET Imaging Can Offer Early Warning," https://www.sciencedaily.com/releases/2008/06/080616124938.htm.

[14] T. P. Hofer, R. A. Hayward, S. Greenfield, E. H. Wagner, S. H. Kaplan, and W. G. Manning, "The unreliability of individual physician report cards for assessing the costs and quality of care of a chronic disease," *Jama*, vol. 281, no. 22, pp. 2098–2105, 1999.

[15] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE transactions on biomedical engineering*, vol. 58, no. 1, pp. 7–11, 2010.

[16] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiological measurement*, vol. 28, no. 3, p. R1, 2007.

[17] L.-G. Lindberg and P. A. Oberg, "Optical properties of blood in motion," *Optical Engineering*, vol. 32, no. 2, pp. 253–257, 1993.

[18] S. Benedetto, C. Caldato, D. C. Greenwood, N. Bartoli, V. Pensabene, and P. Actis, "Remote heart rate monitoring-assessment of the facereader rppg by noldus," *PloS one*, vol. 14, no. 11, p. e0225592, 2019.

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

[20] K. Takazawa, N. Tanaka, M. Fujita, O. Matsuoka, T. Saiki, M. Aikawa, S. Tamura, and C. Ibukiyama, “Assessment of vasoactive agents and vascular aging by the second derivative of photoplethysmogram waveform,” *Hypertension*, vol. 32, no. 2, pp. 365–370, 1998.

[21] A. Reisner, P. A. Shaltis, D. McCombie, H. H. Asada, D. S. Warner, and M. A. Warner, “Utility of the photoplethysmogram in circulatory monitoring,” *The Journal of the American Society of Anesthesiologists*, vol. 108, no. 5, pp. 950–958, 2008.

[22] L. Nilsson, A. Johansson, and S. Kalman, “Respiration can be monitored by photoplethysmography with high sensitivity and specificity regardless of anaesthesia and ventilatory mode,” *Acta anaesthesiologica scandinavica*, vol. 49, no. 8, pp. 1157–1162, 2005.

[23] I. Korhonen and A. Yli-Hankala, “Photoplethysmography and nociception,” *Acta Anaesthesiologica Scandinavica*, vol. 53, no. 8, pp. 975–985, 2009.

[24] A. Avolio, “The finger volume pulse and assessment of arterial properties,” *Journal of hypertension*, vol. 20, no. 12, pp. 2341–2343, 2002.

[25] 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. 15 776–15 790, 2017.

[26] Y. Mendelson and B. Ochs, “Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography,” *IEEE Transactions on Biomedical Engineering*, vol. 35, no. 10, pp. 798–805, 1988.

[27] D. J. Meredith, D. Clifton, P. Charlton, J. Brooks, C. Pugh, and L. Tarassenko, “Photoplethysmographic derivation of respiratory rate: a review of relevant physiology,” *Journal of medical engineering & technology*, vol. 36, no. 1, pp. 1–7, 2012.

[28] T. Tamura, Y. Maeda, M. Sekine, and M. Yoshida, "Wearable photoplethysmographic sensors—past and present," *Electronics*, vol. 3, no. 2, pp. 282–302, 2014.

[29] A. Panda, S. Pinisetty, and P. Roop, "Runtime monitoring and statistical approaches for correlation analysis of ecg and ppg," *arXiv preprint arXiv:2202.00559*, 2022.

[30] Q. Zhan, W. Wang, and G. de Haan, "Analysis of cnn-based remote-ppg to understand limitations and sensitivities," *Biomedical optics express*, vol. 11, no. 3, pp. 1268–1283, 2020.

[31] A. Gudi, M. Bittner, R. Lochmans, and J. van Gemert, "Efficient real-time camera based estimation of heart rate and its variability," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[32] L. V. Coutts, D. Plans, A. W. Brown, and J. Collomosse, "Deep learning with wearable based heart rate variability for prediction of mental and general health," *Journal of Biomedical Informatics*, vol. 112, p. 103610, 2020.

[33] W. Wang, "Robust and automatic remote photoplethysmography," 2017.

[34] H. Singh, J. S. Bhatia, and J. Kaur, "Eye tracking based driver fatigue monitoring and warning system," in *India International Conference on Power Electronics 2010 (IICPE2010)*. IEEE, 2011, pp. 1–6.

[35] A. Mittal, K. Kumar, S. Dhamija, and M. Kaur, "Head movement-based driver drowsiness detection: A review of state-of-art techniques," in *2016 IEEE International Conference on Engineering and Technology (ICETECH)*. IEEE, 2016, pp. 903–908.

[36] J. Mesken, M. P. Hagenzieker, T. Rothengatter, and D. de Waard, "Frequency, determinants, and consequences of different drivers' emotions: An on-the-road study using self-reports, (observed) behaviour, and physiology," *Transportation Research*

Part F: Traffic Psychology and Behaviour, vol. 10, no. 6, pp. 458–475, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1369847807000265

[37] R. Cellan-Jones, "Deepfake videos double in nine months." https://www.bbc.com/news/technology49961089, 2019 (accessed March 6, 2021).

[38] S. Bharadwaj, T. I. Dhamecha, M. Vatsa, and R. Singh, "Computationally efficient face spoofing detection with motion magnification," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2013, pp. 105–110.

[39] J. Hernandez-Ortega, J. Fierrez, A. Morales, and P. Tome, "Time analysis of pulse-based face anti-spoofing in visible and nir," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 544–552.

[40] "NervoHealth, engage your users with real-time health insights," https://nervotec.com.

[41] "BinahAi, The New Way to Measure Health Data Is Software-based," https://www.binah.ai.

[42] I. Anderzén and B. B. Arnetz, "The impact of a prospective survey-based workplace intervention program on employee health, biologic stress markers, and organizational productivity," *Journal of occupational and environmental medicine*, vol. 47, no. 7, pp. 671–682, 2005.

[43] "Vastmindz, Gain health wellness insights, remotely," https://vastmindz.com.

[44] "Vital Signs Camera: Accurate contactless pulse and breathing rate measurement," https://www.philips.com/a-w/about/innovation/ips/ip-licensing/programs/vitalsigns-camera-technology.html.

[45] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision*, vol. 88, no. 2, pp. 303–338, 2010.

[46] M. Kristan, J. Matas, A. Leonardis, T. Vojíř, R. Pflugfelder, G. Fernandez, G. Nebehay, F. Porikli, and L. Čehovin, "A novel performance evaluation methodology for single-target trackers," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 11, pp. 2137–2155, 2016.

[47] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition.* Ieee, 2009, pp. 248–255.

[48] C. Zhao, P. Mei, S. Xu, Y. Li, and Y. Feng, "Performance evaluation of visual object detection and tracking algorithms used in remote photoplethysmography," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[49] B.-F. Wu, P.-W. Huang, D.-H. He, C.-H. Lin, and K.-H. Chen, "Remote photoplethysmography enhancement with machine leaning methods," in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC).* IEEE, 2019, pp. 2466–2471.

[50] C. Wang, T. Pun, and G. Chanel, "A comparative survey of methods for remote heart rate detection from frontal face videos," *Frontiers in bioengineering and biotechnology*, vol. 6, p. 33, 2018.

[51] G. Boccignone, D. Conte, V. Cuculo, A. D'Amelio, G. Grossi, and R. Lanzarotti, "An open framework for remote-ppg methods and their assessment," *IEEE Access*, vol. 8, pp. 216 083–216 103, 2020.

[52] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light." *Optics express*, vol. 16, no. 26, pp. 21 434–21 445, 2008.

[53] Z. Yu, W. Peng, X.-B. Li, X. Hong, and G. Zhao, "Remote heart rate measurement from highly compressed facial videos: An end-to-end deep learning solution with video enhancement," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 151–160, 2019.

[54] Z. Yu, X.-B. Li, and G. Zhao, "Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks," in *BMVC*, 2019.

[55] R. Spetlik, V. Franc, J. ech, and J. Matas, "Visual heart rate estimation with convolutional neural network," in *BMVC*, 2018.

[56] X. Niu, H. Han, S. Shan, and X. Chen, "Synrhythm: Learning a deep heart rate estimator from general to specific," *2018 24th International Conference on Pattern Recognition (ICPR)*, pp. 3580–3585, 2018.

[57] T. Luguev, D. Seuß, and J. Garbas, "Deep learning based affective sensing with remote photoplethysmography," in *2020 54th Annual Conference on Information Sciences and Systems (CISS)*, 2020, pp. 1–4.

[58] M. Kopeliovich, Y. Mironenko, and M. Petrushan, "Architectural tricks for deep learning in remote photoplethysmography," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[59] W. Chen and D. McDuff, "Deepphys: Video-based physiological measurement using convolutional attention networks," *ArXiv*, vol. abs/1805.07888, 2018.

[60] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. S. Bernstein, A. Berg, and L. Fei-Fei, "Imagenet large scale visual

recognition challenge," *International Journal of Computer Vision*, vol. 115, pp. 211–252, 2015.

[61] A. B. Hertzman, "Photoelectric plethysmography of the fingers and toes in man," *Proceedings of the Society for Experimental Biology and Medicine*, vol. 37, no. 3, pp. 529–534, 1937.

[62] M. Lewandowska, J. Rumiński, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity," in *2011 federated conference on computer science and information systems (FedCSIS)*. IEEE, 2011, pp. 405–410.

[63] G. De Haan and V. Jeanne, "Robust pulse rate from chrominance-based rppg," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 10, pp. 2878–2886, 2013.

[64] 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, p. 1913, 2014.

[65] 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.

[66] C. Pilz, S. Zaunseder, J. Krajewski, and V. Blazek, "Local group invariance for heart rate estimation from face videos in the wild," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1335–13 358, 2018.

[67] W. Chen and D. McDuff, "Deepphys: Video-based physiological measurement using convolutional attention networks," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 349–365.

[68] S. K. A. Prakash and C. S. Tucker, "Bounded kalman filter method for motion-robust, non-contact heart rate estimation," *Biomed. Opt. Express*, vol. 9, no. 2, pp. 873–897, Feb 2018. [Online]. Available: http://opg.optica.org/boe/abstract.cfm?URI=boe-9-2-873

[69] Z. Yu, X. Li, and G. Zhao, "Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks," *arXiv preprint arXiv:1905.02419*, 2019.

[70] L. He, K. S. Alam, J. Ma, E. Burkholder, W. C. Chung Chu, A. Iqbal, and S. I. Ahamed, "Remote photoplethysmography heart rate variability detection using signal to noise ratio bandpass filtering," in *2021 IEEE International Conference on Digital Health (ICDH)*, 2021, pp. 133–141.

[71] G. Wang, "Influence of roi selection for remote photoplethysmography with singular spectrum analysis," in *2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 2021, pp. 416–420.

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

[73] S. Bobbia, R. Macwan, Y. Benezeth, A. Mansouri, and J. Dubois, "Unsupervised skin tissue segmentation for remote photoplethysmography," *Pattern Recognition Letters*, vol. 124, pp. 82–90, 2019.

[74] R. M. Sabour, Y. Benezeth, P. De Oliveira, J. Chappe, and F. Yang, "Ubfc-phys: A multimodal database for psychophysiological studies of social stress," *IEEE Transactions on Affective Computing*, 2021.

[75] X. Niu, H. Han, S. Shan, and X. Chen, "Vipl-hr: A multi-modal database for pulse estimation from less-constrained face video," in *Asian Conference on Computer Vision*. Springer, 2018, pp. 562–576.

[76] G. Heusch, A. Anjos, and S. Marcel, "A reproducible study on remote heart rate measurement," *arXiv preprint arXiv:1709.00962*, 2017.

[77] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 42–55, 2011.

[78] 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*, 2018, pp. 1254–1262.

[79] M. Kopeliovich, M. Petrushan, and D. Shaposhnikov, "Approximation-based transformation of color signal for heart rate estimation with a webcam," *Pattern Recognition and Image Analysis*, vol. 28, no. 4, pp. 646–651, 2018.

[80] A. Pai, A. Veeraraghavan, and A. Sabharwal, "Hrvcam: robust camera-based measurement of heart rate variability," *Journal of Biomedical Optics*, vol. 26, no. 2, p. 022707, 2021.

[81] H. Svoren, V. Thambawita, P. Halvorsen, P. Jakobsen, E. Garcia-Ceja, F. M. Noori, H. L. Hammer, M. Lux, M. A. Riegler, and S. A. Hicks, "Toadstool: a dataset for training emotional intelligent machines playing super mario bros," in *Proceedings of the 11th ACM Multimedia Systems Conference*, 2020, pp. 309–314.

[82] Z. Zhang, J. M. Girard, Y. Wu, X. Zhang, P. Liu, U. Ciftci, S. Canavan, M. Reale, A. Horowitz, H. Yang *et al.*, "Multimodal spontaneous emotion corpus for human

behavior analysis," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3438–3446.

[83] X. Li, I. Alikhani, J. Shi, T. Seppanen, J. Junttila, K. Majamaa-Voltti, M. Tulppo, and G. Zhao, "The obf database: A large face video database for remote physiological signal measurement and atrial fibrillation detection," in *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*, 2018, pp. 242–249.

[84] A. of American Medical Colleges, "The complexities of physician supply and demand: Projections from 2019 to 2034," 2021.

[85] A. Lewis, V. Gandhi, B. Venugopal, J. Bowen, M. Polkey, and O. Gibson, "Non contact patient monitoring and management system (oxevision) in patient management," *Physiotherapy*, vol. 114, p. e13, 2022.

[86] "Remote measurement with FaceReader Online," https://www.noldus.com/facereader.

[87] "Nuralogix, Anuara.ai Health Monitoring," https://www.anura.ai.

[88] "AI Based Wellness Selfie," https://wellfie.in.

[89] C. Hurter and D. McDuff, "Cardiolens: remote physiological monitoring in a mixed reality environment," in *ACM siggraph 2017 emerging technologies*, 2017, pp. 1–2.

[90] "Package pyVHR, Python framework for Virtual Heart Rate," https://github.com/phuselab/pyVHR.

[91] "A MATLAB Toolbox for Photoplethysmography Imaging," https://github.com/partofthestars/PPGI-Toolbox.

[92] "The Irissometry, rvPPG implementation using rPPG toolbox for MATLAB. ," https://github.com/marnixnaber/rPPG.

[93] "The Way to my Heart is through Contrastive Learning: Remote Photoplethysmography from Unlabelled Video," https://github.com/ToyotaResearchInstitute/RemotePPG/tree/master/iccv.

[94] "MTTS-CAN: Multi-Task Temporal Shift Attention Networks for On-Device Contactless Vitals Measurement," https://github.com/xliucs/MTTS-CAN.

[95] "Robust Heart rate estimation from facial videos," https://github.com/nasir6/rPPG.

[96] "Video-based Inter-beat Interval Estimation Code," https://github.com/ymonno/RemotePPG.

[97] "The Artificial Light Texture (ALT) technology for non-contact vital signs (heartbeats, respiration) monitoring." https://github.com/lvetech/ALT.

[98] "Heart-rate-measurement-using-camera," https://github.com/habom2310/Heart-rate-measurement-using-camera.

[99] "Heartbeat: Measuring heart rate using remote photoplethysmography (rPPG)," https://github.com/prouast/heartbeat.

[100] "Python Video Pulse Monitor, PulserateV2," https://github.com/MartinChristiaan/PythonVideoPulserateV2.

[101] "iPhys: An Open Non-Contact Imaging-Based Physiological Measurement Toolbox," https://github.com/danmcduff/iphys-toolbox.

[102] "Remote Heart Rate Detection through Eulerian Magnification of Face Videos," https://github.com/rohintangirala/eulerian-remote-heartrate-detection.

[103] "Webcam-pulse-detector, stand-alone (no dependancy) precompiled application," https://github.com/thearn/webcam-pulse-detector.

[104] "Pulse, browser-based, non-contact heartrate detection application," https://github.com/camilleanne/pulse.

[105] E. M. Nowara, D. McDuff, and A. Veeraraghavan, "A meta-analysis of the impact of skin tone and gender on non-contact photoplethysmography measurements," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 284–285.