A Contactless and Non-Intrusive System for Driver's Stress Detection

Muhammad Salman
Korea Institute of Energy Technology
Naju-si, Republic of Korea
salman@kentech.ac.kr

Hyunkyu Jang
Korea Institute of Energy Technology
Naju-si, Republic of Korea
hkjang@kentech.ac.kr

Youngtae Noh
Korea Institute of Energy Technology
Naju-si, Republic of Korea
ytnoh@kentech.ac.kr

Seungwan Jin
Hanyang University
Seoul, Republic of Korea
seungwanjin@hanyang.ac.kr

Dayoung Jeong
Hanyang University
Seoul, Republic of Korea
dayoungjeong@hanyang.ac.kr

Hoyoung Choi
Hanyang University
Seoul, Republic of Korea
choiho00@hanyang.ac.kr

Kyungsik Han
Hanyang University
Seoul, Republic of Korea
kyungsikhan@hanyang.ac.kr

Hyangmi Kim
Hyundai Motor Group
Hwaseong-si, Republic of Korea
kimhm@hyundai.com

ABSTRACT

Stress plays a significant role in fatal accidents, highlighting the importance of timely monitoring of driver stress to facilitate effective interventions and reduce road accidents. However, monitoring driver stress presents numerous challenges in the context of driving. First, state-of-the-art techniques such as self-stress evaluation and periodic cortisol level checks are not suitable for the driving scenario. Second, existing unimodal solutions does not provide a comprehensive and holistic assessment of the driver's stress. Although some research utilizes multimodal features, the use of wearables attached to the driver's body in real-life situations is impractical and highly discomforting. Our proposed solution tackles these challenges by offering a contactless and non-intrusive approach that prioritizes the driver's comfort during the collection of multimodal data, which includes capturing heart rate variability (HRV), respiration rate, and microfacial expressions. Through feature-level data fusion, we combine and integrate these diverse modalities to generate comprehensive insights. These insights are then utilized by the multimodal learning pipeline to predict the driver's stress levels in real driving scenarios.

CCS CONCEPTS

• **Human-centered computing → HCI theory, concepts and models**.

KEYWORDS

Vital Signs, mmWave Radar, Heart Rate Variability (HRV), Driver's Stress, multimodal learning

UbiComp/ISWC '23 Adjunct, October 08–12, 2023, Cancun, Quintana Roo, Mexico

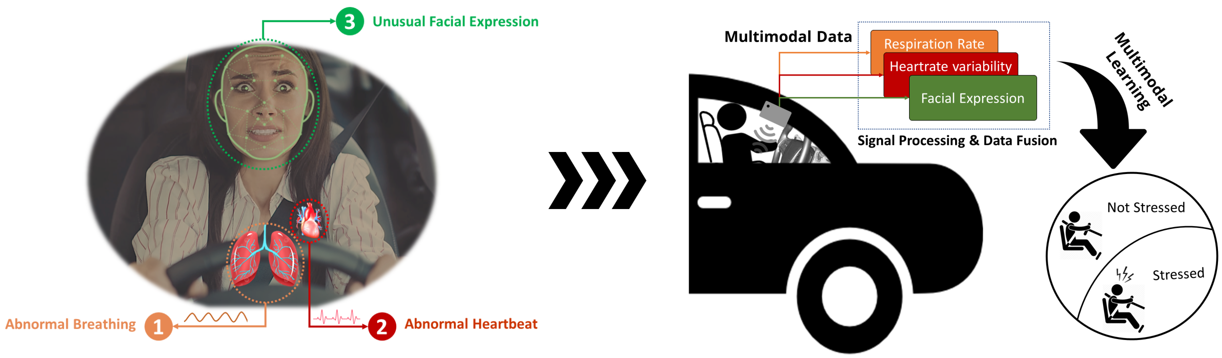
ACM ISBN 979-8-4007-0200-6/23/10.
https://doi.org/10.1145/3594739.3610691

ACM Reference Format:
Muhammad Salman, Hyunkyu Jang, Youngtae Noh, Seungwan Jin, Dayoung Jeong, Hoyoung Choi, Kyungsik Han, and Hyangmi Kim. 2023. A Contactless and Non-Intrusive System for Driver's Stress Detection. In *Adjunct Proceedings of the 2023 ACM International Joint Conference on Pervasive and Ubiquitous Computing & the 2023 ACM International Symposium on Wearable Computing (UbiComp/ISWC '23 Adjunct), October 08–12, 2023, Cancun, Quintana Roo, Mexico.* ACM, New York, NY, USA, Article 4, 5 pages. https://doi.org/10.1145/3594739.3610691

1 INTRODUCTION

According to a road safety study by the World Health Organization (WHO), traffic-related accidents result in over 1.3 million fatalities annually [17]. Negative stress (hereinafter, stress) has been identified as a significant factor contributing to frequent road accidents [18]. Research from the Virginia Tech Transportation Institute reveals that stress can increase the chance of an accident by nearly 10 times [2]. Similarly, Australian national accident reports indicate that stressed condition is an important contributor to fatal car accidents [1]. Another study assessing the European Commission's data estimates the economic cost of car accidents in Europe to be around 160 billion euros, with a significant portion attributed to the psychophysical health of drivers [21]. In summary, stress impairs cognitive abilities and compromises driving performance, thereby increasing the risk of road accidents [20]. The precise diagnosis of driver stress and the implementation of effective interventions have the potential to significantly reduce the rate of road accidents. These proactive measures serve to enhance road safety and prevent accidents from occurring.

However, measuring the driver's stress is a challenging task for two main reasons. Firstly, the existing methods are not suitable for the driving scenario. Self-stress reporting via smartphone applications, although widely used for general cases, can be intrusive for a driver [10]. Similarly, while detecting cortisol levels using saliva samples is a proven technique for stress identification, however, it is impractical for driving scenarios [4]. Secondly, the current methods for assessing driver stress have limitations either in terms



(a) Notable signs of stress in a driver.

(c) Multimodal learning pipeline.

Figure 1: (a) Signs of a stressed driver; (b) System model for contactless stress detection.

of being unimodal or uncomfortable. Unimodal methods, such as detecting stress through the examination of dilated pupils [15] or analyzing facial expressions using the valence and arousal plane [19], are subjective and lack detailed stress-related information. They may not provide a comprehensive understanding of the driver's stress levels. On the other hand, multimodal data obtained from wearables that capture physiological indicators like heart-rate variability, respiration rate, and skin conductivity offer potential for stress diagnosis [9, 14]. However, these wearables are often considered obtrusive and uncomfortable, which may affect the driver's comfort and willingness to use them consistently.

In this paper, we present a novel approach for detecting driver stress that is non-intrusive, contactless, and designed to ensure the driver's comfort. Our method aims to reliably detect and assess the driver's stress levels using multimodal features collected in real-world driving scenarios. To achieve this, we leverage mmWave radar to non-invasively detect vital signs such as heart-rate variability (HRV) and breathing (or respiration) rate. Additionally, we utilize a commodity camera to capture micro-facial expressions of the driver. By combining and analyzing the stress features from both modalities through data fusion techniques, we derive a binary decision indicating the user's stress level.

2 PROMINENT SIGNS OF STRESS

Stress is a physiological and psychological response to a mixture of internal worries and external demands. Although different people perceive stress differently, however, their response to a stress condition referred to as the "fight-or-flight" is quite similar [13]. During the fight-or-flight response, stress hormones such as cortisol and adrenaline are released from the adrenal glands. This process begins when the brain detects stress, leading the hypothalamus to release hormones that stimulate the pituitary gland. The pituitary gland, in turn, releases hormones that reach the adrenal glands. Consequently, the adrenal cortex of adrenal glands releases cortisol [8]. Simultaneously, the sympathetic nervous system detects cognitive load (due to stress) and activates the adrenal glands to release adrenaline (and noradrenaline) hormones [7]. These hormones collaborate to elevate heart rate, respiration rate, and can manifest as noticeable emotional changes, including facial expressions. In our stress prediction approach, we utilize these physiological responses exhibited by the driver[1], as illustrated in Figure 1(a).

3 DEMONSTRATION SYSTEM

Our driver stress detection system consists of three components: multimodal data acquisition, signal processing and data fusion, and a multimodal learning pipeline. These components work together to gather comprehensive driver data, integrate valuable information from different sources, and derive insights related to driver stress. We describe each component below. The testbed setup is shown in Figure 2.

3.1 Multimodal Data Acquisition

The stress-related data collection system includes two main capturing devices: mmWave Radar and a commodity camera. The mmWave Radar captures the driver's heart and breathing rates, while the commodity camera captures facial expressions. In the following sections, we will provide detailed descriptions of each data-capturing device.

Data captured by mmWave radar. To capture the user's vital signs in a contactless manner[16], we utilize the IWR1642 mmWave sensor (or radar) from Texas Instruments (TI), operating at a frequency range of 77 ∼ 81 GHz. This radar system consists of four transmitting and two receiving antennas. To handle the high data transmission rate (approximately 600 Mbps in our case), we employ the DCA1000EVM high-speed transmission adapter by TI. It uses the LVDS data interface to efficiently collect the data from the IWR1642 radar and transport it to the computer via an Ethernet interface. The transferred raw data is saved in binary format on the PC. Since the device does not come with a built-in cover, therefore

[1]Please be aware that while there are additional symptoms such as dilated pupils [15], sweating, and suppressed digestion, they are either less prominent or necessitate long-term medical monitoring. In this research, we concentrate on significant observable stress signs that can be monitored without causing inconvenience to the user.

Figure 2: Car testbed setup: (a) mmWave device linked to laptop via Ethernet for vital sign capture; (b) Front glass-mounted video camera for facial expression capture; and (c) Polar-H10 device for acceleration-based heart rate and breathing capture.

we have created a 3D-printed case to ensure device protection and proper holding in the vehicle[2].

Data captured by camera. For capturing the driver's facial expressions, we utilize a standard commodity camera positioned to record the driver's field of view, ensuring clear visibility of their face from a front-facing angle. To optimize resource usage, we deliberately choose to record in standard definition (480p), even though higher resolutions were supported. The recording is set at a frame rate of 30 frames per second (fps) to maintain smoothness and accuracy in capturing the driver's facial expressions.

3.2 Signal Processing and Data Fusion

The raw data obtained from the mmWave radar requires signal processing to extract precise vital signs, while the video footage necessitates processing for emotion detection and classification. Since the data from the radar and camera differ significantly, data fusion is employed to create a comprehensive and accurate representation of the combined data. The following section provides a detailed explanation of this process.

mmWave Signal Processing. The raw data acquired from the TI mmWave radar consists of ADC samples representing the Intermediate Frequency (IF), which is obtained by calculating the difference between the transmitted and received FMCW chirp signals. To process and analyze the IF data from the binary file effectively, we need to organize it in a properly structured format. Initially, we separate the data for each received channel. Then, for each channel, we arrange the data by grouping the chirps per frame[3]. Once the data is arranged, we focus on tracking the specific chirp or range-bin associated with minute movements such as breathing and heartbeat, while discarding the irrelevant bins. This is accomplished by performing a Fast Fourier Transform (FFT) on the time-domain

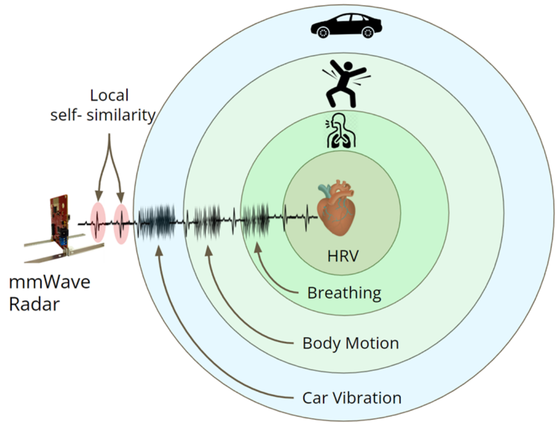


Figure 3: Challenges in measuring heart rate.

data and selecting the chirp with the highest magnitude variations. We particularly leverage the phase information of that specific bin, as it has shown effectiveness in capturing subtle movements like heartbeat motion, as expressed in the following equation:

$$\phi(t) = 2\pi \frac{d(t)}{\lambda} \tag{1}$$

where $d(t)$ represents the sensor's distance from the human body (i.e., chest). This distance varies over time as a result of the chest vibrations induced by respiration and heartbeat.

By analyzing the phase information, we extract the breathing and heart rate (HR). The process of obtaining the breathing rate is relatively straightforward compared to the heart rate. For the breathing rate, we apply a bandpass elliptic filter with frequency limits of 0.12 Hz to 0.6 Hz, as described in [11]. This range effectively captures breathing rates ranging from 7 to 36 breaths per minute. In contrast, the detection of HR presents two main challenges, as illustrated in Figure 3. The first challenge arises from the motion resulting from breathing, which can effectively mask the minute vibrations of the heartbeat. To address this, we employ a second-order differential FIR filter to convert the distance information into acceleration. This allows us to distinguish the distinct patterns of heartbeats from breathing.

The second challenge stems from the motion originating from both the human body and car vibrations, which significantly contaminate the HR signals. To overcome this challenge, we utilize a 1D Convolutional Neural Network (CNN) to accurately identify the recurring peaks that correspond to the heart rate. This is achieved by leveraging the local similarity property, where the motion caused by the target individual and car vibrations is considered as noise and discarded. The 1D CNN computes the HR based on the identified recurrent peaks, allowing for an accurate assessment of the heart rate variability.

Emotion detection for facial expression. Previous research has demonstrated that facial expressions can reveal emotions related to stress [6]. However, there are two main challenges in detecting the driver's stress-related emotions. Firstly, there is ambiguity in expressions, for instance, emotions like "fear" can be similar to "disgust" or "surprise". Secondly, cultural biases play a role, as the

[2]This allows us to properly attach the device to the mirror or the front glass using a holder.
[3]Note that each chirp contains multiple ADC samples.

expression and interpretation of emotions can vary across different cultures. To address the first challenge, we can simplify the classification by categorizing emotions into positive and negative. Positive emotions like smiles, satisfaction, and laughter can be associated with non-stress situations, while negative emotions like anger, disgust, sadness, and fear can be classified as signs of stress. As for the second challenge, while it is indeed difficult, our initial focus is to train the model on diverse datasets from Asian regions such as [3, 12] to ensure accuracy for local regions. In the future, we plan to expand the model's training to include other regions as well.

Data Fusion. As aforementioned, we have data from different modalities, including mmWave data for vital signs features (such as heart rate variability and breathing rate) and video data for extracting facial expressions of the driver during driving. To integrate these modalities effectively, we plan to employ feature-level fusion. This approach involves extracting relevant features from each modality and then combining them to create a joint feature representation. The fused feature representation will be utilized in the subsequent phase of multimodal learning, enabling a more comprehensive analysis of the data.

3.3 Multimodal data learning

The fused feature representation is fed into a deep neural network model for multimodal learning, which is trained on a labeled dataset with ground truth annotations for stress and non-stress instances. Extensive model evaluation will be conducted, and the model parameters will be optimized through techniques like fine-tuning hyperparameters, applying regularization, and exploring ensemble methods. Once the model achieves satisfactory performance, it will be deployed in the wild for stress detection, where new data instances can be classified as stress or non-stress, providing valuable insights for stress monitoring in driving scenarios.

4 DISCUSSION AND FUTURE WORK

In this paper, we introduce a prototype system for non-contact measurement of driver stress to assess it in real-time and offer interventions to prevent potential road accidents. Our system utilizes off-the-shelf commodity devices, including mmWave radar for capturing vital signs and a camera for recording facial expressions displayed by the driver. Through data fusion and multimodal training, our system accurately determines the driver's stress state. It is important to acknowledge that the experiment involves human subjects, and as such, we intend to seek approval from the Institutional Review Board (IRB).

In the subsequent parts, we will address the limitations and challenges associated with our system, as well as outline our future plans for overcoming them.

Impact of the road surface. In scenarios where the vehicle is traveling on a rough surface road, the significant vibrations present can have a detrimental effect on the accuracy of vital sign measurements obtained using the mmWave device. These vibrations are external factors that are beyond our control and can potentially result in false positives in stress detection. To tackle this challenge, our future work will concentrate on implementing a shock-absorbent cover, such as a cover with dampers, for the radar device. By employing this measure, we aim to minimize the disruptive effects of vibrations and enhance the reliability of vital sign measurements.

Vibration from the Car. In certain instances, the mmWave device may vibrate as a result of car vibrations, which can potentially disrupt vital sign measurements due to the device's higher intensity of vibration compared to the vital signs themselves. To enhance system resilience to car vibrations and improve efficiency in detecting vital signs, we aim to adopt the concept of Self-Similarity[4], e.g., [5].

Driver's Stress Inducing Scenarios. Stress can arise from internal concerns, such as negative feedback from a boss, as well as external factors like heavy traffic or unfavorable road conditions. Designing appropriate scenarios for data collection is challenging, and we plan to carefully craft scenarios that induce both types of stress in the driver.

Impact of driver's clothes. Thanks to the efficient climate control systems in modern vehicles, drivers often prefer to wear comfortable clothing during long drives. However, in certain scenarios where drivers wear heavy and thick outfits, such as thick leather jackets with metallic tags, it can partially obscure the detection of vital signs. Addressing this challenge remains a focus for future research and development.

Impact of body wearing accessories. Some drivers, especially females, wear accessories like metallic ear tops that can introduce additional contamination to the mmWave signals. The vibrations caused by these accessories can be filtered out if they have a non-repetitive pattern. However, if their vibrations exhibit a periodic pattern, it may lead to errors in vital signs measurement. Addressing this issue is part of our future work.

ACKNOWLEDGMENTS

This study was supported by Hyundai Motor Group under grant number T004600122120023 as part of the Academy-Industry Joint Research initiative and the National Research Foundation of Korea (NRF) under grant number NRF-2021M3A9E4080780, funded by the Ministry of Education.

REFERENCES

[1] Vanessa Beanland, Michael Fitzharris, Kristie L Young, and Michael G Lenné. 2013. Driver inattention and driver distraction in serious casualty crashes: Data from the Australian National Crash In-depth Study. *Accident Analysis & Prevention* 54 (2013), 99–107.

[2] Thomas G Brown, Marie Claude Ouimet, Manal Eldeb, Jacques Tremblay, Evelyn Vingilis, Louise Nadeau, Jens Pruessner, and Antoine Bechara. 2016. Personality, executive control, and neurobiological characteristics associated with different forms of risky driving. *PloS one* 11, 2 (2016), e0150227.

[3] Haoyu Chen, Henglin Shi, Xin Liu, Xiaobai Li, and Guoying Zhao. 2023. SMG: A Micro-gesture Dataset Towards Spontaneous Body Gestures for Emotional Stress State Analysis. *International Journal of Computer Vision* 131, 6 (2023), 1346–1366.

[4] Sheldon Cohen, Ronald C Kessler, and Lynn Underwood Gordon. 1997. *Measuring stress: A guide for health and social scientists.* Oxford University Press on Demand.

[5] Debidatta Dwibedi, Yusuf Aytar, Jonathan Tompson, Pierre Sermanet, and Andrew Zisserman. 2020. Counting out time: Class agnostic video repetition counting in the wild. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.* 10387–10396.

[6] Hua Gao, Anil Yüce, and Jean-Philippe Thiran. 2014. Detecting emotional stress from facial expressions for driving safety. In *2014 IEEE International Conference on Image Processing (ICIP).* 5961–5965. https://doi.org/10.1109/ICIP.2014.7026203

[4]For finding the local similarity property of heart rate in mmWave, which would automatically discard the corrupted data segments belonging to vibrations.

[7] David S Goldstein. 1987. Stress-induced activation of the sympathetic nervous system. *Bailliere's clinical endocrinology and metabolism* 1, 2 (1987), 253–278.
[8] Emma R Jakoi. 2004. Hypothalamus and pituitary gland. (2004).
[9] Zachary D King, Judith Moskowitz, Begum Egilmez, Shibo Zhang, Lida Zhang, Michael Bass, John Rogers, Roozbeh Ghaffari, Laurie Wakschlag, and Nabil Alshurafa. 2019. Micro-stress EMA: A passive sensing framework for detecting in-the-wild stress in pregnant mothers. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 3 (2019), 1–22.
[10] Susan Levenstein, Cosimo Prantera, Vilma Varvo, Maria L Scribano, Eva Berto, Carlo Luzi, and Arnaldo Andreoli. 1993. Development of the Perceived Stress Questionnaire: a new tool for psychosomatic research. *Journal of psychosomatic research* 37, 1 (1993), 19–32.
[11] Gen Li, Yun Ge, Yiyu Wang, Qingwu Chen, and Gang Wang. 2022. Detection of Human Breathing in Non-Line-of-Sight Region by Using mmWave FMCW Radar. *IEEE Transactions on Instrumentation and Measurement* 71 (2022), 1–11. https://doi.org/10.1109/TIM.2022.3208266
[12] Xin Liu, Henglin Shi, Haoyu Chen, Zitong Yu, Xiaobai Li, and Guoying Zhao. 2021. iMiGUE: An identity-free video dataset for micro-gesture understanding and emotion analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10631–10642.
[13] R McCarty. 2016. The fight-or-flight response: A cornerstone of stress research. In *Stress: Concepts, cognition, emotion, and behavior*. Elsevier, 33–37.
[14] Varun Mishra, Sougata Sen, Grace Chen, Tian Hao, Jeffrey Rogers, Ching-Hua Chen, and David Kotz. 2020. Evaluating the reproducibility of physiological stress detection models. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 4, 4 (2020), 1–29.
[15] Marco Pedrotti, Mohammad Ali Mirzaei, Adrien Tedesco, Jean-Rémy Chardonnet, Frédéric Mérienne, Simone Benedetto, and Thierry Baccino. 2014. Automatic stress classification with pupil diameter analysis. *International Journal of Human-Computer Interaction* 30, 3 (2014), 220–236.
[16] Muhammad Salman and Youngtae Noh. 2023. Contactless Vital Signs Tracking with mmWave RADAR in Realtime. In *2023 IEEE International Conference on Big Data and Smart Computing (BigComp)*. IEEE, 389–390.
[17] Violet Anne Sauerzapf. 2012. *Road traffic crash fatalities: An examination of national fatality rates and factors associated with the variation in fatality rates between nations with reference to the World Health Organisation Decade of Action for Road Safety 2011-2020.* Ph. D. Dissertation. University of East Anglia.
[18] Dawid Konrad Ścigała and Elżbieta Zdankiewicz-Ścigała. 2019. The role in road traffic accident and anxiety as moderators attention biases in modified emotional stroop test. *Frontiers in psychology* 10 (2019), 1575.
[19] Thi-Dung Tran, Junghee Kim, Ngoc-Huynh Ho, Hyung-Jeong Yang, Sudarshan Pant, Soo-Hyung Kim, and Guee-Sang Lee. 2021. Stress analysis with dimensions of valence and arousal in the wild. *Applied Sciences* 11, 11 (2021), 5194.
[20] Sergio A Useche, Viviola Gómez Ortiz, and Boris E Cendales. 2017. Stress-related psychosocial factors at work, fatigue, and risky driving behavior in bus rapid transport (BRT) drivers. *Accident Analysis & Prevention* 104 (2017), 106–114.
[21] Roberto Vivoli, Margherita Bergomi, Sergio Rovesti, Pamela Bussetti, GM Guaitoli, et al. 2006. Biological and behavioral factors affecting driving safety. *Journal of preventive medicine and hygiene* 47, 2 (2006), 69–73.