

Literature Review on Multi-model method based on Remote Photoplethysmography(rPPG) for Driving Fatigue Detection

Name: Yinqiao Li

Student Number: 8455569

April 15, 2024

1 Introduction

It's anticipated that road traffic accidents will rank third in terms of leading causes of death. The societal and economic impacts of such incidents are continuously intensifying for the majority of global populations. Annually, driver exhaustion is implicated in a significant number of mishaps. Specifically, around 20% of total collisions are in relation to driving fatigue. Hence, devising a mechanism for early detection of driver fatigue has emerged as a crucial challenge to curb these escalating traffic accidents.

Several years back, the emergence of computer vision techniques has been proving effective in many scenarios. Vision-based methods were proposed to detect fatigue by capturing specific features with one or more cameras and image processing. Many researchers conducted the operations by recognizing certain characteristics via image processing on one or multiple cameras. Researchers also, faced with the challenge of extracting visible light in nighttime conditions, have attempted to use infrared light cameras instead of traditional RGB cameras. In the early years, a driver monitor system with feature extraction functions including the eyes blinking, yawning, and the head angle variance, was proposed and proven to be effective [8, 9]. Previously, Singh posited that driver fatigue can be identified when the duration of eye closure surpasses a specified threshold[15]. Road camera images are also used for driver fatigue detection, based on the location between vehicles and lanes[17]. However, these methods rely entirely on RGB cameras, which pose many disadvantages in low-light environments. This issue is particularly noticeable at night, which is a period when driver fatigue often occurs. More importantly, they activate an alert only post the occurrence of fatigue. To address these issues, physiological-based procedures have been recommended. These use physiological signals to determine fatigue status. An additional advantage is that a physiological signal directly communicates subject information and is minimally affected by diverse driving environments.

2 Methodology

In addressing the challenge of driver fatigue detection, this study employs a multifaceted approach that combines both vision-based and physiological signal-based methods to ensure robustness across diverse driving conditions. Our methodology is structured into two primary phases: data collection and feature analysis, followed by fatigue detection and alert mechanism.

2.1 Data Collection

Data collection is conducted using two parallel systems. The first system utilizes an infrared camera setup to capture the driver's facial expressions and head movements continuously. This setup is designed to function optimally both during the day and in low-light conditions at night. The second system comprises wearable sensors that monitor various physiological

signals, including heart rate variability (HRV), electroencephalogram (EEG), and galvanic skin response (GSR).

2.2 Feature Analysis

The video data from the infrared camera is processed to extract features indicative of fatigue, such as eye blink frequency, yawn frequency, and head pose angles. These features are extracted using advanced computer vision algorithms that include facial landmark detection and tracking. Concurrently, physiological data is analyzed to extract features such as changes in HRV, EEG patterns associated with drowsiness, and increased GSR, which are known correlates of fatigue.

3 Literature Review

3.1 Remote Photoplethysmography (rPPG)

Physiological information, including heart rate(HR), pulse rate variability, and respiration rate, are widely utilized to evaluate the physical state of individuals. In these physiological sources, heart rate signal is the most widely used. However, the traditional method for heart rate monitoring has many disadvantages. Photoplethysmography (PPG) is a non-invasive method for heart state monitoring, extracting physiological signals by the color variance to measure the blood vessel under the skin, which has been widely used in the medical field to detect and monitor various physiological indicators[14]. The non-invasive nature and low cost of PPG have led to its widespread use. Although many PPG-related products have been applied to medical and other vital signs monitoring fields in recent years, PPG still has limitations for some fields. Based on the exploration of PPG, remote photoplethysmography (rPPG) is a remote method of PPG signal extraction.

rPPG technology extracts blood volume pulse signals based on subtle color variations in facial skin under visible or non-visible light conditions. Utilizing standard signal processing techniques, alongside cutting-edge machine learning and deep learning approaches, it is possible to derive a variety of physiological indicators from these signals, including heart rate variability, pulse rate variability, heart rate, and respiration, among others[6]. As shown in Figure 1. Since

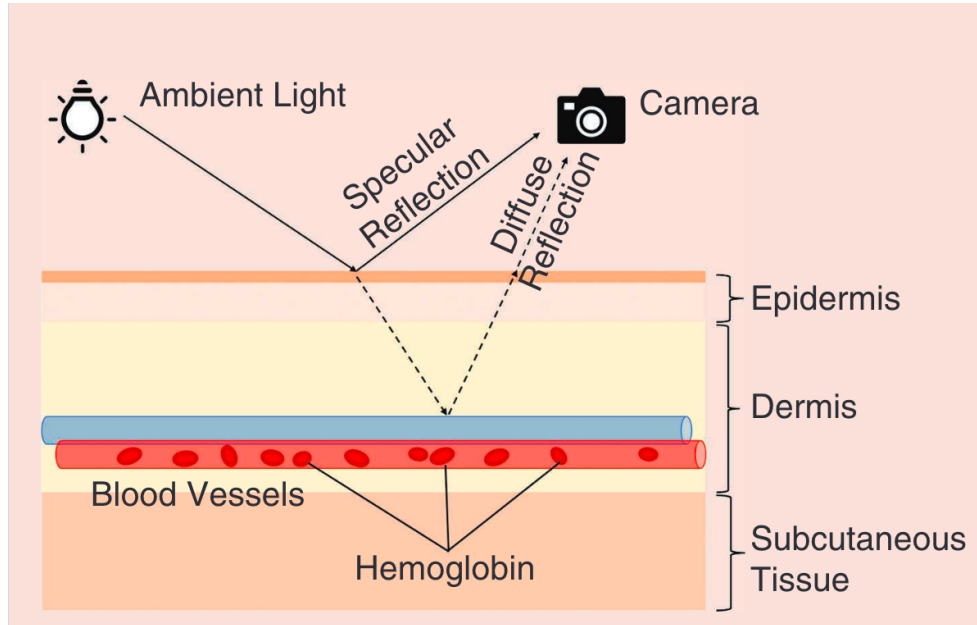


Figure 1: Principle of rPPG

Verkruyse et al. first proposed the use of rPPG technology for physiological measurement, researchers have introduced numerous methods to extract stable Blood Volume Pulse (BVP) signals and perform robust physiological measurements[18]. The extraction methods for rPPG can be broadly categorized into traditional vision-based signal processing techniques and deep learning approaches.

3.1.1 Non-contact, automated cardiac pulse measurements using video imaging and blind source separation

Early work in the field of rPPG was based on visual methods that involved using the subtle color changes in the skin caused by the flow of blood pulses in subcutaneous vessels under ambient light as the input signal. The blood volume pulse signal was obtained through signal processing techniques such as filtering. Traditional image processing methods typically involve multiple steps, such as facial tracking, ROI extraction, RGB transformation, feature selection, and noise filtering. Analysis of the relationship and characteristics between peaks in the signal was conducted to extract physiological information. Poh et al. were the pioneers in applying principal component analysis (PCA) and independent component analysis (ICA) for dimensionality reduction of the extracted parameters, identifying components that were useful for extracting physiological data[12].

3.1.2 Bounded Kalman filter method for motion-robust, non-contact heart rate estimation

The approach developed by Prakash and Tucker, VideoVitals, utilizes a Bounded Kalman Filter (BKF) coupled with denoising techniques to minimize the effects of motion and refine the tracking of features. When subjected to testing on individuals executing a range of head motions, the BKF algorithm outperformed traditional techniques by enhancing accuracy despite movement. It's important to note that the evaluation of BKF's effectiveness was carried out with participants in a controlled lab environment within a university context, paralleling the testing conditions of other advanced methodologies in the field[13].

3.1.3 Deepphys: Video-based physiological measurement using convolutional attention networks

3.1.4 HR-RCNN: Hierarchical relational reasoning for object detection

3.1.3 and 3.1.4 are two similar research which are both based on 1DCNN.

DeepPhys is a network based on deep learning methods that include two smaller networks for ROI detection and extracting heart rate signals through differential frames[4].

HR-CNN is a convolutional neural network model based on two different loss functions, where the two networks function similarly to DeepPhys. However, the difference lies in the configuration of different loss functions for the two networks within the model[3].

3.1.5 Bigsmall: Efficient multi-task learning for disparate spatial and temporal physiological measurements

Girish et al. proposed BigSmall, an efficient physiological and behavioral measurement architecture inspired by the principles of human visual perception. It enhances computational efficiency and accuracy through a dual-branch network and temporal shift modules. The two branches, the big branch processes the high-resolution input to extract AU, and the small branch utilizes the low-resolution input to extract the rPPG signal. The different resolution input is inspired by the different signal input requirements because the resolution quality is not necessary for rPPG extraction. The two branches leverage the characteristic that rPPG requires only low-resolution input to simplify the model complexity of rPPG extraction, As shown in Figure 2 [10].

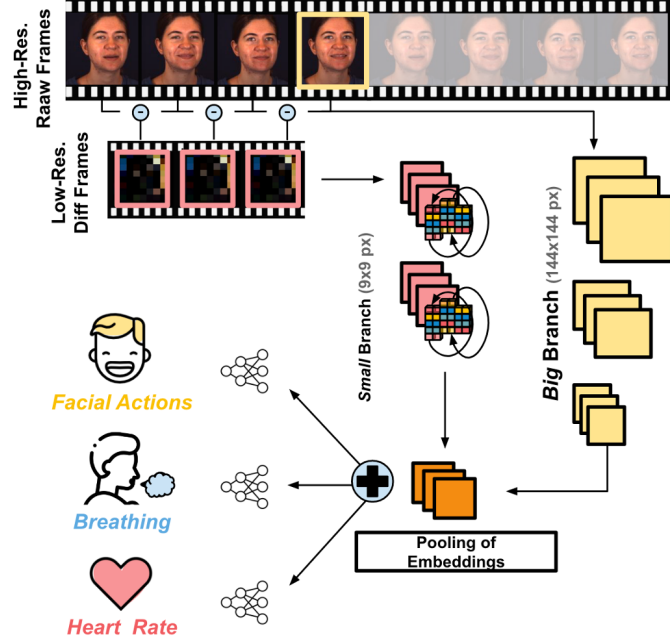


Figure 2: Architecture of BigSmall Branches

However, there are still limitations in this research. This research paper does not specify the performance improvement of a multi-task dual-branch network compared to multiple independent networks. Additionally, it does not elucidate how expressions can be categorized within its emotion recognition task. Future work may focus on quantifying the benefits of the integrated network and detailing the classification mechanism for expression recognition to provide a clearer understanding of the network’s capabilities.

3.1.6 rPPG-Toolbox: Deep Remote rPPG Toolbox

Liu et al. contributed an open-source rPPG toolbox based on deep learning methods, which includes several mainstream deep learning networks. It standardizes steps such as preprocessing, modeling, and postprocessing, providing benchmark testing for the horizontal comparison of various models[7].

This toolbox is quite good, incorporating models that are widely used and recognized in the rPPG field. However, a problem is that it is written for Linux systems, and the environment configuration on Windows is too confusing. If there were a version suitable for Windows configuration, it would be much easier to use.

3.1.7 Privacy-Phys: Facial Video-Based Physiological Modification for Privacy Protection

The extraction of remote photoplethysmography (rPPG) has become remarkably convenient with the development of computer vision technologies in recent years. In today’s era where obtaining video streaming media is convenient, the privacy protection of rPPG is crucial. In this Literature Review, the authors propose Privacy-Phys, a method based on a pre-trained 3D convolutional neural network that introduces distortions without altering the video faces, thus modifying the rPPG signal in the facial video to preserve privacy. The limitation of this study is that the privacy protection method does not modify or perturb the original rPPG signal, but rather adds a visually imperceptible rPPG signal, so the privacy protection method

still carries the risk of privacy leakage[16].

3.2 Driving Fatigue Detection

In the field of fatigue driving detection, in addition to the facial action unit detection methods focused on in early research, there are two important features: one is percentage of eyelid closure over the pupil over time(PERCLOS), and the other is heart rate signals.

Among the numerous facial action units, yawning is the most relevant indicator associated with fatigue driving. Bhandari utilized the yawning action unit to detect driver fatigue[2]. Researchers have explored a variety of methods for monitoring driver fatigue. However, among the numerous metrics, PERCLOS has been proven to be the most effective one after testing. The sole use of PERCLOS for detection is susceptible to interference from illumination, shadows, and moderate movement changes during driving in the driving environment, thus making fatigue detection based only on PERCLOS lack reliability. Integrating the rPPG method for multimodal detection is a promising approach. Multimodal fusion methods can make full use of visual information, and the specificity of the rPPG detection method means that this modality is less affected by ambient light and driver movement, providing an effective complement to the PERCLOS method.

3.2.1 Neural Network Based Luminance Variation Resistant rPPG for Driver's Heart Rate Monitoring

The rPPG method is less disruptive to drivers in a driving context compared to other contact or invasive sensors. However, research in this area has been lacking regarding outdoor driving environments, making this paper the first to focus on reducing the impact of ambient light on rPPG extraction in real-world outdoor driving conditions. In literature review [19], this study is the first to apply personalized parameter extraction for the ANN model, achieving an absolute error of 10 bpm in outdoor scenarios.

3.2.2 Near-Infrared Imaging Photoplethysmography During Driving

In this paper, Nowara et al. designed a narrowband near-infrared system for extracting the remote photoplethysmography (rPPG) signal and identified the optimal wavelength range for extracting rPPG in a driving environment. They found that the wavelength of 940nm can significantly reduce the influence of ambient light changes during driving. Furthermore, they released the first fatigue-driving dataset equipped with rPPG physiological data[11].

The narrow-band near-infrared light modality for nighttime fatigue driving detection proposed in this article is very insightful, but I believe it might be better if models with different parameters could be configured for both daytime and nighttime ambient light conditions.

3.2.3 A Multi-Stage, Multi-Feature Machine Learning Approach to Detect Driver Sleepiness in Naturalistic Road Driving Conditions

Bakker et al. integrated bimodal data, combining driver facial characteristics with vehicle road driving features. They fused facial image features with lane departure data, extracted features using deep learning methods like MTCNN (Multi-Task Cascaded Convolutional Networks) and CE-CLM (Constrained Local Model with Convolutional Experts), and employed various machine learning techniques to achieve a fatigue driving detection accuracy rate of 92.0%[1].

3.2.4 A Multimodal Fusion Fatigue Driving Detection Method Based on Heart Rate and PERCLOS

Du et al. extracted PERCLOS eye features and rPPG signals from RGB videos, and trained fatigue monitoring models for PERCLOS and rPPG signals separately using two one-dimensional CNN networks. Ultimately, they achieved enhanced robustness and accuracy by fusing the two

models through boosting ensemble methods and emphasized the need for continued research to enhance the detection framework further [5].

The PERCLOS and rPPG modalities can effectively detect driver fatigue, but the 1D-CNN used by the authors could perhaps be improved. Perhaps employing a WTSM module to reduce complexity when fusing the two modalities could be beneficial.

4 Discussion and Conclusion

Combining traditional methods of fatigue driving detection with rPPG as a multimodal detection approach is a good strategy. Sole reliance on facial action units and PERCLOS does not have strong robustness for fatigue driving detection. The rPPG modality can detect physiological information under the human face earlier than traditional methods and has certain predictive advantages. Overall, multimodal methods have good application prospects in the field of fatigue driving detection, allowing for the full utilization of multiple image sensors in the driving environment and the computational power available in vehicles. Indeed, privacy concerns associated with video sensors remain an issue that needs to be addressed. Possibly, implementing lightweight fatigue monitoring along with effective privacy-preserving methods would be a good solution to this problem. This could involve anonymizing the data, ensuring that monitoring is consensual and transparent, and exploring non-intrusive sensors that are less privacy-invasive yet provide meaningful data for fatigue detection.

References

- [1] BAKKER, B., ZABŁOCKI, B., BAKER, A., RIETHMEISTER, V., MARX, B., IYER, G., ANUND, A., AND AHLSTRÖM, C. A multi-stage, multi-feature machine learning approach to detect driver sleepiness in naturalistic road driving conditions. *IEEE Transactions on Intelligent Transportation Systems* 23, 5 (2021), 4791–4800.
- [2] BAMIDELE, A. A., KAMARDIN, K., ABD AZIZ, N. S. N., SAM, S. M., AHMED, I. S., AZIZAN, A., BANI, N. A., AND KAIDI, H. M. Non-intrusive driver drowsiness detection based on face and eye tracking. *International Journal of Advanced Computer Science and Applications* 10, 7 (2019).
- [3] CHEN, H., AND SHRIVASTAVA, A. Hr-rcnn: Hierarchical relational reasoning for object detection. *arXiv preprint arXiv:2110.13892* (2021).
- [4] CHEN, W., AND MCDUFF, D. Deepphys: Video-based physiological measurement using convolutional attention networks. In *Proceedings of the european conference on computer vision (ECCV)* (2018), pp. 349–365.
- [5] DU, G., ZHANG, L., SU, K., WANG, X., TENG, S., AND LIU, P. X. A multimodal fusion fatigue driving detection method based on heart rate and perclos. *IEEE transactions on intelligent transportation systems* 23, 11 (2022), 21810–21820.
- [6] LEWANDOWSKA, M., AND NOWAK, J. Measuring pulse rate with a webcam. *Journal of Medical Imaging and Health Informatics* 2, 1 (2012), 87–92.
- [7] LIU, X., NARAYANSWAMY, G., PARUCHURI, A., ZHANG, X., TANG, J., ZHANG, Y., SENGUPTA, R., PATEL, S., WANG, Y., AND MCDUFF, D. rppg-toolbox: Deep remote ppg toolbox. *Advances in Neural Information Processing Systems* 36 (2024).
- [8] LUO, X.-Q., HU, R., AND FAN, T.-E. The driver fatigue monitoring system based on face recognition technology. In *2013 Fourth International Conference on Intelligent Control and Information Processing (ICICIP)* (2013), IEEE, pp. 384–388.

- [9] MBOUNA, R. O., KONG, S. G., AND CHUN, M.-G. Visual analysis of eye state and head pose for driver alertness monitoring. *IEEE transactions on intelligent transportation systems* 14, 3 (2013), 1462–1469.
- [10] NARAYANSWAMY, G., LIU, Y., YANG, Y., MA, C., LIU, X., McDUFF, D., AND PATEL, S. Bigsmall: Efficient multi-task learning for disparate spatial and temporal physiological measurements. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (2024), pp. 7914–7924.
- [11] NOWARA, E. M., MARKS, T. K., MANSOUR, H., AND VEERARAGHAVAN, A. Near-infrared imaging photoplethysmography during driving. *IEEE transactions on intelligent transportation systems* 23, 4 (2020), 3589–3600.
- [12] POH, M.-Z., McDUFF, D. J., AND PICARD, R. W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics express* 18, 10 (2010), 10762–10774.
- [13] PRAKASH, S. K. A., AND TUCKER, C. S. Bounded kalman filter method for motion-robust, non-contact heart rate estimation. *Biomedical optics express* 9, 2 (2018), 873–897.
- [14] SCHRUMPF, F., FRENZEL, P., AUST, C., OSTERHOFF, G., AND FUCHS, M. Assessment of deep learning based blood pressure prediction from ppg and rppg signals. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 3820–3830.
- [15] SINGH, S., AND PAPANIKOLOPOULOS, N. P. Monitoring driver fatigue using facial analysis techniques. In *Proceedings 199 IEEE/IEEEJ/JSAP International Conference on Intelligent Transportation Systems (Cat. No. 99TH8383)* (1999), IEEE, pp. 314–318.
- [16] SUN, Z., AND LI, X. Privacy-phys: Facial video-based physiological modification for privacy protection. *IEEE Signal Processing Letters* 29 (2022), 1507–1511.
- [17] TSAI, Y.-C., LAI, P.-W., HUANG, P.-W., LIN, T.-M., AND WU, B.-F. Vision-based instant measurement system for driver fatigue monitoring. *IEEE Access* 8 (2020), 67342–67353.
- [18] VERKRUYSE, W., SVAASAND, L. O., AND NELSON, J. S. Remote plethysmographic imaging using ambient light. *Optics express* 16, 26 (2008), 21434–21445.
- [19] WU, B.-F., CHU, Y.-W., HUANG, P.-W., AND CHUNG, M.-L. Neural network based luminance variation resistant remote-photoplethysmography for driver’s heart rate monitoring. *IEEE Access* 7 (2019), 57210–57225.