ADVANCING SAFETY IN ROADWAY WORK ZONES WITH WORKER-CENTRED AUGMENTED REALITY: ASSESSING THE FEASIBILITY, USABILITY, AND EFFECTIVENESS OF AR-ENABLED WARNING SYSTEMS

by

Sepehr Sabeti

A dissertation submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Infrastructure and Environmental Systems

Charlotte

2023

Approved by:

____

Dr. Omidreza Shoghli

____

Dr. Hamed Tabkhi

____

Dr. Nichole Morris

____

Dr. Zachary Wartell

____

Dr. Mona Azarbayjani

ABSTRACT

SEPEHR SABETI. Advancing Safety in Roadway Work Zones with Worker-Centred Augmented Reality: Assessing the Feasibility, Usability, and Effectiveness of AR-Enabled Warning Systems. (Under the direction of DR. OMIDREZA SHOGHLI)

In recent years, there has been a concerning increase in serious injuries and fatalities within highway work zones, deviating from a previously declining trend. While innovative technologies offer potential for improving safety, the specific needs of the highway work zone community have often been overlooked, resulting in a lack of essential information for designing user-centered safety systems. Current safety practices primarily rely on reactive technologies that may not provide timely responses to intrusions and environmental risks. To bridge this gap, there is a need for transformative technologies that integrate real-time predictive safety systems and extensive user research to develop worker-centered Augmented Reality (AR)-based safety systems specifically tailored for highway work zone safety. This dissertation aims to address these challenges and outlines our efforts in designing such a worker-centered AR-based safety system.

Chapter 2 explores the feasibility, requirements, and challenges associated with integrating AI capabilities into AR systems to enhance the safety of highway work zones. This chapter delves into the feasibility, requirements, and challenges associated with incorporating AI capabilities into the AR system to develop a predictive safety system that can proactively identify potential hazards and issue timely warnings to workers. The outcomes of this chapter indicate that the real-time communication latency and AI execution latency meet the tight timing constraints of a real-time safety system. The early user research demonstrates positive reception and acceptance of the proposed safety framework and interface by highway maintenance and operation professionals across multiple states in the US.

Chapter 3 focuses on conducting a mixed-method usability investigation of the proposed AR-based safety system using a high-fidelity prototype. The investigation assesses aspects such as user interface design, interaction patterns, and user feedback to evaluate the overall usability and effectiveness of the technology in enhancing roadway work zone safety. The findings indicate that participants rated the usability of the system above average in both indoor and outdoor settings and perceived a reasonable level of mental effort. Perceived trust was found to be significantly correlated with usability, underscoring its importance in user experience.

Chapter 4 examines the impact of different sensory modalities on worker reaction times in augmented reality warnings within roadway work zones. The analysis of data from experiments provides insights into the effectiveness of various warning modalities, including visual, audiovisual, haptic visual, and combined haptic audiovisual cues, in improving worker reaction times. The findings indicate that the haptic visual design triggered the fastest response on average among the participants, and its performance was statistically comparable to that of the audio haptic visual design. Furthermore, both of these designs demonstrated significantly faster reaction times compared to visual and audiovisual warnings. The results also indicate that reaction times to augmented reality warnings in real-world outdoor scenarios were generally longer and exhibited greater variability compared to baseline desktop warnings and simulated AR in virtual reality. Surprisingly, VR simulated warnings did not show statistically significant shorter reaction times compared to their real-world counterparts. These observations suggest that simulating AR in virtual reality may not accurately replicate the reaction times observed in real-world situations.

Collectively, the results from these chapters demonstrate the usability, perceived safety benefits, and potential for timely notifications offered by the proposed AR-based safety system. This research also contributes to establishing best practices for designing time-sensitive safety systems, prioritizing situational awareness, and

implementing worker-centered design principles in AR safety systems. Ultimately, the findings have the potential to significantly enhance the safety of highway workers and the broader workforce operating in roadway work zones.

DEDICATION

This dissertation is dedicated to my beloved wife and family, with infinite love, gratitude, and admiration. Throughout this challenging journey, they have been my constant source of support, inspiration, and strength.

To my dear wife, Yeganeh, your belief in my abilities, your patience during the long hours spent on research, and your consistent encouragement have been my guiding light. Your unlimited love and understanding have supported me during the most difficult moments, and your constant presence has been my greatest motivation.

To my loving family, your constant support, encouragement and understanding have formed the foundation of my achievements. Your selfless sacrifices, both large and small, have allowed me to pursue this academic endeavor. Your enduring faith in me has instilled the confidence and determination needed to overcome obstacles and reach this significant milestone.

May this dedication serve as a humble expression of my deepest gratitude and love for each of you. You are the pillars of my life and I am eternally thankful for your presence and continued support.

ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who contributed to the completion of this doctoral dissertation. Their guidance, support, and encouragement have been priceless throughout this journey. I am deeply indebted to the following individuals and organizations:

My supervisor, Dr. Omidreza Shoghli, whose expertise, wisdom, and belief in me have shaped this research work. His insightful comments and constructive criticism have been remarkable in driving this project forward.

The members of my dissertation committee, Dr. Hamed Tabkhi, Dr. Nichole Morris, Dr. Zachary Wartell, and Dr. Mona Azarbayejani, for their invaluable feedback, suggestions, and scholarly insights. Your collective experience and thoughtful inputs have greatly enriched this research and helped me refine my ideas.

I would like to express my appreciation for the financial and intellectual support provided by multiple organizations throughout the course of this research. I wish to acknowledge the contributions of the National Science Foundation, Virginia Department of Transportation, Leidos, and the UNC Charlotte INES program and Graduate School. Their support has been significant in enabling the execution of this research project.

I am and will be forever indebted to my family for their love, encouragement, and understanding. Your belief in my abilities and continuous support, both emotionally and practically, have been my pillars of strength throughout this challenging endeavor.

I would like to sincerely thank Dr. Adrian Burde for his invaluable support and assistance throughout my doctoral journey. I also would like to express my deep appreciation to Dr. Jay Wu, Monica Rasmussen, current and previous members of the Smart Infrastructure Asset Management (SIAM) research lab, and other individuals who have been outstanding in supporting me throughout this journey. I am truly grateful for their contributions.

I would like to acknowledge the immensely significant role of participants in the human subject studies of this research, particularly the Minnesota Department of Transportation, who generously dedicated their time and experiences to this research. Without their invaluable contributions, this study would not have been possible.

Although the list above highlights the significant contributors, I extend my gratitude to everyone who has played a part, no matter how big or small, in the completion of this PhD dissertation. Your support has been invaluable, and I am grateful for your contributions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI Artificial Intelligence

AR Augmented Reality

ARROWS Augmented Reality for Roadway Work Zone Safety

AV Audio Visual

BLS Basic Life Support

BLS Bureau of Labor Statistics

CDC Centers for Disease Control and Prevention

CGS Curved Grading Scale

CNN Convolutional Neural Networks

CRT Choice Reaction Time

CSP Cross Stage Partial

DOT Department of Transportation

FHWA Federal Highway Administration

FPN Feature Pyramid Networks

FPS Frame Per Second

GPU Graphical Processing Unit

HAV Haptic Audio Visual

HCI Human Computer Interaction

HMD Head Mounted Display

HV Haptic Visual

IoT Internet of Things

IP Internet Protocol

IQR Interquartile Range

IRB Internal Review Board

MnDOT Minnesota Department of Transportation

MS Millisecond

MUTCD Manual on Uniform Traffic Control Devices

RSME Rating Scale Mental Effort

RT Reaction Time

RTT Round Trip Time

SD Standard Deviation

SRT Simple Reaction Time

SUS System Usability Scale

TAM Technology Acceptance Model

TCP Transfer Control Protocol

UI User Interface

V Visual

VR Virtual Reality

VR-WOT Virtual Reality Without Traffic scenario

VR-WT Virtual Reality With Traffic scenario

WOZ Wizard of Oz Methodology

WT Wearable Technology

CHAPTER 1: INTRODUCTION

Highway work zones are vital for inspecting, maintaining and upgrading transportation infrastructure. However, a hazardous combination of factors, including speeding and careless driving, night shifts, and limited maneuvering space, poses significant risks to the safety of workers. The Centers for Disease Control and Prevention reports that between 2003 and 2017, 1844 workers lost their lives at road construction sites, averaging 123 fatalities per year [1]. These incidents not only result in death, but also cause severe injuries and mental or physical health issues for many workers due to intrusions by drivers. Current safety practices in highway work zones primarily rely on basic measures such as portable signs, flaggers, alarms, lights, and rudimentary intrusion alert systems. Unfortunately, these reactive technologies are activated after or just before an intrusion occurs, providing insufficient reaction time for workers to respond adequately to imminent dangers [2, 3]. As a result, workers remain vulnerable to intrusions and their associated risks.

Recent advances in Augmented Reality (AR) have presented unique opportunities to address safety challenges in various domains, including the construction industry [4, 5, 6]. However, there exists a disparity in the research landscape, with the highway work zone community being underrepresented, limiting the availability of crucial information for the development of user-centered system designs [2, 7, 8]. This lack of focus on the highway work zone community hinders the progress of innovative solutions with potential impacts on acceptance models and future applications [9, 10]. Therefore, it is imperative to address this gap and recognize the importance of studying the community of the highway work zone to drive advances in safety technologies and practices.

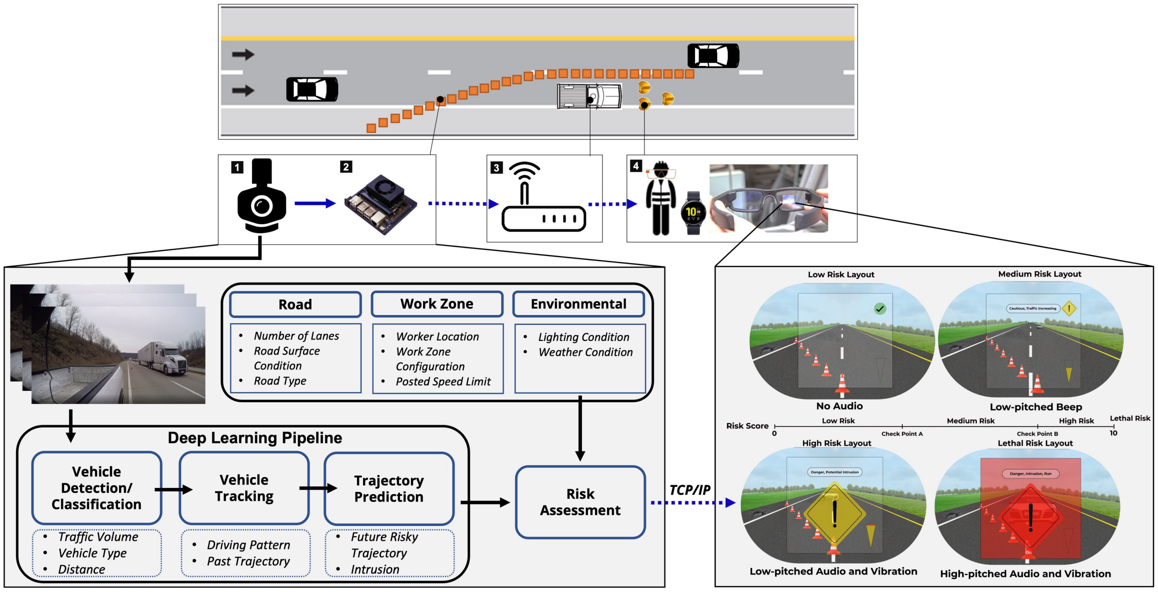


Figure 1.1: System design of the proposed technology and its major components

This dissertation documents our comprehensive efforts in conceptualizing, designing, prototyping, and researching an innovative AI-enabled Augmented Reality safety system specifically tailored for highway work zone safety. The proposed system architecture is depicted in Figure 1.1. In this context, the term "worker" encompasses various individuals involved in highway work zones, such as road inspectors, law enforcement personnel, and first responders. The backend of the system incorporates an AI bundle comprising a camera and an embedded GPU, responsible for processing real-time traffic feeds and detecting potential intrusions or hazardous maneuvers in advance and from a distance. In the front-end, the system comprises two components: AR smart glasses and a smartwatch. Each worker is equipped with this package, which facilitates a real-time multimodal notification mechanism designed to effectively alert workers within the challenging and noisy environment of highway work zones.

The subsequent sections of this dissertation provide a detailed account of our comprehensive efforts in designing, developing, and researching the AR safety system for highway work zone safety. The objectives outlined in this document guide our exploration and investigation of various aspects related to the proposed technology.

The first objective, as detailed in Chapter 2, focuses on studying the feasibility and addressing the technical challenges associated with implementing the AI-enabled AR/WT safety system. We delved into the intricate details of the system's architecture and consider its practical implementation. This chapter provides a holistic view of the framework, examining its technical aspects, potential limitations, and the steps taken to mitigate any challenges that arise. In this chapter, our objectives were:

- Defining the functional and technical requirements of an AI-enabled Augmented Reality (AR) system to enhance the safety of highway workers.
- Developing a holistic design framework that integrates real-time AI processing and edge communication with AR user interface design to provide multimodal notifications to highway workers.
- Demonstrating the feasibility of AI-enabled worker-in-the-loop technologies in improving the safety of highway work zones through a proof-of-concept model.

In Chapter 3 of this dissertation, our focus shifts to user research and usability testing, aiming to quantify the usability of the AI-enabled AR safety system while documenting user experience benchmarks. To facilitate this evaluation, we developed a high-fidelity prototype of the technology, which serves as a crucial tool for conducting a comprehensive usability study. The usability study was administered in both indoor and outdoor settings to provide a comprehensive understanding of the system performance. Adopting a mixed-method approach, we delved into the perspectives of end-users and seek insights and feedback regarding the system's usability. Specifically, we collaborate with 13 participants who are members of the Minnesota Department of Transportation (MnDOT), ensuring that their expertise and experiences are effectively captured. To further assess the usability of the technology, we conducted a thorough experiment in a temporary work zone environment. This environment mimicked the real-world conditions encountered during the maintenance and operation of highway work zones. Participants were assigned specific tasks that are commonly performed in such scenarios, allowing us to evaluate their interactions with the system and collect valuable data on usability, efficiency, and user experience. To measure usability, we employed the System Usability Scale (SUS), a widely recognized tool for evaluating the usability of interactive systems. Additionally, we utilized trust measurements to assess user perception and confidence in the system's performance and reliability. These measures provide us with valuable quantitative data to gauge the effectiveness of the AI-enabled AR and WT safety system and establish benchmarks for future comparisons and improvements. In specific, our research objectives were:

- Evaluating the usability of the proposed AR system by incorporating the perspectives of highway workers in both indoor and outdoor settings.
- Establishing usability benchmarks and identifying factors that contribute to the user experience, informing the development of future AR solutions for the highway workforce.
- Gathering in-depth feedback from experienced highway workers regarding their use of the proposed technology.

Chapter 4 of this dissertation delves into the quantification of reaction times to multimodal AR warnings within the context of highway work zone safety. Our study adopts a comprehensive mixed-method research framework that includes a high-fidelity prototype of the AR system, virtual reality simulations to emulate real-world scenarios, and the Wizard of Oz methodology for synchronized user journeys during experiments. To assess reaction times, we employed two distinct methods: the simple reaction time (SRT) approach and a vision-based metric that leverages real-time pose estimation and upper body joint displacement. The primary objective of our research was to comprehensively evaluate the impact of multimodal design on users' reaction times across various environments. Furthermore, we aimed to investigate the potential of VR simulations in replicating real-world settings by comparing reaction times obtained from both VR and real-world scenarios. By subjecting users to various combinations of these stimuli, we can analyze their reaction times and evaluate the system's effectiveness in delivering timely and appropriate warnings.

- Evaluating the reaction time to different multimodal AR warning designs and determining which design triggers the fastest reaction time.

- Comparing the reaction times between VR-simulated AR warnings and real-world AR warnings in an outdoor environment to determine if they show statistically similar response times.

- Investigating the feasibility of using real-time pose estimation as an indicator of the reaction time to AR warnings.

Throughout this document, our research journey unfolds systematically, addressing different aspects and stages of the project. By undertaking these investigations, we aim to contribute to the field of highway work zone safety and lay the foundation for future advancements in AI-enabled AR applications. Our work strives to improve the safety and well-being of workers in these critical environments, promoting innovation and driving progress in the domain.

CHAPTER 2: TOWARD AI-ENABLED AUGMENTED REALITY TO ENHANCE THE SAFETY OF HIGHWAY WORK ZONES: FEASIBILITY, REQUIREMENTS, AND CHALLENGES

2.1 Abstract

Highway work zones are considered among the most hazardous working environments. In 2018 alone, 124 workers lost their lives to fatal accidents. The lack of predictive safety systems that notify workers of upcoming dangers in advance is a major reason to blame in the highway maintenance and operation community. This article presents an integrative design framework for bringing recent advances in Augmented Reality (AR) and Artificial Intelligence (AI) to enhance the safety of highway workers through real-time multimodal notifications on-spot. To this end, this article conceptualizes and co-designs three major pillars: (1) AR user interface design for multimodal notification, (2) real-time AI at the edge for vehicle detection/ classification from distance, and (3) real-time wireless communication in work zone setting to enable latency-aware operation between AI and AR components. Our early results demonstrate that we can achieve 24.83 FPS end-to-end execution latency on the Xavier AGX Jetson board with 48.7% on BDD100K dataset, and a real-time communication covering 120 meters with an average latency of 5.1 milliseconds at the farthest distance. Our mixed-method user research also reveals an acceptable level of excitement and engagement from the body of highway workers toward both the proposed technology and the designed user interface. Overall, this article provides a proof-of-concept toward AI-enabled AR safety systems in highway work zones.

2.2 Introduction

Highway work zones are considered among the most hazardous working environments. In 2018 alone, Federal Highway Administration (FHWA) reported that 124 workers lost their lives at road construction sites [11]. In the meantime, previous studies documented that the annual average fatality rate of highway workers between 2016 and 2018 was around 135. Moreover, a total of 158,000 crashes and 42,000 corresponding injuries were reported in work zones in 2016. FHWA also highlighted that every 15 hours, one fatality in highway work zones took place in the US [11]. With the growing potential of massive investments in infrastructure in the coming years, it can be only assumed that highway workers will be even more exposed to these fatal risks due to the forthcoming increase in the number of work zones. Therefore, securing safety of highway work zones is one of the most pressing challenges that the highway construction, maintenance ,and operation community is facing [2, 12].

In the past years, researchers have developed new technologies to mitigate some of the fatal risks and severe injuries that highway workers encounter. However, the majority of the developed technologies are reactive, meaning that they are triggered only after intrusion or when threats are in a close proximity of work zones [2, 13, 14, 15]. These reactive aspects limit the ability of workers in showing a timely and proportional reaction to the extent of upcoming safety risks. At the same time, recent technological advances in Artificial Intelligence (AI) have enabled researchers in different disciplines to leverage visual analytic in solving some of the safety-oriented problems [16, 17]. However, the majority of the previous research efforts related to the safety of construction workers were mainly centered around building construction discipline, and highway construction has only received marginal attention [2]. Furthermore, the nature of risk factors in building and highway constructions are inherently different. For instance, highway workers are required to be present near high-speed traffic, and their working environment offers less space for maneuvering, which makes their needs different from that of building construction workers [18, 19, 20].

One major challenge in developing safety systems for workers is designing a notification mechanism with appropriate modalities that could play out well in the noisy and distracting environment of highway work zones. To this end, recent studies have unveiled the leverage of multimodal over unimodal cueing in risk communication [21]. Several researchers also reported that the reaction time to a visual cue combined with auditory or vibratory stimuli were shorter than each modality alone [22, 23]. This growing evidence has encouraged researchers to design their warning mechanisms around multimodal notification mechanisms in different contexts [24, 25]. With this, visual cue coupled with auditory and vibratory modalities through Augmented Reality (AR) smart glasses will help workers to better handle a specific situation or task in a more informed fashion with minimal vision obstruction [26]. Such devices also provide hands-free interactions and can overlay information on what users naturally see, which makes them suitable for many applications [27, 28, 29]. These attributes have made AR a desirable technology to be leveraged in assisting workers in different contexts [30, 31, 5]. However, designing context-aware and user-centered interfaces that maximize the benefits of this technology have been proven as a critical step in increasing end users' engagement [32, 33].

This chapter proposes a novel framework that incorporates the benefits of real-time AI and AR to enhance the situational awareness of highway workers by providing a real-time communication infrastructure. The proposed framework has three major pillars: (1) AR user interface design for multimodal notification, (2) real-time deep learning for vehicle detection/classification from distance, and (3) real-time wireless communication between AR and AI hardware components. Figure 2.1 provides a holistic view of the proposed framework and illustrates two examples of its applications. The worker in this context is a broad term that represents individuals involved in highway work zones (e.g. road inspectors, law enforcement, and first responders).

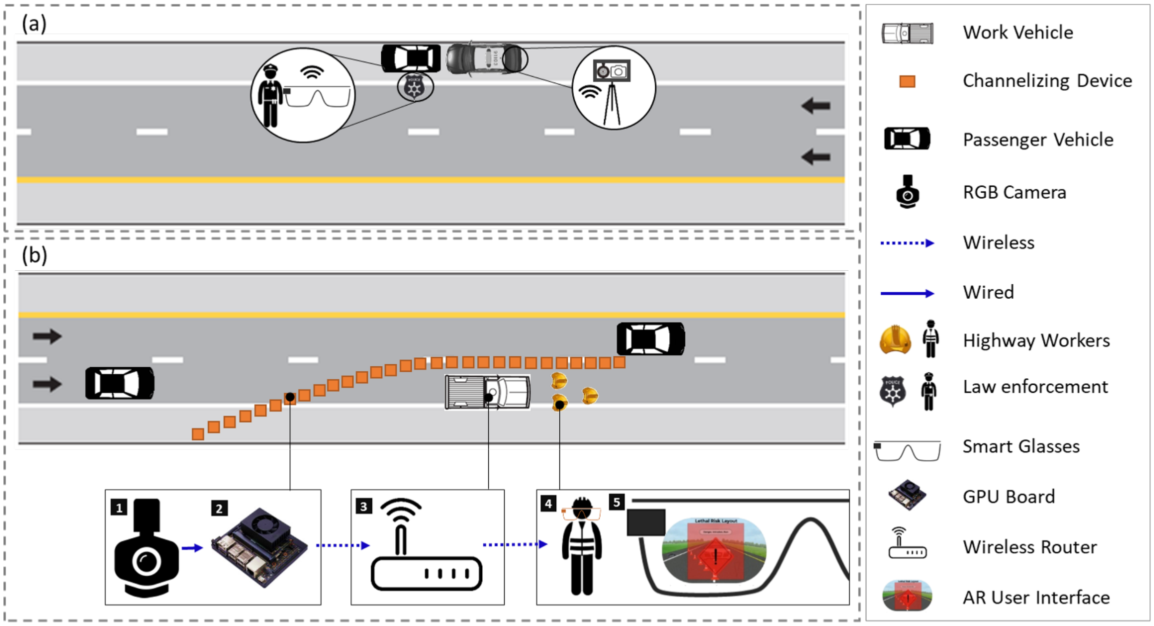


Figure 2.1: Holistic View of the Proposed Framework - its Application in (a) Law Enforcement, (b) Highway Work Zones

Figure 2.1(a) demonstrates how this framework can be used for securing the safety of law enforcement in highways. Figure 2.1(b) also visualizes how the proposed framework functions in a highway work zone setting. Within this framework, a camera (1) that is wired to the embedded Graphics Processing Units (GPUs) (2) provides real-time information about the potential threats. Workers (4) are equipped with a pair of smart glasses (5) for receiving notification enabled by a local WiFi network (3) within the work zone area and across the utilized smart devices. With the aim of developing a user-centered technology, this chapter also presents the results of an early mixed-method user research that investigates highway workers' perception toward the proposed framework through a devised mixed-method user research methodology based a low-fidelity prototype, which includes a semi-structured interviews and a quantitative questionnaire.

In summary, the results of this chapter demonstrate that the trained AI model achieved 48.7% mAP for detecting vehicles from distance with 24.83 Frame Per Seconds (FPS) execution latency on a Nvidia Jetson Xavier embedded platform. The outcomes also indicate that the real-time end-to-end wireless communication latency between the embedded board and a pair of AR smart glasses combined with the AI execution latency is within 46 milliseconds margin, which provides a reliable foundation for meeting the tight timing constraints of a real-time safety system. The early mixed-method user research also reveals that the proposed safety framework and the conceptualized interface are positively welcomed by the body of the highway maintenance and operation community. Participants were selected from multiple states in the US, which provides a reasonable and cross-regional understanding of workers' perception toward the proposed technology, and also fuels future worker-centered developments.

Overall, in addressing the identified research gaps, our contributions to the body of knowledge are:

- This chapter is the first research work that conceptualizes and formalizes functional and technical requirements of an AI-enabled Augmented Reality (AR) system to enhance the safety of highway workers.
- This paper also is the first research effort that presents a holistic design framework integrating real-time AI processing and communication at the edge with AR user interface design for multimodal notification to highway workers.
- This chapter provides a proof-of-concept model for AI-enabled worker-in-the-loop technologies in increasing the safety of highway work zones.
- The outcomes of the user research of this chapter also help future designers in developing worker-centered technologies and customizing user experience for highway workers and broader worker-in-the-loop safety systems.

2.3 Related Work

In this section, we are exploring the related works of this paper from the literature. First, we will explore the existing technologies in highway work zones safety. Then, we will go through the background of user research and its importance in the successful development of the technology. Next, we will move on to the history of deep learning for vehicle detection/classification in highways. Finally, we will conclude this section by exploring the application of AR in the context of safety.

2.3.1 Highway Work Zone Safety and Technology

Highway work zone safety is one of the major concerns in transportation agencies. The job description of maintenance and construction highway crew mostly requires their presence near or adjacent to traffic flow, which fuels the risk of a fatal environment for both workers and road users [14]. Several studies in the literature tried to identify the major roots of risks in highway work zones. For instance, in a comprehensive work, [20] analyzed the major risk factors in highway work zones. They identified that besides at-fault drivers, lighting condition and vehicle types also play a significant role in causing risks in highway work zones. In another study, [34] investigated the significance of lighting condition by comparing nighttime and daytime crashes. Finally, [35] developed an Ordered Probit Model to identify the factors that affect severity of work zone crashes under weather adverse condition.

Even though the risky environment of highway work zones have been highlighted in several research studies, the safety practice is yet to rise to this challenge by enriching the current practice of highway work zone safety by implementing new technologies. Several research studies have suggested that the current safety technologies used in highway work zones are still limited to portable signs, automated flagger, directional alarms, warning lights. They have also cited the fact that the current intrusion systems are mainly reactive, meaning that the triggering mechanism only fires after intrusion or when the intruding objects are in a close proximity of work zones [36, 12, 37, 38]. On the contrary, the dynamic environment of highway work zones accentuates the need for more robust and adaptive safety systems that can quickly adapt to the changes in the working environment of highway workers.

In a comprehensive study, [2] systematically reviewed 147 work zone safety technology-related articles and identified that the idea of departing from old-fashioned technologies toward smart automated technologies have exponentially grown In the past years. The emergence of more advanced sensing and wearable technologies have also energized more researchers in the pursuit of this trend. For instance, [3] designed a proximity-based alerting system leveraging only tactile cue as the sole mode of communication with users. However, highway work zones are noisy and taxing, and the job description of highway workers is demanding, which could weaken the effectiveness of this system in alerting workers. In another research effort, [39] proposed a wearable-based hazard proximity warming system for increasing the awareness of construction workers, which still relies on proximity-based triggering mechanisms. In yet another study, [40] developed a novel system based on AR in alerting workers of the orientations and the proximity of potential hazards. Despite the promising potential of this system, given the fact that it generates warnings only based on the Field-of-View (FOV) of workers, it may not be efficient in preventing hazards that are not within the visual scope of workers.

2.3.2 User Research and Successful Technology Adoption

User acceptance is one of the most crucial factors in technology development. Whether the intended users of a technology accept it or not plays an important role in the ultimate success of a product [41]. With the overflow of new technologies in recent years, workers are poised to witness a revolution in the way they work [42]. Even though researchers have been struggling to maximize the benefits of new technologies for workersâ safety and health, several studies showed that it is not uncommon for the end-users to resist adopting new technologies regardless of their benefits. Therefore, before forcing and massively investing in new technologies, it is of utmost importance for the developers to always take account of potential technology resistance and user acceptance while developing a new technology [43]. To this end, a few researchers investigated how construction workers react to new technologies. For instance, the authors of [44] proposed an extended Technology Acceptance Model (TAM) that is capable of reflecting the possible changes in the attitude of workers in time. In another study, [45] studied the factors that impact the implementation of a mobile-based computing device systems in the construction industry from usersâ perspective. The authors extended the TAM model to better understand the relationship between user satisfaction and pierced performance of the devices. In yet another similar study, [46] explored the general perception of occupational groups in construction companies toward the use of Information and Communication Technology (ICT) and how it impacts the post-adoption stage of the management process. In summary, it is critical for us to investigate early on the perception of our users (highway workers) toward our proposed system. While our main goal in this paper is not to come up with an extensive Technology Acceptance Model for highway workers, we wanted to obtain a preliminary idea about their perception toward the proposed framework and designed interface early on.

2.3.3 Deep Learning for Vehicle Detection in Highways

With swift progress in recent years, AI has enabled researchers to expand their capabilities to new horizons that sounded far-fetched years ago. Deep neural networks and particularly, Convolutional Neural Networks (CNN) have demonstrated near-human and in some cases even beyond-human capabilities in object detection and classification [47]. VGGNet, GoogleNet, and ResNet are some of the algorithms that have been widely used in feature extraction, creating objectsâ bounding box and object classification [48, 49, 50]. Such algorithms have been recently utilized in a wide range of vehicle detection and challenging challenges. For example, Wang et al. [51] used Restricted Boltzman Machine, which is a type of deep neural network to detect vehicles using online transfer learning. Authors in [52] use Inception network [53] with a multiscale feature fusion network to achieve good accuracy for vehicle detection. In [54], Yolov3 [55] was used to detect vehicles on the UA-Detrac [56] dataset, achieving good results. Arinaldi et al. in [57] used Faster R-CNN [58] object detector to detect vehicles. Authors in [59] also adopted Faster R-CNN with an anchor proposal network outperforming existing object detectors in terms of accuracy. In [60], Rujikietgumjorn et al. further divided vehicle classes in UA-Detrac into sub-classes and used Faster R-CNN. Zhu et al. in [61] used Faster R-CNN with a computationally efficient method for feature extraction while getting highly accurate results on the UA-Detrac dataset. [62] also employed Faster R-CNN with a feature fusion module to get state-of-the-art results on the UA-Detrac dataset.

In the context of vehicle tracking, earlier works like [63] used classical methods like background and foreground subtraction for vehicle detection and tracking. Authors in [51] used HOGs with adaboost gradient for vehicle localization and classification respectively. They applied kalman filter [64], which is a simple generative model, to track these detected vehicles. Peng et al. [65] used R-CNN [66] to detect vehicles and a gaussian distribution for tracking them. Zou et al. in [67] used Yolov3 as their detection framework with a CNN based siamese network to re-identify them. Markov decision process was used to track the vehicles using pre-defined policies.

2.3.4 Augmented Reality for Safety Benefits

The idea of smart wearable devices has a long history [68]. The recent advances in machine learning, deep learning and visual analytic have energized scientists and industry leaders for new generation of smart wearable devices to enhance the cognitive capability and situational awareness of humans and assisting them in decision making [69, 70]. The benefits of smart glasses with real-time visual feedback are many -

from training and education, to manufacturing and maintenance and information assistance [29]. Scientists have already demonstrated the benefits of AR smart glasses systems for military training [71], oil refinery training [72], nuclear plant maintenance [73], equipment status check for industrial services [74], facility management [75] asset management [76], enhancing surgeons' vision [77], and step-by-step guidance for assembly and manufacturing [78, 79, 80].

In recent years, the research community has witnessed a growing interest in using AR in safety-related issues. As an example, the authors of [81] used an in-vehicle head-up display to notify car drivers of any possible collision with pedestrians in advance . Authors in [82] investigated the applicability of AR in enhancing occupational safety workers in industrial environments. They concluded that AR provides a suitable and effective platform for offering instructions at industrial workplaces. In another similar study, a few researchers in [83] studied the perspectives and challenges of AR in industry workplaces for workersâ safety and health. They conducted multiple interviews with actual workers and reported that their interviewees believed that AR could potentially contribute to their safety and health.

We also have identified a trend of using AR smart glasses in construction work zones with different applications. Some examples include integrated smart glasses into helmets for augmented reality safety applications [84], underground infrastructure positioning and layout display [85, 86]. Several other applications of augmented reality in the construction safety domain are studied with the purpose of safety risk identification, training, and inspection [87, 88]. Finally, [89] presented a research agenda for AR in construction. They provided vital information for practitioners and researchers about possible directions and trends of AR in construction industry in near future.

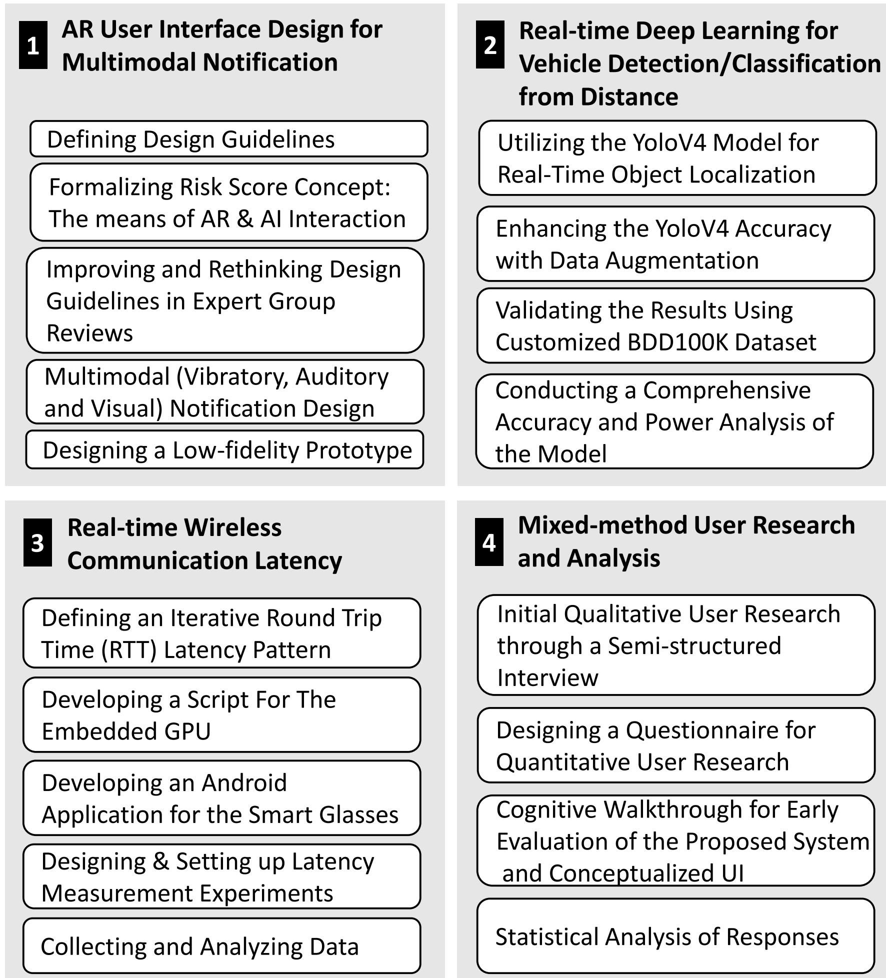


Figure 2.2: Our Designed Methodology for Investigating the Pillars of the Proposed Framework

2.4 Methodology

In this section, we will explain our methodology in investigating the main research pillars of our proposed framework which are (1) AR user interface design for multimodal notification, (2) real-time vehicle detection/classification from distance, and (3) real-time wireless communication. Furthermore, we will also present our mixed-method user research approach for studying end users' perspective toward the designed framework and interface. Figure 2.2 demonstrates the taken steps in this study.

2.4.1 AR User Interface Design for Multimodal Notification

In the first step, we need to conceptualize and design the AR user interface. To this end, we first defined some guidelines to be used in the design process. For this purpose, we followed the usability heuristics provided by [90] when pertinent. These heuristics have been traditionally used in designing and evaluating UIs in different disciplines [91]. Furthermore, we set the MUTCD manual [92] as the reference for choosing the colors and signs to be used in the design to increase the familiarity of users with the UI following the "consistency and standards" heuristic. In addition, we tried to include the effectiveness of prewarning in our design [93]. Finally, we decided to aggregate different risks into one parameter, which will be called "Risk Score" hereinafter. RS specifies the severity of the current safety threat to the workers, aggregates all risk

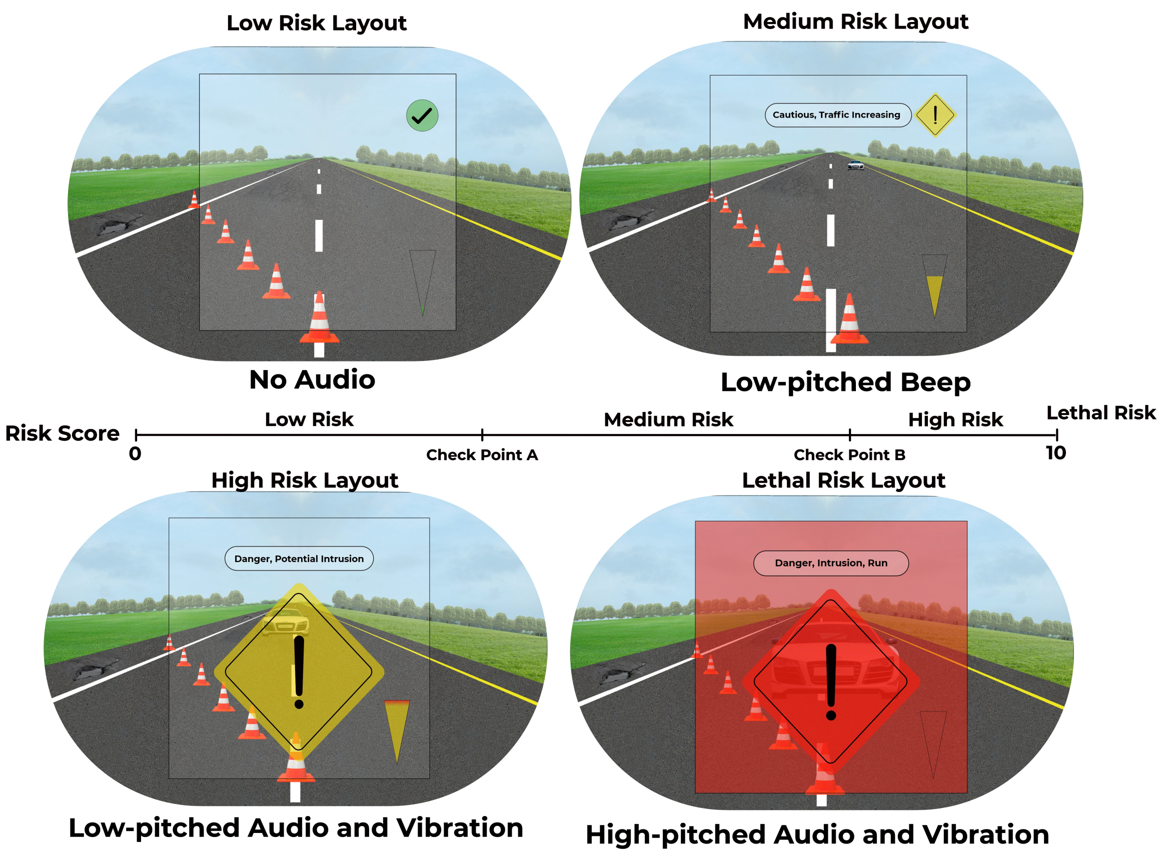


Figure 2.3: Designed Interface for Multimodal Notification Mechanism for the AR Smart Glasses

origins, such as traffic, weather, etc. This value determines transition among the layouts. RS varies between 0 (not any risk at all) to 10 (definite intrusion in a few seconds), which is calculated in the backend and is sent to the frontend in real-time.

After brainstorming and preparing our preliminary sketches, we improved our initial design through expert group reviews and came up with the layouts shown in Figure 2.3. The proposed interface consists of four different layouts: Low Risk, Medium Risk, High Risk, and Lethal Risk. The occurrence of each layout depends on the value of RS. One of the strengths of the proposed framework is its adaptability. The utilized checkpoints in the framework specify the severity of the identified risk on workers, and in turn, the corresponding layout. These checkpoints could customize the framework based on the different types of work zones. For instance, the Low Risk interval in a short-duration work zone could be less than that of a permanent work zone, allowing the framework to go in High Risk margin sooner.

It should be also noted that in Figure 2.3 we considered a square display for the smart goggles, and the background shows a hypothetical user view in a hypothetical work zone. In Low Risk, Medium Risk, and High Risk layouts, we are using an RS bar on the bottom right to show the current status of the user, under the "visibility of system status" heuristic. Within the proposed framework, we envisioned the High Risk layout to work as the prewarning stage in our UI. In the first two layouts, the ratio of the occupied display is so much less than the last two ones, where the whole display is occupied. Therefore, we design the High Risk layout to be the prewarning stage that facilitates drawing users' attention to the to-go-off warning. Also, Figure 2.3 schematically illustrates the transparency of the designed UI as well. While in the first two layouts, the display is completely transparent, it becomes partially opaque in the last two ones so that the effectiveness of warming increases. Figure 2.3 also demonstrates that we leveraged multimodal cueing consisting of visual, vibratory and auditory modalities in the proposed mechanism. Depending on the phase, different tools would be used within this framework. In this context, different message contents could also be sent to the users. Therefore, any other special update that requires the immediate attention of the workers, such as weather advisory, could also be communicated with the users in different scenarios. Therefore, the application of the designed interface is not just limited to intrusions into the work zones. Yet, it can go beyond that and reflect other scenarios as well.

2.4.2 Real-time Deep Learning for Vehicle Detection/Classification from Distance

In this section, we introduce our deep learning framework for vehicle detection and classification from distance in highways. First, we will explain our the network model and the training framework based on YoloV4 [94]. Then, we will go into the details of fine-tuning our baseline model for custom vehicle detection/classification dataset. This includes enhancing the accuracy of the vehicle detection from distance based on data augmentation and transfer learning.

2.4.2.1 Vehicles Localization

This part aims to briefly introduce the architecture of the model utilized to detect and locate objects in a frame. Based on the needs we identified in this research, we customized YoloV4 structruce [94] as shown in Figure 2.4. The YoloV4 network architecture consists of *backbone*, *neck*, and the *head*. The main task of the backbone network is to extract the essential features from the input image. The backbone of YoloV4 can be selected from *CSPResNext50* [95], *CSPDarknet53* [95], and *EfficientNet-B0/B7* [96]. In line with the work of [94], we used the CSPDarknet53 as a backbone. This is because it has equal or higher accuracy than the other two networks while having lower operational complexity. Cross Stage Partial (CSP) [95] strategy, applied to the CSPDarknet53, concatenates the previous output features from the previous layers and passes them on to the next layers. A CSP-based network architecture lessens vanishing gradients problem in deep networks, and it also

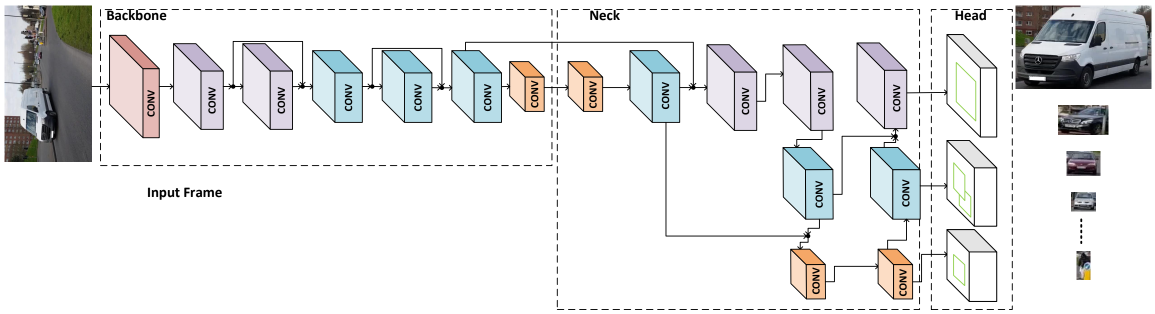


Figure 2.4: Baseline Model for General Object Detection/Classification

minimizes network parameters by reusing previously extracted features.

The task of the neck is to aggregate features with different resolutions from the backbone. YoloV4 uses the Feature Pyramid Networks (FPN) method to have the prediction at different scales. YoloV4uses the features from the top-down stream and adds them to the next neighbor in bottom-top stream. YoloV4 also uses a modified version of Path Aggregation Network (PANet) [97] in its neck to flow the information in the neck layers. Contrary to the original PANet, YoloV4 uses a concatenation layer rather than reduction to create a new vector. As a result, spatial information from the features will be better preserved by the network.

In the final stage, the head makes predictions about objects and their location based on features received from the neck. YoloV4 uses the same header used in YoloV3. The output of the head is the confidence per each class and its location in the form of $\langle x, y, w, h \rangle$, where $\langle x, y \rangle$ is the center of the object and w, h are width, and height of the object from its center, respectively.

2.4.3 Real-time Wireless Communication Latency

Real-time communication latency is one of the most important pillars that needs to be considered within the proposed framework. This latency refers to the time that takes for any decision inferred by the backend to be sent to frontend of the proposed framework. For measuring this latency, we first designed a Server-Client software pattern illustrated in Figure 2.5. In this context, Round Trip Time or RTT is the time that takes for the server to send a message to the client and receive the same message back, divided by two. Therefore, RTT is the average time from a back and forth communication between the server and the client, and is considered as the average communication latency [98]. We also desisgned this pattern around Transfer Control Protocol (TCP). TCP is a connection-oriented transition protocol, which is capable of sending payloads in multiple packets and replicating the process at the receiving end. This enables TCP to be able to support large payloads [99].

In the designed pattern, the server first initializes a server socket and starts listening on a port (e.g. 8888). Next, using the IP address given to the server through the access point, the client tries to connect to the server. Upon the establishment of the connection, the server awaits sending a message to the client. In the meantime, the client awaits receiving this message in full. As soon as the data is completely received, the client sends the same message back to the server and awaits for that to be delivered. Server, in turn, awaits for the message to be fully received. Finally, the client will close the socket.

Finally, we developed an Android application to be run on the AR smart glasses using ServerSocket and Socket classes in Android [100, 101] based on the designed methodology. We also used the Socket class in Python to develop another software for the embedded GPU [102]. In order to collect reliable data, we measured the communication latency 500 times in an iterative manner, and recorded the results. In the end, we used $1.5 * IQR$ (interquartile range) rule to identify and remove outliers in our dataset, and removed all observations with a value more than $Q3 + 1.5 * IQR$ or less than $Q1 - 1.5 * IQR$ [103].

Now that we have a methodology for measuring the latency, it is time to investigate what type of messages should be sent to the end-users from the board in a real scenario. According to [104] and [105], presenting full sentences to the users might increase the cognitive workload. The physical activities of highway workers are taxing

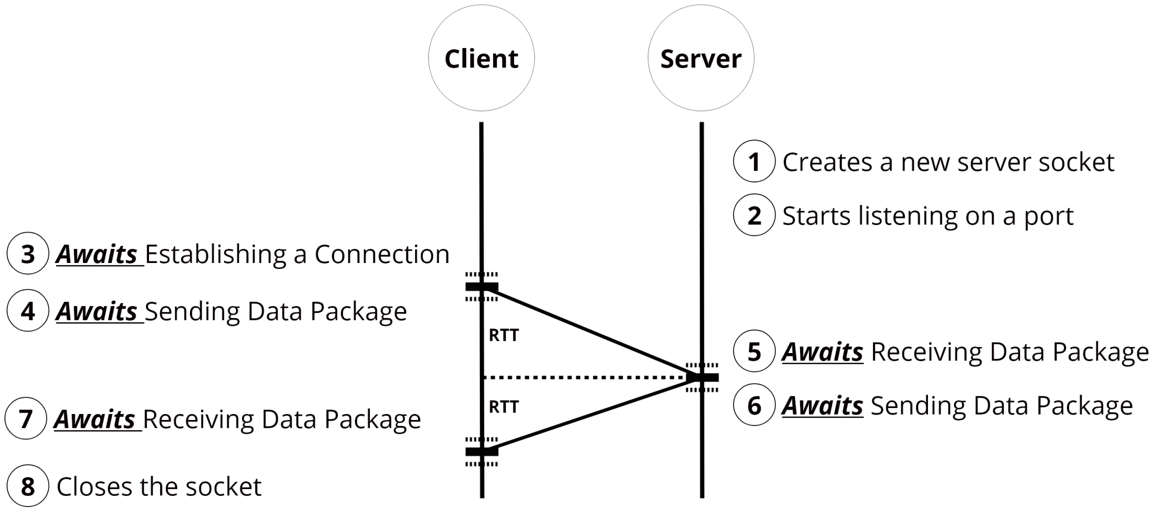


Figure 2.5: Round Trip Time (RTT) Concept for Measuring Wireless Communication Latency

and it makes sense to reduce the load as much as possible. Therefore, we designed a concise pattern of "Context, Command, Time" for the messages that will be sent to the users from the board. For instance, "Possible intrusion, Run, 2 seconds".

2.4.4 Mixed-method User Research and Analysis

To investigate the perception of our end users toward the proposed framework and the designed UI, we devised a mixed-method user research approach. Due to the unique needs of highway workers and as the qualitative part of our mixed methodology, we first conducted an in-depth interview with an experienced highway maintenance worker to have his initial input about the proposed framework and the designed UI. Our interviewee was a senior and former member of a state Department of Transportation (DOT) with more than 20 years of experience. He was directly involved in the division of maintenance and operation of a state DOT during his tenure. Our interview was semi-structured. We started with a typical set of demographic questions and had an open discussion about the proposed concept and its pros and cons. Our interviewee believed that the proposed framework can be efficient in highway work zones.

In the next step and for investigating users' idea about the proposed framework and designed UI, we leveraged a quantitative questionnaire coupled with a cognitive walkthrough. Cognitive walkthrough is a well-known and widely used methodology in the literature that is usually utilized in the early stages of product development [106, 107]. Based on the results of our interview, we decided to separate the participants of our questionnaire into two different groups: highway maintenance crew and affiliated participants. The first group only includes workers, and the latter includes state DOT members, private consultants and managers, researchers, and other affiliated members who are often present in highway work zones. The reasons behind these categorizations are: (1) highway maintenance crew are more frequently present in a work zone than other members due to their job description, and (2) highway maintenance crew might have a different set of expectations from safety systems than affiliated members. As an example, financial aspects might not be an influencing factor for workers, whereas it is an important factor for the management team. It should be noted that hereinafter and for the sake of brevity, we only will be using the terms "maintenance crew" and "affiliated participants" to refer to the considered groups in the survey, respectively.

We used Google Forms [1] for hosting and performing a survey. The survey started with some demographic questions. We asked our participants to report their working districts, age, experience, and the frequency of their presence in highway work zones. Then, we asked the participants to watch a video uploaded on YouTube[2]. In this video, we walked the participants through our proposed technology and debriefed them about how the framework would work in a real-world application. We tried to provide a comprehensive yet concise video in order to minimize the number of drop-offs from our participants. After watching the video, we asked our participants to answer the three designed Likert-scale questions about (i) the practicality of the proposed framework (Question (a)), (ii) the likelihood of them using this framework (Question

[1]https://forms.gle/sReS6h1CM4DJ2qj98
[2]https://www.youtube.com/watch?v=Rtih3rh1q0s

(b)), and (iii) and the likelihood of them recommending this framework to others (Question (c)). As the final question, we asked the participants to share with us their concerns about the proposed technology. We asked them to select from "impractical for operation in highway work zones, unpleasant experience with devices", "influence on the performance such as vision obstruction", "slow and painful adaptation to the technology as a routine", "the unreliability of the devices in identifying potential dangers", "repetitive false alarms and loss of your trust in devices", and "none". They also had the option to select multiple options from the provided list.

Finally, we used another prerecorded video uploaded on YouTube[3] as part of our qualitative research to survey our participants' opinion about the designed UI. In this video, we briefly discussed the interface design procedure. We also used an animated low-fidelity prototype of the UI to further describe how the UI would work in a real-world application. After watching the video, the participants were asked to report their first impression of the UI (Question (d)). Finally, we also asked them to share with us their thoughts on strengths and weaknesses of the designed UI.

2.5 Results and Evaluation

Here, we are reporting the early results and exploration of our proposed system. Below, we will start first by explaining our experiment setup, the devices and the other material that we used in this study. Then, we will move on to explaining our results from the defined pillars.

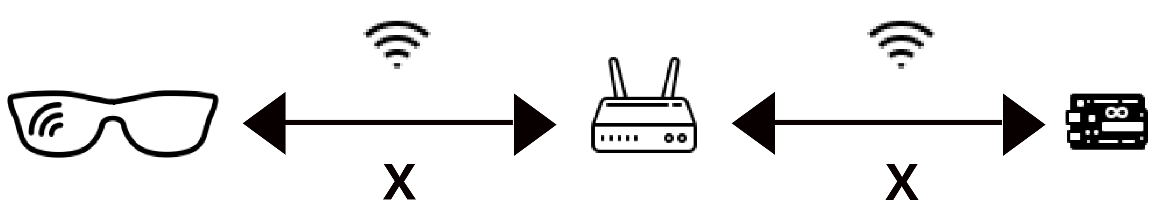


Figure 2.6: Our Experiment Setup

[3]https://www.youtube.com/watch?v=pEVSS2aRLaA

2.5.1 Experiment Setup and Specification

Experiment Design In a real work zone setting, we envision the access point to stay in the vehicle (truck) of the deployed team, as visualized in Figure 2.1. In addition, we already mentioned that the board would be installed next to the camera and set up at the beginning of the work zone. Therefore, in order to create the worst-case scenario, we assumed that the worker, wearing the glasses, would be on the other side of the work zone. That leaves us with an experiment in which the access point is centered between the board and the glasses as the length of the work zone expands. While the length of work zones vary case by case, we measured the latency over multiple distances to investigate how the real-time communication latency differs as the distance between end nodes grows. We also ran our experiment in an outdoor environment and next to the traffic flow to mimic the environment of a work zone and the possible impact of moving objects on the magnetic field. We considered different distances between the end points and the router (X in Figure 2.6), ranging from 10 meters to 60 meters with an increasing 10-meter interval.

Smart Glasses Among the abundance of advanced and modern options available in the market, we used Vuzix Blade [108]. Vuzix Blade smart glasses are a see-through display that supports Android as their operating system (OS) and offers an auto-focus 8-megapixel camera, built-in stereo speakers, and advanced voice control. Its right eye monocular optic display provides a platform for removing distractions and minimizing vision occlusion. It is also equipped with a 2.4 GHz bandwidth Wi-Fi module that enables wireless communication with other devices. One of the benefits of this technology is its compatibility with the American National Standards Institute (ANSI) Z87.1 certification that makes it a viable option for industry usage.

Embedded GPU In 2017, the leading manufacturer of Graphics Processing Units (GPUs), Nvidia, manufactured Jetson TX2. In 2018, it released Jetson Xavier, which is the state-of-the-art embedded GPU for mobile processing. Xavier is a System-On-Module(SoM) device with a 512 core volt GPU, 8 ARM cores, 16 GB memory and consumes 30 Watts of power at most. We used this device because it offers all computational capabilities required for executing object detection algorithms while keeping up with the real-time output expectations. Besides, it is a low-power and light-weight device that can be easily mounted on a tripod, which makes it a suitable candidate.

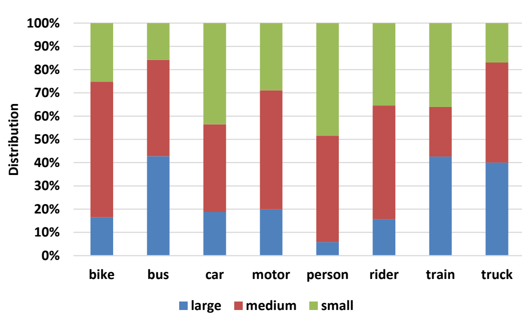
Wi-Fi Access Point We built our Wi-Fi network using a 2.4 GHz bandwidth router by TP-Link. This access point offers communication in compliance with IEEE 802.11n/b/g standard and advanced wireless technology for delivering wireless communication speed up to 300 Mbps [109].

The Summary of the Community Participation in the Survey At the time of preparing this manuscript, we received 76 responses from the maintenance crew and 52 from affiliated participants with the body of highway work zone community. Almost half of our participants were at least 45 years old. Moreover, we realized that our participants were fairly experienced with 71 percent of them having at least 10 years of experience in the field. 68 participants from the maintenance crew group mentioned that they are present in highway work zones on either daily or weekly basis. Furthermore, 82.7 percent of affiliated participants reported that they have been to a highway work zone at least 50 times. It should be noted that the results provided here are part of a larger survey that was performed.

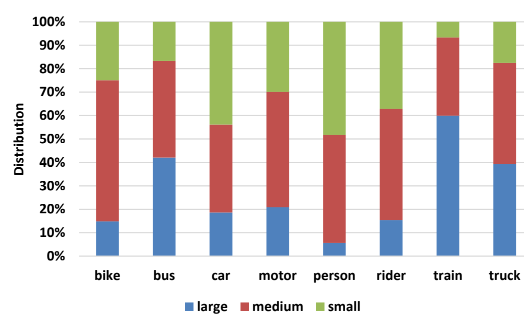
Dataset for Vehicle Detection we selected BDD100K [110] dataset due to its direct view-point. The dataset was collected in four different areas: San Francisco, Berkeley, Bay Area, and New York. The object localization segment of the dataset consists of ten objects: bike, bus, car, motor, person, rider, traffic light, traffic sign, train, and truck. For our study, we removed traffic light and traffic signs categories from the dataset. Table 2.1 summarizes the distribution of each class in both validation and training segments of the customized dataset. The car class has the highest distribution and it comprised approximately 83% of both training and validation subsets. We also visualized the distribution of each class for each subset based on its area in Figure 2.7. We used COCO [111] dataset evaluation metrics for defining the size of the objects. The COCO (Common Objects in Context) dataset is a widely used large-scale dataset for object detection, segmentation, and captioning tasks in computer vision research. It provides a comprehensive collection of images with detailed annotations, making it valuable for training and evaluating algorithms related to visual recognition tasks. Based on COCO dataset evaluation metrics, an object with an area of fewer than 32^2 pixels is small, between 32^2 and 96^2 pixels is medium, and any area larger than that is referred to as large objects.

Table 2.1: Breakdown of the customized BDD100K

Subsets	Bike	Bus	Car	Motor	Person	Rider	Train	Truck
Train	0.84%	1.36%	82.83%	0.35%	10.61%	0.52%	0.02%	3.48%
Val	0.81%	1.29%	82.84%	0.37%	10.72%	0.52%	0.01%	3.43%



(a) Train subset



(b) Validation subset

Figure 2.7: Distribution of area scale per each category for training and validation sets.

2.5.2 Real-time Deep Learning for Vehicle Detection/Classification from Distance

In this section, we will provide the accuracy of YoloV4 on the customized BDD100K dataset and their network latency on the Xavier platform.

2.5.2.1 Accuracy

We analyzed our model's performance using the mean Average Precision (mAP) metric. The mAP of a network over queries, $\mathcal{Q}$, is given by: mAP $=$ $\sum_{i=1}^{|\mathcal{Q}|} AP(q_i) \frac{}{|\mathcal{Q}|, q_i \in \mathcal{Q} AP(q) = \frac{1}{TP_{gt}} \sum_{j}^{|\mathcal{G}|} \frac{TP_{detected}}{j}}$ where TP_{gt} is the number of ground truth true positives, and $TP_{detected}$ is the number of true positives detected by the network. Following the PASCAL-VOC [112] standard for object detection, we calculated the mAP when Intersection-over-Union (IoU) is higher than 50%. It means that an IoU of 0.5 or higher between detection and the actual ground-truth of the object would result in a positive detection. We also provide the mAP of different categories at different area scales. Table 2.2 summarized the mAP of all objects and per each category for the input size of 608×352 pixels. In our case, the model provides the mAP of 48.7% for all classes at different area sizes. The accuracy of the model increases as the objects get closer to the camera (larger area). Therefore, for medium and large objects, the mAP stands at 52.7% and 73.3%, respectively. For car objects, the model shows the highest mAPof 74.9%, as it has the highest portion of the training dataset. Bus and truck objects have the next highest accuracy (∼59%), as they are typically larger than the other categories. We also provided some qualitative results in Figure 2.8. These images are sampled from different resources, including YouTube, videos taken from the industry partners and the validation portion of the BDD100K dataset. These qualitative results illustrate the capability of our AI model in detecting and classifying vehicles/incoming vehicles in highway/urban settings. However, it should be noted that our AI model, as is, does not perform any velocity/acceleration and traffic status measurements.

2.5.2.2 Execution Latency

In comparison with three different power modes, we analyzed the latency of processing every single frame on the Xavier embedded GPU in Table 2.3. Our results

Table 2.2: Mean Average Precision (mAP) of defined classes at different area scales

Category	$\boldsymbol{m}$**AP**[0.5]	$\boldsymbol{m}$**AP**$_{small}$[0.5]	$\boldsymbol{m}$**AP**$_{medium}$[0.5]	$\boldsymbol{m}$**AP**$_{large}$[0.5]
All	48.70%	20.90%	52.70%	73.30%
Bike	46.40%	14.50%	51.20%	78.80%
Bus	58.90%	17.80%	48.30%	81.00%
Car	74.90%	49.20%	88.60%	97.60%
Motor	41.90%	20.30%	48.20%	63.00%
Person	58.40%	33.70%	78.00%	88.90%
Rider	43.90%	10.60%	54.20%	86.70%
Truck	58.30%	20.70%	53.20%	78.80%

Figure 2.8: Results of the Trained Model

demonstrate that we achieved the highest end-to-end FPS of 24.83 while burning 25.83 Watts in MAXN power mode. The Xavier reaches to highest energy efficiency for 15W power mode when it processes 13.63 images per second by consuming 10.11 Watts; therefore, its energy efficiency stands at 1.35 FPS/Watts.

Table 2.3: Execution latency, power consumption, and energy efficiency of Yolov4 on Xavier Jetson AGX board

Power Mode	Pre-Processing (ms)	Inference (ms)	Post-Processing (ms)	End-to-End (ms)	End-to-End (FPS)	Power (W)	Energy Efficiency (FPS/Watts)
MAXN	10.2	24.1	6	40.3	24.83	25.83	0.96
15W	18	45.6	9.8	73.4	13.63	10.11	1.35
10W	23.7	92.2	10.3	126.2	7.92	7.77	1.02

2.5.3 Real-time Wireless Communication Latency

In the following, we present the communication RTTs between the embedded GPU and the smart glasses using the wireless network through the utilized access point and developed standalone softwares. We considered four different message contents to be sent to the highway workers in our experiment and measured the corresponding RTT latencies that we measured outdoor in 38 degrees Fahrenheit in an outdoor environment. These contents are provided in Table 2.4. It should be noted that the message contents that we used mimicked multiple scenarios in which the workers can be informed via our framework, whether that is a possible intrusion or a weather advisory condition.

In order to better analyze the data, we first used a two-way ANOVA test to investigate the impact of distance and message size on the RTT latencies. We considered distance and size without any interaction as our inputs to the test. The obtained p-values reported in Table 2.5 demonstrate that only distance plays a statistically important role in the variation of RTT latencies. Therefore, we aggregated the collected data regardless of the message size, cleaned them using the 1.5 IQR rule, and visualized them in Figure 2.9.

In our experiment, we were able to cover up to 60 meters of distance between each

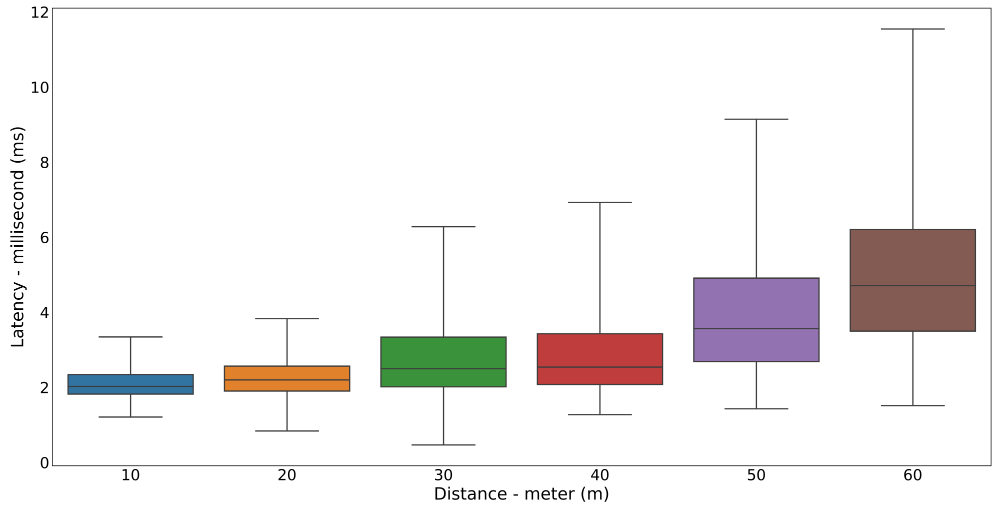


Figure 2.9: Round Trip Time (RTT) Latencies Over Different Distances

end point (smart glasses and the board) to the router, or 120 meters between end points. However, we were not able to establish a stable connection between end nodes when they were placed farther than 60 meters from the router. In addition, during our experiment, we also noticed that RTT latencies were noticeably impacted by the battery level of the smart glasses, weather temperature, and heavy moving objects. Measuring actual impacts of these obstacles require further investigation. However, we ensured that the RTTs were collected under similar condition, when the smart glasses were not in low-battery mode.

Table 2.4: The Considered Message Contents and Their Corresponding Memory Size

Message	Memory Size (Bytes)
Danger , run, now !	68
Strong storm coming , pack up and leave , 5 minutes	100
Barrier removed , fix it , now !	81
Staying too close to the border, move, now !	93

Table 2.5: Two-way Anova Test on Collected RTT Latencies

	Sum of Squared	Degree of Freedom	F-Statistic	p-value
Distance	10685.44	1	6142.193	0.000
Size	2.04	1	1.1728	0.278

Our results demonstrate that we received an acceptable real-time communication latency between the embedded GPU and the smart glasses through an 2.4 GHz band access point. The 120-meter coverage distance with an average of 5.1 milliseconds at the farthest scenario gives us a promising coverage area for highway work zones. In addition, Table 2.5 suggests that in our context, as long as the designed pattern (i.e. Context, Command, Time) is utilized, the size of the message content should not play a statistically significant role in RTT latencies, leaving the distance as the most important factor. However, Figure 2.9 indicates that as the distance between

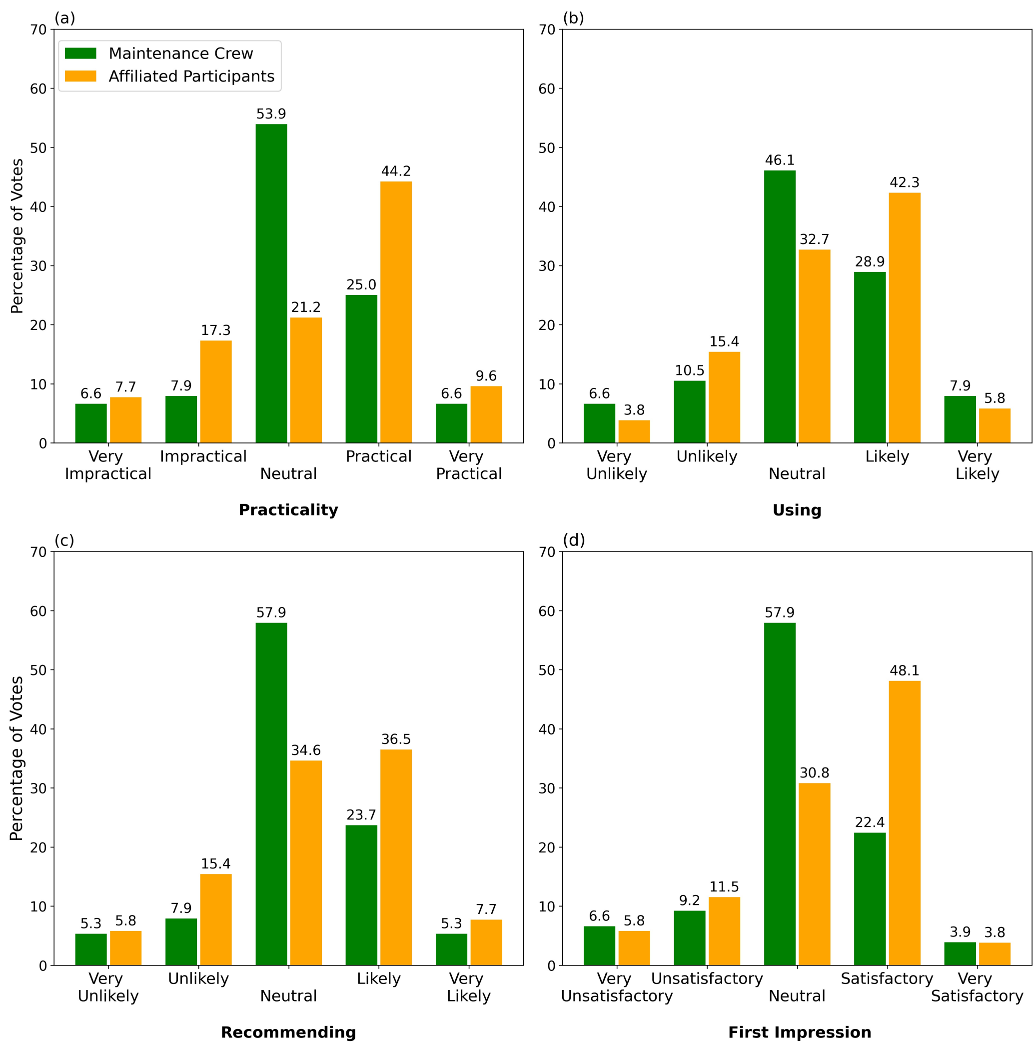


Figure 2.10: Participants' Response to the Questions of (a) Practicality of the framework (b) Likelihood of Them Using the framework and (c) Likelihood of Them Recommending the framework to Others (d) Their First Impression of the Designed UI

Table 2.6: Our Participants' Concerns about the Proposed framework

Concern	**Maintenance Crew Votes**	**Affiliated Participants Votes**
Impractical for operation in highway work zones	12	12
Unpleasant experience with devices	4	5
Influence on the performance such as vision obstruction	18	21
Slow and painful adaptation to the technology as a routine	8	8
Unreliability of the devices in identifying potential dangers	25	24
Repetitive false alarms and loss of your trust in devices	23	29
None	24	7
Other	6	13

two endpoints increased, the IQR and the average of RTTs increased. It means that in farther distances, the latencies are seemed to become both less effective and less reliable. However, the collected RTTs within the 120-meter coverage distance provide a promising foundation for a timely notification to workers from the board in highway work zone setting.

2.5.4 Mixed-method User Research and Analysis

The collected responses from our participants to questions a to d are summarized in Figure 2.10. Moreover, Figure 2.11 visualizes the opinions of our participants about the strengths and weaknesses of the designed interface. Finally, their concerns about the proposed technology are summarized in Table 2.6. For better representing the collected data, Table 2.7 provides some central tendency parameters for questions a to d. Since the collected data in these questions are ordinal, we assigned numeric values to each category from 1 to 5, each one corresponding to a category from the farthest left category to farther right category in order. Then, we identified the first quartile (Q1), median (Q2), third quartile (Q3), and interquartile variability (IQR) of the collected data from both groups, separately.

In the next step, we investigated whether (1) the sample data come from a population with a uniform distribution over the considered categories and (2) affiliated participants and maintenance crew always show different views. To this end, we used both goodness of fit and independence versions of Chi-square test[113]. The results

Table 2.7: Statistical Summary of Responses from our Participants to the Questions a to d

Question	Group	Mode	Q1	Median (Q2)	Q3	IQR
Practicality of the framework	Maintenance Crew	Neutral	3	3	4	1
	Affiliated Participants	Practical	3	4	4	1
Likelihood of using the framework	Maintenance Crew	Neutral	3	3	4	1
	Affiliated Participants	Practical	3	3	4	1
Likelihood of recommending the framework	Maintenance Crew	Neutral	3	3	4	1
	Affiliated Participants	Practical	3	3	4	1
First impression of the UI	Maintenance Crew	Neutral	3	3	4	1
	Affiliated Participants	Practical	3	4	4	1

of these tasks are summarized in 2.8.

Table 2.6 indicates that affiliated participants' main concern about our proposed technology is repetitive false alarms whereas maintenance crew are more concerned about the unreliability of devices. In addition, the participants who chose Others in Table 2.6 mentioned that they were mostly concerned about the durability, comfort, integrity and cost of the technology. They also iterated that they had some concerns

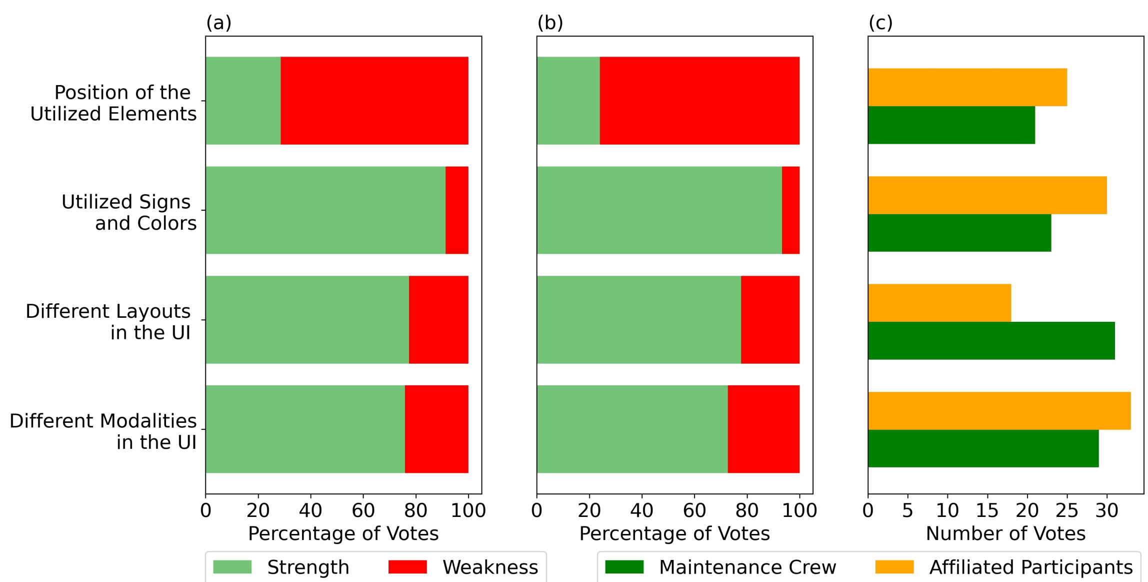


Figure 2.11: Strengths and Weaknesses of the Designed UI - (1) The Percentage of Maintenance Crew Votes (2) The Percentage of the Affiliated Participants Votes (3) Total Number of Votes

Table 2.8: The Results of Chi-square Test in Questions a to d

Question	Groups	Goodness of fit p-value	Independence p-value
a	Maintenance Crew	$4.17\text{x}10^{-13}$	0.0061
	Affiliated Participants	0.00018	
b	Maintenance Crew	$4.68\text{x}10^{-9}$	0.384
	Affiliated Participants	$5.55\text{x}10^{-6}$	
c	Maintenance Crew	$6.96\text{x}10^{-16}$	0.13
	Affiliated Participants	0.00016	
d	Maintenance Crew	$1.32\text{x}10^{-15}$	0.023
	Affiliated Participants	$1.47\text{x}10^{-7}$	

toward the reliability of the proposed framework in emitting timely notifications and whether it could successfully notify them in a timely manner when they are engaged in labor-intensive tasks such as drilling. They also highlighted their lack of prior experience with such technologies and technical challenges such as federal and state level safety requirements and prescription glasses as other challenges and concerns of theirs. Moreover, Table 2.7 demonstrates that the IQRs in responses are relatively low. In other words, the responses from both groups are fairly cluttered around the median. Furthermore, Table 2.8 shows that for questions a to d, we have enough statistical evidence to conclude that the null hypothesis of the goodness of fit Chi-square test is rejected. This means that the samples do not come from uniformly-distributed populations, and the differences between the number of responses in different categories are statistically significant. On the other hand, the test of independence reveals that while affiliated participants and maintenance crew have statistically different opinions about the likelihood of them using and recommending the framework to others, they might have dependent opinions in terms of the practicality of the framework and their first impression of the designed UI. This negates our first assumption in separating the responses of affiliated participants and maintenance crew in question a and d. Therefore, we re-performed the IQR and the Chi-square goodness of fit test for the cumulative responses of both groups for questions a and d. In this case, we again found a relatively low IQR of 1 with the mode and median of "Neutral" in both questions. In addition, the results of Chi-square goodness of fit tests in these questions are $9.11 * 10^-13$ and $5.99 * 10^-19$, respectively. It means that the null hypothesis of the goodness of fit Chi-square test is still rejected in both questions and the differences between the number of votes in each category are statistically significant.

Overall, our results show that the highway maintenance and operation community seems to be fairly interested in this framework and the designed interface. Considering the results of Chi-square test and the central tendencies in samples, we can conclude that in almost all of the questions, the majority of maintenance crew seem to feel neutral about the framework and the affiliated participants feel more positively. When we couple these results with the concerned mentioned in Table 2.6, it can be discussed that our end-users' major problem is with their perception about the reliability and functionality of the technology, rather than its practicality. Our results also demonstrate that affiliated participants are more likely to use our technology while maintenance crew staying neutral. This can be attributed to either their technology resistance or lack of prior experience with such devices.Figure 2.11 also indicates that even though the number of votes for choices were not identical, a similar trend was observed in both groups regarding the strengths and weaknesses of the designed interface. In specific, both groups believed that utilized signs and colors, multimodality of the designed notification mechanism, and different layouts for different risk levels are the major strengths of the proposed UI. These results prove that our considered guideline to set MUTCD as a resource for choosing colors and signs was on point, and this document can be considered as a suitable starting point for leveraging metaphors in the design in order to increase the level of familiarity of the end-users with the interface.

2.5.5 Overall Latency Measurement

Based on early results and evaluation, the proposed framework seems to be able to meet the tight latency requirements of emitting timely and effective notifications in highway work zones. The work by Sternberg [114] suggested that the reaction time of human beings under normal condition varies between 0.5 up to 1.5 seconds. Under the assumption that the anomalous cars travel at the speed of 70 MPH and 90 MPH, detecting cars every 30 and 50 feet, and a 10-second notification time for the highway workers, the FPS latency only varies between 2.05 and 4.4 FPS. That leaves us with the need of detecting cars from 1320 feet ahead of the work zone. Our results show that the execution latency of the developed AI model was 40.3 ms on a Xavier AGX Jetson embedded board. In addition, we measured the real-time end-to-end communication latency between the embedded platform and AR smart glasses to be 5.1 ms at the farthest distance. With that and given the aforementioned assumptions, we can conclude that our framework provides promising results in terms of real-time vehicle detection and communication. Comparing the given 10 seconds for workers to respond and the calculated RTT latencies, we also can conclude that the framework can emit multiple timely notifications to workers seconds ahead of the intrusion if needed. Workers, needing 1.5 seconds for reaction, can, in turn, respond to the danger properly.

2.5.6 Risk Assessment

As discussed earlier in this chapter and also highlighted by [115], multiple factors could contribute to causing risks to highway workers. This includes (1) vehicle's type, speed, and distance from the work zone, (2) accelerating characteristics of approaching vehicles, (3) vehicles' trajectories including drifting, swerving, and weaving, (4) maneuver pattern of vehicles before and around work zones, (5) work zone type, and (6) weather and lighting conditions. The proposed vehicle detection from the distance approach enables identifying and estimating risks associated with the incoming traffic. However, the extracted information should be augmented by the real-time feed from other contributors such as the real-time location of workers, geometrical properties of the work zone, and weather conditions. In this chapter, we primarily focused on the foundations of worker-in-the-loop safety systems for highway work zones. The risk estimation will be part of our future work which demands additional data collection from work zones with diverse properties. This also includes further research work for improving the current AI model for an accurate risk estimation with respect to the different types and number of incoming vehicles and corresponding traffic status.

2.6 Conclusion

This article conceptualizes and co-designs an integrative framework levering Artificial Intelligence (AI) and Augmented Reality (AR) to address the highway work zone safety concern. To this end, it presents and investigates three major pillars of the proposed framework which are: (1) AR user interface design for multimodal notification, (2) real-time deep learning for vehicle detection/classification from distance, and (3) real-time wireless communication. This article also presents the results of an early mixed-method user research that investigated end users' perception toward the proposed framework and the conceptualized interface through a cognitive walkthrough using a low-fidelity prototype. Overall, the early results demonstrate that the trained AI model achieved 48.7% mAP for detecting vehicles from distance with 24.83 Frame Per Seconds (FPS) execution latency on Nvidia Jetson Xavier embedded platform. The outcomes also indicate that the real-time execution and communication latencies combined are within 46 ms margin on average, which provides the foundation for emitting on-spot notification to highway workers and enabling them to show a timely reaction to the identified dangers. The early user research also reveals that the proposed safety framework and the designed interface were positively welcomed by the body of the highway maintenance and operation community.

CHAPTER 3: MIXED-METHOD USABILITY INVESTIGATION OF ARROWS: AUGMENTED REALITY FOR ROADWAY WORK ZONE SAFETY

3.1 Abstract

The emergence of novel technologies has provided promising safety applications. However, in the context of highway work zone safety, there is a need for user research to facilitate designing usable, worker-centered technologies. This paper investigates the usability of AI-enabled Augmented Reality (AR) and Wearable Technology (WT), and documents design opportunities and challenges in worker-centered AR/WT development of such technologies for the highway workforce. To this end, we designed a mixed-method approach and leveraged a high-fidelity prototype mimicking the interaction that a designed AR/WT-based framework offers. At the same time, we used different measurements to quantify usability and document some other pertinent design metrics. Our results suggest strong worker acceptance and perceived safety benefits of the AR/WT-backed system and highlight a critical need for developing novel technologies that secure workers' safety while accommodating their unique needs. Moreover, participants rated the usability of the technology as slightly above average while requiring reasonable mental effort. The outcomes demonstrate a significant correlation between perceived trust and usability, reinforcing the need for worker-centered design guidelines.

3.2 Introduction

Highway work zones play a critical role in inspecting, maintaining, and upgrading roadway infrastructure. However, speeding and careless driving, night shifts, and limited maneuver space have long created a dangerous situation that makes workers vulnerable to serious and fatal injuries. According to the Centers for Disease Control and Prevention, between 2003 and 2017, 1,844 workers lost their lives at road construction sites, averaging 123 causalities per year [1]. Furthermore, many workers have suffered life-changing severe injuries and other mental or physical health issues attributed to drivers intruding into their working environment. The current practice of safety in highway work zones is typically limited to portable signs, flaggers, directional alarms, warning lights, and reactive alert technologies [2, 3]. Such systems usually activate only after an intrusion has happened or is imminent and do not provide enough reaction time for workers to properly respond to the upcoming dangers, leaving workers vulnerable to intrusions. Meanwhile, the latest technological advances in Augmented Reality (AR) and Wearable Technology (WT) have paved the way for tackling complex real-world safety problems. In recent years, the workforce safety domain has witnessed transformative efforts toward developing the next generation of safety systems by leveraging AR and WT in different contexts and applications [4, 5, 6]. However, the majority of previous studies - especially in the highway construction, maintenance, and operation sector - have mainly focused on how such technologies can be designed in a holistic view with limited attention to user experience. These studies often lack accompanying user research outcomes and usability benchmarks that contextualize users' preferences and perspectives on the proposed systems [116, 117]. This has resulted in user experience and usability being less investigated across products and services for this domain, which has contributed to the slow progress of AR/WT for worker safety in the transportation industry despite its documented potential and benefits in safety research [118, 119, 120]. Limited user research on workers-AR/WT interaction coupled with restricted user information that stems from few user studies conducted on highway workforce could potentially hinder future AR/WT-backed innovations [2, 7]. At the same time, recent trends in Human-Computer Interaction have further highlighted the importance of human-centerd technology design and how it contributes to the future of interactive systems [121, 122]. Therefore, extensive user research that facilitates designing worker-centerd, usable, and practical solutions and promotes workers' unique needs has become further critical to the future of work and safety systems [9, 123]. In response, this paper chronicles our efforts in studying the end-to-end usability and functionality of ARROWS â Augmented Reality for Roadway Work Zone Safety â and documents design opportunities and user experience contributors from workers' perspectives. For this purpose, we designed a two-step mixed-method methodology to study the usability of the safety system proposed by the authors [124]. We designed this approach to both quantitatively and qualitatively study usability, and identify the key contributors to usability, user experience, and trust. To this end, we leveraged the Wizard of Oz technique in our research approach to create a "wizardâ and supplemented that with a high-fidelity prototype to emulate the end-to-end interaction of the system for participants. Our methodology featured two complimentary experiments that were conducted in two different settings, indoor and outdoor. The indoor experiment was centerd around qualitative data collection and was administered in one of the truck stations of the Minnesota Department of Transportation (MnDOT) with participation from 13 experienced highway workers. The outdoor experiment was designed around quantitative data collection and was conducted in a temporary work zone where 30 participants actively engaged in a routine highway maintenance task. In both steps, the user journey included direct interaction with the AR and WT hardware components of the designed prototype through a pre-devised mock-up scenario executed through our programmed "wizard". The participants were then asked to express their perspectives through standard usability measurements such as System Usability Scale (SUS) [125], Rating Scale Mental Effort (RSME), and trust questionnaires that we used to benchmark usability, study user experience, and identify design opportunities and challenges. This study contributes to the body of knowledge in several ways. Primarily, it stands out as one of the first studies to examine the usability of an innovative AR/WT-backed system specifically designed for safety in highway work zones, integrating the perspectives of highway workers. The developed system, Augmented Reality for Roadway Work Zone Safety (ARROWS), is at the forefront of safety solutions for highway work zones. It leverages low-latency connections to proactively communicate safety risks to workers through multimodal notifications enabled by AR and WT. By providing usability benchmarks and identifying user experience factors, this paper offers valuable and user-centerd insights for those intending to implement AR/WT solutions for the highway workforce. This was made possible by executing a robust two-step mixed-method research plan that includes both indoor and outdoor settings. In addition, the qualitative feedback acquired from experienced highway workers regarding their use of ARROWS reveals the unique and often overlooked needs of highway workers. These findings could fuel further innovations in AR/WT applications within the transportation infrastructure maintenance and operation industry. Ultimately, this study promotes further exploration and improvement in advanced worker-centerd safety systems, bringing us closer to creating safer and more accommodating work environments for highway workers.

3.3 Background and Context

This section begins with exploring the related works by looking at the history of advanced technologies in the context of safety and highway workers. Then, we investigate the importance of usability studies in technology development and how the Wizard of Oz approach has been utilized in similar studies.

3.3.1 Highway Workers, Wearable Technology and Augmented Reality

In recent years, research communities across disciplines have witnessed a flow of efforts in developing new technologies that could solve some of their complex unanswered problems [126, 127, 128, 129]. Safety has been among the topics that have attracted a great deal of attention from researchers and practitioners in different domains. To this end, the application of AR/VR has been studied in different contexts, and topics from fire safety to the safety of cyclists [130, 131, 132]. Similarly, building construction safety has been among the disciplines that researchers have focused on in recent years, where researchers investigated how AR and WT technologies can be leveraged in reducing environmental risks imposed on workers [133, 134, 135, 136]. However, developers and practitioners in highway maintenance and construction domains have not kept up with other disciplines and have been left behind in terms of technological advances and research efforts.

Recently, some research activities have targeted the application of AR in highway construction, with a greater focus on infrastructure and management [137, 138]. However, safety has still been underrepresented in the literature, especially for the maintenance and operation section. To this date, the current safety of practice in highway work zones is only limited to reactive technologies, and the lack of advanced and predictive systems is quite noticeable [36, 12, 38]. After systematically reviewing 147 papers, [2] highlighted the importance of departing from old-fashioned technologies toward smart advanced technologies and modernizing the safety practice of highway work zones. To this end, in recent years, the community has seen new research efforts that have proposed novel technologies for highway work zone safety [117, 139]. However, what they lack is appropriate user research that investigates the interaction between highway workers and AR/WT technologies, identifies the unique needs and preferences of workers, and highlights design guidelines and directions for worker-centered technology design. Such information will be vital in the smooth transition and adoption of these technologies in the future [140, 141].

3.3.2 Usability Test and Technology Development

Usability has been established as a focal factor in product development. Although usability alone is not enough to guarantee/strengthen the user-product relationship, it is widely accepted as an important player in the success of a product or interface [142]. Usability has traditionally been defined as a quality feature that examines how easy it is to use a given product. The word "usability" itself refers to the mechanisms that could improve ease of use during the product design process [143, 144, 145]. The importance of early and often iterative usability tests in discovering potential bugs and continuous improvement of the product has long been highlighted in different studies [146, 145, 147]. To this end, several researchers and practitioners have published their efforts to evaluate the usability of their novel products [148, 149, 150]. With the rapid evolution of AR and WT and their current widespread application, several researchers have studied the usability of such devices in different populations to address the growing need for user-centered design guidelines and best practices for such technologies [151]. For example, [152] conducted a usability study to examine the feasibility and best practices of AR for Basic Life Support (BLS) and Defibrillation training purposes. In another study, [153] explored how different types of head mounted display (HMD) and user interface designs could affect perceived workload, usability, visual discomfort, and job performance of workers during a simulated warehouse job. In yet another study, [147] leveraged a mixed-method usability test to provide a series of principles that can elevate the design of AR applications based on interactive mobile applications to learn basic English in early childhood. The same trend can also be observed in the WT landscape and other novel technologies [154, 155].

3.3.3 Wizard of Oz Methodology

The Wizard of Oz (WOZ) technique is and has been one of the most widely used tools in the toolkit of Human Computer Interaction researchers. In this methodology, a person or a predefined experimenter remotely takes control of the system and its real-time performance, creating the illusion of a working system [156]. In recent years, this methodology has been widely used in different disciplines and topics. For example, Faas et al. [157] studied the application of an external human-machine interface that displays the automated driving mode and how it affects the response of pedestrians to different types of drivers (attentive, tinted and distracted) using the WOZ methodology. In another study, Palmeiro et al. [158] investigated pedestrian crossing decisions in the presence of an autonomous vehicle compared to a traditional vehicle using WOZ and a within-subject experiment. This technique has also been used in several studies in the context of AR and WT. For example, Alce et al. [159] provided a pilot study of a wearable AR interaction framework that simulates an AR city tour. The authors in the pilot study used WOZ to collect 21 participants' data to analyze precision, relevance, responsiveness, technical stability, visual fidelity, general user experience, and human operator performance of their designed system. In another study, Billah et al. [160] explored the application of using off-the-shelf and available smartwatches that are paired with smartphones to help visually impaired people read and write papers and checks. For this purpose, they devised a WOZ experiment that included different custom interactions and documented their results.

3.4 Methodology

This section will begin by explaining the prototype design that will be utilized in the upcoming usability test. Following that, we will delve into the details of the designed surveys, procedures, and other specifications of the conducted usability test.

3.4.1 Design and Functionality of ARROWS

3.4.1.1 System Functionality

The comprehensive functionality of the ARROWS involves the integration of a camera and a GPU edge device, powered by deep learning. This combination actively monitors incoming vehicles, assesses their speed, predicts their trajectory, and determines their proximity to the work zone. The system leverages this data to predict potential work zone intrusions and assigns a risk score accordingly. This score then

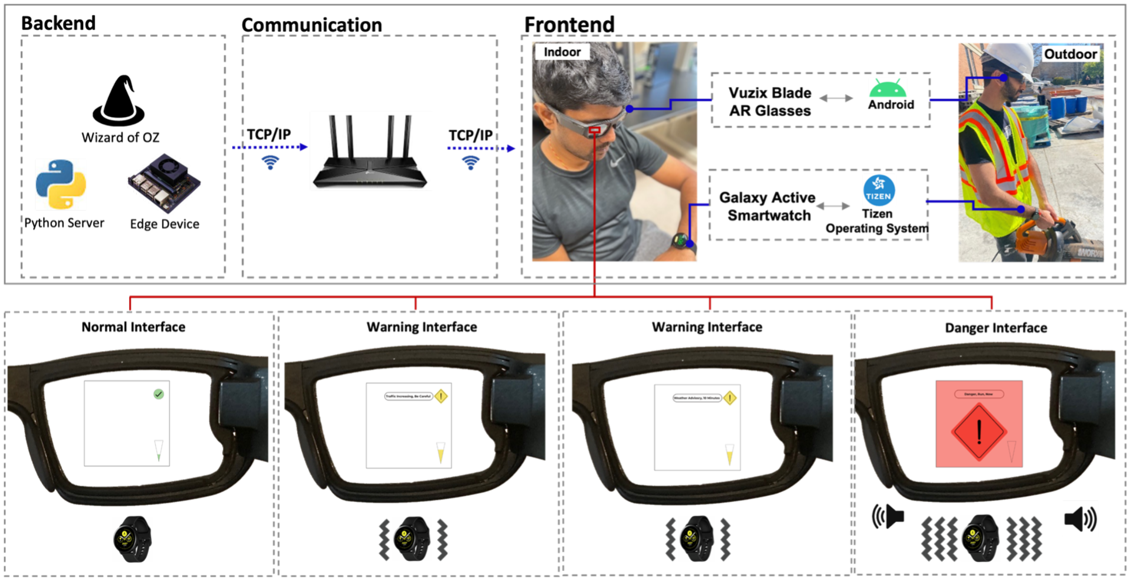


Figure 3.1: The elements of the utilized high-fidelity prototype of ARROWS and its corresponding user interface design

triggers the generation of real-time safety notifications appropriate to the identified risk level. Workers within the work zone will be equipped with a pair of AR glasses, which serve as the interface for delivering visual notifications and audio warnings. In addition, the system employs a smartwatch to deliver haptic notification, which varies in intensity based on the assessed level of risk. The connection between the AR glasses, smartwatch, and the backend GPU is facilitated through a local network. This framework uses three distinct modes of notifications to intuitively convey the level of risk to the workers: Normal Interface Warning Interface, and Danger Interface, as shown in Figure 3.1, Each layout is triggered based on the risk score value. For instance, a vehicle approaching above the speed limit and on a trajectory predicted to intrude into the work zone would trigger a Danger Interface. Through a comprehensive combination of visual, auditory, and haptic notifications, ARROWS is designed to enhance safety within work zones, ensuring workers are consistently aware and responsive to potential hazards.

3.4.1.2 Prototype Design

Conducting a comprehensive evaluation of a concept is an essential and fundamental step that should be undertaken before initiating the process of prototyping. This step serves as a critical foundation to ensure that the concept is thoroughly assessed, analyzed, and understood from various perspectives. This process involves conducting in-depth analyzes to assess the concept's feasibility, viability, and potential limitations from technical standpoints. For this purpose, a comprehensive evaluation of ARROWS technical requirements and challenges has been conducted before this study [124]. This included exploring communication protocols and evaluating latency benchmarks, creating an AR/WT interface that minimizes vision obstruction for end-users, selecting appropriate technologies that adhere to safety requirements, and actively involving potential end-users to obtain preliminary feedback. Engaging with end-users throughout the design and development stages allowed for a deeper understanding of their needs, preferences, and pain points. This constant communication has resulted in multiple rounds of improvements and conceptual design refinement in ARROWS [124]. After finalizing the design, we developed a high-fidelity prototype of ARROWS and used the Wizard of Oz methodology in lieu of AI backend to mimic comprehensive coverage of ARROWS' capabilities during the testing phase, particularly under conditions of unlikely intrusion. By adopting this approach, we successfully developed and implemented the complete end-to-end functionality of ARROWS, which was then subjected to rigorous testing in this study. This includes the entire process of transmitting notification commands from the backend, receiving them in the frontend, and visualizing the corresponding risk scores using the ARROWS interface based on the received risk score. For this purpose, we programmed a "wizard" to take over the backend and create an end-to-end functionality of ARROWS. The programmed wizard enabled the prototype to emulate the intelligence and features of the real-world AI-powered backend in the controlled environment where this study was performed. We used NVIDIA Jetson AGX Xavier as the wizard hardware in this study . We programmed a server script using TCP/IP protocol in Python programming language for this hardware, alongside the mock-up scenarios, and leveraged this for the communication between the server and frontend. The frontend of the system consists of a pair of AR smart glasses and a smartwatch, as illustrated in Figure 3.1. The Vuzix Blade AR smart glasses and Galaxy smartwatch were utilized as the frontend hardware for this prototype. They operate in conjunction to provide real-time multimodal notifications, which include vibration through the smartwatch, and audio and visualization through the AR glasses. These notifications are based on the preprogrammed scenarios and the instructions provided by the wizard in real-time. We selected the Vuzix blade because of its compatibility with the ANSI Z87.1 certificate, prescribed glasses/shades friendliness, and built-in audio modules, which make it a favourite option for industrial applications. We incorporated our design of real-time notifications in the display located on the right lens of the Vuzix Blade AR glasses with a marginal impact on the workers' natural vision. As shown in Figure 3.1, the AR glasses only contain a display on the right lens, and the left lens is completely see-through. In our user interface design, the three first examples that represent normal and warning notifications (normal and warning interfaces) feature a transparent background. However, the danger notification has a semi-transparent background to increase the effectiveness of the warning. This danger warning will only appear in less likely dangerous situations and will maintain sufficient transparency, allowing users to navigate without obstructing their vision. Because of this user interface design and AR glasses configuration, the end-user will constantly engage with a fully transparent background on the left lens, and in most cases, they will also interact with a completely transparent display on the right lens, without their natural visions being obstructed. We used TCP/IP protocol for real-time communication and developed the pertinent software in Android. For the smartwatch, we developed the required software in the .NET framework and implemented similar communication protocols and software patterns.

3.4.2 Experiment Design

Our study was driven by three primary objectives: (i) quantifying the usability of ARROWS, (ii) identifying the major factors affecting user experience from the worker's perspective, and (iii) exploring the significant pain points of integrating ARROWS in the everyday practice of highway workers. Our experimental strategy was designed with the aim of acquiring a comprehensive understanding of the challenges and requirements associated with designing AR/WT-focused solutions for the highway workforce. Moreover, by engaging workers from diverse backgrounds and experiences within the construction industry, we aimed to establish a robust foundation for usability. This inclusive approach allows us to gather data-driven insights while exploring and studying the user experience of ARROWS for highway work zone safety applications.

To achieve our defined goals, we developed a research strategy that utilized a dual-experiment mixed-method design. This approach involved conducting two separate experiments, each with its own unique objectives, and carried out in indoor and outdoor settings. The indoor experiment was designed to have an increased focus on collecting qualitative data from experienced highway workers who had extensive backgrounds in highway maintenance and operation. Its primary objective was to identify the unique safety needs of highway workers and conduct a rigorous thematic analysis using high-quality data from end-users with an extensive field background. The purpose of the outdoor experiment was to create an additional quantitative platform for conducting usability testing in an outdoor setting that resembles real-world highway work zones, with participants actively involved in maintenance activities. Our goal was to broaden participation by including individuals with construction backgrounds. The outdoor experiment also provided additional data points for numerical analysis and benchmarking of the technology's usability, RSME, and trust score.

3.4.3 Indoor Experiment

3.4.3.1 Test Procedure and Participants

We conducted the indoor experiment in partnership with the Minnesota Department of Transportation (MnDOT). Overall, 13 individuals ($N = 13$), 12 male and one female, participated in the experiment. The sample size in this study provides a suitable basis for an early usability test and is in line with domain-specific usability experiments conducted in the literature [90, 161]. The average age of the participants was 50.83 ($SD = 11.12$) years old. It should be noted that one participant did not disclose their age, and the missing data point was replaced with the average of other values. In addition, on average, our participants had 9.50 ($SD = 6.01$) years of maintenance work experience. The test involved the participation of workers with various job titles, such as general laborers, sign maintenance personnel, traffic controllers, and transportation experts. The methodology of the indoor experiment consisted of three main steps: Pre-Experiment Surveys, the Wizard of Oz Experiment, and Post-Experiment. When applicable, we used 6-scale Likert questions throughout our surveys. The study protocol was also reviewed and approved by the Institutional Review Board (IRB No. 21-0357) of the University of North Carolina at Charlotte. The usability test was conducted at a MnDOT truck station located in the Minneapolis metro, and the duration of each test ranged from 45 minutes to an hour. In the following sections, we will provide a detailed description of each step involved in the testing process.

3.4.3.2 Pre-Experiment Survey

The initial survey was developed to gather basic background information about the participants before the test. This survey consisted of (1) age, (2) work experience, (3) self-reported tech-savviness, (4) level of safety concerns in highway work zones,

(5) perceived helpfulness of the current safety systems, (6) perceived practicality of AR/WT in highway work zones and (7) likelihood of using AR/WT for safety purposes. These features have been proven to be important in accepting new technologies and improving usability [161, 162, 163, 164, 165].

3.4.3.3 Wizard of Oz User Journey

As specified earlier, we leveraged WOZ methodology to mimic the real-world interaction of the proposed system in a controlled environment based on a mock-up scenario programmed in the backend (i.e., the wizard). Each participant interacted with the technology for 3-5 minutes. During this interaction, a series of notifications, as illustrated in Figure 3.1, were sent to the frontend, and participants were asked to think aloud and share their experience with the administrator. The interaction started with the normal interface for 50 seconds. As illustrated in Figure 3.1, this interface consists of a triangle on the bottom right and a checkmark on the top right corner of the AR display. Then, the first warning interface with the textual content of "Traffic Increasing - Be Careful" showed up for 5 seconds and was accompanied by a light vibration stimulus on the smartwatch. In this interface, a yellow warning sign was rendered on the top right corner of the display and the color of the bottom right triangle changed to yellow. Next, the normal interface for the next 20 seconds was displayed. Then, the second warning interface with the textual interface of "Weather Advisory - 10 Minutes" with a light vibration stimulus was delivered for another 5 seconds. The display went back to the normal interface for another 20 seconds. Finally, the danger interface with the textual content of "Danger - Careful - Now" appeared and that was combined with both strong vibratory and auditory modalities for 5 seconds. In this interface, the background color of the display changed to transparent red, and a large red warning sign was also rendered in the center of the display. The normal interface was used for the remainder of the user journey. Moreover, participants were asked to highlight the pros and cons of the system, as well as its appropriateness for the context and their inputs on the utilized devices (smart glasses and smartwatch) as the interaction went on. The scenario was similarly executed for all participants. We also included a cognitive walkthrough in this step where the research team explained to participants how the concept of this system would use AI for detecting and predicting safety risks in highway work zones when fully developed.

3.4.3.4 Post-Experiment Surveys

Following the interaction, we administered the Rating Scale Mental Effort (RSME) to measure the subjective mental workload. First introduced by Zijlstra and Van Doorn [166] and Zijlstra [167], RSME has been considered one of the widely used subjective tools in measuring mental workload [168, 169]. Next, we used System Usability Scale (SUS) to quantify the usability of the proposed technology. SUS was first introduced by Brooke [125] and ever since has been used in numerous studies for a quick and easy assessment of the usability of a given product or service [170, 171, 145, 172]. We also included a subjective trust questionnaire to benchmark the perceived trust of participants. We also collected the likelihood of using AR/WT technologies after the experiment to conduct a statistical comparison to determine if there was a significant change in their attitudes toward these technologies after the usability test.

3.4.4 Outdoor Experiment

3.4.4.1 Test Procedure and Participants

We created a short-duration work zone on the University of North Carolina at Charlotte campus to carry out the outdoor experiment. For this purpose, we followed the specifications of this type of work zone as provided by the Manual on Uniform Traffic Control Devices (MUTCD) [92]: a temporary working area for maintenance activities that takes less than an hour according to the. We also designed a real-world routine maintenance task commonly performed by highway workers in short-duration work zones and included that in the experiment. The task involves removing leaves from drop inlets. This routine maintenance activity helps prevent flooding and facilitates drainage. The setup and experiment specifications are illustrated in Figure 3.2. Of the 34 participants who participated in the study, 4 were unable to complete the experiment due to unexpected circumstances, including unexpected fire drills and internet outages which disrupted the data collection process. Therefore, the study analyzed data from 30 participants ($N = 30$), 20 men, and 10 women, who completed the experiment. These participants had an average age of 25.93 ($SD =$ 5.20) years. On average, they reported having 2.58 ($SD = 2.70$) years of experience in the construction industry. Participants also rated their self-reported tech savviness on average at 5.03 ($SD = 0.74$) on a 6-point Likert scale in response to a question about their tech proficiency. The test process lasted between 30 and 45 minutes for each participant, during which they were instructed to wear AR glasses and the smartwatch and perform the leaf blowing task to clean a drop inlet within the work zone, as shown in Figure 3.2. The experiment was carried out under regular lighting conditions between 11 am and 4 pm. We kept both the weather and the lighting conditions in a similar range throughout the experiment. Similarly to the indoor experiment, the outdoor experiment was also divided into three main stages: The Pre-Experiment Survey, the Wizard of Oz Experiment, and the Post-Experiment Survey. The study was carried out with the approval of the institutional review board (IRB No. 21-0357) of the University of North Carolina at Charlotte.

3.4.4.2 Pre-Experiment Survey

Similar to the indoor experiment, we collected (1) age, (2) experience, (3) safety concerns, and (4) self-reported tech-savviness as part of the pre-experiment data collection in the outdoor experiment.

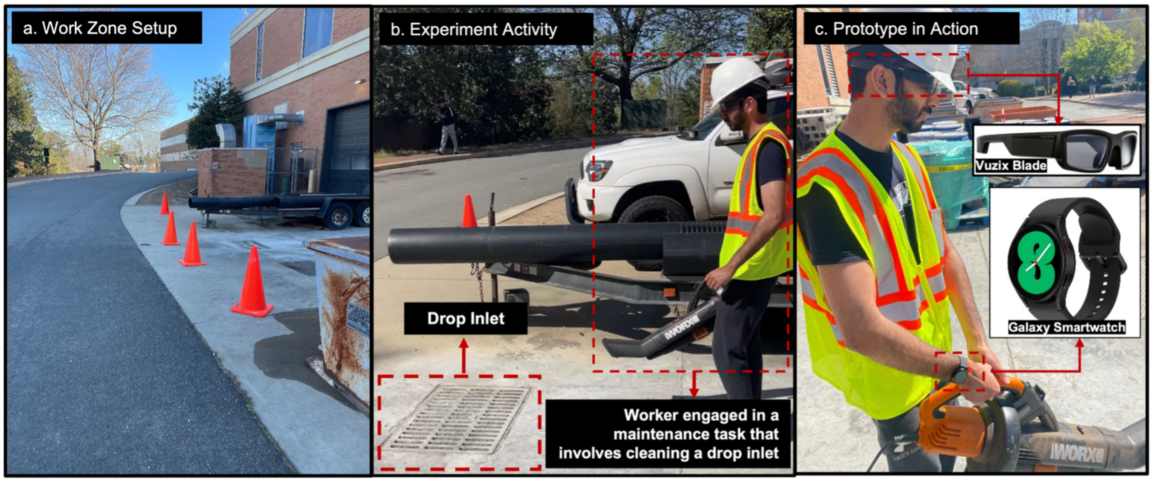


Figure 3.2: Details of the outdoor experiment: (a) temporary work zone configuration (b) the designed physical activity included in the outdoor experiment (c) prototype in action during the outdoor experiment

3.4.4.3 Wizard of Oz User Journey

We used a similar user journey that was used in the indoor experiment. Each participant interacted with the technology for 3-5 minutes. At the same time, participants were also instructed to engage in the physical activity that we designed as well as shown in Figure 3.2 The interaction started with the normal interface lasting for 50 seconds. Next, the first warning interface displaying the textual content of "Traffic Increasing - Be Careful" was delivered for 5 seconds, accompanied by a light vibration on the smartwatch. Next, the normal interface for the next 20 seconds was displayed and was followed by the second warning interface with the textual interface of "Weather Advisory - 10 Minutes" with a light vibration stimulus for another 5 seconds. The screen went back to the normal interface for another 20 seconds. Finally, the danger interface carrying the textual content of "Danger - Careful - Now" appeared on the display in synchrony with both strong vibratory and auditory cues for 5 seconds.

3.4.4.4 Post-Experiment Surveys

Similarly to the indoor experiment, we used RSME, SUS, and Trust to quantify usability and study user experience. After the experiment was ended, we immediately administered the questionnaires and instructed the participants to fill them out.

3.5 Results and Discussion

In this section, we will analyse the results of our experiments. We will first explore the necessity and challenges associated with augmented reality technologies. We will then focus on the implications of usability, mental load, and trust on user experience design, highlighting the importance of these factors in ensuring the successful adoption and use of these technologies. Finally, we will conduct a correlation analysis to examine the link between usability, trust, and mental load, shedding light on the interplay between these attributes. Through these discussions, we aim to provide valuable insights for researchers, practitioners, and policymakers seeking to enhance the effectiveness and usability of AR/WT-based safety technologies.

3.5.1 Necessity and Challenges of AR/WT-based Safety Technologies

Figure 3.3 summarizes the responses to the questions asked in the Pre-Experiment Surveys of the indoor experiment. As shown in Figure 3.3(a), the results suggest that our participants were quite concerned with their safety in highway work zones. Furthermore, when asked what their concerns were in specific, our participants highlighted speeding traffic, and careless and distracted drivers as their major issues. They also gave the helpfulness of current safety strategies in highway work zones on average a score of 4 out of 6, as illustrated in Figure 3.3(b). Furthermore, the participants in the indoor experiment also expressed their concerns about the existing safety mechanisms in highway work zones and pointed out the need for improved and updated safety technologies in this context. Our qualitative data, as quoted in the following, points out the same trend:

Table 3.1: Participants' concerns toward the application of AR/WT in highway work zones expressed in the indoor experiment

Concern	**Percentage of Participants**
Incompatibility with tasks and environment	69%
Affecting movement and performance	61%
Lack of Trust	30%
Poor Usability	30%
Other (Dependability)	7%
Other (Adverse Weather)	7%

P1: "The more we could have for safety, I am all for it. You know, and I like to still make it home to them." P3: "The best thing that we ever found to help us above and beyond the crash trucks was having a state trooper sitting there with their flashing lights on. That seemed to get everybody's attention. But without that, all bets are off."

P8: "And there's really no safety systems in place on work, so well, our cones, that's the only thing out there is. Well, we got lights, two lights flashing, and cones. So yeah, something has got to enhance our system."

Negative perception and limited prior experience with these technologies seem to be two possible obstacles to the acceptance of AR/WT-based safety systems for this group. On average, indoor experiment participants rated their self-reported tech-savviness to be 4.23 on a 6-score Likert question with a median of 4 on the same score. Moreover, as shown in Table 3.1, when asked about their concerns in specific, 9 participants mentioned the incompatibility of such devices with their tasks and working environment as their major issue. Eight participants also mentioned the potential impact on their mobility as another major barrier to such technologies in the context of highway work zone safety. This was followed by a lack of trust and poor usability with four votes. It shows that customizing the AR and WT user experience to be compatible with the harsh outdoor environment while meeting the unique needs of workers coupled with easing usability are broader issues in the eyes of end-users in deploying AR and WT technologies in this context.

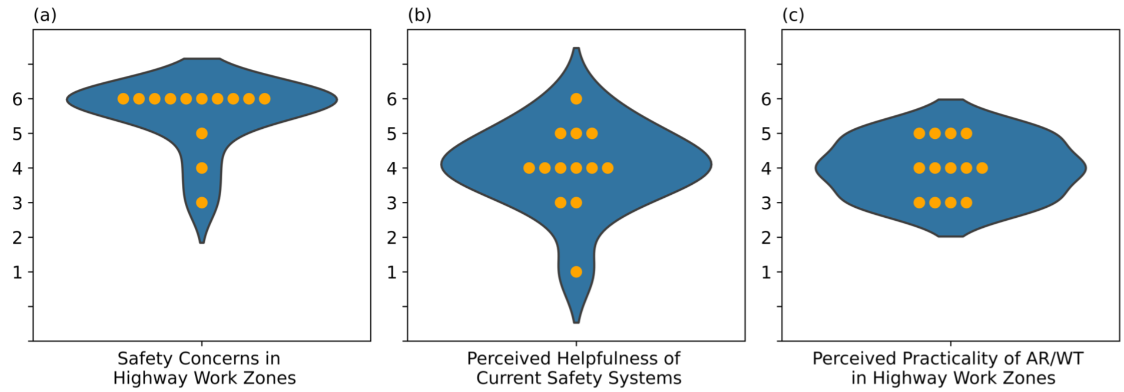


Figure 3.3: Participants' responses to the questions asked prior to the indoor usability test

3.5.2 Implications of Usability, Mental Load, and Trust on User Experience Design

3.5.2.1 Usability

Figure 3.4(a) illustrates that the average SUS score collected from the first round of this study was 75, with a *SD* of 14.47. In 2016, Sauro and Lewis [173] provided an alternative Curved Grading Scale (CGS) based on 446 SUS studies for interpreting SUS results. According to this scale, our system's usability is above the average (i.e., 68) and is given a B grade, while the traditional usability scale gives this technology a C. While this grade highlights some room for improvement, it also provides encouraging potential toward the application of AR/WT in highway work zone safety, given the novelty of the system and the fidelity of the prototype. Moreover, Figure 3.4(d) also demonstrates the SUS scores that we collected in the second round of the experiment in the outdoor setting. The collected SUS scores add up to an average of 67 ($SD = 17.85$). While the average score of outdoor SUS scores is less than the indoor experiment, the t-test results reveal that with a p of 0.13, these averages are not statistically different, and the indoor experiment resulted in statistically similar usability results. Moreover, the average of the outdoor usability metrics is around the average of 68, which indicates fairly acceptable usability. Some participants still cited their concerns about the potential incompatibility of AR/WT to their working environment after interaction with the system in our qualitative data collection as follows:

P2: "These are probably not safety glasses. They don't wrap around, they don't cover the sides of our eyes. So as nice as this part is, it has a screen that's really pretty cool. But you would have to design it so that it wraps around, you know, we have to protect all sides of our eyes because we're cutting and we got flying metal, We got all kinds of stuff. And so you got gaps underneath here."

P4: "My biggest concern is whether they fall off. When we're moving around and we are bent over doing something and then they fall off."

P5: "You get the heat, you got the humidity and then you got your body heat. So it is just about what you are wearing. I try not to wear my safety glasses too much because it fogs up because it is a concern where I cannot see."

P9: "The watch was probably one of the best because you can really feel the vibration. But also, like I said, there could be kind of hiccups with that. If you're going through and using power equipment and you don't feel that vibration because your hands are already in motion."

Therefore, our results suggest that worker-centred technologies that offer customized user experience and enhanced integration with already existing safety gears would be critical for the broader deployment of AR/WT in sensitive contexts such as safety. In recent years, there have been some research efforts toward utilizing more advanced AR hardware such as HoloLens 2 that is specifically manufactured for labour-intensive contexts. However, there remain some questions in assimilating such technologies with the everyday tasks of workers given its cost and impact on Field of View (FOV) and depth perception. In addition, even though the Vuzix Blade AR glasses that we used here comes with an ANSI Z87.1 certificate and are compatible

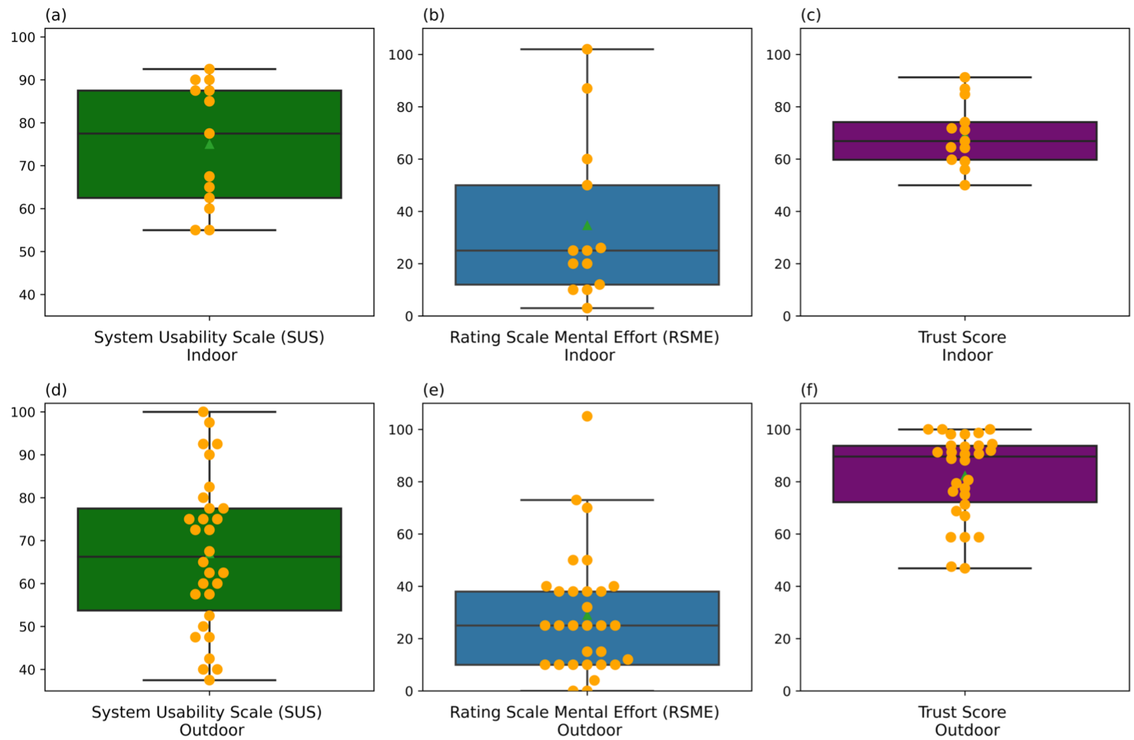


Figure 3.4: Summary of the usability (0-100), trust (0-100), and mental load (0-150) responses in indoor (a-c) and outdoor (e-f) experiments

with safety regulations in industrial settings, some of the participants still raised some concern about the suitability of replacing existing safety glasses with AR technologies.

3.5.2.2 Mental Load

Mental load is a critical factor with a significant potential impact on usability and user experience. Figures 3.4 (b) and (e) illustrate the RSME values that we received in our indoor and outdoor experiments. As shown in Figure 3.4(b), the results indicate that the average RSME is 34.62 (SD = 31.06). This suggests that the system did not cause a noticeable mental burden on participants in our indoor experiment. Moreover, the average RSME score in our outdoor experiment, as shown in Figure 3.4(e), is 28.93 (SD = 23.56), which is on par with indoor results, indicating a reasonable mental load on our participants. However, the quite large SD of this metric in both settings suggests that there could be a potential need for individual adjustments to support inclusive technology developments in terms of interaction and user interface design, given the diversity among workers in terms of task descriptions and demographic. Additionally, a t-test between these data collections in the RMSE indoor and outdoor settings shows that there is no statistically significant difference between the indoor and outdoor measurements with a p of 0.562.

3.5.2.3 Trust

We used a trust questionnaire that consisted of 8 questions and asked participants to rate their level of agreement with the questions on a 1-100 scale. In response to this questionnaire that is provided in the appendix, Figure 3.4(c) and (f) illustrate the trust scores recorded in indoor and outdoor experiments, respectively. As illustrated in Figure 3.4(c), our results suggest that our participants rated their trust in the designed technology on average to be around 69.25 (SD = 12.45) in the indoor experiment. This points out that our participants fairly seemed to be receptive to the proposed technology. While the general perception is that broader labour seems to view new technologies negatively, our participants quite welcomed this proposal. However, we still observed some differing perspectives toward AI among our participants and how it impacts their acceptance. During the indoor experiment, we asked our participants if they had heard of AI and what was their perception of it; our participants responded:

P7: "Well, personally, I would say it depends on what kind of program you are using for that AI. I mean, some people, they just think of it as a machine trying to take over the world from movies and such. But it is deeper than that. It depends on what kind of programming you use and how it's coded and what its primary function is."

P12: "Oh, just computers. You know, becoming more intelligent is the big thing with artificial intelligence kind of computers being able to almost think like a human being."

P13: "It is good and bad. It has its good sides, obviously, because it's always continually trying to think of new ways to fix things, but I don't know, maybe, maybe it's paranoia in me like everybody else, but it seems like the algorithms on your phone pick up what you are saying. It's like, who is really listening? And, you know, I don't know if that's exactly where we're going with it, but I mean, the concept is good."

At the same time, as shown in Figure 3.4(f), participants in the outdoor experiment rated their trust in the technology on average 82.29 out of 100 ($SD = 16.08$). The results of a t-test between trust scores in indoor and outdoor experiments suggest that these two means are statistically different, with a p of 0.007. The results indicate that participants had a higher level of trust in the technology in the outdoor experiment, which better simulated the context compared to the indoor experiment. This difference can be attributed to factors such as age, experience, or personal preferences, but no conclusive statements can be made based solely on the collected data.

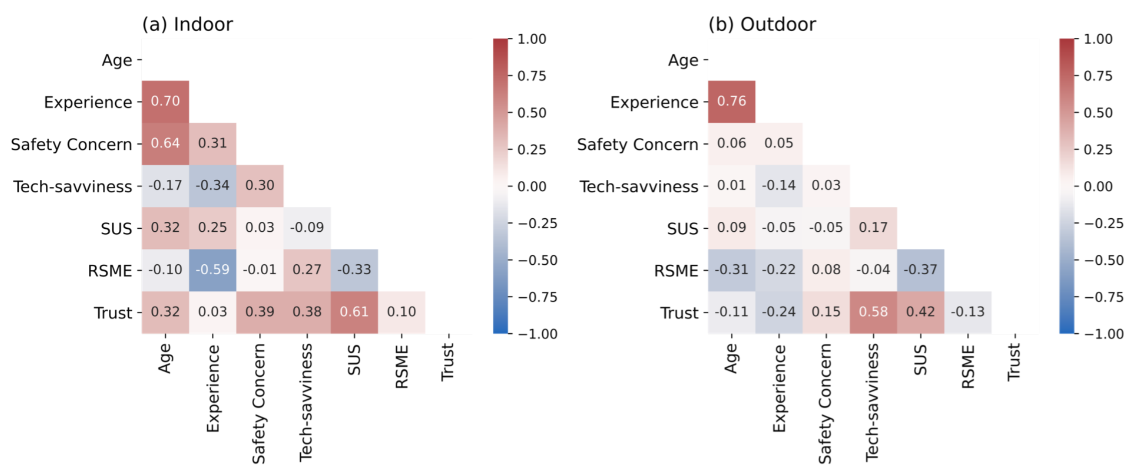


Figure 3.5: Correlation analysis between usability, mental load, trust, and demographics results collected in the (a) indoor and (b) outdoor experiments

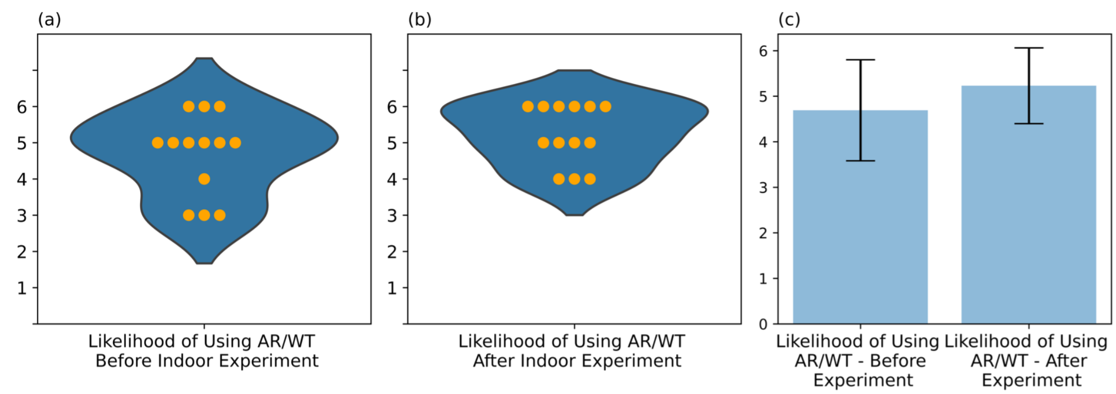


Figure 3.6: Comparison analysis between the collected likelihood of using wearable and augmented reality technologies among participants in the indoor experiment (a) before and (b) after the usability test and (c) the average and standard deviation of results

Table 3.2: Correlation and regression analysis between SUS, Trust, and RSME in indoor and outdoor experiments

Indoor Experiment								
Variable 1	**Variable 2**	**Coefficient**	**Correlation**	**Standard Error**	**t-Stat**	***p***	**Lower 95%**	**Upper 95%**
Trust	SUS	0.71	**0.61**	0.27	2.56	**0.021**	0.1	1.32
Experience	RSME	-3.05	**-0.59**	1.25	-2.43	**0.033**	-5.81	-0.29
Outdoor Experiment								
Variable 1 (X)	**Variable 2 (Y)**	**Coefficient**	**Correlation**	**Standard Error**	**t-Stat**	***p***	**Lower 95%**	**Lower 95%**
Trust	SUS	0.46	**0.42**	0.19	2.42	**0.024**	0.07	0.85
SUS	RMSE	-0.49	**-0.37**	0.23	-2.09	**0.044**	-0.96	-0.01
Tech-savviness	Trust	9.33	**0.58**	2.47	3.76	**0.000**	4.26	14.41

3.5.2.4 Link Between Usability, Trust, and Mental Load:Correlation and Statistical Analysis

To better understand the dynamic between SUS, trust, and RSME, we conducted a correlation analysis between the collected metrics for both indoor and outdoor experiments. The results are visualized in Figure 3.5. Our results suggest that Trust and SUS are highly and positively correlated in both indoor and outdoor settings. This dynamic has already been identified in different contexts and concepts in the literature [174, 175, 176]. Therefore, one of the areas that could potentially increase the usability and acceptance of AR and WT among general labour is increasing their trust in advanced technologies such as AR/WT and AI. Table 3.2 also summarizes the regression analysis on the metrics with high correlation. The results suggest that SUS and Trust correlation is significant in both settings, backing up our hypothesis (i.e., SUS and Trust implications) in both indoor and outdoor settings. Table 3.2 also demonstrates that trust and tech-savviness are positively correlated, and their correlation is statistically significant. This further highlights the role of workforce education that could result in higher acceptance of technologies as noted in other studies as well [177, 178, 179]. Therefore, increasing investments in public outreach, education, and other means of community engagement - that could potentially elevate workers' perception by increasing awareness about the application and usefulness of advanced technologies - could also increase the usability of AR and WT technologies in this context [180]. Moreover, Figure 3.6 indicates that the average likelihood of using AR/WT technologies for safety purposes among our participants after conducting the indoor usability test increased. Table 3.3 provides the results of a paired t-test on the collected values before and after conducting the indoor usability test. The outcomes of this analysis suggest that this increase is statistically significant ($p = 0.02$). Therefore, it can be concluded that more interaction and exposure to novel technologies could potentially result in higher acceptance among workers. While this alone cannot completely justify the increase in trust, nor guarantee acceptable usability, a combination of investment in public outreach and facilitating more interaction and education among workers on this topic could potentially increase the trust of the workforce in novel technologies and increase usability and acceptance.

3.5.3 Limitations and Future Directions

The proposed technology has great potential when it comes to improving accessibility and inclusion for highway maintenance workers and construction laborers operating in highway work zones. With the help of augmented reality technology, digital overlays are created that are particularly useful for some individuals with disabilities. The proposed multimodal notification system can assist individuals with

Table 3.3: Paired t test between the collected likelihood of using AR/WT technologies before and after usability test in indoor experiment

	Pre-Usability Test Step		**Post-Usability Test Step**
Mean	4.69		5.23
Variance	1.23		0.69
Observations	13		13
Correlation		**0.71**	
Degree of Freedom		12	
t-Statistics		-2.50	
P(T<=t) two-tail		**0.02**	
t-Critical two-tail		2.17	

disabilities in reacting more promptly and navigating safely to secure locations. Even though no specific disabled groups were included in this study, it is recommended for future studies to broaden the participation by including disabled and other minority groups to further analyse the impact of disruptive technologies such as AR/WT in the broader highway workforce community. Another direction for future studies is studying the impact of learning curves. Users who are not familiar with advanced technologies such as AR may struggle with the learning curve, which can affect their willingness to adopt and utilize the technology. In this study, we solely focused on usability, and how mental load and perceived trust could impact user experience. However, investigating the dynamic between the learning curve and usability is one of the interesting topics recommended for future studies. A different intriguing future direction involves examining the effects of fatigue and ergonomic factors on usability. Past studies have emphasized the importance of including fatigue and ergonomics when introducing new products to ensure they do not cause undue discomfort [181]. Specifically, the impact of fatigue on usability has been investigated in the context of Augmented Reality applications in different contexts [182]. Therefore, a compelling direction for future research would be to assess the impact of fatigue and ergonomics on the usability and user experience of AR and WT technologies in the context of highway work zone safety.

3.6 Conclusion

The focus of this paper was to examine the usability of a new safety system that leverages Augmented Reality (AR) and Wearable Technology (WT) in the context of safety of highway work zones (ARROWS). The study aimed at identifying key factors that contribute to user experience as perceived by workers and documenting major pain points in including ARROWS in the operation of the highway work zone industry. For this purpose, we devised two complementary experiments in indoor and outdoor settings using the Wizard of Oz technique. A high-fidelity prototype of ARROWS was utilized to study usability, research user experience, and highlight design implications and future directions of AR and WT technologies in highway operation discipline through the lens of user-centred design best practices. The mixed-method outcomes of this study provide a series of usability benchmarks and user experience contributors that can be leveraged in customizing AR/WT technologies for the highway workforce. Our examination of the collected qualitative insights and responses to the pre-experiment questions of the surveyed workers demonstrate that participants rated their perception of the practicality of AR/WT solutions as 4 on a scale of 1 to 6 and cited their openness to trying novel technical solutions that address their safety challenges. This suggests that participants recognized the potential of AR/WT technologies in enhancing safety and are receptive to adopting such technologies. Working under strict time constraints and direct exposure to traffic differentiate highway work zones from other work environments. Consequently, there is a need for safety solutions that are both usable and functional within these specific limitations. Our findings exhibit participantsâ consensus in recognizing the critical need for novel technologies that enhance workersâ safety with a strong emphasis on the accommodation of their unique needs in the field through customized user experience. This includes incorporating design strategies that adjust haptic intensity to match the specific characteristics of highway maintenance activities, account for potential external impacts on the operation of the technology such as excessive heat and cold, and tailor functionality of the technology for addressing the diverse working conditions encountered by highway workers. Furthermore, participants rated the usability of ARROWS above average in both indoor and outdoor settings while reporting a reasonable mental effort during the end-to-end functionality that the high-fidelity prototype offered. It is anticipated to see similar trends in the early stages of research and development of other novel AR/WT technologies for highway workforce safety. The outcomes also point out a significant correlation between perceived trust and usability. This association highlights the importance of leveraging trust to enhance the usability and user experience of multimodal augmented reality in highway work zone safety applications.

CHAPTER 4: AUGMENTED REALITY WARNINGS IN ROADWAY WORK ZONES: EVALUATING THE EFFECT OF MODALITY ON WORKER REACTION TIMES

4.1 Introduction

The economic growth and prosperity of nations are heavily reliant on the presence of robust and efficient transportation infrastructure. In this regard, roadway work zones play a critical role in the inspection, maintenance, and upgrading of roadways to ensure their effectiveness, facilitate their development, and maintain continuous operation. However, these work zones pose significant dangers and risks to workers due to various factors, including direct exposure to traffic, extended work hours, and lack of adequate safety technologies [183]. As a result, they are recognized as one of the most hazardous work environments, globally. An examination of workplace fatalities and injuries around the world underscores the severity of the hazardous environment in work zones. According to the Centers for Disease Control and Prevention (CDC), between 2003 and 2017, a staggering 4,444 deaths were reported at road construction sites in the United States alone, with an average of 123 fatalities per year [1]. Furthermore, the US Bureau of Labor Statistics reported 135 work-related fatalities in 2019, further highlighting the ongoing dangers associated with this line of work [184]. Similarly, Great Britain also experienced significant challenges in workplace safety within roadway construction areas. In 2022, accidents involving workers being struck by moving vehicles ranked as the second most common cause of fatal accidents in workplaces, underscoring the critical need for enhanced safety measures in this sector [185]. Further, the transportation and storage sector accounted for 15.0% of fatal accidents in the European Union, highlighting the broader European context and the urgency of addressing safety concerns in this industry [186].

In parallel with the growing recognition of the dangers inherent in work zones, there has been a notable surge in investments directed toward the construction of efficient transport infrastructure. These investments, guided by initiatives such as Infrastructure and Investment Jobs Acts in the US and the multi-bilion Euro projects in the European Union under strategic plans for investment in sustainable, safe and efficient transport infrastructure, are projected to result in a substantial increase in the number of roadway work zones in various countries [187, 188]. As the expansion of work zones becomes an imminent reality, there arises an urgent and critical need to prioritize the improvement of safety systems within roadway work zones. This imperative stems from the recognition that workers in these zones face numerous risks, including intrusion by passing vehicles and exposure to other safety hazards.

To address these pressing concerns, concerted efforts must be made to strengthen the existing safety measures in place. It is paramount to improve and implement effective safety systems that can mitigate the risks faced by workers in work zones. To this end, in recent years, there has been an increasing emphasis within the roadway work zone community on utilizing technology to mitigate safety concerns [2, 3, 189]. To address this objective, limited studies have explored the potential of warnings, focusing on evaluating how different designs can contribute to improving safety measures for workers [190, 139, 191]. The design of an effective warning is crucial in ensuring worker safety by alerting workers to potential hazards and quickly disseminating critical information. A well-designed warning system should be able to capture the attention of workers and convey information clearly and effectively. However, most of these technologies primarily concentrated on delivering haptic signals through wearable devices [117, 3].

Meanwhile, AR has gained traction in numerous industries as a powerful tool to improve safety and productivity by superimposing virtual information on the real-world environment [82, 192]. AR technology has already begun to penetrate the transportation and construction industry [193, 5]. Therefore, this technology is expected to continue to expand its impact and influence on the safety domain of roadway work zones in the near future. Yet, current understanding of the impact of different warning designs in augmented reality systems on worker reaction times is limited. Previous research studies have not thoroughly examined the effects of different multimodal designs of AR warnings on worker response, particularly within the context of safety systems. The design of safety systems for roadway work zones cannot rely solely on general reaction time assumptions. The distinctive characteristics of these work zones, including the cognitive load associated with physical activities and the sensory taxing environment, can have a substantial impact on workers' reaction time performance. Therefore, it is essential to consider these factors when developing time-critical safety warnings, taking into account the unique challenges faced by workers in roadway work zones. Therefore, this knowledge gap has significant implications for the development of efficient real-time warning designs in roadway work zones.

Roadway work zones have also distinct characteristics that distinguish them from other construction areas, mainly due to their exposure to traffic and limited maneuver space [34, 194]. Exposure to traffic introduces an additional layer of complexity in roadway work zones [14]. Workers must not only manage their tasks, but also navigate and interact with moving vehicles [195]. Furthermore, the limited space for maneuvering in roadway work zones restricts workers' mobility and may affect their reaction times to warnings [196]. Meanwhile, the complex nature of replicating work zone environments has presented significant challenges in conducting high-fidelity experiments to assess reaction times. Such experiments often require substantial resources, involve intricate setups, and carry inherent risks [197]. Simultaneously, Virtual Reality (VR) has emerged as a viable alternative to recreate scenarios that are costly and logistically challenging to replicate. In the context of roadway work zone safety, VR has been used in different studies to investigate worker behavior in different contexts and technological settings [198, 197, 199]. In particular, AR simulation in VR is one of the research directions that researchers have considered in different applications, such as healthcare and military [200, 201, 202, 203]. However, a crucial question that still needs to be addressed is whether VR-simulated AR can effectively simulate the complexities of real-world dynamics and accurately capture reaction times in the context of safety in roadway work zones. This is an essential consideration for future simulation-based training applications, as it directly impacts the reliability and validity of using virtual reality simulations to train workers in real-world scenarios.

This paper documents our research endeavors to evaluate the reaction time for multimodal Augmented Reality-based warnings within the domain of roadway work zone safety. To achieve this, we developed an integrative research framework that incorporates an innovative concept of an AR-based system proposed by the authors [124]. The designed methodology is rooted in three key elements: a high-fidelity prototype of the AR system, virtual reality that simulates the functionality of the AR system in its context, and the Wizard of OZ methodology that synchronizes the user journeys across different experiments. Furthermore, our methodology employs two approaches to quantify the reaction time: (i) a simple reaction time (SRT) methodology, and (ii) a vision-based strategy that we devised. In the SRT approach, we designed two mechanisms using physical interfaces, including a keyboard and VR controllers, to capture participants' reactions. These interfaces provided intuitive means for participants to respond by pressing keys or buttons when prompted by stimuli. In the vision-based approach, we used a pose estimation model to examine whether warnings communicated through the simulated AR system in VR elicited any observable physical responses to cumulative joint displacement of the participants' upper body. By analyzing the speed of the movements and positions of the participants' pose, we aimed to detect any changes or patterns that could indicate a physical response triggered by the warnings. We then developed a mixed-method approach that incorporated both between-subject and within-subject designs to investigate the impact of different conditions on reaction times and assess the effectiveness of AR-based warnings. Our study involved conducting five experiments, including an outdoor experiment using a real-world prototype in a temporary work zone, as well as multiple indoor experiments utilizing a simulated version of the prototype within a virtual reality environment. Using this integrated approach, we were able to gather comprehensive data on the performance of different multimodal warning designs in both real-world and controlled environments. In our experimental design, our aim was to address the following research questions as described below:

- **RQ1.** Which multimodal AR warning design generates the quickest reaction time in real-world setting?
- **RQ2.** Does the reaction time to VR-simulated AR warnings statistically match that of the outdoor environment?
- **RQ3.** Can real-time pose tracking serve as an indicator of the reaction time to AR warnings?

Our study holds significant implications for enhancing safety in roadway work zones through the incorporation of AR-based warning systems and vision-based reaction time measurement. It provides a comprehensive range of reaction time metrics and benchmarks specifically tailored to roadway work zone safety, serving as a valuable reference for the development of real-time safety systems that utilize AR technology. This study represents a pioneering effort in quantifying reaction times to multimodal AR-based warnings in the context of roadway work zone safety. Furthermore, our research contributes to the existing body of knowledge by shedding light on the use of virtual reality as a simulation tool for AR and its application in reaction time safety research. By establishing a benchmark to compare the reaction time of VR-simulated AR with real-world multimodal AR scenarios, we improve our understanding of the feasibility and effectiveness of using VR simulations to measure reaction times within AR safety systems. Furthermore, this study offers information on the design of multimodal warning mechanisms for safety and evaluates their efficacy in triggering timely responses. This investigation increases our understanding of the impact of warning mechanisms on workers' reaction times, which, in turn, facilitates the optimization of safety systems, particularly when integrating AR technology. Furthermore, the introduction of the vision-based reaction time measurement strategy brings forth exciting possibilities for utilizing innovative data acquisition techniques to monitor safety hazards and responses in roadway work zones. By capturing and analyzing visual data, we can obtain valuable insights into the reactions and response times of workers when they encounter different safety hazards within roadway work zones. What makes this approach particularly advantageous is that it can be implemented in a non-intrusive manner, respecting the natural work environment and minimizing disruptions to the workers' tasks. Overall, by establishing reaction time benchmarks and laying the foundations for further research, our study paves the way for the development of advanced real-time safety systems for roadway workers.

4.2 Related Work

Reviewing previous work, we first evaluate safety measures in roadway work zones, then investigate trends in warning design and methods of measuring reaction times. Finally, we explore the role of virtual reality in simulating augmented reality across multiple fields.

4.2.1 Safety Measures in Roadway Work Zones

Despite acknowledging the risks associated with roadway work zones in various studies, the implementation of new technologies to improve safety in these environments has been limited. Existing safety measures primarily rely on reactive approaches, triggering alarms or alerts only after an intrusion occurs or when intruding objects are in a close proximity [36, 12, 37, 38]. However, there is a growing shift toward smarter and more proactive safety systems in roadway work zones. A systematic review conducted by Nnaji et al. [2] highlighted the growing need for adopting smart automated technologies and a departure from traditional approaches, a shift further propelled by the emergence of advanced sensing technologies.

One example of recent innovative research studies in this field is the work by Sakhakarmi et al. [3], who developed a proximity-based alerting system that utilizes tactile cues as the primary mode of communication with users. However, the effectiveness of this system in roadway work zones can be compromised due to the high levels of noise and cognitive demands placed on workers. In another study, Chan et al. [39] proposed a wearable-based hazard proximity warning system to improve the awareness of construction workers. Although this system relies on proximity-based triggering mechanisms, it still shares the limitations of reactive systems that only activate warnings when hazards are in close proximity. Similarly, Kim et al. [40] developed a novel system using AR to alert workers about potential hazards based on orientation and proximity. However, the functionality of this system may be limited to hazards within the workers' field of view, overlooking hazards outside their visual scope. Furthermore, Kim et al. [204] proposed an IoT-based proximity warning system that alerts workers when they are in close proximity to heavy equipment

These studies highlight ongoing efforts to develop advanced safety systems for work zones on roads. However, there are challenges to address, such as the noisy and demanding nature of these work environments. Future research should aim to overcome these limitations and strive for comprehensive solutions that take into account multiple modes of communication and multimodal warning, incorporate real-time data, and effectively address the unique challenges and demands faced by workers in the work zones.

4.2.2 Warnings Modality and Design

Warnings are signals generated by a device in the form of visual, auditory, or haptic stimulus, which are designed to attract users' attention and provide proactive information [205]. Warnings facilitate an interactive process between users and a device or system, encouraging them to react, respond, and take specific actions based on the content and delivery of the warning warnings [206, 207]. Previous studies have highlighted that the presentation and delivery of warnings play a crucial role in their effectiveness [208, 207]. Research suggests that effective presentation and delivery of warnings depend on various factors, including user characteristics, task at hand, device being used, and the surrounding environment [209].

Although there is a wealth of literature on digital warnings in general, research specifically addressing warning delivery in the AR technology, particularly in the context of real-time safety systems, is limited [207]. The majority of previous studies have focused mainly on in-vehicle AR warnings, leaving a gap in understanding how to effectively design AR warnings for workers and other vulnerable road users [210, 211, 212, 213]. However, recent studies have started to explore AR-based warnings in different contexts related to VRU. For example, Matviienko et al. [214] investigated the safety enhancement of e-scooters with unimodal warnings, including AR warnings, vibrotactile feedback on the handlebar, and auditory signals, to prevent collisions with other road users. Similarly, Von et al. [215] explored potential approaches for AR applications to enhance cyclist safety and conducted a pilot study. These studies have considered multimodal designs, which incorporate visual, audio, and haptic feedback in warnings. However, there is still a lack of enough research to specifically address the effects of AR-based warnings on roadway workers and their reaction time in different settings and contexts.

4.2.3 Reaction Time: Measures, Influences, and Applications to Roadway Worker Safety

Reaction time measurement has been an important tool in psychology and neuroscience for more than a century [216]. RT measurements have been widely used by researchers to explore a diverse range of cognitive processes, study perception, and examine how quickly individuals can detect and interpret sensory stimuli from their environment [217, 218, 219]. To this end, Simple Reaction Time (SRT) task is widely recognized as the most common method to measure reaction time [220]. In this approach, participants are instructed to respond as rapidly as they can to a single stimulus. The stimulus can take various forms, including visual cues, auditory signals, or somatosensory stimuli. Participants typically execute their response by pressing buttons, pressing keys, or performing vocal reactions [221, 222, 223]. The review of literature suggest that reaction time can be influenced by various factors such as age [224], sex [225, 226, 227], attention [228], fatigue [229], arousal levels [230], and task complexity [231].

Meanwhile, understanding workers' reaction time to safety warnings plays a vital role in the development of effective alert technologies. This significance is particularly accentuated in the context of roadway work zones, where the complex environment and the presence of fast-moving vehicles necessitate timely and fast response from workers in case of intrusions [41]. To this end, several studies have been conducted to investigate workers' reaction time in various systems and working environments related to roadway work zones. For example, Thapa et al. examined the optimal configuration of a work zone intrusion alert technology and explored the relationship between sensor placement and alerting modules, considering workers' naturalistic reaction [232]. In another research work, Nnaji et al. provided guidelines for adopting different commercially available work zone technologies for roadway work zones, taking into account workers' reaction time and response rate as essential metrics in their framework [41]. In another study, Awolusi et al. quantified the reaction time of roadway workers to two commercially available intrusion alert technologies specifically designed for roadway work zones [14]. Finally, in a recent study, Yang et al. [191] conducted three experiments to assess the viability of using vibrotactile signals as warnings for road workers. The experiments aimed to assess the perception and performance of the generated signals at different body locations and to examine the usability of various warning strategies. Our review suggest that the existing literature does not provide sufficient evidence to provide insights into reaction times specifically related to AR-based warnings in this particular field.

4.2.4 Virtual Reality Simulations for Evaluation of AR Warnings

The rapid development and widespread adoption of AR technology have propelled it to the forefront of various domains and industries, offering a multitude of applications and possibilities [5, 233]. However, the accelerated growth of AR has created a demand for efficient methods to prototype and evaluate AR interfaces and interactions in a timely manner [234, 235]. Traditional approaches such as wire-framing and paper sketching, while useful in certain design contexts, often fall short of capturing the true essence of the user experience in AR. These methods struggle to convey the immersive and context-sensitive aspects of AR applications, making it challenging to evaluate user interactions and gather meaningful feedback on usability and functionality. To overcome these limitations, researchers and practitioners have turned to alternative approaches that take advantage of virtual environments, specifically VR, for AR prototyping and evaluation. Using virtual reality simulations, designers and developers can create virtual representations of AR interfaces and interactions that closely mimic the real-world user experience. This immersive and interactive environment allows for more realistic user testing and evaluation, allowing stakeholders to gain a deeper understanding of the usability, effectiveness, and user satisfaction of AR applications.

In order to achieve this specific objective, several studies have been conducted. For example, in the field of surgical applications, Hettig et al. used VR simulations to mimic AR environments and investigated individual parameters for surgical procedures [200]. By simulating the AR experience, they were able to explore different scenarios and optimize the application of AR technology in this novel context. In another study, Terrier et al. focused on the impact of registration errors between virtual and real objects in AR [236]. They used VR simulations to control experimental conditions and examine the effects of registration errors. In the realm of public safety applications, Grandi et al. proposed a framework that used virtual reality to evaluate AR interfaces in traffic stops and firefighting search and rescue scenarios [237]. Using virtual reality simulations, they were able to simulate realistic situations and gather feedback to improve the design and effectiveness of AR interfaces in these critical public safety contexts. Furthermore, Zaman et al. conducted a study focused on the usability of AR technology in combat missions. They used virtual reality simulations during subterranean operations to investigate the usability aspects of AR interfaces [202]. Finally, Burova et al. [238] used VR AR simulation along with gaze tracking to evaluate the effectiveness of AR guidance and safety awareness features for elevator maintenance. Through an iterative development-evaluation process, industry experts participated in testing and providing feedback on the AR simulation and gaze tracking system. The study also included a survey that used actual gaze data from the evaluation to collect comments and insights from industry experts regarding the usefulness of the AR simulation and gaze tracking approach.

4.3 Methodology

4.3.1 Study Overview

To achieve the objectives of this study, we incorporated a multi-faceted approach to measure reaction times across distinctive setups. Each of these setups was designed to gauge different aspects of reaction time and to evaluate the efficacy of various modes of warnings. For this purpose, we conducted a desktop-based simple reaction time measurement which served as our baseline. In this experimental setup, participants were instructed to respond as quickly as possible to a multimodal warning, with the visual sign presented on a computer monitor. The objective here was to create a standard benchmark against which we could compare the reaction times observed in other scenarios. Next, we conducted a real-world test in a controlled outdoor work zone setting. Here, participants were equipped with AR glasses that delivered warnings about potential intrusions or hazards. The aim was to evaluate the impact of different sensory modes of AR warnings on the reaction time of workers in a realistic environment. Finally, we conducted a series of experiments within an immersive virtual reality environment that mimicked a roadway work zone. The intent of this setup was twofold. Firstly, we sought to validate the fidelity of our VR work zone simulations by comparing the participants' reaction times in the VR environment with those from the real-world test. Second, we used this setup to measure reaction times using vision-based pose-tracking algorithms by capturing and analyzing the movements of the participants in response to the simulated warnings. In this section, we will detail

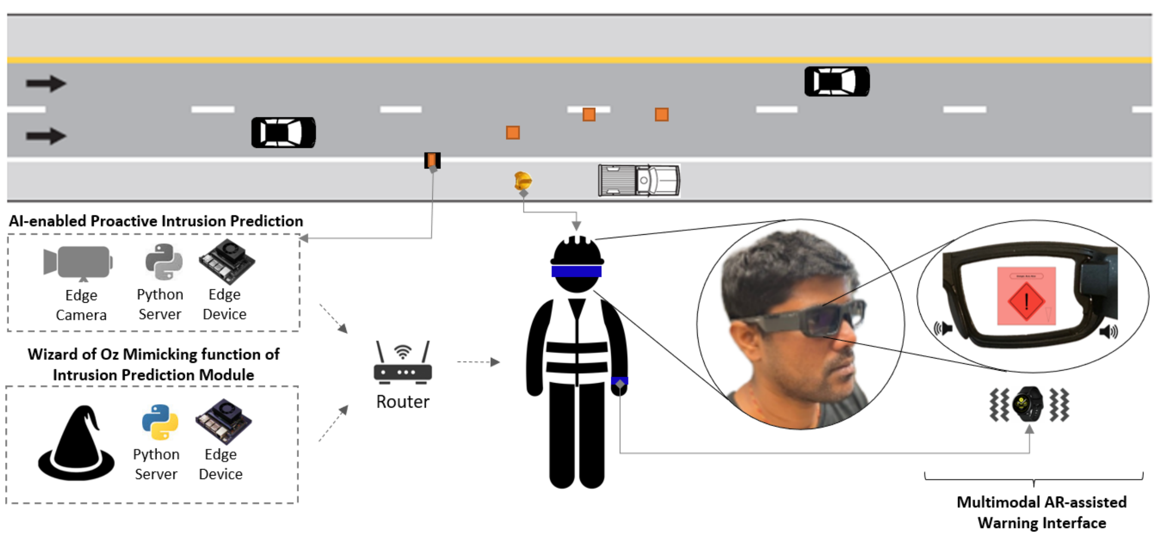


Figure 4.1: Overview of the Augmented Reality-Based Safety Technology and Its Warning Interface Features

our methodological framework, discuss the AR and VR technologies used, explain the design and execution of our experiments, and elaborate on other significant aspects integral to addressing the research questions. Table 4.1 also provides comprehensive information on the details, specifications, apparatus, and warning designs of the experiments conducted in this study.

4.3.2 Experimental Apparatus and Setup

4.3.2.1 Utilized Augmented Reality Technology

This study utilized an innovative AR warning technology, conceptualized and developed by the authors' team, as detailed in our previous work [124]. This technology, shown in Figure 4.1, consists of two main components: the AI-powered back-end, which processes real-time video data to predict potential intrusions in the work zone, and the front-end, which uses a multimodal AR interface to alert workers in real time about immediate risks. This AR warning interface incorporates visual, auditory, and haptic cues and provides workers with timely warnings regarding possible intrusions or hazards, improving their situational awareness, and facilitating quick and appropriate responses to ensure worker safety. In this study, we used the Wizard of Oz methodology (WOZ) to replicate the functionality of the back-end component without using an actual AI module. We replaced the role of AI with a pre-definied scripted "wizard" written in Python that runs on the edge device. We adopted this approach to focus on specific intrusion scenarios that activate the warning interface within a controlled environment, as these occurrences are infrequent in real-world situations. Using this methodology, we could iteratively record participants' reaction times and responses to AR warnings. Additionally, the approach ensured that participants encountered similar and synchronized scenarios with consistent triggering points across multiple experiments. The back and front ends were connected within a local network facilitated by a router.

Figure 4.2(b) provides an overview of the hardware and software components of AR

Table 4.1: Details and Specifications of the Designed Warnings and Experiments

<table>
<tr><th rowspan="2">Experiments</th><th rowspan="2">Contextual Environment</th><th colspan="2">Visual (V) Stimulus</th><th colspan="2">Audio (A) Stimulus</th><th colspan="2">Haptic (H) Stimulus</th><th rowspan="2">Warning
Modalities Tested</th><th rowspan="2">Reaction Time
Measurement</th></tr>
<tr><th>Delivery</th><th>Design</th><th>Delivery</th><th>Design</th><th>Delivery</th><th>Design</th></tr>
<tr><td>A - Baseline</td><td>-Indoor Desktop-based</td><td>Computer Monitor</td><td rowspan="5">Warning
Sign as
visualized
in Figure ??</td><td>Computer
Speakers</td><td rowspan="5">-Stereo
High-pitched
Beep
-44100 Hz
Frequency
-0.2 milliseconds</td><td rowspan="2">Right-hand
Galaxy
Smartwatch</td><td rowspan="5">Haptic
Feedback
Through
Tizen
Native API</td><td rowspan="4">Four
Combinations
(V, AV, HV, HAV)</td><td>-Keyboard Click
-Simple Reaction
Time</td></tr>
<tr><td>B - AR</td><td>-Outdoor
-Controlled Roadway
Work Zone Testbed
-No Traffic
-Natural Ambient Noise</td><td>-AR Glasses
Display on
Right Lens
-Semi-transparent</td><td>Bluetooth
Earbuds</td><td>-VR Right-hand
Controller
"A" Button
-Simple Reaction
Time</td></tr>
<tr><td>C - VR-WOT</td><td>-Indoor
-VR Simulated
Roadway Work Zone
-No Traffic
-No Ambient Noise</td><td rowspan="3">-VR Environment
Mimicking AR
-Semi-transparent</td><td rowspan="3">Built-in VR
Headset
Speakers</td><td rowspan="3">Right-hand
Galaxy
Smartwatch
in Reality
(User sees
the virtual
watch in the
VR environment)</td><td rowspan="2">-VR Rright-hand
Controller
"A" Button
-Simple Reaction
Time</td></tr>
<tr><td>D - VR-WT</td><td>-Indoor
-VR Simulated
Roadway Work Zone
-Simulated Traffic
-Ambient Noise</td></tr>
<tr><td>E - Vision-based</td><td>-Indoor
-VR Simulated
Roadway Work Zone
-Simulated Traffic
-Worker Engaged
in Maintenance Activity
- Ambient Noise</td><td>One
Combination
(HAV)</td><td>-Vision-based
Pose Tracking</td></tr>
</table>

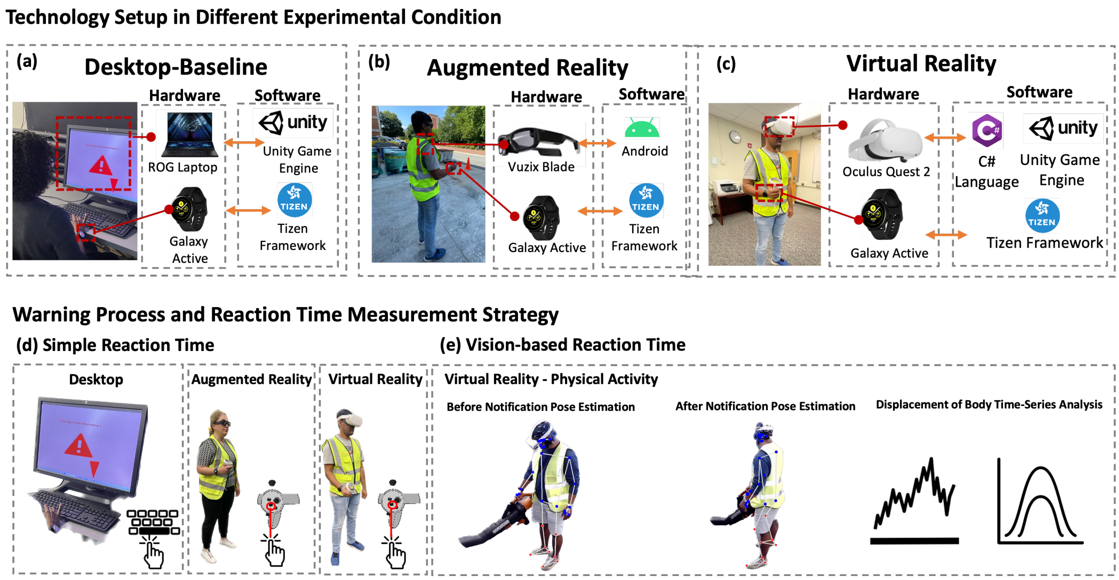


Figure 4.2: Proposed Framework for Quantifying Reaction Time to Multimodal AR Warnings: Experimental Setup, Hardware and Software Development, and Utilization Mechanisms

technology used in this study. We used Vuzix Blade AR smart glasses and a Galaxy smartwatch for this particular front-end design. We selected Vuzix Blade as they comply with ANSI Z87.1 standards and can be comfortably paired with prescription eyewear or shades. Featuring built-in audio modules and a display located on the right lens, these glasses ensured minimal obstruction of natural vision while being suited for industrial applications. The glasses, which operate on the Android OS, facilitated the programming of the networking software and the design of our study-specific warning layout, shown in Figure 4.2(b). In terms of audio cues, we employed Bluetooth-connected earbuds in tandem with AR glasses to deliver audio cues directly to workers' ears. For haptic cues, the Tizen framework was utilized to develop the relevant software for the Galaxy smartwatch, enabling the delivery of haptic cues within the prototype's ecosystem. As a Samsung-specific platform, Tizen facilitated programming networking functionalities and other necessary algorithmic elements for seamless back-end and front-end communication in augmented reality glasses.

4.3.2.2 Virtual Reality Simulation

The development of this simulation was centered on assessing the suitability of virtual reality as a means to simulate AR warnings. For this purpose, the virtual reality simulation was designed to closely resemble the AR prototype developed. Similar layout design, visual cues, and audio frequency were implemented in the VR simulation to maintain consistency with the AR interface. We used the Oculus Quest 2 VR headset shown in Figure 4.2(c) in the VR simulation. Oculus Quest 2 provided an immersive and interactive VR experience for participants. Unity was utilized to develop pertinent software. Unity is a popular game development engine that supports VR development and provides a wide range of tools and resources for creating virtual environments and interactions. Furthermore, we employed the identical smartwatch component from the AR prototype in the VR simulation. During the simulation, users were able to simultaneously observe the virtual smartwatch on the screen while wearing the physical smartwatch on their wrist.This allowed us to simulate the exact interactions and user experience as those offered by AR technology.

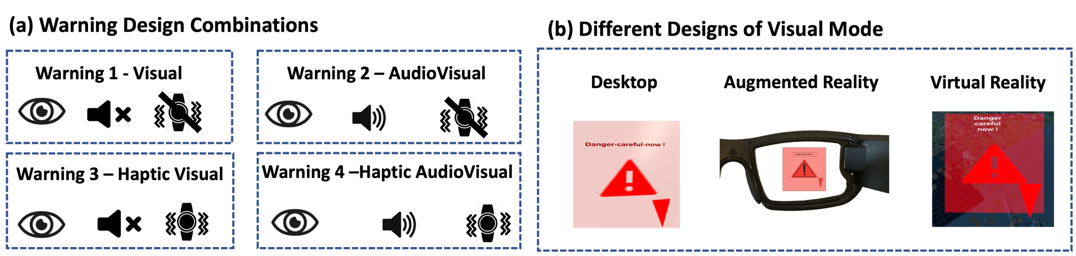


Figure 4.3: Multimodal Warning Design Specifications and Delivery Means of Visual Cue in Different Setups

4.3.2.3 Desktop-based Setup

In addition to the AR and VR settings, we also developed a desktop-based replica of the AR warning to quantify the reaction times to provide a baseline for comparison with the AR and VR interfaces. Using the software developed for the VR simulation in Unity, we made the necessary adjustments to create a desktop version of the AR warning mechanism. Figure 4.2(a) illustrates the hardware and software setup for this desktop replica. The main difference between the desktop interface and the AR/VR settings is that visual and audio cues are delivered through the desktop display and speaker, rather than the AR glasses or VR headset. Algorithmic designs, developed software, and haptic feedback through the smartwatch remained exactly the same as in AR and VR technologies.

4.3.3 Study Design

In this section, we will outline the design and details of the experiments conducted. A total of five experiments were carried out. These experiments can be categorized into two distinct groups on the basis of the reaction time measurement strategies. The first group, consisting of experiments A to D, aimed to quantify the reaction time to the designed warnings in various settings and conditions using a simple reaction time approach. However, experiments E used a vision-based metric combined with a naturalistic task design to quantify the reaction time. The task was carefully designed to simulate real-world scenarios and capture naturalistic reactions of the participants. Further details of each set of experiments are provided in the following. This study was conducted with the approval of the institutional review board (IRB No. 21-0357) of the University of North Carolina at Charlotte.

4.3.4 Warning Design

Figure **??**(a) depicts the schematic representation of the various multimodal warning designs utilized in the study, namely Visual (V), Audio Visual (AV), Haptic Visual (HV), and Haptic Audio Visual (HAV) designs. Table 4.1 provides further information on each design, specifying the visual, audio, and haptic cues that were used synchronously in each warning design. To ensure consistency in different settings, the layout of visual stimuli in all setups was intentionally designed to be identical, as illustrated in Figure **??**(b). This uniform layout implementation aimed to minimize any potential variations or confounding factors introduced by differences in visual presentation across the different environments and experiments. Haptic stimuli were implemented using the Tizen Native API framework, leveraging a predefined pattern available in the API [239]. This framework facilitated the generation of haptic cues synchronized with the other modalities. For the audio component, a high-pitched beep with a frequency of 44100 Hz and a duration of 0.2 milliseconds was used as the auditory signal. It is important to note that all warnings were intentionally designed to trigger simultaneously, without any intentional delays, upon activation by the back-end system. This simultaneous triggering ensured that participants experienced the multimodal warning stimuli in a synchronized manner, allowing consistent evaluation of their reaction times in different designs.

4.3.5 Experiment Procedure and Specifications

A consistent WOZ scenario was designed and executed in experiments A through D to ensure uniformity in the study. The duration of the scenario was set to 45 seconds, during which five iterations of the warning trigger occurred. The scripted scenario was standardized and followed in all experiments, incorporating the five iterations of the SRT task. The trigger points for the warnings remained constant across all designs, guaranteeing consistency across the experiments. Experiment B was the first one conducted, while the order of conducting Experiments A, C, and D was randomized to minimize the impact of a learning curve. Additionally, to address potential bias or confounding factors, the measurement order of reaction times for each design in experiments A to D was also randomized. This randomization ensured a fair and unbiased assessment across the board. For instance, participant X had their reaction times measured in order AV, V, HAV, HV, while participant Y had their reaction times quantified in order V, AV, HV, HAV.

Experiment E was conducted as the final experiment primarily for logistical reasons. A WOZ-programmed scenario was carefully designed and executed consistently for all participants. The scenario had a duration of 1 minute and consisted of two iterations where the warning was triggered. The trigger points for the warning remained constant throughout the experiment. In particular, for Experiment E, the HAV warning design was used exclusively. Figure 4.2 offers a comprehensive overview of the experiments conducted in this study, providing a visual representation of the different configurations and designs used.

In the following, we provide further details and specifications of the experiments conducted.

4.3.5.1 Experiment A: Desktop-based Baseline

This experiment was carried out using the developed desktop-based interface. The participants performed the experiment on a 21-inch display, sitting at a distance of 20 inches from the display, which was placed in front of a black background. A total of 32 participants (N = 32) completed the study, with an average age of 28.7 years (SD = 5.5) and an average experience of 3.4 years (SD = 0.9) in the construction industry. Among the participants, 20 identified as male and 12 as female. The experiment was carried out in the Advanced Infrastructure Management Lab at UNC Charlotte.

4.3.5.2 Experiment B: AR Warnings in the Real World

This experiment was carried out outdoors and in a temporary work zone that was specifically created for this study on the campus of UNC Charlotte. The design of the work zone followed the guidelines outlined in the Manual of Uniform Traffic Control Devices (MUTCD) for short-duration work zones. A total of 34 participants (N = 34) participated in this experiment, with an average age of 25.9 years (SD = 5.1) and an average experience of 2.3 years (SD = 2.6) in the construction industry. Among the participants, 21 identified as male and 13 as female. The duration of the experiment for each participant ranged from 30 to 45 minutes and was conducted during daylight hours between 11 am and 4 pm. Efforts have been made to keep the weather and lighting conditions consistent throughout the experiment to minimize their impact on the results.

4.3.5.3 Experiment C & D: VR Testbed W/WO Traffic

Both experiments were carried out using the simulated AR interface developed in virtual reality. The VR environment was specifically designed to replicate a short-duration highway work zone based on the guidelines provided by the MUTCD. Figures 4.4(d) provide examples of the design of the virtual work zone used in the experiments. A total of 32 participants ($N = 32$) completed the study, with an average age of 28.7 years ($SD = 5.5$) and an average of 3.4 years ($SD = 0.9$) of experience in the construction industry. Among the participants, 20 identified as male and 12 as female. The study was carried out in the Advanced Infrastructure Management lab at UNC Charlotte.

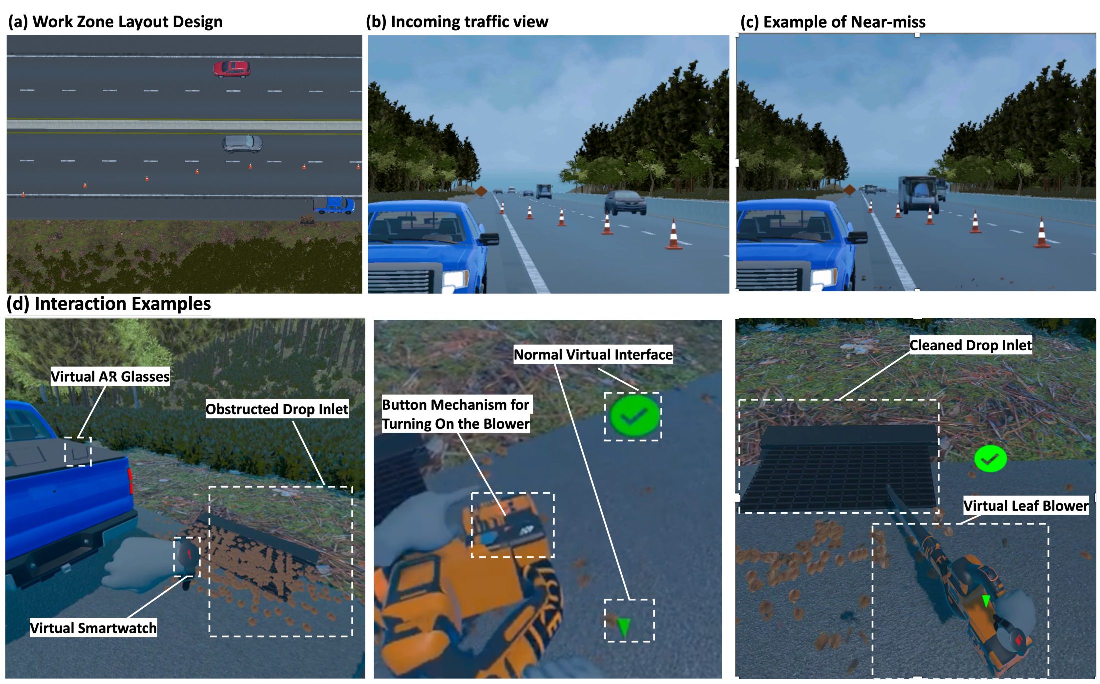


Figure 4.4: Examples of the Developed Virtual Reality Environment and Designed Interactions

4.3.5.4 Experiment E : Immersive Mixed Reality Task Through VR Simulation

A total of 28 participants ($N = 27$) participated in the study, with an average age of 28.7 ($SD = 5.6$) and an average of 3.4 years of experience ($SD = 0.9$) in the construction industry. However, a total of 6 participants in the activity task encountered technical difficulties, simulation sickness, and other logistical problems during data collection and were unable to complete the experiment. As a result, the final count of participants in the experiment was 21 ($N = 21$), with an average age of 28.2 years ($SD = 5.8$) and an average industry experience of 3.1 years ($SD = 1.1$).

In this experiment, our main goal was to replicate real-world scenarios commonly encountered in highway work zones. Our focus was specifically on the task of removing obstructions from obstructed drop inlets. To design our study, we considered the existing literature that discusses the influence of physical activity intensity and cognitive load on reaction time [240, 241]. Based on this knowledge, we developed an obstruction removal task that required participants to engage in higher levels of physical exertion compared to other routine maintenance activities. Our intention was to simulate a task that is frequently encountered in the maintenance and operation industry. This task was chosen due to its practical significance and its common occurrence in the industry, allowing us to recreate real-world scenarios and evaluate the impact of warnings on participants' performance.

In order to conduct this study, we developed a virtual work zone that adhered to the instructions outlined in MUTCD. This virtual work zone, depicted in Figure 4.4(a), served as the backdrop for our research. To enhance the realism and interactivity of the study, we adopted a mixed reality interaction approach for executing Experiment E. This approach allowed participants to engage with both physical objects in the real world and virtual objects within the virtual environment simultaneously. By combining elements from both realms, we aimed to create a unique and immersive experience for the participants.

One key aspect of the mixed reality approach was the implementation of a leaf blowing effect within the virtual environment. This effect was designed to simulate the action of using a leaf blower to clear leaves that obstruct a drop inlet. As participants entered the simulation, they were equipped with a physical leaf blower in their hands, mirroring the position and movements of the virtual leaf blower shown in Figure 4.4(d).

The task began with participants activating the leaf blower and directing it towards the obstructed drop inlet within the virtual environment. As they did so, the virtual reality environment featured a carefully designed blowing effect that effectively cleared the leaves positioned on top of the drop inlet. This dynamic and interactive task continued until all necessary warnings were delivered, and the administrator signaled the completion of the task.

4.3.6 Reaction Time Measurement

Experiments A through D focused on capturing the reaction time of the participants using an SRT strategy under various conditions. Each experiment consisted of five iterations of an SRT task in which participants were required to respond by pressing keys or buttons, as shown in Figure 4.2(d). However, the objective of Experiment E was to investigate whether multimodal AR warning, when presented, triggers any observable physical response that can be captured by pose tracking in a naturalistic setting, as illustrated in Figure 4.2(e). Our goal was to go beyond the traditional SRT approach and develop tasks that closely mimic real-world scenarios. By doing so, we aim to assess the influence of real-time warnings on body motion and analyze that impact on reaction time. In the following sections, we present a comprehensive description of the utilized strategies.

4.3.6.1 Simple Reaction Time

We used two similar approaches to measure reaction time in Experiments A-D. The methodology consisted of recording the time interval between receiving the warning-triggering command from the server and pressing the designated capturing button. For the desktop interface, the designated button was the space button, while for AR and VR experiments it was the A button of the right controller of the VR headset. In the context of AR prototype, we enhanced the IoT network infrastructure and created custom software to facilitate the integration of virtual reality controllers for the purpose of capturing reaction times. To achieve this, we integrated the VR headset as an additional end point within the existing network. This configuration allowed us to utilize the controller in tandem with the AR prototype, facilitating the capture of reaction times. Communication latency between these endpoints was estimated to be less than 10 milliseconds [242], which is considerably shorter than the average reaction time observed in humans [243]. Therefore, the impact of this latency on the overall accuracy of the measurement was considered negligible.

4.3.6.2 Vision-based Reaction Time

We utilized a Logitech camera, capable of capturing data at a rate of 30 frames per second (fps), to record the experiment and participants reaction to warnings. To analyze body pose, we used the pose tracking API of ML Kit, an open source tool provided by Google [244]. This API offers a lightweight and flexible solution for the real-time detection and tracking of body poses from video streams and images. It provides a comprehensive 33-point skeletal map of the entire body, including facial landmarks, hands, and feet. In our analysis, we focused on the upper body landmarks, as we hypothesized that significant movement would occur primarily in this region. This hypothesis was derived from our observations and the relevant literature [245]. Figure 4.5 illustrates the landmarks used in our analysis, where the blue dots represent the included landmarks and the red dots represent the excluded ones. These dots represent landmarks 0 to 24 according to [244].

To quantify the overall upper body movement, we followed several steps. Firstly, we extracted the raw coordinates of the landmarks from the output of the model applied to the video recordings. These landmarks were then filtered to select the desired ones for analysis. Next, using the coordinates of the selected landmarks, we calculated the total pairwise displacement between consecutive frames. This involved measuring the distance in the image space between each landmark and its previous location. By summing all these distances, we obtained the delta displacement, which provides an indication of the overall movement of the upper body. To determine the speed of the upper body movement, we divided the delta displacement by the corresponding frame duration. This calculation yielded the speed in terms of pixel distance per frame, allowing us to quantify the rate of movement.

Figure 4.5: Examples of the Outcomes of the Utilized Pose Estimation Algorithm on the Developed Task, and Included (blue) /Excluded (red)Landmarks in the Analysis

After calculating the upper body movement velocity, it was necessary to establish the reaction time pattern within the time series to identify reaction time patterns. In our study, we defined the reaction time as the duration between the delivery of the warning and the onset of the reaction pattern exhibited by the participants. We used a Gaussian kernel to represent the reaction pattern within the velocity time series of subjects. This specific kernel has been used in the literature as an indicator of rapid reactions in the human body [246, 247, 248]. Equation 4.1 illustrates the Gaussian kernel that is specifically used in our research. In this equation, $K(t)$ represents the Gaussian kernel at time t. The parameters μ and σ control the shape and width of the Gaussian curve. μ represents the mean or center of the kernel, indicating the time at which the reaction pattern is expected to peak, while σ represents the standard deviation, which determines the spread or width of the Gaussian curve.

$$K(t) = amplitude \cdot \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) \tag{4.1}$$

To analyze the collected time-series data and identify reaction-time patterns in the recorded body movements, we employed two pattern recognition techniques: convolution and wavelet analysis. Convolution plays a vital role in pattern recognition tasks within signal processing. It is commonly utilized to detect and extract patterns or features from signals by convolving a signal with a predefined pattern or filter [249]. The convolution operation highlights regions in the velocity signal where a pattern similar to the Gaussian kernel is observed, indicating the presence of the reaction pattern. Analyzing the resulting convolution output allows us to extract relevant features and information pertaining to the reaction time. Mathematically, the convolution of the time series with the kernel can be expressed as Equation 4.2:

$$y(t) = \int_{-\infty}^{\infty} x(t-\tau)k(\tau)\, d\tau \tag{4.2}$$

To account for individual differences and ensure a personalized analysis, our study takes a within-subject approach. This involves determining the duration of each kernel based on the participant's prior performance in the HAV warning design recorded in virtual reality conditions with the traffic scenario. By setting the kernel duration as the recorded reaction time, we aimed to accurately capture the temporal dynamics of each participant's response, individually. Through this customization, our aim is to capture the subtle nuances of each participant's reaction pattern. Additionally, to account for the anticipated agility in participants' reactions following the delivery of the warning, we further refined the kernel by setting its width to be one eighth of the total duration. This choice results in a steeper pulse shape, allowing us to capture the anticipated rapid changes in participants' response immediately after receiving the warning. The amplitude of the Gaussian kernel was standardized to a value of 1. Finally, to determine the time that corresponds to the maximum convolution value, we evaluated the function $y(t)$ over the desired time range and identified the time t_{max} that satisfies the condition in 4.3. Using this approach, we can accurately identify the exact time at which the reaction pattern reaches its maximum intensity. The reaction time was then calculated as the duration between the delivery of the warning stimulus and the onset of the observed reaction pattern.

$$t_{max} = \arg\max_{t} y(t) \tag{4.3}$$

In addition to convolution analysis, we also used wavelet analysis to further investigate patterns in the collected data. Previous studies have suggested the application of wavelet analysis, specifically the Gaussian wavelet, for analyzing body movements [246, 247, 248]. Wavelet analysis decomposes a signal into simpler components using an algorithm similar to Fourier analysis. However, wavelet analysis is particularly effective in capturing transient behavior and discontinuities commonly observed in human movement signals. It enables a more accurate characterization of anomalies, pulses, and other transient events within the signal [250]. The equation for wavelet analysis is given by Equation 4.4. In this equation, $W(a, b)$ represents the wavelet transform of the signal $f(t)$ at scale a and translation b. The symbol ψ denotes the complex conjugate of the mother wavelet function ψ, and $\psi^(t)$ represents the complex conjugate of the scaled and translated wavelet function $\psi(t)$. The integral is computed over the entire real line from $-\infty$ to ∞. The scaling factor $1/a$ ensures the appropriate normalization of the wavelet transform. In our study, we utilized 2$^{\text{nd}}$-order Gaussian wavelets, similar to our kernel analysis, with the aim of uncovering any underlying patterns hidden within the data.

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\Psi^*\left(\frac{t-b}{a}\right) dt = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t)\psi^*\left(\frac{t-b}{a}\right) dt \tag{4.4}$$

4.4 Results and Discussions

In this section, we present the results obtained from our experiments, which are categorized into two main groups. The first group focuses on analyzing the variations in reaction time observed across different prototypes and experimental conditions. We explore how factors such as the warning design and the experiment condition influence the participants' reaction times. Second, we investigate the reaction time after the presentation of warnings using our vision-based approach. This analysis aims to understand how participants physically respond to warnings. We examine the patterns in the physical responses of participants and statistically analyze the relationship between vision-based reaction time and simple reaction time across different subjects.

Table 4.2: Summary of Reaction Times Recorded for Different Warnings Designs Across Different Conditions (AR: Augmented Reality, VR-WT: Virtual Reality With Traffic, VR-WOT: Virtual Reality Without Traffic)

		AR	Baseline	VR-WT	VR-WOT
V	Average	0.597	0.410	0.493	0.489
	SD	0.232	0.105	0.177	0.159
AV	Average	0.627	0.422	0.477	0.483
	SD	0.249	0.122	0.108	0.162
HV	Average	0.530	0.359	0.411	0.41
	SD	0.245	0.145	0.149	0.184
AHV	Average	0.574	0.365	0.438	0.410
	SD	0.273	0.149	0.154	0.127

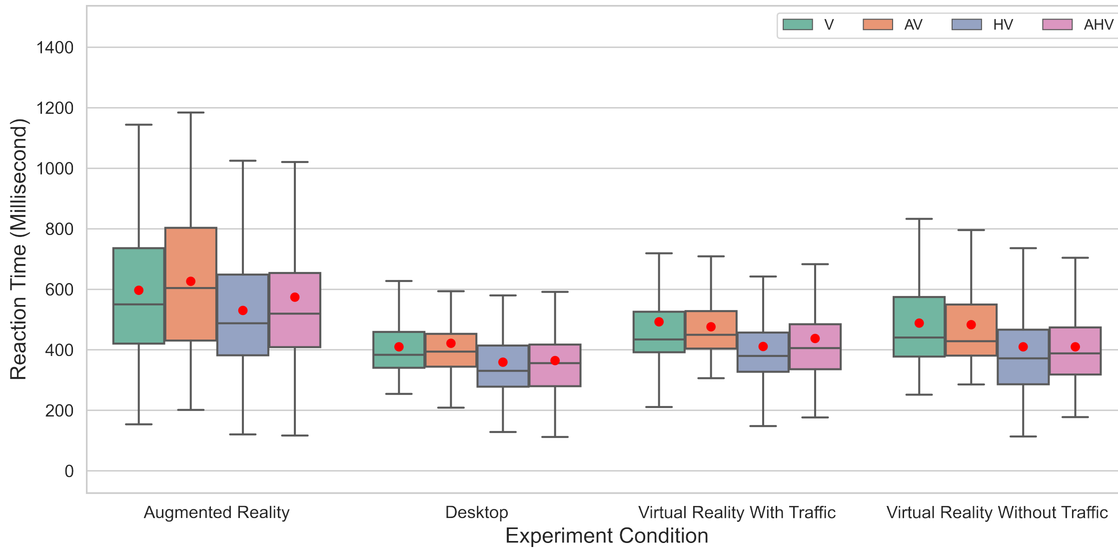


Figure 4.6: Comparative Analysis of Reaction Times for Different Warning Mechanisms (V: Visual, AV: AudioVisual, HV: Haptic Visual, AHV: Haptic AudioVisual) across Experimental Conditions (WT: With Traffic, WOT: Without Traffic)

4.4.1 Impact of Warning Design and Experiment Condition on Reaction Time

Figure 4.6 presents the results of experiments A to D, showcasing reaction times to different warning designs under the experimental conditions defined. Several patterns and trends can be observed from this graph. The findings suggest that, on average, the reaction times to AR warnings in the real world exhibited a longer duration and higher variability compared to the baseline of desktop warnings, which AHVe shorter duration and lower variability. These results are in line with our expectations, as the real-world environment involves more dynamic situations and cognitive distractions compared to the controlled indoor environment. Furthermore, reaction times to simulated AR warnings in virtual reality, both with and without traffic presence, exhibited similar patterns. The recorded values fall between the RTs of the real-world AR and the controlled desktop environment. This suggests that the controlled virtual environment in VR allows for more efficient and focused interactions than in the real world, resulting in reduced reaction times compared to real-world scenarios.

Figure 4.7 shows the results of the t-tests conducted to compare reaction times for

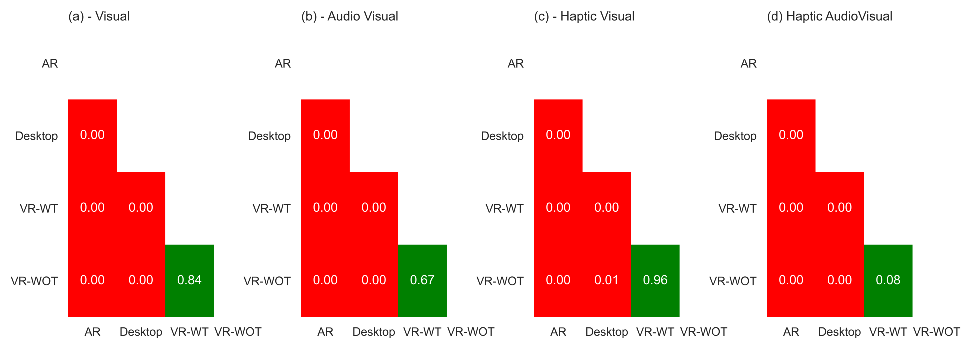


Figure 4.7: The Results of t-test Conducted Between Each Experiment for Different warning Designs

different warning designs under various conditions. The figure indicates that there were no significant differences in the reaction time to simulated AR warnings in VR when traffic was present or absent. This implies that, within the scope of our study, the presence of traffic did not noticeably affect participants' reaction times when interacting with AR warnings in the simulated environment. Furthermore, the figure suggests that the mean values collected under the real-world and simulated environments are statistically different, indicating that RTs to VR-simulated warnings are not equivalent to real-world measurements. This highlights the importance of conducting real-world testbeds for evaluating safety-related aspects, as simulation studies

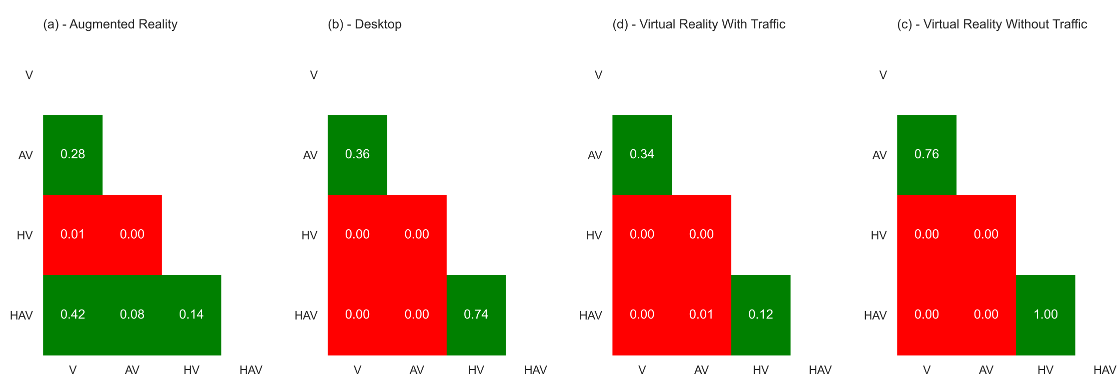


Figure 4.8: The Results of t-test Conducted Between Each Warning Design (V:Visual, AV:AudioVisual, HV:Haptic Visual, AHV:Haptic AudioVisual) for Different Experiment Conditions

cannot fully replicate real-world scenarios in terms of reaction time. Furthermore, the figure demonstrates that, on average, both real-world AR warnings and simulated AR warnings resulted in higher reaction times compared to the baseline of the desktop in all types of warning design. Therefore, designing AR-oriented solutions should not be based on the assumption of similarity between AR and desktop reaction times.

Figure 4.6 also suggests that, on average, HV warnings resulted in lower reaction times under different conditions. Table 4.2 provides a comprehensive summary of the reaction time measurements collected in the experiments. Figure 4.8 indicates that this difference is statistically significant compared to V and AV warnings. However, there is no significant difference between HV and haptic AV warnings. This suggests that HV alone can trigger similar response times in participants similar to those of AHV. This observation is consistent with the overall findings, where AV warnings generally resulted in longer reaction times compared to other warning designs, including visual-only warnings, except for the case of virtual reality. Furthermore, the trends also indicate that, on average, haptic visual warnings lead to lower reaction times in all different prototypes. This suggests that, in this particular context, the audio component may not be as influential in capturing participants' attention, possibly due to the noisy environment of highway work zones. However, it should be noted that the incorporation of additional audio frequencies characteristic of the audio module may alter this trend and warrant further investigation.

4.4.2 Potential of Vision-based Metric for Reaction Time Analysis

In this section, we present an evaluation of the results obtained from our proposed vision-based approach for quantifying reaction time. We set the duration of the Gaussian kernel as the reaction time of each participant to the AHV warning design recorded under the simulated AR condition in VR with traffic in the previous steps. This choice was made because it closely resembled the conditions of the current round of experiments. We then applied our convolution analysis approach

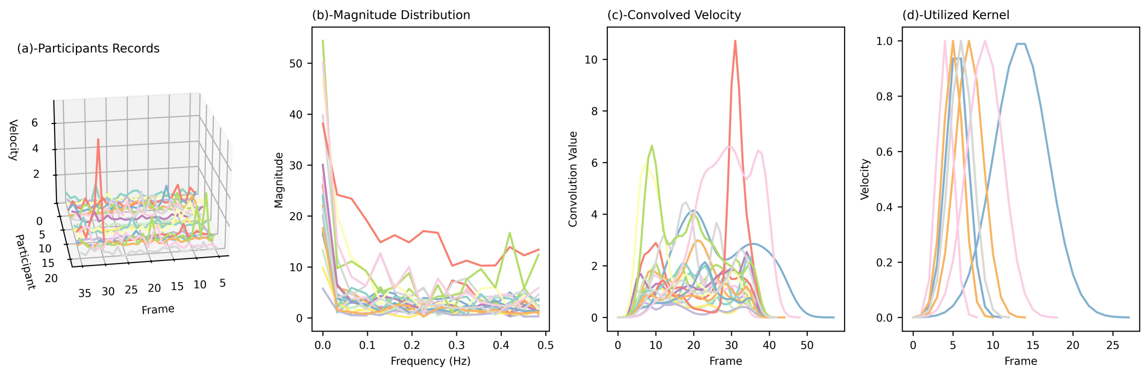


Figure 4.9: Specifications of the Adopted Time-series Analysis in the First warning: (a) Velocity Time-series of Upper Body Cumulative Joint Movement in Participants, (b) Magnitude Distribution Per Frequency (c) Convolution Results of Gaussian Kernel on Time-series and (d) Utilized Individual Kernels for Each Participant Based on the Recorded Baselines

and summarized the results in Figures 4.9 and 4.10, which depict the outcomes for warnings 1 and 2, respectively. To calculate the strength of the frequency content of the velocity time-series, we used Fast Fourier Transform (FFT), and decomposed the signal into frequency domain, and calculated the magnitude of each frequency strength. The results are summarized in Figures 4.9(b) and 4.10(b). These figures highlight a consistent trend in the strength distribution across different frequencies in

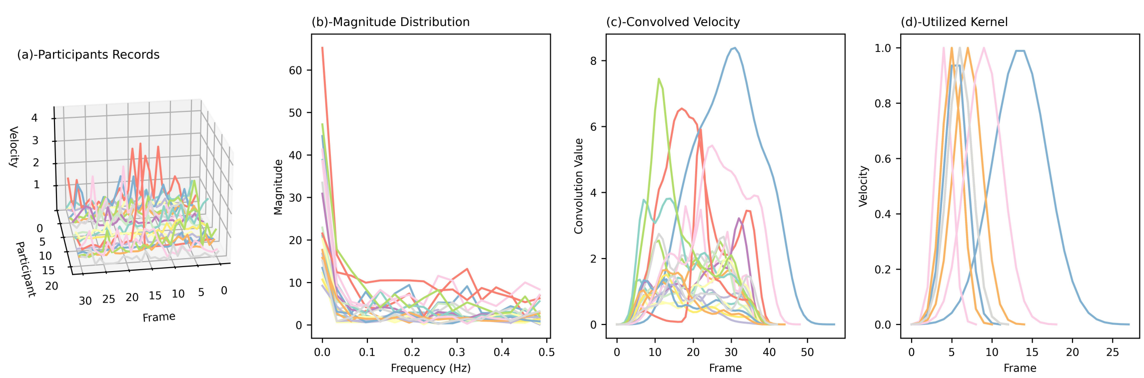


Figure 4.10: Specifications of the Adopted Time-series Analysis in the Second warning: (a) Velocity Time-series of Upper Body Cumulative Joint Movement in Participants, (b) Magnitude Distribution Per Frequency (c) Convolution Results of Gaussian Kernel on Time-series and (d) Utilized Individual Kernels for Each Participant Based on the Recorded Baselines

the post-warning velocity of upper body movement. Additionally, the results indicate a similar one-peak convolution distribution among different participants.

Furthermore, we performed wavelet analysis and presented the results in Figures 4.11 and 4.12. These figures illustrate the coefficients of the frequencies, revealing the similar one-peak distribution in almost all participants with very few exceptions in both post-warning upper body movement velocity. These findings demonstrate

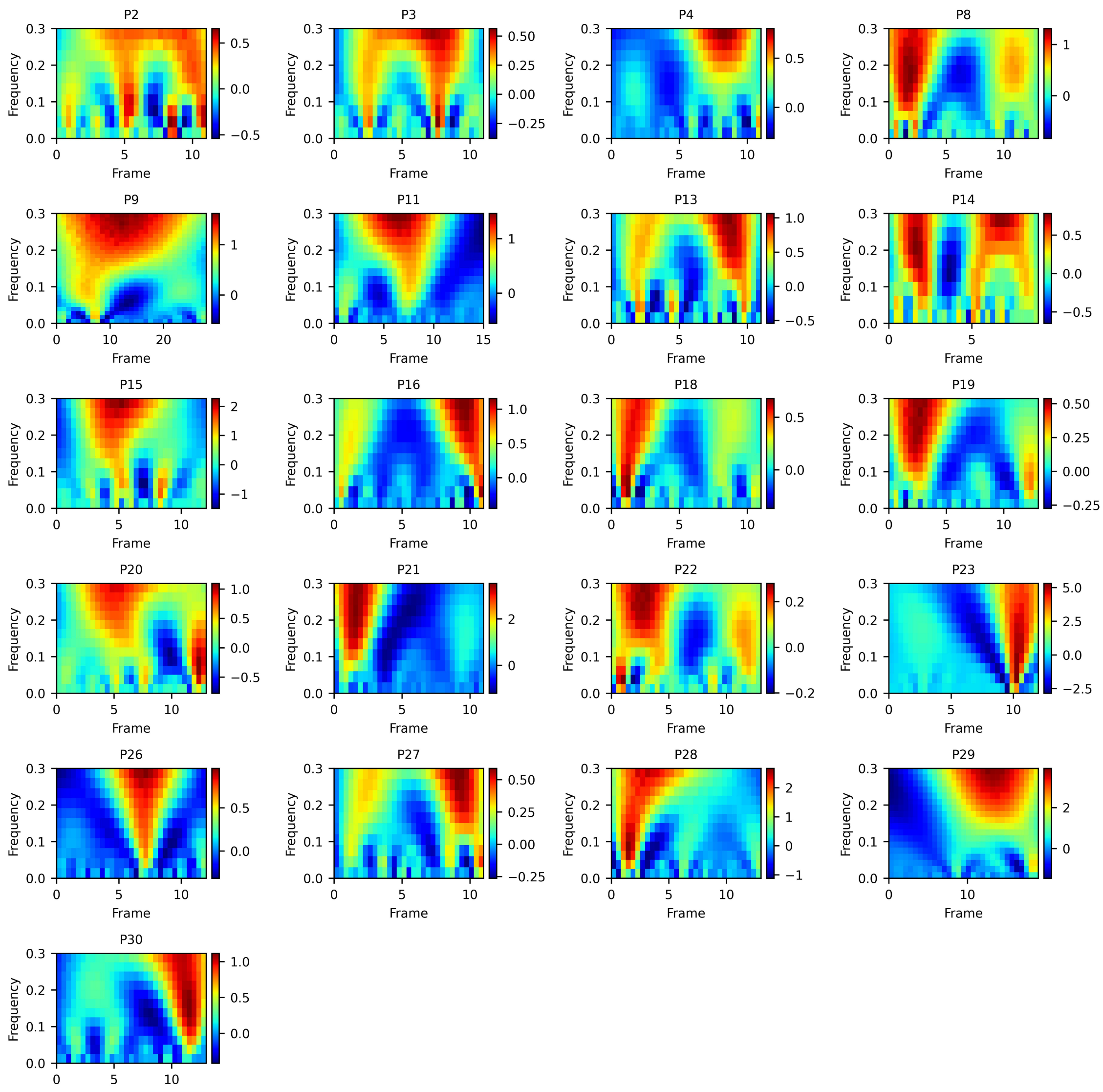


Figure 4.11: Wavelet Analysis Results on Velocity Time-series of Cumulative Upper Body Joint Movement in the First warning for Each Participant

the efficacy of our vision-based approach in capturing and analyzing reaction times patterns in different participants. The consistent trend observed in the energy distribution and coefficients across different frequencies and participants supports the validity of the proposed method.

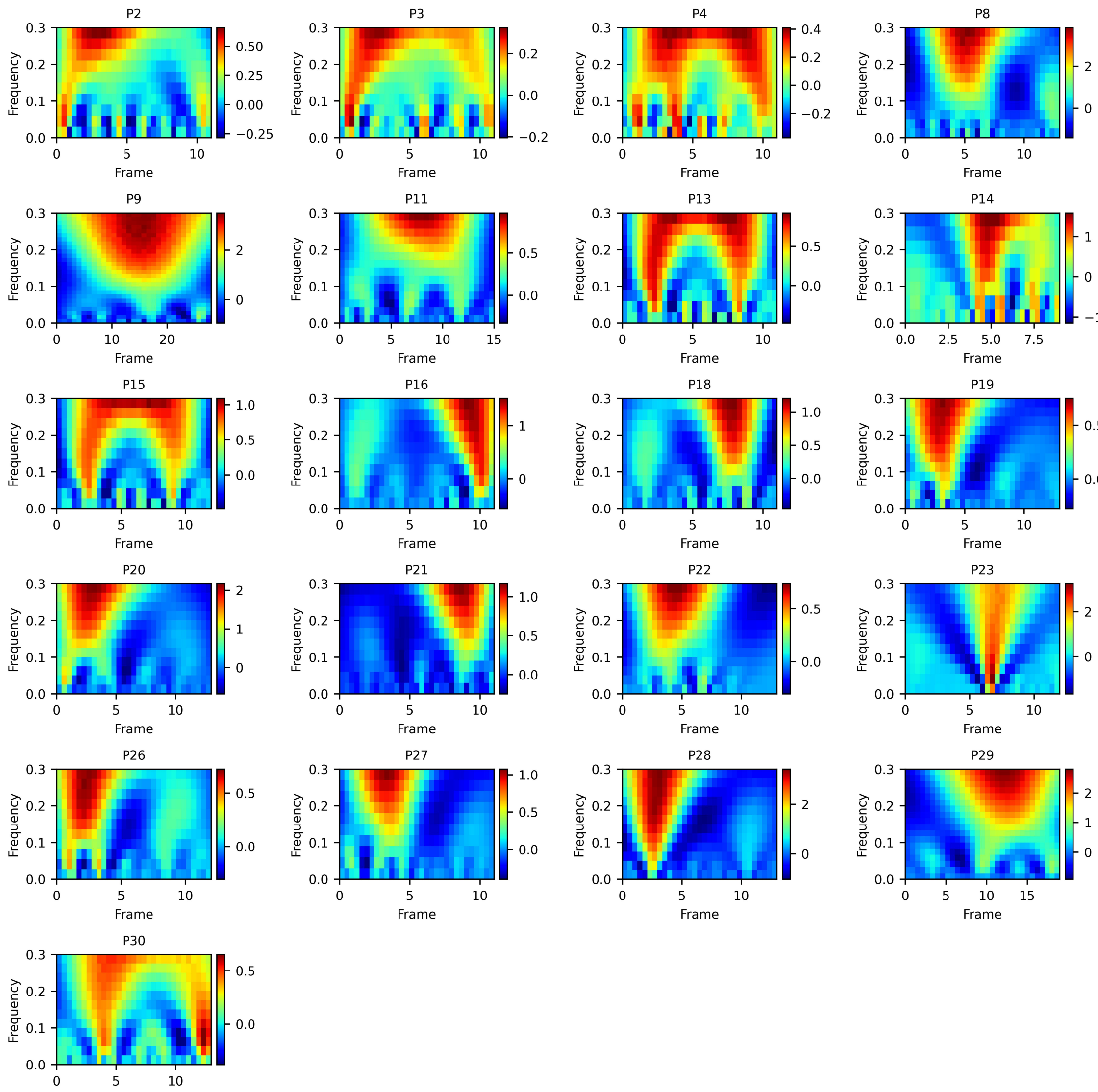


Figure 4.12: Wavelet Analysis Results on Velocity Time-series of Cumulative Upper Body Joint Movement in the Second warning for Each Participant

Finally, Table 4.3 provides a summary of the results obtained from our vision-based approach for quantifying the reaction time. This table presents the average and standard deviation of the reaction times calculated for each participant in the first

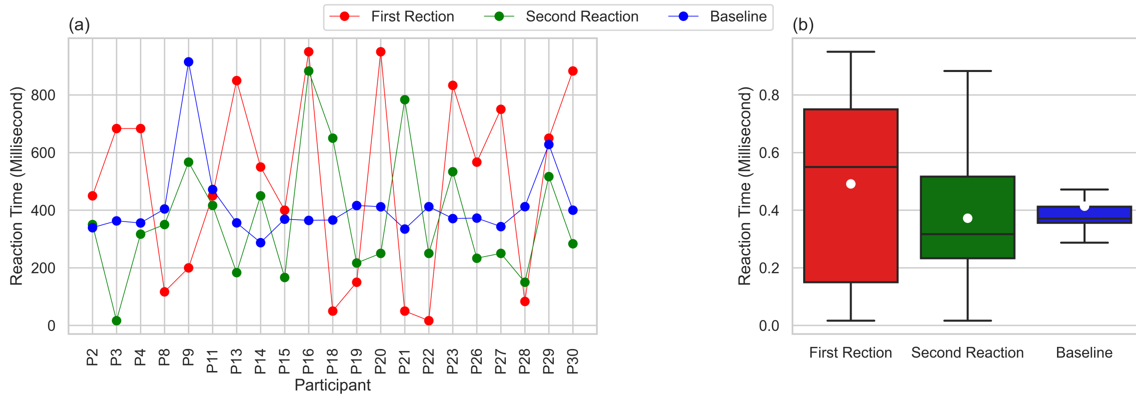


Figure 4.13: (a) Recorded Reaction Times for Each Participant using the Vision-Based Metric in the First and Second warning along with the Recorded Baseline and (b) Box plot of The Recorded Values

and second warning. To validate the effectiveness of our strategy, we performed paired t-tests between these reaction time values and the reaction time to AHV warnings in the VR-simulated warnings, which served as our baseline. The results of the t-tests indicate that the vision-based reaction times are not statistically different from the baseline, suggesting that our vision-based approach yields results comparable to the SRT metrics. Furthermore, we performed paired t-tests between the reaction time measurements in the first and second rounds of vision-based metrics. The results of these t-tests revealed that our strategy resulted in comparable reaction times in both rounds. These findings are visualized in Figure 4.13, where the reaction times collected in the first and second rounds of the vision-based strategy, as well as the baseline, are presented. It is worth noting that although the standard deviation of the vision-based metric was higher than the baseline, the average reaction times were statistically comparable. The consistency of results between the vision-based metric and the SRT baseline, along with the stability of vision-based measurements across multiple iterations, further highlights the potential of computer vision and pose estimation techniques to investigate real-time reaction times. This development has significant implications for real-time safety monitoring applications in highway work zones.

4.5 Conclusion

This paper presents our research conducted to quantify the reaction time to multimodal augmented reality warnings in the context of highway work zone safety. We designed five experiments using Simple Reaction Time (SRT) and a vision-based approach to thoroughly investigate the reaction time triggered by different warning designs in real-world, indoor baseline, and Virtual Reality simulated settings. Our rigorous experimentation yielded a series of results that provide insights into the complex relationship between reaction time and multimodal augmented reality warning design. Specifically, our findings indicate that the haptic visual design triggered the fastest response on average among the participants and produced measurements comparable to those of the audio haptic visual design. Moreover, both of these designs significantly outperformed visual and audio visual warnings in terms of reaction time.

Our findings also reveal that, on average, the reaction time to augmented reality warnings in real-world scenarios was longer with greater variability compared to the baseline of desktop warnings and simulated AR in virtual reality. The results of our statistical comparisons indicated that VR simulated warnings resulted in not statistically significant shorter reaction times than their real-world counterparts. Interestingly, the presence of traffic did not AHVe a significant impact on narrowing the gap between real-world and simulation measurements. This observation suggests that simulating AR in VR may not produce comparable reaction times to those observed in real-world scenarios. Furthermore, we observed a noticeable difference in reaction times between AR warnings and the baseline desktop version under different design

Table 4.3: Summary of Collected Reaction Times Using Vision-based Metric

	First Warning	Second Warning
Average	0.49	0.37
SD	0.33	0.22
Paired t-test with Baseline	0.36	0.42
Paired t-test with Each Other	0.19	

conditions. This emphasizes the importance of considering and accounting for this difference when designing AR-oriented safety systems.

We also developed and proposed a vision-based metric, using real-time pose tracking, to quantify the reaction time. Our approach involved applying a Gaussian kernel and analyzing the velocity of cumulative upper body displacement across frames within a naturalistic task design inspired by the typical operations in highway work zones. By utilizing the argmax function and individual participant records, we identified the initiation time of the Gaussian kernel pattern in the participants' upper body motion velocity time-series through convolution techniques. We then used a within-subject approach to compare these results with the baselines obtained from previous experiments. Our findings demonstrated the statistical comparability of the vision-based metric with the Simple Reaction Time based metrics at an individual level.

Overall, the study's findings offer valuable information on the effectiveness and efficiency of different warning designs to improve safety within highway work zones. Through a systematic analysis of the experimental data, patterns, trends, and correlations were identified, providing a comprehensive understanding of reaction time in multimodal augmented reality scenarios. This research contributes to the field by presenting a systematic approach to studying reaction time in multimodal augmented reality, with a specific focus on highway work zone safety. The insights gained from this study can inform the design and implementation of augmented reality systems in work zone environments, ultimately improving the safety and well-being of workers in these crucial settings.

This study has identified limitations and avenues for further research. While it was found that audio contribution may not be necessary in the warning design, this conclusion is based on the assumption of a constant frequency for the audio. Future studies could explore how variations in the design of the audio module, such as different frequencies or patterns, could affect reaction times and user responses. Additionally, future research can expand on the task design used in this study and apply it to other common activities or scenarios. This could involve investigating the relationship between cognitive load, physical engagement, and reaction time in different contexts. By exploring these factors, researchers can gain a deeper understanding of the complex interplay between various variables and their influence on reaction times.

CHAPTER 5: CONCLUSIONS

Over the past few years, Augmented Reality has garnered increasing attention and recognition as a promising solution for addressing safety issues in various domains. The unique capabilities of AR, which overlay digital information onto the real world, have opened new avenues to enhance safety measures and mitigate risks in different contexts. An area where AR has shown significant potential is in improving the safety of roadway work zones. These working environments present numerous challenges and risks to workers, including traffic hazards, heavy machinery, and the need for effective communication and situational awareness. Traditionally, ensuring the safety of workers in these environments has relied on conventional methods such as signage, barriers, and safety protocols. However, these methods may have limitations in effectively conveying critical information or addressing rapidly changing and dynamic situations.

AR technology offers a transformative approach by augmenting the physical environment with digital information in real time. By overlaying visual cues and warnings directly into the worker's field of view, AR can enhance their situational awareness and provide timely guidance and alerts. This real-time information can help workers navigate complex work zones, identify potential hazards, and make informed decisions to mitigate risks. Moreover, AR can leverage the power of Artificial Intelligence to analyze data from various sources, such as traffic patterns, weather conditions, and lighting conditions, to provide intelligent insights and recommendations. AI algorithms can process large amounts of data and generate actionable information, enabling AR systems to provide customized and context-specific safety warnings to workers. This integration of AI and AR creates a dynamic and adaptive safety system that can respond to changing conditions and provide tailored support to each worker.

The development of high-fidelity prototypes and the co-design of AI and AR frameworks, as mentioned in the previous chapters, are crucial steps in harnessing the potential of AR for highway work zone safety. By thoroughly investigating the feasibility and technical challenges associated with each pillar of the integrative framework, this research attempts to provide a holistic understanding of the proposed AR-based safety concept in highway work zones. This knowledge serves as a solid foundation for the subsequent phases of the study, including the evaluation of usability and user experience in real-world and simulated environments. The design and administration of experiments to evaluate reaction times to different notification designs further contribute to the understanding of effective safety systems. By employing a mixed-method research methodology that combines Simple Reaction Time and a vision-based approach, the approach can gather comprehensive data on participants' reaction times and analyze the influence of various notification modalities.

Overall, in Chapter 2, the focus of this dissertation is on integrating AI capabilities into AR systems to enhance highway work zone safety. The chapter explores the feasibility, requirements, and challenges of incorporating AI to develop a predictive safety system. The outcomes indicate that real-time communication and AI execution meet the timing constraints. Early user research shows positive reception by highway maintenance professionals. Chapter 3 focuses on conducting a mixed-method usability investigation of the proposed AR-based safety system. The evaluation considers user interface design, interaction patterns, and feedback to assess usability and effectiveness. Participants rated the system's usability above average in both indoor and outdoor settings, with perceived trust significantly correlated with usability. Finally, Chapter 4 examines the impact of different sensory modalities on worker reaction times in AR warnings within roadway work zones. The findings reveal that haptic visual and audio haptic visual designs elicited faster responses compared to visual and audiovisual warnings. Reaction times in real-world outdoor scenarios were longer and more variable, with VR simulated warnings showing no significant advantage over real-world counterparts. These results suggest limitations in replicating real-world reaction times using VR simulations.

Ultimately, the insights gained from this research effort provide valuable information on the usability, effectiveness, and user experience of AR in highway work zones. By establishing worker-centered design guidelines and highlighting the potential of new technologies to meet the unique needs of the workforce, this document contributes to the body of knowledge in the field of roadway work zone safety. The findings and recommendations presented in this work have the potential to inform the design and implementation of AR systems in work zone environments, ultimately improving the safety and well-being of workers in these crucial settings.

REFERENCES

[1] Centers for Disease Control and Prevention, "Highway Work Zone Safety," 2019. https://www.cdc.gov/niosh/topics/highwayworkzones/.

[2] C. Nnaji, J. Gambatese, H. W. Lee, and F. Zhang, "Improving construction work zone safety using technology: A systematic review of applicable technologies," *Journal of traffic and transportation engineering (English edition)*, vol. 7, no. 1, pp. 61–75, 2020.

[3] S. Sakhakarmi, J. Park, and A. Singh, "Tactile-based wearable system for improved hazard perception of worker and equipment collision," *Automation in Construction*, vol. 125, p. 103613, 2021.

[4] C. R. Ahn, S. Lee, C. Sun, H. Jebelli, K. Yang, and B. Choi, "Wearable sensing technology applications in construction safety and health," *Journal of Construction Engineering and Management*, vol. 145, no. 11, p. 03119007, 2019.

[5] X. Li, W. Yi, H.-L. Chi, X. Wang, and A. P. Chan, "A critical review of virtual and augmented reality (vr/ar) applications in construction safety," *Automation in Construction*, vol. 86, pp. 150–162, 2018.

[6] L. Hou, H. Chen, G. K. Zhang, and X. Wang, "Deep learning-based applications for safety management in the aec industry: A review," *Applied Sciences*, vol. 11, no. 2, p. 821, 2021.

[7] B. Fischer, A. Peine, and B. Östlund, "The importance of user involvement: a systematic review of involving older users in technology design," *The Gerontologist*, vol. 60, no. 7, pp. e513–e523, 2020.

[8] H. Taherdoost, "Importance of technology acceptance assessment for successful implementation and development of new technologies," *Global Journal of Engineering Sciences*, vol. 1, no. 3, 2019.

[9] A. L. Ármannsdóttir, P. Beckerle, J. C. Moreno, E. H. van Asseldonk, M.-T. Manrique-Sancho, A. J. Del-Ama, J. F. Veneman, and K. Briem, "Assessing the involvement of users during development of lower limb wearable robotic exoskeletons: a survey study," *Human factors*, vol. 62, no. 3, pp. 351–364, 2020.

[10] T. Grundgeiger, J. Hurtienne, and O. Happel, "Why and how to approach user experience in safety-critical domains: The example of health care," *Human factors*, p. 0018720819887575, 2020.

[11] "Facts and statisticsâwork zone safety," *Federal Highway Administration (FHWA)*, 2018.

[12] J. A. Gambatese, H. W. Lee, C. A. Nnaji, *et al.*, "Work zone intrusion alert technologies: Assessment and practical guidance," tech. rep., Oregon. Dept. of Transportation. Research Section, 2017.

[13] J. Hourdos, "Portable, non-intrusive advance warning devices for work zones with or without flag operators," 2012.

[14] I. Awolusi and E. D. Marks, "Active work zone safety: preventing accidents using intrusion sensing technologies," *Frontiers in built environment*, vol. 5, p. 21, 2019.

[15] P. B. Fyhrie, A. H. Maintenance, and C. T. R. C. (Calif.), *Work Zone Intrusion Alarms for Highway Workers*. Caltrans Division of Research, Innovation and System Information, 2016.

[16] Y. Gheraibia, S. Kabir, K. Aslansefat, I. Sorokos, and Y. Papadopoulos, "Safety+ ai: A novel approach to update safety models using artificial intelligence," *IEEE Access*, vol. 7, pp. 135855–135869, 2019.

[17] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," *Automation in Construction*, vol. 112, p. 103085, 2020.

[18] Y. A. Kim, B. Y. Ryoo, Y.-S. Kim, and W. C. Huh, "Major accident factors for effective safety management of highway construction projects," *Journal of construction engineering and management*, vol. 139, no. 6, pp. 628–640, 2013.

[19] A. I. Glendon and D. K. Litherland, "Safety climate factors, group differences and safety behaviour in road construction," *Safety science*, vol. 39, no. 3, pp. 157–188, 2001.

[20] Y. Li and Y. Bai, "Highway work zone risk factors and their impact on crash severity," *Journal of Transportation engineering*, vol. 135, no. 10, pp. 694–701, 2009.

[21] E. C. Haas and J. B. Van Erp, "Multimodal warnings to enhance risk communication and safety," *Safety science*, vol. 61, pp. 29–35, 2014.

[22] C. Geitner, F. Biondi, L. Skrypchuk, P. Jennings, and S. Birrell, "The comparison of auditory, tactile, and multimodal warnings for the effective communication of unexpected events during an automated driving scenario," *Transportation research part F: traffic psychology and behaviour*, vol. 65, pp. 23–33, 2019.

[23] C. L. Baldwin, C. Spence, J. P. Bliss, J. C. Brill, M. S. Wogalter, C. B. Mayhorn, and T. K. Ferris, "Multimodal cueing: The relative benefits of the auditory, visual, and tactile channels in complex environments," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, pp. 1431–1435, SAGE Publications Sage CA: Los Angeles, CA, 2012.

[24] H. Yun and J. H. Yang, "Multimodal warning design for take-over request in conditionally automated driving," *European transport research review*, vol. 12, pp. 1–11, 2020.

[25] H. S. Park and K.-H. Kim, "Adaptive multimodal in-vehicle information system for safe driving," *ETRI Journal*, vol. 37, no. 3, pp. 626–636, 2015.

[26] A. Syberfeldt, O. Danielsson, and P. Gustavsson, "Augmented reality smart glasses in the smart factory: Product evaluation guidelines and review of available products," *Ieee Access*, vol. 5, pp. 9118–9130, 2017.

[27] W. Barfield, *Fundamentals of wearable computers and augmented reality*. CRC press, 2015.

[28] G. Kipper and J. Rampolla, *Augmented Reality: an emerging technologies guide to AR*. Elsevier, 2012.

[29] D. Chatzopoulos, C. Bermejo, Z. Huang, and P. Hui, "Mobile augmented reality survey: From where we are to where we go," *IEEE Access*, vol. 5, pp. 6917–6950, 2017.

[30] J. Zhang, S.-K. Ong, and A. Y. Nee, "A multi-regional computation scheme in an ar-assisted in situ cnc simulation environment," *Computer-Aided Design*, vol. 42, no. 12, pp. 1167–1177, 2010.

[31] S. Ong, A. Yew, N. Thanigaivel, and A. Nee, "Augmented reality-assisted robot programming system for industrial applications," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101820, 2020.

[32] C. Merenda, M. Smith, J. Gabbard, G. Burnett, and D. Large, "Effects of real-world backgrounds on user interface color naming and matching in automotive ar huds," in *2016 IEEE VR 2016 Workshop on Perceptual and Cognitive Issues in AR (PERCAR)*, pp. 1–6, IEEE, 2016.

[33] W. P. Chan, G. Hanks, M. Sakr, T. Zuo, H. M. Van der Loos, and E. Croft, "An augmented reality human-robot physical collaboration interface design for shared, large-scale, labour-intensive manufacturing tasks," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11308–11313, IEEE, 2020.

[34] K. Zhang and M. Hassan, "Crash severity analysis of nighttime and daytime highway work zone crashes," *PLoS one*, vol. 14, no. 8, p. e0221128, 2019.

[35] A. Ghasemzadeh and M. M. Ahmed, "Exploring factors contributing to injury severity at work zones considering adverse weather conditions," *IATSS research*, vol. 43, no. 3, pp. 131–138, 2019.

[36] H. Brown, C. Sun, and T. Cope, "Evaluation of mobile work zone alarm systems," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2485, pp. 42–50, 2015.

[37] Y. Huang and Y. Bai, "Effectiveness of graphic-aided portable changeable message signs in reducing vehicle speeds in highway work zones," *Transportation research part C: emerging technologies*, vol. 48, pp. 311–321, 2014.

[38] J. Park and S. Sakhakarmi, "Embedded safety communication system for robust hazard perception of individuals in work zones," 2019.

[39] K. Chan, J. Louis, and A. Albert, "Incorporating worker awareness in the generation of hazard proximity warnings," *Sensors*, vol. 20, no. 3, p. 806, 2020.

[40] K. Kim, H. Kim, and H. Kim, "Image-based construction hazard avoidance system using augmented reality in wearable device," *Automation in construction*, vol. 83, pp. 390–403, 2017.

[41] C. Nnaji, A. A. Karakhan, J. Gambatese, and H. W. Lee, "Case study to evaluate work-zone safety technologies in highway construction," *Practice Periodical on Structural Design and Construction*, vol. 25, no. 3, p. 05020004, 2020.

[42] D. Brougham and J. Haar, "Employee assessment of their technological redundancy," *Labour & Industry: a journal of the social and economic relations of work*, vol. 27, no. 3, pp. 213–231, 2017.

[43] A. Naweed and J. Rose, "Assessing technology acceptance for skills development and real-world decision-making in the context of train driving," *Transportation research part F: traffic psychology and behaviour*, vol. 52, pp. 86–100, 2018.

[44] D. Liu, W. Lu, and Y. Niu, "Extended technology-acceptance model to make smart construction systems successful," *Journal of Construction Engineering and Management*, vol. 144, no. 6, p. 04018035, 2018.

[45] H. Son, Y. Park, C. Kim, and J.-S. Chou, "Toward an understanding of construction professionals' acceptance of mobile computing devices in south korea: An extension of the technology acceptance model," *Automation in construction*, vol. 28, pp. 82–90, 2012.

[46] M. Jacobsson and H. C. Linderoth, "User perceptions of ict impacts in swedish construction companies:âitâs fine, just as it isâ," *Construction management and economics*, vol. 30, no. 5, pp. 339–357, 2012.

[47] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[48] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[49] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[50] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9, 2015.

[51] H. G. C. X. J. Z. Zhang, Xinyu and Y. Liu, "Real-time vehicle detection and tracking using improved histogram of gradient features and kalman filters," *International Journal of Advanced Robotic Systems*, vol. 15, 2019.

[52] H. Fu, H. Ma, G. Wang, X. Zhang, and Y. Zhang, "Mcff-cnn: Multiscale comprehensive feature fusion convolutional neural network for vehicle color recognition based on residual learning," *Neurocomputing*, pp. 178–187, 2019.

[53] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Computer Vision and Pattern Recognition (CVPR)*, 2015.

[54] K. Kim, P. Kim, Y. Chung, and D. Choi, "Multi-scale detector for accurate vehicle detection in traffic surveillance data," *IEEE Access*, vol. 7, pp. 78311–78319, 2019.

[55] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv*, 2018.

[56] L. Wen, D. Du, Z. Cai, Z. Lei, M. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu, "Ua-detrac: A new benchmark and protocol for multi-object detection and tracking," *Computer Vision and Image Understanding*, vol. 193, Apr. 2020.

[57] A. Arinaldi, J. Pradana, and A. Gurusinga, "Detection and classification of vehicles for traffic video analytics," *Procedia Computer Science*, vol. 144, pp. 259–268, 01 2018.

[58] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *CoRR*, vol. abs/1506.01497, 2015.

[59] C. Yingfeng, L. Ze, S. Xiaoqiang, and Y. Zou, "Selecting optimal proposal number for image-based object detection," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3797–3801, May 2019.

[60] S. Rujikietgumjorn and N. Watcharapinchai, "Vehicle detection with sub-class training using r-cnn for the ua-detrac benchmark," in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–5, 2017.

[61] F. Zhu, Y. Lu, N. Ying, and G. Giakos, "Fast vehicle detection based on evolving convolutional neural network," pp. 1–4, 10 2017.

[62] W. Guan, Y. Zou, and X. Zhou, "Multi-scale object detection with feature fusion and region objectness network," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2596–2600, April 2018.

[63] Q. Li, H. Dang, Y. Zhang, and D. Wu, "Video vehicle detection and tracking system," in *Advances in Computer Science and Education Applications* (M. Zhou and H. Tan, eds.), (Berlin, Heidelberg), pp. 24–31, Springer Berlin Heidelberg, 2011.

[64] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice with MATLAB*. Wiley-IEEE Press, 4th ed., 2014.

[65] J. Peng, T. Shen, Y. Wang, T. Zhao, J. Zhang, and X. Fu, "Continuous vehicle detection and tracking for non-overlapping multi-camera surveillance system," in *Proceedings of the International Conference on Internet Multimedia Computing and Service*, ICIMCSâ16, (New York, NY, USA), p. 122â125, Association for Computing Machinery, 2016.

[66] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR â14, (USA), p. 580â587, IEEE Computer Society, 2014.

[67] Y. Zou, W. Zhang, W. Weng, and Z. Meng, "Multi-vehicle tracking via real-time detection probes and a markov decision process policy," *Sensors*, vol. 19, p. 1309, 03 2019.

[68] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, AFIPS '68 (Fall, part I), (New York, NY, USA), pp. 757–764, ACM, 1968.

[69] M. Satyanarayanan, "From the editor in chief: Augmenting cognition," *IEEE Pervasive Computing*, vol. 3, pp. 4–5, April 2004.

[70] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for vm-based cloudlets in mobile computing," *IEEE Pervasive Computing*, vol. 8, pp. 14–23, Oct 2009.

[71] S. J. Yohan, S. Julier, Y. Baillot, M. Lanzagorta, D. Brown, and L. Rosenblum, "Bars: Battlefield augmented reality system," in *In NATO Symposium on Information Processing Techniques for Military Systems*, pp. 9–11, 2000.

[72] M. Träskbäack and M. Haller, "Mixed reality training application for an oil refinery: User requirements," in *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry*, VRCAI '04, (New York, NY, USA), pp. 324–327, ACM, 2004.

[73] G. Klinker, O. Creighton, A. H. Dutoit, R. Kobylinski, C. Vilsmeier, and B. Brugge, “Augmented maintenance of powerplants: a prototyping case study of a mobile ar system,” in *Proceedings IEEE and ACM International Symposium on Augmented Reality*, pp. 124–133, Oct 2001.

[74] S. Goose, S. Güven, X. Zhang, S. Sudarsky, and N. Navab, “Paris : Fusing vision-based location tracking with standards-based 3 d visualization and speech interaction on a pda,” 2004.

[75] S. White, S. Feiner, and J. Kopylec, “Virtual vouchers: Prototyping a mobile augmented reality user interface for botanical species identification,” in *3D User Interfaces (3DUI’06)*, pp. 119–126, March 2006.

[76] S. Deffeyes, “Mobile augmented reality in the data center,” *IBM Journal of Research and Development*, vol. 55, pp. 5:1–5:5, Sep. 2011.

[77] F. K. Wacker, S. K. Vogt, A. Khamene, J. A. Jesberger, S. G. Nour, D. R. Elgort, F. Sauer, J. L. Duerk, and J. S. Lewin, “An augmented reality system for mr image-guided needle biopsy: initial results in a swine model.,” *Radiology*, vol. 238 2, pp. 497–504, 2006.

[78] M. Hakkarainen, C. Woodward, and M. Billinghurst, “Augmented assembly using a mobile phone,” in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 167–168, Sep. 2008.

[79] S. J. Henderson and S. K. Feiner, “Augmented reality in the psychomotor phase of a procedural task,” in *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, pp. 191–200, Oct 2011.

[80] A. Tang, C. Owen, F. Biocca, and W. Mou, “Comparative effectiveness of augmented reality in object assembly,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’03, (New York, NY, USA), pp. 73–80, ACM, 2003.

[81] H. Kim, A. Miranda Anon, T. Misu, N. Li, A. Tawari, and K. Fujimura, “Look at me: Augmented reality pedestrian warning system using an in-vehicle volumetric head up display,” in *Proceedings of the 21st International Conference on Intelligent User Interfaces*, pp. 294–298, 2016.

[82] D. Tatić and B. Tešić, “The application of augmented reality technologies for the improvement of occupational safety in an industrial environment,” *Computers in Industry*, vol. 85, pp. 1–10, 2017.

[83] S. Kim, M. A. Nussbaum, and J. L. Gabbard, “Augmented reality âsmart glassesâ in the workplace: industry perspectives and challenges for worker safety and health,” *IIE transactions on occupational ergonomics and human factors*, vol. 4, no. 4, pp. 253–258, 2016.

[84] S. Helmet, "Smart helmet - daqri." `https://daqri.com/`. (Accessed on 04/10/2019).

[85] G. Schall, E. Mendez, E. Kruijff, E. Veas, S. Junghanns, B. Reitinger, and D. Schmalstieg, "Handheld augmented reality for underground infrastructure visualization," *Personal and ubiquitous computing*, vol. 13, no. 4, pp. 281–291, 2009.

[86] S. A. Talmaki, S. Dong, and V. R. Kamat, "Geospatial databases and augmented reality visualization for improving safety in urban excavation operations," in *Construction Research Congress*, vol. 2010, pp. 91–101, 2010.

[87] X. Li, W. Yi, H.-L. Chi, X. Wang, and A. P. Chan, "A critical review of virtual and augmented reality (vr/ar) applications in construction safety," *Automation in Construction*, vol. 86, pp. 150 – 162, 2018.

[88] C.-S. Park and H.-J. Kim, "A framework for construction safety management and visualization system," *Automation in Construction*, vol. 33, pp. 95–103, 2013.

[89] J. M. D. Delgado, L. Oyedele, P. Demian, and T. Beach, "A research agenda for augmented and virtual reality in architecture, engineering and construction," *Advanced Engineering Informatics*, vol. 45, p. 101122, 2020.

[90] J. Nielsen, "Ten usability heuristics." `http://www.nngroup.com/articles/ten-usability-heuristics/(acc-essed~âĂę`, 2005. [Online; accessed 2022].

[91] D. Quiñones and C. Rusu, "How to develop usability heuristics: A systematic literature review," *Computer Standards & Interfaces*, vol. 53, pp. 89–122, 2017.

[92] T. MUTCD, "Manual on uniform traffic control devices," *Texas Department of Transportation, Austin*, 2006.

[93] J. Kim, Y. Jo, and J. P. Jeong, "Design and evaluation of a smartphone-based alarming system for pedestrian safety in vehicular networks," in *International Conference on Internet of Vehicles*, pp. 221–233, Springer, 2015.

[94] A. Bochkovskiy, C. Wang, and H. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *CoRR*, vol. abs/2004.10934, 2020.

[95] C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "Cspnet: A new backbone that can enhance learning capability of cnn," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1571–1580, 2020.

[96] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proceedings of the 36th International Conference on Machine Learning* (K. Chaudhuri and R. Salakhutdinov, eds.), vol. 97 of *Proceedings of Machine Learning Research*, (Long Beach, California, USA), pp. 6105–6114, PMLR, 09–15 Jun 2019.

[97] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 8759–8768, 2018.

[98] P. Karn and C. Partridge, "Improving round-trip time estimates in reliable transport protocols," *ACM SIGCOMM Computer Communication Review*, vol. 17, no. 5, pp. 2–7, 1987.

[99] D. Schoonwinkel, *Practical measurements of Wi-Fi Direct in content sharing, social gaming android applications*. PhD thesis, Stellenbosch: Stellenbosch University, 2016.

[100] Android, "Socket class. retrieved from: `https://developer.android.com/reference/java/net/Socket`,"

[101] Android, "Serversocket class. retrieved from: `https://developer.android.com/reference/java/net/ServerSocket`,"

[102] Python, "Socket class. retrieved from: `https://docs.python.org/3/library/socket.html`,"

[103] A. Garre, M. C. Ruiz, and E. Hontoria, "Application of machine learning to support production planning of a food industry in the context of waste generation under uncertainty," *Operations Research Perspectives*, p. 100147, 2020.

[104] J. L. Harbluk, Y. I. Noy, P. L. Trbovich, and M. Eizenman, "An on-road assessment of cognitive distraction: Impacts on driversâ visual behavior and braking performance," *Accident Analysis & Prevention*, vol. 39, no. 2, pp. 372–379, 2007.

[105] Y.-C. Liu, "Comparative study of the effects of auditory, visual and multimodality displays on drivers' performance in advanced traveller information systems," *Ergonomics*, vol. 44, no. 4, pp. 425–442, 2001.

[106] J. Rieman, M. Franzke, and D. Redmiles, "Usability evaluation with the cognitive walkthrough," in *Conference companion on Human factors in computing systems*, pp. 387–388, 1995.

[107] T. Mahatody, M. Sagar, and C. Kolski, "State of the art on the cognitive walkthrough method, its variants and evolutions," *Intl. Journal of Human–Computer Interaction*, vol. 26, no. 8, pp. 741–785, 2010.

[108] V. Blade, "Vuzix Blade." `https://www.vuzix.com/products/blade-smart-glasses-upgraded`, 2021. [Online; accessed 2022].

[109] tp link, "https://www.tp-link.com/us/home-networking/wifi-router/archer-ax50/," 2020.

[110] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "Bdd100k: A diverse driving dataset for heterogeneous multitask learning," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2633–2642, 2020.

[111] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *Computer Vision – ECCV 2014* (D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, eds.), (Cham), pp. 740–755, Springer International Publishing, 2014.

[112] M. Everingham, L. Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *Int. J. Comput. Vision*, vol. 88, p. 303â338, June 2010.

[113] M. L. McHugh, "The chi-square test of independence," *Biochemia medica: Biochemia medica*, vol. 23, no. 2, pp. 143–149, 2013.

[114] S. Sternberg, "Reaction times and the ex-gaussian distribution: When is it appropriate," 2014.

[115] H. Yang, K. Ozbay, O. Ozturk, and K. Xie, "Work zone safety analysis and modeling: a state-of-the-art review," *Traffic injury prevention*, vol. 16, no. 4, pp. 387–396, 2015.

[116] Y.-J. Chen, Y.-S. Lai, and Y.-H. Lin, "Bim-based augmented reality inspection and maintenance of fire safety equipment," *Automation in Construction*, vol. 110, p. 103041, 2020.

[117] S. Wu, L. Hou, G. K. Zhang, and H. Chen, "Real-time mixed reality-based visual warning for construction workforce safety," *Automation in Construction*, vol. 139, p. 104252, 2022.

[118] K. E. Schein and P. A. Rauschnabel, "Augmented reality in manufacturing: exploring workersâ perceptions of barriers," *IEEE Transactions on Engineering Management*, 2021.

[119] K. Savolainen, "User-centred design without involving users: A longitudinal case study in a human-centred-design–mature company," *The Design Journal*, vol. 24, no. 6, pp. 887–905, 2021.

[120] M. Schumann, C. Fuchs, C. Kollatsch, and P. Klimant, "Evaluation of augmented reality supported approaches for product design and production processes," *Procedia CIRP*, vol. 97, pp. 160–165, 2021.

[121] T. A. Schoonderwoerd, W. Jorritsma, M. A. Neerincx, and K. Van Den Bosch, "Human-centered xai: Developing design patterns for explanations of clinical decision support systems," *International Journal of Human-Computer Studies*, vol. 154, p. 102684, 2021.

[122] B. A. Kadir and O. Broberg, “Human-centered design of work systems in the transition to industry 4.0,” *Applied ergonomics*, vol. 92, p. 103334, 2021.

[123] G. Sorensen, J. T. Dennerlein, S. E. Peters, E. L. Sabbath, E. L. Kelly, and G. R. Wagner, “The future of research on work, safety, health and wellbeing: A guiding conceptual framework,” *Social science & medicine*, vol. 269, p. 113593, 2021.

[124] S. Sabeti, O. Shoghli, M. Baharani, and H. Tabkhi, “Toward ai-enabled augmented reality to enhance the safety of highway work zones: Feasibility, requirements, and challenges,” *Advanced Engineering Informatics*, vol. 50, p. 101429, 2021.

[125] J. Brooke, “Usability evaluation in industry, chap. sus: a âquick and dirtyâ usability scale,” 1996.

[126] A. F. Borges, F. J. Laurindo, M. M. Spínola, R. F. Gonçalves, and C. A. Mattos, “The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions,” *International Journal of Information Management*, vol. 57, p. 102225, 2021.

[127] N. Ashtari, A. Bunt, J. McGrenere, M. Nebeling, and P. K. Chilana, “Creating augmented and virtual reality applications: Current practices, challenges, and opportunities,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.

[128] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. Hancock, “The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis,” *Human factors*, vol. 63, no. 4, pp. 706–726, 2021.

[129] M. Grabowski, A. Rowen, and J.-P. Rancy, “Evaluation of wearable immersive augmented reality technology in safety-critical systems,” *Safety science*, vol. 103, pp. 23–32, 2018.

[130] H. Chen, L. Hou, G. K. Zhang, and S. Moon, “Development of bim, iot and ar/vr technologies for fire safety and upskilling,” *Automation in Construction*, vol. 125, p. 103631, 2021.

[131] J. Bang, Y. Lee, Y.-T. Lee, and W. Park, “Ar/vr based smart policing for fast response to crimes in safe city,” in *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 470–475, IEEE, 2019.

[132] R. Lovreglio and M. Kinateder, “Augmented reality for pedestrian evacuation research: promises and limitations,” *Safety science*, vol. 128, p. 104750, 2020.

[133] I. Awolusi, E. Marks, and M. Hallowell, "Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices," *Automation in construction*, vol. 85, pp. 96–106, 2018.

[134] M. Akinlolu, T. C. Haupt, D. J. Edwards, and F. Simpeh, "A bibliometric review of the status and emerging research trends in construction safety management technologies," *International Journal of Construction Management*, pp. 1–13, 2020.

[135] K. Dhalmahapatra, J. Maiti, and O. Krishna, "Assessment of virtual reality based safety training simulator for electric overhead crane operations," *Safety science*, vol. 139, p. 105241, 2021.

[136] F. Dai, A. Olorunfemi, W. Peng, D. Cao, and X. Luo, "Can mixed reality enhance safety communication on construction sites? an industry perspective," *Safety Science*, vol. 133, p. 105009, 2021.

[137] A. Harikrishnan, A. S. Abdallah, S. K. Ayer, M. El Asmar, and P. Tang, "Feasibility of augmented reality technology for communication in the construction industry," *Advanced Engineering Informatics*, vol. 50, p. 101363, 2021.

[138] J. Mallela, K. Gilson, P. M. Goodrum, *et al.*, "Leveraging augmented reality for highway construction," tech. rep., United States. Federal Highway Administration, 2020.

[139] S. Sakhakarmi and J. Park, "Improved intrusion accident management using haptic signals in roadway work zone," *Journal of safety research*, vol. 80, pp. 320–329, 2022.

[140] D. Alfaro-Serrano, T. Balantrapu, R. Chaurey, A. Goicoechea, and E. Verhoogen, "Interventions to promote technology adoption in firms: A systematic review," *Campbell Systematic Reviews*, vol. 17, no. 4, p. e1181, 2021.

[141] C. Nnaji, J. Gambatese, A. Karakhan, and C. Eseonu, "Influential safety technology adoption predictors in construction," *Engineering, Construction and Architectural Management*, vol. 26, no. 11, pp. 2655–2681, 2019.

[142] Z. Mack and S. Sharples, "The importance of usability in product choice: A mobile phone case study," *Ergonomics*, vol. 52, no. 12, pp. 1514–1528, 2009.

[143] J. Nielsen, "Usability inspection methods," in *Conference companion on Human factors in computing systems*, pp. 413–414, 1994.

[144] S. C. Lee, C. Nadri, H. Sanghavi, and M. Jeon, "Eliciting user needs and design requirements for user experience in fully automated vehicles," *International Journal of Human–Computer Interaction*, vol. 38, no. 3, pp. 227–239, 2022.

[145] H. Kivijärvi and K. Pärnänen, "Instrumental usability and effective user experience: Interwoven drivers and outcomes of human-computer interaction," *International Journal of Human–Computer Interaction*, pp. 1–18, 2021.

[146] C. W. Turner, J. R. Lewis, and J. Nielsen, "Determining usability test sample size," *International encyclopedia of ergonomics and human factors*, vol. 3, no. 2, pp. 3084–3088, 2006.

[147] N. Tuli and A. Mantri, "Evaluating usability of mobile-based augmented reality learning environments for early childhood," *International Journal of Human–Computer Interaction*, vol. 37, no. 9, pp. 815–827, 2021.

[148] T. Luger, M. Bär, R. Seibt, M. A. Rieger, and B. Steinhilber, "Using a back exoskeleton during industrial and functional tasksâeffects on muscle activity, posture, performance, usability, and wearer discomfort in a laboratory trial," *Human Factors*, p. 00187208211007267, 2021.

[149] K. Kous, M. Pušnik, M. Heričko, and G. Polančič, "Usability evaluation of a library website with different end user groups," *Journal of Librarianship and Information Science*, vol. 52, no. 1, pp. 75–90, 2020.

[150] A. S. Mueller, J. B. Cicchino, J. Singer, and J. W. Jenness, "Effects of training and display content on level 2 driving automation interface usability," *Transportation research part F: traffic psychology and behaviour*, vol. 69, pp. 61–71, 2020.

[151] A. Dey, M. Billinghurst, R. W. Lindeman, and J. Swan, "A systematic review of 10 years of augmented reality usability studies: 2005 to 2014," *Frontiers in Robotics and AI*, vol. 5, p. 37, 2018.

[152] P. L. Ingrassia, G. Mormando, E. Giudici, F. Strada, F. Carfagna, F. Lamberti, and A. Bottino, "Augmented reality learning environment for basic life support and defibrillation training: usability study," *Journal of medical Internet research*, vol. 22, no. 5, p. e14910, 2020.

[153] S. Kim, M. A. Nussbaum, and J. L. Gabbard, "Influences of augmented reality head-worn display type and user interface design on performance and usability in simulated warehouse order picking," *Applied ergonomics*, vol. 74, pp. 186–193, 2019.

[154] K. A. Mackintosh, S. E. Chappel, J. Salmon, A. Timperio, K. Ball, H. Brown, S. Macfarlane, and N. D. Ridgers, "Parental perspectives of a wearable activity tracker for children younger than 13 years: Acceptability and usability study," *JMIR mHealth and uHealth*, vol. 7, no. 11, p. e13858, 2019.

[155] A. Keogh, R. Argent, A. Anderson, B. Caulfield, and W. Johnston, "Assessing the usability of wearable devices to measure gait and physical activity in chronic conditions: a systematic review," *Journal of neuroengineering and rehabilitation*, vol. 18, no. 1, pp. 1–17, 2021.

[156] L. D. Riek, "Wizard of oz studies in hri: a systematic review and new reporting guidelines," *Journal of Human-Robot Interaction*, vol. 1, no. 1, pp. 119–136, 2012.

[157] S. M. Faas, V. Stange, and M. Baumann, "Self-driving vehicles and pedestrian interaction: Does an external human-machine interface mitigate the threat of a tinted windshield or a distracted driver?," *International Journal of Human–Computer Interaction*, pp. 1–11, 2021.

[158] A. R. Palmeiro, S. van der Kint, L. Vissers, H. Farah, J. C. de Winter, and M. Hagenzieker, "Interaction between pedestrians and automated vehicles: A wizard of oz experiment," *Transportation research part F: traffic psychology and behaviour*, vol. 58, pp. 1005–1020, 2018.

[159] G. Alce, M. Wallergård, and K. Hermodsson, "Wozard: a wizard of oz method for wearable augmented reality interactionâa pilot study," *Advances in human-computer interaction*, vol. 2015, 2015.

[160] S. M. Billah, V. Ashok, and I. Ramakrishnan, "Write-it-yourself with the aid of smartwatches: A wizard-of-oz experiment with blind people," in *23rd International Conference on Intelligent User Interfaces*, pp. 427–431, 2018.

[161] A. Tarhini, K. Hone, X. Liu, and T. Tarhini, "Examining the moderating effect of individual-level cultural values on usersâ acceptance of e-learning in developing countries: a structural equation modeling of an extended technology acceptance model," *Interactive Learning Environments*, vol. 25, no. 3, pp. 306–328, 2017.

[162] J. E. Chung, N. Park, H. Wang, J. Fulk, and M. McLaughlin, "Age differences in perceptions of online community participation among non-users: An extension of the technology acceptance model," *Computers in Human Behavior*, vol. 26, no. 6, pp. 1674–1684, 2010.

[163] S. Sun, P. C. Lee, R. Law, and S. S. Hyun, "An investigation of the moderating effects of current job position level and hotel work experience between technology readiness and technology acceptance," *International Journal of Hospitality Management*, vol. 90, p. 102633, 2020.

[164] C. E. Porter and N. Donthu, "Using the technology acceptance model to explain how attitudes determine internet usage: The role of perceived access barriers and demographics," *Journal of business research*, vol. 59, no. 9, pp. 999–1007, 2006.

[165] S. C. Peres, T. Pham, and R. Phillips, "Validation of the system usability scale (sus) sus in the wild," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, pp. 192–196, SAGE Publications Sage CA: Los Angeles, CA, 2013.

[166] F. Zijlstra and L. Van Doorn, *The construction of a scale to measure perceived effort*. University of Technology, 1985.

[167] F. Zijlstra, “Efficiency in work behaviour: A design approach for modern tools(ph. d. thesis),” 1993.

[168] A. Widyanti, A. Johnson, and D. de Waard, “Adaptation of the rating scale mental effort (rsme) for use in indonesia,” *International Journal of Industrial Ergonomics*, vol. 43, no. 1, pp. 70–76, 2013.

[169] D. A. Peterson and D. Kozhokar, “Peak-end effects for subjective mental workload ratings,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, pp. 2052–2056, SAGE Publications Sage CA: Los Angeles, CA, 2017.

[170] J. R. Lewis, “Measuring user experience with 3, 5, 7, or 11 points: Does it matter?,” *Human factors*, vol. 63, no. 6, pp. 999–1011, 2021.

[171] A. M. Braly, B. Nuernberger, and S. Y. Kim, “Augmented reality improves procedural work on an international space station science instrument,” *Human factors*, vol. 61, no. 6, pp. 866–878, 2019.

[172] A. Liapis, C. Katsanos, N. Karousos, M. Xenos, and T. Orphanoudakis, “User experience evaluation: a validation study of a tool-based approach for automatic stress detection using physiological signals,” *International Journal of Human–Computer Interaction*, vol. 37, no. 5, pp. 470–483, 2021.

[173] J. Sauro and J. R. Lewis, *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2016.

[174] C. Z. Acemyan and P. Kortum, “The relationship between trust and usability in systems,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 56, pp. 1842–1846, SAGE Publications Sage CA: Los Angeles, CA, 2012.

[175] C. Flavián, M. Guinalíu, and R. Gurrea, “The role played by perceived usability, satisfaction and consumer trust on website loyalty,” *Information & management*, vol. 43, no. 1, pp. 1–14, 2006.

[176] D. Salanitri, C. Hare, S. Borsci, G. Lawson, S. Sharples, and B. Waterfield, “Relationship between trust and usability in virtual environments: An ongoing study,” in *International Conference on Human-Computer Interaction*, pp. 49–59, Springer, 2015.

[177] K. Othman, “Public acceptance and perception of autonomous vehicles: a comprehensive review,” *AI and Ethics*, vol. 1, no. 3, pp. 355–387, 2021.

[178] E. Holdack, K. Lurie-Stoyanov, and H. F. Fromme, "The role of perceived enjoyment and perceived informativeness in assessing the acceptance of ar wearables," *Journal of Retailing and Consumer Services*, p. 102259, 2020.

[179] E. Park, "User acceptance of smart wearable devices: An expectation-confirmation model approach," *Telematics and Informatics*, vol. 47, p. 101318, 2020.

[180] S. A. Kamal, M. Shafiq, and P. Kakria, "Investigating acceptance of telemedicine services through an extended technology acceptance model (tam)," *Technology in Society*, vol. 60, p. 101212, 2020.

[181] V. Patel, A. Chesmore, C. M. Legner, and S. Pandey, "Trends in workplace wearable technologies and connected-worker solutions for next-generation occupational safety, health, and productivity," *Advanced Intelligent Systems*, vol. 4, no. 1, p. 2100099, 2022.

[182] B. Munsinger and J. Quarles, "Augmented reality for children in a confirmation task: Time, fatigue, and usability," in *Proceedings of the 25th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–5, 2019.

[183] C. Nnaji, J. Gambatese, and H. W. Lee, "Work zone intrusion: Technology to reduce injuries & fatalities," *Professional Safety*, vol. 63, no. 4, pp. 36–41, 2018.

[184] US Bureau of Labor Statistics, "National Census of Fatal Occupational Injuries in 2021." Publication No. USDL-22-2309, 2020.

[185] Unite the Union, "Workplace fatal injuries in great britain, 2022." `https://resources.unitetheunion.org`. Published 6 July 2022.

[186] Eurostat, "European Statistics on Accidents at Work." `https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Accident_at_work`, 2018.

[187] U.S. Congress, "H.R.3684 - Infrastructure Investment and Jobs Act," 2021.

[188] European Commission - Mobility and Transport, "EU invests in sustainable, safe, and efficient transport infrastructure." `https://transport.ec.europa.eu/news-events/news/eu-invests-eu62-billion-sustainable-safe-and-efficient-transport-infrastructure-2023-06-22_en`, July 2023.

[189] E. National Academies of Sciences, Medicine, *et al.*, "Use of smart work zone technologies for improving work zone safety," 2022.

[190] N. Roofigari-Esfahan, E. White, M. Mollenhauer, and J. P. Talledo Vilela, "Development of a connected smart vest for improved roadside work zone safety," 2021.

[191] X. Yang and N. Roofigari-Esfahan, “Vibrotactile alerting to prevent accidents in highway construction work zones: An exploratory study,” *Sensors*, vol. 23, no. 12, p. 5651, 2023.

[192] T. A. Sitompul and M. Wallmyr, “Using augmented reality to improve productivity and safety for heavy machinery operators: State of the art,” in *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry*, pp. 1–9, 2019.

[193] J. Ramos-Hurtado, F. Muñoz-La Rivera, J. Mora-Serrano, A. Deraemaeker, and I. Valero, “Proposal for the deployment of an augmented reality tool for construction safety inspection,” *Buildings*, vol. 12, no. 4, p. 500, 2022.

[194] Z. Zou, S. D. Bernardes, A. Kurkcu, S. Ergan, K. Ozbay, *et al.*, “An integrated approach to capture construction workersâ response towards safety alarms using wearable sensors and virtual reality,” 2020.

[195] D. Arditi, M. Ayrancioglu, and J. Jingsheng Shi, “Worker safety issues in nighttime highway construction,” *Engineering, Construction and Architectural Management*, vol. 12, no. 5, pp. 487–501, 2005.

[196] J. W. Hinze and J. Teizer, “Visibility-related fatalities related to construction equipment,” *Safety science*, vol. 49, no. 5, pp. 709–718, 2011.

[197] S. Ergan, Z. Zou, S. D. Bernardes, F. Zuo, and K. Ozbay, “Developing an integrated platform to enable hardware-in-the-loop for synchronous vr, traffic simulation and sensor interactions,” *Advanced Engineering Informatics*, vol. 51, p. 101476, 2022.

[198] C. Li, Z. Qing, P. Edara, C. Sun, B. Balakrishnan, and Y. Shang, “Semi-automatic construction of virtual reality environment for highway work zone training using open-source tools,” in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 489–492, IEEE, 2023.

[199] N. Roofigari-Esfahan, C. Porterfield, T. Ogle, T. Upthegrove, M. Jeon, and S. W. Lee, “Group-based vr training to improve hazard recognition, evaluation, and control for highway construction workers,” in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 513–516, IEEE, 2022.

[200] J. Hettig, S. Engelhardt, C. Hansen, and G. Mistelbauer, “Ar in vr: assessing surgical augmented reality visualizations in a steerable virtual reality environment,” *International journal of computer assisted radiology and surgery*, vol. 13, no. 11, pp. 1717–1725, 2018.

[201] R. Horikawa, M. Ito, K. Komiya, T. Nakajima, and H. Yamana, “Comparing augmented reality-based display methods to present guiding information,” in

2022 IEEE 4th Global Conference on Life Sciences and Technologies (LifeTech), pp. 22–25, IEEE, 2022.

[202] F. Zaman, W. Drake, J. Intriligator, A. Gardony, M. Natick, and J. Rife, "Investigating a virtual reality based subterranean scenario examining augmented reality implications for military operators," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 65, pp. 1129–1133, SAGE Publications Sage CA: Los Angeles, CA, 2021.

[203] C. Merenda, C. Suga, J. Gabbard, and T. Misu, "Effects of vehicle simulation visual fidelity on assessing driver performance and behavior," in *2019 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1679–1686, IEEE, 2019.

[204] K. Kim, I. Jeong, and Y. K. Cho, "Signal processing and alert logic evaluation for iot–based work zone proximity safety system," *Journal of Construction Engineering and Management*, vol. 149, no. 2, p. 05022018, 2023.

[205] S. T. Iqbal and B. P. Bailey, "Oasis: A framework for linking notification delivery to the perceptual structure of goal-directed tasks," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 17, no. 4, pp. 1–28, 2010.

[206] D. Gavilan, S. Fernández-Lores, and G. Martinez-Navarro, "Vividness of news push notifications and usersâ response," *Technological Forecasting and Social Change*, vol. 161, p. 120281, 2020.

[207] M. J. Lazaro, S. Kim, J. Lee, J. Chun, and M.-H. Yun, "Interaction modalities for notification signals in augmented reality," in *Proceedings of the 2021 International Conference on Multimodal Interaction*, pp. 470–477, 2021.

[208] M. Pielot and L. Rello, "Productive, anxious, lonely: 24 hours without push notifications," in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pp. 1–11, 2017.

[209] L. G. Morrison, C. Hargood, V. Pejovic, A. W. Geraghty, S. Lloyd, N. Goodman, D. T. Michaelides, A. Weston, M. Musolesi, M. J. Weal, *et al.*, "The effect of timing and frequency of push notifications on usage of a smartphone-based stress management intervention: an exploratory trial," *PloS one*, vol. 12, no. 1, p. e0169162, 2017.

[210] F. Schwarz and W. Fastenmeier, "Augmented reality warnings in vehicles: Effects of modality and specificity on effectiveness," *Accident Analysis & Prevention*, vol. 101, pp. 55–66, 2017.

[211] G. Wiegand, C. Mai, K. Holländer, and H. Hussmann, "Incarar: A design space towards 3d augmented reality applications in vehicles," in *Proceedings of the 11th international conference on automotive user interfaces and interactive vehicular applications*, pp. 1–13, 2019.

[212] P. Zhou, W. Zhang, T. Braud, P. Hui, and J. Kangasharju, "Arve: Augmented reality applications in vehicle to edge networks," in *Proceedings of the 2018 Workshop on Mobile Edge Communications*, pp. 25–30, 2018.

[213] A. Calvi, F. DâAmico, C. Ferrante, and L. B. Ciampoli, "Effectiveness of augmented reality warnings on driving behaviour whilst approaching pedestrian crossings: A driving simulator study," *Accident Analysis & Prevention*, vol. 147, p. 105760, 2020.

[214] A. Matviienko, F. Müller, D. Schön, R. Fayard, S. Abaspur, Y. Li, and M. Mühlhäuser, "E-scootar: Exploring unimodal warnings for e-scooter riders in augmented reality," in *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1–7, 2022.

[215] T. Von Sawitzky, P. Wintersberger, A. Löcken, A.-K. Frison, and A. Riener, "Augmentation concepts with huds for cyclists to improve road safety in shared spaces," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2020.

[216] M. I. Posner, "Timing the brain: Mental chronometry as a tool in neuroscience," *PLoS biology*, vol. 3, no. 2, p. e51, 2005.

[217] M. Green, "" how long does it take to stop?" methodological analysis of driver perception-brake times," *Transportation human factors*, vol. 2, no. 3, pp. 195–216, 2000.

[218] J. Fernandez-Ruiz, W. Wong, I. T. Armstrong, and J. R. Flanagan, "Relation between reaction time and reach errors during visuomotor adaptation," *Behavioural brain research*, vol. 219, no. 1, pp. 8–14, 2011.

[219] S.-C. Li, U. Lindenberger, B. Hommel, G. Aschersleben, W. Prinz, and P. B. Baltes, "Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span," *Psychological science*, vol. 15, no. 3, pp. 155–163, 2004.

[220] M. T. Willoughby, C. B. Blair, L. J. Kuhn, and B. E. Magnus, "The benefits of adding a brief measure of simple reaction time to the assessment of executive function skills in early childhood," *Journal of experimental child psychology*, vol. 170, pp. 30–44, 2018.

[221] J. Gauntlett-Gilbert and V. J. Brown, "Reaction time deficits and parkinson's disease," *Neuroscience & Biobehavioral Reviews*, vol. 22, no. 6, pp. 865–881, 1998.

[222] M. Mueckstein, S. Heinzel, U. Granacher, M. Brahms, M. A. Rapp, and C. Stelzel, "Modality-specific effects of mental fatigue in multitasking," *Acta Psychologica*, vol. 230, p. 103766, 2022.

[223] D. Maslovat, S. T. Klapp, C. J. Forgaard, R. Chua, and I. M. Franks, "The effect of response complexity on simple reaction time occurs even with a highly predictable imperative stimulus," *Neuroscience Letters*, vol. 704, pp. 62–66, 2019.

[224] N. Richer, N. Paquet, and Y. Lajoie, "Impact of age and obstacles on navigation precision and reaction time during blind navigation in dual-task conditions," *Gait & posture*, vol. 39, no. 3, pp. 835–840, 2014.

[225] K. Lu, J. M. Nicholas, S.-N. James, C. A. Lane, T. D. Parker, A. Keshavan, S. E. Keuss, S. M. Buchanan, H. Murray-Smith, D. M. Cash, *et al.*, "Increased variability in reaction time is associated with amyloid beta pathology at age 70," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 12, no. 1, p. e12076, 2020.

[226] G. Der and I. J. Deary, "Age and sex differences in reaction time in adulthood: results from the united kingdom health and lifestyle survey.," *Psychology and aging*, vol. 21, no. 1, p. 62, 2006.

[227] D. L. Woods, J. M. Wyma, E. W. Yund, T. J. Herron, and B. Reed, "Age-related slowing of response selection and production in a visual choice reaction time task," *Frontiers in human neuroscience*, vol. 9, p. 193, 2015.

[228] A. G. Greenwald and K. E. Rosenberg, "Sequential effects of distracting stimuli in a selective attention reaction time task," in *Attention and performance VII*, pp. 487–504, Routledge, 2022.

[229] R. Langner, M. B. Steinborn, A. Chatterjee, W. Sturm, and K. Willmes, "Mental fatigue and temporal preparation in simple reaction-time performance," *Acta psychologica*, vol. 133, no. 1, pp. 64–72, 2010.

[230] R. Ulrich, "Does immediate arousal enhance response force in simple reaction time?," *The Quarterly Journal of Experimental Psychology: Section A*, vol. 49, no. 4, pp. 972–990, 1996.

[231] E. A. Maylor, P. Rabbitt, G. James, and S. Kerr, "Effects of alcohol, practice, and task complexity on reaction time distributions," *The Quarterly Journal of Experimental Psychology Section A*, vol. 44, no. 1, pp. 119–139, 1992.

[232] D. Thapa and S. Mishra, "Using worker's naturalistic response to determine and analyze work zone crashes in the presence of work zone intrusion alert systems," *Accident Analysis & Prevention*, vol. 156, p. 106125, 2021.

[233] E. Bottani and G. Vignali, "Augmented reality technology in the manufacturing industry: A review of the last decade," *IISE Transactions*, vol. 51, no. 3, pp. 284–310, 2019.

[234] A. Kelly, R. B. Shapiro, J. de Halleux, and T. Ball, "Arcadia: A rapid prototyping platform for real-time tangible interfaces," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–8, 2018.

[235] L. Müller, K. Pfeuffer, J. Gugenheimer, B. Pfleging, S. Prange, and F. Alt, "Spatialproto: Exploring real-world motion captures for rapid prototyping of interactive mixed reality," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.

[236] R. Terrier, F. Argelaguet, J.-M. Normand, and M. Marchal, "Evaluation of ar inconsistencies on ar placement tasks: A vr simulation study," in *International Conference on Virtual Reality and Augmented Reality*, pp. 190–210, Springer, 2018.

[237] J. G. Grandi, Z. Cao, M. Ogren, and R. Kopper, "Design and simulation of next-generation augmented reality user interfaces in virtual reality," in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 23–29, IEEE, 2021.

[238] A. Burova, J. Mäkelä, J. Hakulinen, T. Keskinen, H. Heinonen, S. Siltanen, and M. Turunen, "Utilizing vr and gaze tracking to develop ar solutions for industrial maintenance," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.

[239] "Tizen IoT Documentation." `https://docs.tizen.org/iot/api/5.0/tizen-iot-headed/group__CAPI__SYSTEM__FEEDBACK__MODULE.html`. Accessed: [August 1, 2023].

[240] R. Goenarjo, L. Bosquet, N. Berryman, V. Metier, A. Perrochon, S. A. Fraser, and O. Dupuy, "Cerebral oxygenation reserve: The relationship between physical activity level and the cognitive load during a stroop task in healthy young males," *International Journal of Environmental Research and Public Health*, vol. 17, no. 4, p. 1406, 2020.

[241] N. A. Snyder and M. E. Cinelli, "Aperture crossing in virtual reality: Physical fatigue delays response time," *Journal of Motor Behavior*, vol. 54, no. 4, pp. 429–437, 2022.

[242] S. Sabeti, O. Shoghli, and H. Tabkhi, "Toward wi-fi-enabled real-time communication for proactive safety systems in highway work zones: A case study," in *Construction Research Congress 2022*, pp. 1166–1173, 2022.

[243] A. Jain, R. Bansal, A. Kumar, and K. Singh, "A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students," *International Journal of Applied and Basic Medical Research*, vol. 5, no. 2, p. 124, 2015.

[244] "Ml kit pose detection api." `https://developers.google.com/ml-kit/vision/pose-detection`. Accessed on: [Insert Date].

[245] H. S. Legg, C. M. Arnold, C. Trask, and J. L. Lanovaz, "Does functional performance and upper body strength predict upper extremity reaction and movement time in older women?," *Human Movement Science*, vol. 77, p. 102796, 2021.

[246] A. Godfrey, R. Conway, D. Meagher, and G. ÓLaighin, "Direct measurement of human movement by accelerometry," *Medical engineering & physics*, vol. 30, no. 10, pp. 1364–1386, 2008.

[247] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bula, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly," *IEEE Transactions on biomedical Engineering*, vol. 50, no. 6, pp. 711–723, 2003.

[248] M. Z. Uddin, M. M. Hassan, A. Alsanad, and C. Savaglio, "A body sensor data fusion and deep recurrent neural network-based behavior recognition approach for robust healthcare," *Information Fusion*, vol. 55, pp. 105–115, 2020.

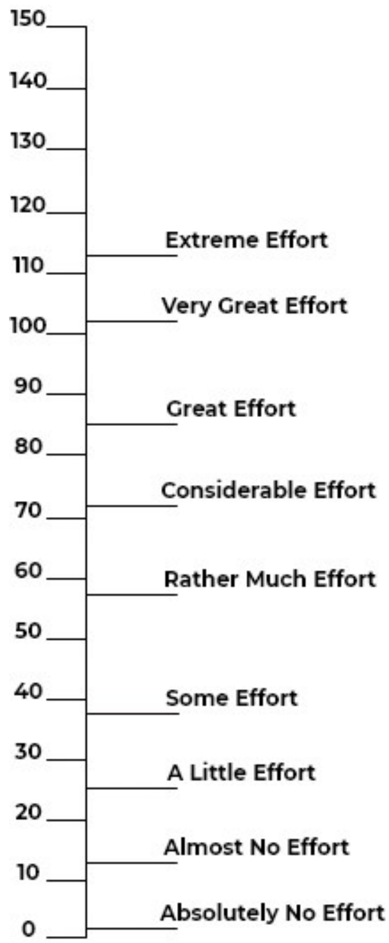
[249] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," *Journal of biomechanics*, vol. 38, no. 3, pp. 401–408, 2005.

[250] H. Sato, N. Tanaka, M. Uchida, Y. Hirabayashi, M. Kanai, T. Ashida, I. Konishi, and A. Maki, "Wavelet analysis for detecting body-movement artifacts in optical topography signals," *Neuroimage*, vol. 33, no. 2, pp. 580–587, 2006.

APPENDIX A: System Usability Scale (SUS)

System Usability Scale Questionnaire	**Strongly Disagree**				**Strongly Agree**
1. I think that I would like to use this product frequently.	1	2	3	4	5
2. I found the product unnecessarily complex.	1	2	3	4	5
3. I thought the product was easy to use.	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this product.	1	2	3	4	5
5. I found the various functions in the product were well integrated.	1	2	3	4	5
6. I thought there was too much inconsistency in this product.	1	2	3	4	5
7. I imagine that most people would learn to use this product very quickly.	1	2	3	4	5
8. I found the product very awkward to use.	1	2	3	4	5
9. I felt very confident using the product.	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this product.	1	2	3	4	5

APPENDIX B: Rating Scale Mental Effort (RSME)



APPENDIX C: Trust Questionnaire

1. This system could improve my safety in highway work zones.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

2. I got familiar with the operation of the system.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

3. I trust this system.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

4. The system is reliable.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

5. The system is dependable.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

6. The system has integrity.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

7. I am comfortable with the intent of the system.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree

8. I am confident that this system would enhance the safety systems already in place in work zones.

   0 — 50 — 100
   Strongly Disagree — Strongly Agree