Advancing Crash Investigation with Connected and Automated Vehicle Data: Insights from a Survey of Law Enforcement

Meredith King
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, 37996
Email: mking63@vols.utk.edu

Sheikh M. Usman
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, 37996
Email: susman1@vols.utk.edu

Muhammad Adeel
Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, 37996
Email: madeel1@vols.utk.edu

Asad J. Khattak, PhD
Beaman Professor and Transportation Program Coordinator
Department of Civil and Environmental Engineering
University of Tennessee, Knoxville, 37996
Email: akhattak@utk.edu

Word Count: 5741 words + 7 tables (250 words per table) = 7491 words

Submitted [7/31/2023]

ABSTRACT

Some of the 6 million plus police-reported crashes in the US require detailed investigations. As connected and automated vehicles (CAVs) gain popularity and are also involved in crashes, new data from CAVs can be used to improve the accuracy of crash investigations. In addition to information from Event Data Recorders, CAV data on vehicle trajectories and pre-crash conditions can enhance the understanding of causation factors. This study investigates how CAV data can advance crash investigations by answering research questions about what pertinent information is lacking in crash investigations and how law enforcement can use CAV data. This study addresses these questions by performing a survey of 61 law enforcement officials working in crash investigations in the state of Tennessee. The survey included respondents' vehicle crash investigation experience, training exposure, and familiarity with automated vehicle technologies. The survey showed that respondents have previous experience working with video camera footage in crash investigations, and responses indicate that camera footage would provide the most helpful new information when investigating a crash. The survey also revealed law enforcement officials' moderate familiarity with lidar and low familiarity with millimeter-wave radar and ultrasound sensors. Survey respondents also indicated that data such as vehicle trajectories made available through automated vehicle sensors would be useful during crash investigations. The survey revealed a need for standardization in CAV data retrieval and training processes. It resulted in a list of pertinent training topics for law enforcement that prioritizes a thorough understanding of CAV technology and data retrieval processes.

Keywords: Automated Vehicle Sensor Data, Crash Investigation, Law Enforcement Training, Survey

1. INTRODUCTION

Understanding the contributing factors in more than 6 million vehicle crashes that occur annually in the US is very challenging. Some crashes require detailed crash investigation. Police officers investigating crashes need a plethora of tools to reconstruct the crash. Given that the Connected and Automated Vehicle (CAV) era is rapidly unfolding, this study seeks to leverage newly available CAV data to improve crash investigation procedures and obtain input from stakeholders, specifically law enforcement.

Currently, law enforcement relies on Event Data Recorders (EDRs), which store vehicle kinematics during a crash. EDRs lack information such as vehicle trajectories, the behavior of surrounding vehicles and pedestrians, the behavior of the driver, and roadway conditions. Information gathered by Automated Driving System (ADS) technologies such as radar, cameras, LiDAR, infrared, and ultrasonic could help fill some data gaps in a crash investigation. This detailed data could improve the fidelity of future crash investigations, with potential new information such as driver/operator state, vehicle automation capabilities, location, objects and people in the immediate area, performance and diagnostic data, and environmental factors. This study addresses two research questions: what pertinent information is lacking in crash investigations, and how can law enforcement be prepared to utilize CAV data in crash investigations? Through a survey with law enforcement officials, this study contributes by further understanding how CAV data can be harnessed to advance a crash investigation. Further, the study explores law enforcement involvement in training for using and applying CAV data and assesses their knowledge of automated vehicle technology data. This research also aims to produce a list of training topics to inform the curation of curriculums for law enforcement training in CAV technology.

2. LITERATURE REVIEW

The current body of literature provides a variety of techniques and technologies that are used in crash reconstruction. To properly investigate a crash, factors such as local conditions, series and sequence of harmful events, contributing circumstances from the roadway, driver, and vehicle actors, vehicle speed, vehicle information and condition, Emergency Medical Service (EMS) information, etc., must be considered (*1*). Conventional accident reconstruction methods rely on event data recorders (EDRs) to recreate and understand the pre-crash conditions that led to the accident (*1*, *2*). EDRs typically store data beginning five seconds before the triggering event occurs (*1*). Triggering events occur when data that is sent from the sensors to the EDRs indicate that an impact exceeds a certain threshold, such as the deployment of the seatbelt pre-tensioner or airbags or significant accelerations (*1*). The National Highway Traffic Safety Administration (NHTSA) (*3*) provides a “top ten” list of data elements stored in EDRs, which includes acceleration and direction of force, crash location, number and location of occupants, seatbelt status, pre-crash data, rollover sensor, yaw rate, time of crash, braking, traction, and stability information, and airbag information.

Detailed vehicle sensor data can be very valuable when recreating a vehicular accident. However, EDRs do not provide audio or visual data, which can greatly enhance a crash investigation. Multiple studies cite the usefulness of dashboard cameras during accident reconstruction (*4*), (*5*), (*6*). A study showed that dashboard camera sound was useful when calculating vehicle speed, especially when the vehicle’s torque converter slip is not severe or when checking for engine RPM changes (intentional accidents) (*4*). Another study showed that dashboard camera footage was pivotal in accurately reconstructing a pedestrian and tractor-trailer crash (*5*). Before analyzing the camera footage, the crash appeared to be accidental. However, the footage showed that the pedestrian suddenly darted in front of the moving vehicle, revealing that the collision resulted from suicide (*5*).

Although expensive, LiDAR sensors can help reconstruct crash events. LiDAR is a range-finding environmental sensor commonly used to model terrain, and it can also be used for adaptive cruise control, collision avoidance, and object recognition (*1*). LiDAR can be used in collaboration with UAV photogrammetry and FARO simulation software to recreate crash events (*7*). However, it is noted that UAV photogrammetry is more accurate than LiDAR (*7*). LiDAR is used in automated vehicles for obstacle detection (*1*). Obstacles 1.36-m wide can be reliably detected at distances less than 10.82 meters using LiDARs. Reliable classification of obstacles, however, is only achieved at distances of less than 5.42 meters (*8*). There is a gap in the research regarding how CAV LiDAR data can be harnessed after crash events for accident reconstruction.

In addition to LiDAR, radar holds value for crash reconstruction. Several studies in the literature cite radar for use by automated vehicles in pedestrian/obstacle detection (*9*), (*1*), (*10*). Radar is highly effective at detecting pedestrians even in a complex background (*1*,*9*). Radar can accurately track moving and fixed objects, which can be used during crash reconstruction (*10*). There is a need in the literature, however, to evaluate the effectiveness of radar in crash reconstruction using moving hosts and targets (*10*).

A data source unique to connected vehicles is Basic Safety Messages (BSMs), which are exchanged between connected vehicles using onboard units (OBUs). Several studies have discussed how these messages can provide pertinent vehicle information. BSM data set was used to capture variations in vehicle control (*11*). By capturing these variations and finding meaningful relationships between control variations and crash data, the usefulness of BSM data in crash investigations becomes worth further investigation.

The literature generally reveals that automated vehicles provide many potential sources for post-crash data, including LiDAR, radar, cameras, and OBUs. While EDRs and crash simulation software are frequently used by crash reconstructionists, CAV data is not yet being harnessed in industry practices to reconstruct accidents. There is a significant gap in the research for testing and using this data to reconstruct real or simulated crashes. While researchers are still studying CAV penetration rates and deployment predictions, automobile manufacturers such as Tesla, Waymo, Ford, and General Motor Company, among others, are producing vehicles with increasing automatic capabilities. Moreover, most new cars on the market include Advanced Driver Assistance Systems (ADAS) such as adaptive cruise control, hands-on-lane-centering steering, and hands-free steering. Thus, as vehicular networks become increasingly automated, crash investigators and law enforcement can benefit by updating crash reconstruction processes to include CAV sensor data.

Another gap in the research is understanding how law enforcement and crash investigators should be trained to handle CAV data and recreate AV-involved crashes. This study addresses this gap by delivering a questionnaire geared toward law enforcement to develop a list of pertinent CAV training topics.

3. METHODS

The survey is conducted through the Qualtrics online survey platform. The study population includes officials from city police departments, sheriff's offices, and Tennessee Highway Patrol. The respondents are Tennessee officials who specifically work in vehicle crash investigations. This includes 61 officials from a database of 326 local police departments and 95 Sheriff's offices. All respondents are over the age of 21.

The respondents were contacted through a third party, with assistance from the Tennessee Highway Safety Patrol Office. A survey link was emailed a survey link to the Tennessee law enforcement officials involved in crash investigations. The survey responses were delivered directly to the researchers. All responses are anonymous, and the researchers collected no personally identifying information. The survey begins with a statement of informed consent, and if the respondent chooses to consent to fill out the survey, they are then allowed to proceed to a brief background section. After reading the background, the respondents answered 26 questions: 11 multiple-choice questions, 9 short answer questions, and 6 Likert scale matrix questions. Data cleaning procedures enhanced data reliability. This was done through outlier detection, response validation (e.g., respondents' age is above 21), analysis of skipping patterns, and realism of responses. Exploratory factor analysis was conducted to reduce the dimensionality of the data by obtaining factors that reasonably explain the data's variation.

4. RESULTS

The research team performed data analysis for the survey using Qualtrics. The results are reported in several sections below. Descriptive statistics are reported for the survey, supplemented by graphical representations and text mining.

4.1 Survey Respondents – Work Context

A set of questions in the survey were used to understand the organizations and work experiences of the respondents. Descriptive statistics of responses are reported in Tables 1-2, with Table 1 reporting response statistics for multiple-select questions and Table 2 reporting that for multiple-choice questions. Multiple-select is distinct from multiple-choice as multiple-select allows respondents to select multiple responses, whereas multiple-choice requires the respondent to select only one response. Therefore, response percentages for each question in Table 1 sum to a value greater than 100%, whereas percentages for each question in Table 2 are exactly 100%.

The introductory questions relate to the respondents' roles, organizational structures, and work experiences. The organizations contained an average of 435.5 sworn officers, with a range of 1460 spanning from 5 to 1465 officers. Each respondent has worked on an average of 160 fatal and/or prosecutable crashes, ranging from 0 to 1700 individual crashes. Another question asks respondents to select multiple occupational roles that apply to themselves (Crash Reconstructionist, Traffic Division, Patrol, and Other); Table 1 shows that 91.1% selected crash reconstructionist, 39.3% selected traffic division, and 14.3% selected patrol. Another 10.7% of respondents chose the "other" category, where they generally inputted into the text box that they were a commander or supervisor.

Furthermore, Table 2 shows that 78.3% of the respondents' organizations have a separate division charged with investigating crashes. In response to the question, "Does your organization have a separate division charged with crashes?", 8.9% of respondents chose the "other" option, and they provided supplemental information in the provided text box. These "other" responses generally indicated that organizations have specialized training or a specialized unit for more serious crashes, but all patrol officers investigate basic crashes. Overall, the respondents represent law enforcement officers, and most investigate crashes. However, the respondents' roles, work experiences, and organizational structures vary widely.

The next set of questions relates to the police officers' handling of crash data. As provided in Table 2, a survey question gauges the level of automated vehicle training the officers have available. Of the 43 responses to this question, only 16.3% indicated they had the opportunity for Automated Vehicle (AV) sensor training for crash reconstruction. Nearly one-half of respondents (48.8%) indicated that while they do not currently have a plan for CAV training, they expect to have a plan to implement this training in the future. Some respondents (32.2%) indicated no plans to implement this training. Also provided in Table 2, the survey assesses whether officers have access to the processing or managing of crash data. Most respondents (75.5%) answered "yes," with another 20.4% stating that they had a plan for or expected to have a plan for managing crash data.

The next set of questions discusses the use of EDRs. As shown in Table 2, in response to whether officers have used EDRs, 90.2% responded "yes." Next, officers were asked what information they received from EDRs, and responses included vehicle speed (97.9%), brake status (89.4%), seatbelt usage (87.2%), throttle position (76.6%), engine RPM (76.6%), steering input (68.1%), and "other" (23.4%). The "other" responses included change in velocity ("delta-V"), friction, and vehicle roll angle ("roll over"). Next, the survey asks respondents what level(s) of EDR training they have completed. Only 47.5% of the total survey respondents answered the question. The majority of these respondents (58.6%) had completed EDR technician training, 44.8% completed EDR basic, and 34.5% completed EDR advanced. Another 24.1% of respondents selected "other," which generally indicated that the respondent had completed no EDR training.

A set of questions allows the officers to assess the future of automated vehicle data in crash investigations. The officers were asked which information is not usually available today that they would most like to receive from a vehicle following a collision. The majority of officers (54.9%) answered "vehicle and occupant dynamics." Another 31.4% answered "vehicle systems and performance." The officers were asked which CAV data source(s) would provide the most helpful information in crash investigation, and an overwhelming number of officers, 94.1% selected cameras. A majority of officers also chose GPS (58.8%) and LiDAR (51.0%).

Officers were asked about the perceived barriers to using CAV sensor data. Most of the officers selected "data accessibility and availability" (60.8%) and "budget" (56.9%). Another 9.8% of officers chose "other," which included responses such as "ability to translate complex data to a jury," "lack of training," "getting a search warrant for data," and "ability to validate data in crashes with limited crash evidence." Lastly, the officers were queried about how CAV sensor data can enhance the crash investigation. The majority selected each of the options: "improved data accuracy" (88.2%), "increased data availability" (82.4%), "improved understanding of human factors" (74.5%), "enhanced vehicle and occupant safety" (56.9%), and "improved understanding of environmental factors" (49.0%).

Overall, the respondents represent well-experienced law enforcement officials who have had experience and training in investigating roadway crashes. Work experiences and roles are diverse, with some respondents working as supervisors or commanders with experiences in thousands of crashes, while others have only worked on a few. Similarly, a wide range of technology use and training levels exist. This diversity is important to consider when creating an automated vehicle training curriculum for law enforcement.

Table 1: Survey Responses – Work Context (Multiple-Select)

Selection	Checked Percent	Confidence Interval	Checked Count	Sample Size (N)
What is your role in collision investigation? (Select all that apply.)				
Crash Reconstructionist	91.1%	80.7% to 96.1%	51	56
Traffic Division	39.3%	27.6% to 52.4%	22	56
Patrol	14.3%	7.4% to 25.7%	8	56
Other	10.7%	5.0% to 21.5%	6	56
Total	155.4%			
What information have you typically received from the EDR automatically after a collision? (Select all that apply)				
Vehicle speed	97.9%	88.9% to 99.6%	46	47
Brake status	89.4%	77.4% to 95.4%	42	47
Seatbelt usage	87.2%	74.8% to 94.0%	41	47
Throttle position	76.6%	62.8% to 86.4%	36	47
Engine RPM	76.6%	62.8% to 86.4%	36	47
Steering input	68.1%	53.8% to 79.6%	32	47
Other	23.4%	13.6% to 37.2%	11	47
Total	749.5%			

Have you completed any training for EDR data retrieval? If so, please specify which course(s) you have completed.				
EDR Technician	58.6%	40.7% to 74.5%	17	29
EDR Basic	44.8%	28.4% to 62.5%	13	29
EDR Advanced	34.5%	19.9% to 52.7%	10	29
Other	24.1%	12.2% to 42.1%	7	29
Total	162.0%			
Of the available data sources in automated vehicles mentioned, which would provide the most helpful information that is not currently available? (Select all that apply.)				
Cameras	94.1%	84.1% to 98.0%	48	51
Global Positioning System (GPS)	58.8%	45.2% to 71.2%	30	51
LiDAR from vehicles	51.0%	37.7% to 64.1%	26	51
Onboard Units (OBU)	47.1%	34.1% to 60.5%	24	51
Millimeter Wave Radar (MMWR)	27.5%	17.1% to 40.9%	14	51
Infrared	21.6%	12.5% to 34.6%	11	51
Ultrasound	17.6%	9.6% to 30.3%	9	51
Other	3.9%	1.1% to 13.2%	2	51
Total	321.6%			
What are some significant barriers based on your work experience for using automated vehicle sensor data in crash reconstruction? (Select all that apply.)				
Data availability and accessibility	60.8%	47.1% to 73.0%	31	51
Budget	56.9%	43.3% to 69.5%	29	51
Data format and standardization	45.1%	32.3% to 58.6%	23	51
Data analysis	37.3%	25.3% to 51.0%	19	51
Technical complexity	35.3%	23.6% to 49.0%	18	51
Liability and privacy concerns	27.5%	17.1% to 40.9%	14	51
Time	21.6%	12.5% to 34.6%	11	51
Other	9.8%	4.3% to 21.0%	5	51
Total	294.12%			
Based on your work experience, how can automated vehicle sensor data enhance crash investigation? (Select all that apply.)				
Improved data accuracy	88.2%	76.6% to 94.5%	45	51
Increased data availability	82.4%	69.7% to 90.4%	42	51

Improved understanding of human factors	74.5%	61.1% to 84.5%	38	51
Enhanced vehicle and occupant safety	56.9%	43.3% to 69.5%	29	51
Improved understanding of environmental factors	49.0%	35.9% to 62.3%	25	51
Other	2.0%	0.3% to 10.3%	1	51
Total	352.9%			

Table 2: Survey Responses – Work Context (Multiple-Choice)

Choice	**Count**	**Percent of Data**	**Confidence Interval (Percent of Data)**	**Sample Size (N)**
Does your organization have a separate division charged with investigating crashes?				
Yes	36	78.30%	64.4% to 87.7%	56
No	7	12.5%	6.2% to 23.6%	56
Other	5	8.9%	3.9% to 19.3%	56
Total	56	100%		
Has your organization provided the opportunity for training on the use of automated vehicle sensor data for crash reconstruction purposes?				
Yes	7	16.3%	8.1% to 30.0%	43
No, but we have a specific plan to implement this training	1	2.3%	0.4% to 12.1%	43
No, but we expect to have a specific plan to implement this training in the future	21	48.8%	34.6% to 63.2%	43
No, and we do not plan to implement this training	14	32.6%	20.5% to 47.5%	43
Total	43	100%		
Does your organization have access to the processing or managing of crash data from vehicles involved in a collision?				
Yes	37	75.5%	61.9% to 85.4%	49
No, but we have a specific plan for this	3	6.1%	2.1% to 16.5%	49
No, but we expect to have a specific plan for this in the future	7	14.3%	7.1% to 26.7%	49
No, and we don’t plan for this	2	4.1%	1.1% to 13.7%	49
Total	49	100%		
Have you ever used Event Data Recorders (EDRs) for collision investigation?				
Yes	46	90.2%	79.0% to 95.7%	51
No	5	9.8%	4.3% to 21.0%	51
Total	51	100%		
Thinking of the future of collision investigation, what information (not usually available today) would you most like to get from a vehicle automatically after a collision?				

Vehicle and occupant dynamics	28	54.9%	41.4% to 67.7%	51
Environmental data	5	9.8%	4.3% to 21.0%	51
Vehicle systems and performance	16	31.4%	20.3% to 45.0%	51
Other	2	3.9%	1.1% to 13.2%	51
Total	51	100%		

4.2 Crash Investigation Practices by Officers

The next set of questions relates to crash investigations and asks officers to evaluate a series of statements on a 5-point Likert scale. Respondents were asked about how six different aspects of collision investigation are fulfilled in the current process (Table 3). Most respondents chose "adequate" for all six of the statements. However, "training and certification for collision investigation," "Standardization of how collisions are investigated," and "accuracy and reliability of collision investigation" each received 12% "inadequate" responses. "Accuracy and reliability of collision investigations" was the most positively rated aspect, with a majority (56%) of ratings as "Adequate" and another 22% of ratings as "Excellent." Both "Efficiency and speed of collision investigations" and "Data availability during collision investigations" received varied responses, with most respondents ranking these aspects as "somewhat adequate" or "adequate."

The survey also asks officers to rank their familiarity with seven automated vehicle technologies, as shown in Table 4. Most officers responded that they were not at all familiar with Millimeter Wave Radar (MMWR) (64.7%), ultrasound (60.8%), and infrared sensors (54.0%). Another large percentage (41.2%) indicated that they were not at all familiar with onboard units (OBUs). However, officers indicated they were moderately familiar (51.0%) or extremely familiar (5.9%) with cameras in vehicles. The officers were also queried about their familiarity with ten different advanced driver assistance systems (ADAS) technologies. Table 4 shows that a high proportion of officers indicated they were not familiar with Rear Control Traffic Alert (45.1%) and Night Vision (40.0%). Officers were most familiar with Blind Spot Monitoring, with 11.8% ranking the technology as "Extremely Familiar." Many officers were also familiar with Adaptive Cruise Control, Lane Departure Warning, and Forward Collision Warning, with 33.3% of officers ranking each of these technologies as "Moderately Familiar." Lastly, the survey asks officers to rate their familiarity with different CAV training topics. Table 4 shows the topics that were most rated as "not at all familiar" among respondents are "understanding automated vehicle technology" (66.7%), "cybersecurity" (66.7%), and "communication and community engagement" (51.0%). None of the topics received a significant percentage of "Extremely Familiar" ratings. However, officers indicated moderate familiarity with "legal and ethical considerations" (14.0%) and "incidence response and crash investigation" (11.8%).

Table 3: Respondent Rankings of the Adequacy of Current Collision Investigation Practices

		Very inadequate	Inadequate	Somewhat adequate	Adequate	Excellent	Total
Aspects of Collision Investigation	**Accuracy and reliability of collision investigations (N = 50)**	2.0%	0.0%	20.0%	56.0%	22.0%	100%
	Improvement of safety and mitigation of future collision investigations (N = 50)	0.0%	8.0%	46.0%	40.0%	6.0%	100%
	Efficiency and speed of collision investigations (N = 50)	0.0%	8.0%	34.0%	46.0%	12.0%	100%
	Data availability during collision investigations (N = 50)	0.00%	12.0%	38.0%	44.0%	6.0%	100%
	Standardization of how collisions are investigated (N = 50)	2.0%	12.0%	30.0%	48.0%	8.0%	100%
	Training and certification for collision investigation (N = 50)	2.0%	10.0%	28.0%	46.0%	14.0%	100%

Table 4: Respondent Rankings of Familiarity with Technologies and Training Topics

		Not at all familiar	Slightly familiar	Somewhat familiar	Moderately familiar	Extremely familiar	Total
Automated Vehicle Sensors	**Global Positioning System (GPS) (N = 51)**	3.9%	23.5%	15.7%	51.0%	5.9%	100%
	Onboard Units (OBU) (N = 51)	41.2%	33.3%	13.7%	51.0%	0.0%	100%
	Millimeter Wave Radar (N = 51	64.7%	25.5%	9.8%	51.0%	0.0%	100%
	Ultrasound Sensors (N = 51)	60.8%	25.5%	7.8%	51.0%	0.0%	100%
	Infrared Sensors (N = 50)	54.0%	28.0%	10.0%	51.0%	0.0%	100%
	LiDAR (N = 51)	33.3%	27.5%	15.7%	51.0%	2.0%	100%
	Cameras (N = 51)	3.9%	15.7%	15.7%	51.0%	5.9%	100%
Advanced Driver Assistance Systems (ADAS)	**Adaptive Cruise Control (ACC) (N = 51)**	13.7%	17.6%	25.5%	33.3%	9.8%	100%
	Lane Departure Warning (LDW) (N = 51)	13.7%	15.7%	27.5%	33.3%	9.8%	100%
	Blind Spot Monitoring (BSM) (N = 51)	17.6%	13.7%	19.6%	37.3%	11.8%	100%
	Rear Cross Traffic Alert (RCTA) (N = 51	45.1%	11.8%	19.6%	19.6%	3.9%	100%
	Forward Collision Warning (FCW) (N = 51)	11.8%	23.5%	23.5%	33.3%	7.8%	100%
	Automatic Emergency Braking (AEB) (N = 51)	19.6%	21.6%	19.6%	29.4%	9.8%	100%
	Park Assist (N = 51)	15.7%	39.2%	11.8%	25.5%	7.8%	100%

	Night Vision (N = 50)	40.0%	30.0%	14.0%	14.0%	2.0%	100%
	Head-Up Display (N = 50)	22.0%	30.0%	16.0%	24.0%	8.0%	100%
	Driver Monitoring Systems (DMS) (N = 51)	23.5%	33.3%	23.5%	17.6%	2.0%	100%
Law Enforcement Training Topics	**Understanding automated vehicle technology (N = 51)**	66.7%	23.5%	7.8%	2.0%	0.0%	100%
	Legal and ethical considerations (N = 51)	36.0%	32.0%	16.0%	14.0%	2.0%	100%
	Traffic enforcement and regulation (N = 51)	45.1%	27.5%	21.6%	5.9%	0.0%	100%
	Incident response and crash investigation (N = 51)	41.2%	35.3%	11.8%	11.8%	0.0%	100%
	Cybersecurity (N = 51)	66.7%	23.5%	9.8%	0.0%	0.0%	100%
	Human factors (N = 51)	39.2%	43.1%	11.8%	5.9%	0.0%	100%
	Communication and community engagement (N = 51)	51.0%	29.4%	15.7%	3.9%	0.0%	100%

4.3 Qualitative Information Provided by Respondents

Throughout the questionnaire, short answer questions allow the respondents to provide their thoughts via text entry. Because of the varying responses, text analytic methods are preferable to traditional descriptive statistics. The first answer is to a question asking respondents if they have ever used vehicle camera footage in a crash investigation, and if so, how this footage affected their investigation. The comments were classified into the following topics:

1. *Have used cameras:* Comments indicating the respondent has used video camera footage for crash investigation.
2. *Never used cameras:* Comments indicating that the respondent has never used video camera footage.
3. *Confirming evidence:* Comments indicating that video camera footage confirmed evidence collected by other means during the crash investigation.
4. *Conflicting evidence:* Comments indicating that video camera footage refuted evidence collected by other means, such as witness statements.
5. *New data:* Comments indicating that video camera footage brought new data or evidence to the investigation.
6. *Prosecution/fault:* Comments indicating that video camera footage was used to help prosecute or determine fault during a crash.
7. *Surveillance footage:* Comments indicating that video camera footage in the form of surveillance cameras has been used in a crash investigation.
8. *Aftermarket cameras:* Comments indicating that video camera footage in the form of aftermarket dashboard cameras has been used in a crash investigation.

The descriptive statistics of responses to these questions are provided in Table 5, and a Word cloud of responses is shown in Figure 1. More comments indicated that they had used cameras (67.39%) than those that had not (26.09%), and almost a quarter (23.91%) of the comments indicated that the cameras brought new data that was not available through other means during the crash investigations. The Word cloud shows that words such as “determine,” “fault,” “speed,” “evidence,” and “help” are common among the comments, indicating that video cameras have provided valuable evidence to many officers during crash investigations.

The survey asked officers if they used in-vehicle LiDAR and radar in a crash investigation. The question regarding LiDAR received 49 responses, and the question regarding radar received 48 replies. All comments discussing LiDAR indicated that the officers have not used in-vehicle LiDAR for a crash investigation. Nearly all responses discussing radar indicated that in-vehicle radar had not been used for a crash investigation. However, one comment stated that in-vehicle radar was used to pull vehicle information to confirm investigation information. Another question asks officers to list which tools they typically use in a crash investigation. Responses were varied, and comments were classified into topics as follows:

1. *EDR:* Comments including electronic data recorders (EDRs).
2. *Drone Photography:* Comments including drones or drone footage.
3. *Total Station:* Comments including total stations. Makes of total stations mentioned include Leica, Carlson, Nikon, and Faro.
4. *Infotainment Data:* Comments including infotainment data or mention of an infotainment data tool, such as the Berla system tool.
5. *Digital Camera:* Comments mentioning the use of a digital camera.
6. *EDR Retrieval Equipment:* Comments mentioning the use of EDR retrieval equipment, such as the Bosch Crash Data Retrieval Tool.
7. *3D Laser Scanner:* Comments mentioning the use of a 3D laser scanner. The Faro 360 scanner is commonly referenced.
8. *GPS/GNSS:* Mentions of either GPS or GNSS rovers. Leica is a common make for rovers referenced.
9. *Modeling Software:* Comments include modeling software such as Crashzone, Pix4D, Cyclone, and IMS Map360.
10. *Traffic Camera:* Comments mentioning the use of traffic cameras such as Redflex cameras in crash investigations.
11. *Crash Simulation Software:* Comments that mention crash simulation software such as Virtual Crash (VCrash).
12. *Dash Camera:* Comments that mention the use of dashboard cameras.
13. *Motion Performance Instruments:* Comments that mention using motion performance instruments such as Vericom tools and friction testing devices.
14. *Crash Database:* This includes comments mentioning crash databases such as TITAN.
15. *Outsourcing:* Comments that indicate that an organization relies on outside sources, such as the Tennessee Highway Patrol, for crash analysis due to a lack of equipment or other resources.

The most prevalent answers were Total Stations (76.09%), EDR (67.39%), and Drones (69.57%). Figure 2 presents a Word cloud of these answers, which revealed that "rover," "scanner," and "camera" emerge as the commonly mentioned tools.

Table 5: Qualitative Responses by Respondents

Question	Topic	Count	Percentage
Have you ever used vehicle camera footage during a crash investigation? If so, how did this footage impact the process and outcome of the investigation? (N =49)	**Never used cameras**	12	26.09%
	Have used cameras	31	67.39%
	Confirming evidence	8	17.39%
	Conflicting evidence	1	2.17%
	New data	11	23.91%
	Prosecution/fault	10	21.74%
	Surveillance footage	2	4.35%
	Aftermarket cameras	4	8.70%
	Total		171.74%

What software or other tools do you typically use during a crash investigation (e.g., Analysis and Simulation Software, Total Stations, Drone Cameras, Event Data Recorders)? (N = 46)	**EDR**	31	67.39%
	Drone Photography	32	69.57%
	Total Station	35	76.09%
	Infotainment Data	9	19.57%
	Digital Camera	3	6.52%
	EDR Retrieval Equipment	5	10.87%
	3D Laser Scanner	15	32.61%
	GPS/GNSS	13	28.26%
	Modeling Software	8	17.39%
	Traffic Camera	1	2.17%
	Crash Simulation Software	8	17.39%
	Dash Camera	1	2.17%
	Motion Performance Instruments	2	4.35%
	Crash Database	1	2.17%
	Outsourcing	2	4.35%
	Total		360.87%

Figure 1: Word cloud for Vehicle Camera Footage **Figure 2: Word cloud for Equipment Used**

The survey asks officers whether they have completed automated vehicle data retrieval training. Of the 42 comments, 37 answered "No," 3 mentioned EDR training, 1 mentioned infotainment training, and 1 said a Berla system training. With these responses, it can be reasonably assumed that no officers have received any AV-specific training. Notably, this question was directed at determining if any officers have received CAV training.

Officers were asked open-ended questions about providing their thoughts on what could be improved in collision investigations. The comments generally emphasized the need for more standardized, state-funded crash investigation training and certification. The comments regarding the current process of collision investigation emphasize the need for standardized EDRs across vehicle manufacturers, comprehensive data captured from smart vehicles, training in new technologies, and increased availability of crash investigation classes. The final thoughts from the respondents highlight the need for improved understanding and access to evidence in automated vehicle crash investigations, emphasizing the importance of additional training and updated technology. State assistance is requested due to budget constraints, with a focus on advanced training and diverse crash instruction classes. Word clouds, shown in

Figures 3 and 4, were created to get useful insights into the provided comments. Both the word clouds include the word "training" as a central theme of the responses.

Figure 3: Word cloud for Current Processes **Figure 4: Word cloud for Final Thoughts**

4.4 Factor Analysis

Factor Analysis can reduce large dimensional data into a set of distinct factors that explain the most variation in the data and reveal the underlying relationship between key variables. Fifteen variables were extracted from the survey responses for an exploratory factor analysis. The description of these variables is provided below:

1. *Number of sworn officers in the investigation agency (Continuous)*
2. *Presence of a separate AV Crash Investigation Division (Binary outcomes 1,0)*
3. *Training in Crash Reconstruction (Binary outcomes 1,0)*
4. *Access to AV crash data (Binary outcomes 1,0)*
5. *Use of Vehicle Camera footage in AV crash investigation (Binary outcomes 1,0)*
6. *Extracting data from In-Vehicle radar(Binary outcomes 1,0)*
7. *Use of Electronic Data Recorder (EDR) in AV crash investigation (Binary outcomes 1,0)*
8. *Training provided to Officers to use EDRs (Binary outcomes 1,0)*
9. *Training for AV Data Retrieval (Binary outcomes 1,0)*
10. *Adequacy of Current Crash Investigation Methods in Accurate & Reliable AV crash investigation (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*
11. *Adequacy of Current Crash Investigation Methods in Improving AV Safety (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*
12. *Adequacy of Current Crash Investigation Methods in Improving Efficiency of AV crash investigation (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*
13. *Adequacy of Available AV Crash Data for crash investigation (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*
14. *Adequacy of Current Crash Investigation Methods in Standardizing AV Crash Investigation (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*
15. *Adequacy of Training and Certifications in AV Crash Reconstruction and Investigation (1: Very Inadequate, 2: Inadequate, 3: Somewhat Adequate, 4: Adequate, 5: Excellent)*

Missing values in each variable were replaced with the mean of the available values. The final dataset consisted of 15 variables, with 61 entries in each variable. The eligibility of the dataset for factor analysis was determined by conducting Bartlett's test of Sphericity and Keiser-Meyer-Olkin factor adequacy test. Bartlett's test of sphericity tests whether the correlation matrix of the study variables is significantly different from the identity matrix. The correlation matrix of study variables is assumed to be similar to the identity matrix as a null hypothesis. The test result rejects the null hypothesis ($p < 2.22$ e-16). Furthermore, the overall Measure of Sampling Adequacy (MSA) in the Keiser-Meyer-Olkin test for the dataset is 0.72, which indicates the suitability of data for conducting a factor analysis since MSA values higher than 0.70 are usually considered meritorious for applying factor analysis. After successful preliminary tests, factor analysis was conducted for the dataset using "psych" and "REdaS" packages in "R" software.

The analysis provided Factor 1-Digital forensics integration for AV crash reconstruction, Factor 2- AV crash data management and investigation standards, Factor 3- AV crash investigation efficacy and safety assurance, and Factor 4- Advanced training in AV crash investigation. The number of distinct factors was selected according to Keiser's rule, which states that only the factors with Eigenvalues greater than 1 can be considered distinct factors. Results of the factor analysis indicate that only four out of the fifteen factors had eigenvalues greater than 1. The remaining factors were dropped from the analysis, and the analysis was re-run for only four factors. Results from the analysis (shown in Table 6) indicate that Factor 1 explains 35% of the variance in the data, followed by Factor 3 with 23% explained variance, while Factors 2 and 4 explain 20% and 22% variance in the data, respectively. The variable loadings in Table 6 indicate that the variables "Access to AV crash data," "Use of Vehicle Camera Footage in AV Crash Investigation," and "Use of EDR in crash investigation" have higher loadings (more than 0.5) in Factor 1 (above). Similarly, the loadings of other variables in other factors can also be assessed. The factors were named to represent the highly loaded variables coherently.

Finally, tests to assess the sufficiency of these four factors are conducted. The value of the root mean square of residuals (RMSR) for the analysis is 0.05. Theoretically, this value should be close to zero to establish the sufficiency of the factors. Hence, this result is satisfactory. Additionally, the Tucker-Lewis Index (TLI) of factoring reliability was estimated as 0.945. A TLI value higher than 0.90 is usually considered satisfactory for establishing factor sufficiency. Hence, this check was also satisfied. Figure 5 presents the association of variables with their respective factors through loading values.

Table 6: Results of Factor Analysis

Statistics	Factor 1	Factor 2	Factor 3	Factor 4
Eigenvalues	2.31	1.30	1.56	1.49
Proportion of Explained Variance	0.35	0.20	0.23	0.22
Cumulative Proportion of Explained Variance	0.35	0.55	0.78	1.00
Standardized Variable Loadings				
Number of sworn officers in the investigation agency	0.35	-0.08	0.03	0.10
Presence of a separate AV Crash Investigation Division	0.38	0.22	0.25	-0.02
Training in Crash Reconstruction	-0.02	0.01	0.01	1.02
Access to AV crash data	0.71	0.03	0.05	0.10
Use of Vehicle Camera footage in AV crash investigation	0.81	-0.20	0.02	-0.03
Extracting data from In-Vehicle radar	-0.09	0.01	-0.10	-0.03
Use of Electronic Data Recorder (EDR) in AV crash investigation	0.66	0.17	0.03	0.05
Training provided to Officers to use EDRs	0.59	0.18	-0.16	-0.06
Training for AV Data Retrieval	0.15	-0.12	-0.07	0.56
Adequacy of Current Crash Investigation Methods in Accurate & Reliable AV crash investigation	0.10	-0.06	0.70	0.06
Adequacy of Current Crash Investigation Methods in Improving AV Safety	-0.06	-0.15	0.64	-0.10
Adequacy of Current Crash Investigation Methods in Improving Efficiency of AV crash investigation	0.01	0.26	0.31	0.04
AV Crash Data Availability during crash investigation	-0.01	0.82	-0.02	0.02
Standardization of AV Crash Investigation	0.02	0.54	0.06	-0.21
Training and Certifications in AV Crash Reconstruction and Investigation	-0.04	0.19	0.65	0.01

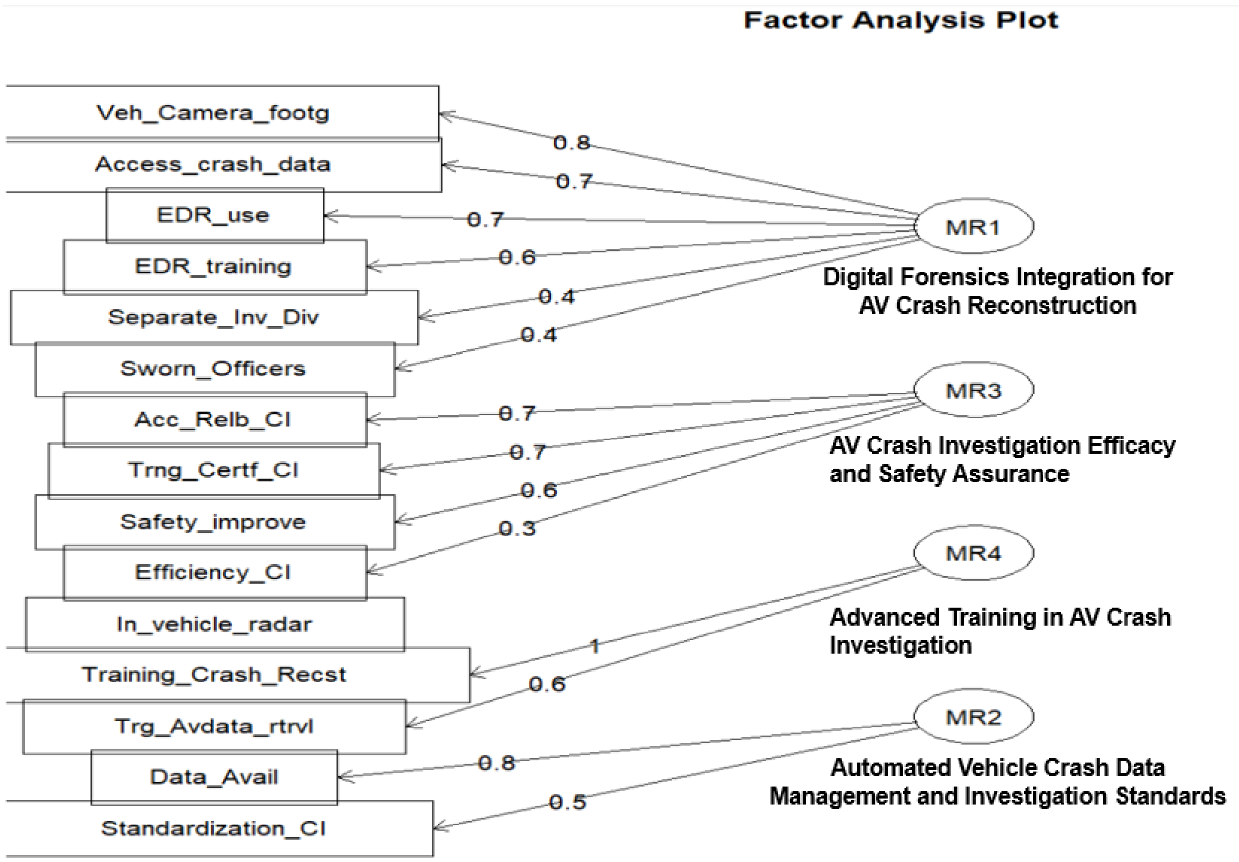

Figure 5: Factor Analysis Plot

5. Discussion

This survey aims to answer two research questions: 1) What pertinent information or processes are lacking in crash investigations? 2) How can law enforcement be prepared to utilize CAV data in crash investigations? The survey highlights several gaps in the current process of collision investigation. Respondents indicated a need for more standardization in EDR training and connections for downloading data. Further, respondents indicated that vehicle and occupant dynamic information, which is not usually available today, would be most valuable after a collision. Vehicle and occupant dynamics include information such as vehicle trajectory or driver behavior, and cameras or other automated vehicle sensors provide this information. Furthermore, both data availability and data format were indicated as significant barriers in crash reconstruction. While automated vehicle sensors can provide a robust dataset for crash investigation, accessing this information from sensors and CAV manufacturers may be more challenging. Respondents also indicate the efficiency and speed of collision investigations could be improved. Access to the information provided by CAV sensors and cameras can allow for more efficient crash investigations by depicting a clearer image of crash events.

The value of vehicle camera footage as a data source for crash investigation emerges as a theme in the responses to multiple survey questions. Out of each automated vehicle sensor, cameras are the most valuable and familiar to the officers. Many respondents already have experience using cameras during a crash investigation, whether through surveillance cameras, aftermarket dashboard cameras, or automated vehicle cameras. Further, cameras in the form of drones and digital cameras are often used during crash investigations to capture the crash scene accurately. With the growing prevalence of interior and exterior cameras in modern vehicles, the abundance of crash footage will significantly augment the available information to crash investigators.

Question two asks how law enforcement can be better prepared to utilize CAV data in crash investigations, and this study responds by providing a list of training topics. The need for additional training is emphasized throughout the multiple-choice responses. Specific areas of unfamiliarity include understanding automated vehicle technology, cybersecurity, communication and community engagement, and traffic enforcement. The following list of training topics, shown in Table 7, is curated based on survey results. While an understanding of all these training topics is pertinent, they are prioritized both by respondents' unfamiliarity and the potential to advance crash investigations.

Table 7: List of Training Topics

Pertinent Automated Vehicle Training Topics for Crash Investigators in Law Enforcement		
1.	**Understanding Automated Vehicle Technology**	This is a broad topic that includes multiple facets of learning, including what sensors are used in different makes and models of automated vehicles, what data can be collected from these sensors, and how this new technology can impact human and roadway factors. According to the survey, the most unfamiliar automated vehicle technologies are Ultrasound Sensors, Millimeter Wave Radar, Infrared Sensors, and Onboard Units. Cameras and GPS are both relatively familiar to survey respondents; however, it is necessary to train officers to access the cameras and GPS sensors equipped in automated vehicles.
2.	**Accessing Automated Vehicle Data**	Accessing automated vehicle sensor data may require coordination with vehicle manufacturers and the use of data retrieval equipment. Crash investigators must be properly trained in the processes by which data is retrieved.
3.	**Applying Automated Vehicle Sensor Data to Crash Investigation**	Once data is retrieved, crash investigators must be trained in how to properly analyse and apply crash data as evidence. This will require familiarity with handling multiple data types from various sensors and using different software programs for analysis.
4.	**Cybersecurity Concerns**	While automated vehicles can increase safety and mobility, automated vehicles can be subjected to cybersecurity threats, which introduce new hazards on the roadways. Crash investigators must be made aware of these threats and learn how to properly mitigate and respond to cybersecurity concerns.
5.	**Traffic Enforcement and Regulation**	Automated vehicles operate differently than conventional vehicles, which may lead to shifting traffic enforcement and regulation practices in the near future. Law enforcement will need to be trained in local traffic regulations regarding AVs.
6.	**Communication and Community Engagement**	Law enforcement, once trained in automated vehicle technology, should also be trained in how to raise public awareness to new CAV technology, regulations, and potential risks.
7.	**Legal and Ethical Implications**	Automated vehicles introduce new driver-vehicle relationships to the roadway, and with these shifting relationships, the ethical and legal landscape also evolves. It is not always immediately clear how all automated vehicle crashes should be handled and who should be faulted for a crash. Therefore, crash investigators should be trained in the ethical and legal principles that guide crash culpability.

6. LIMITATIONS

Since all the respondents are from Tennessee, the results may not be generalizable in other jurisdictions with differing crash investigation procedures. Furthermore, this survey has a small sample size. The methodology of this survey can be reproduced in future studies with a larger sample size to enable greater generalization of the results. Other typical biases associated with survey research are recognized, e.g., non-coverage bias, response bias, social desirability bias, and self-reporting biases that can lead to inaccurate data and results.

7. CONCLUSION

This survey of 61 crash investigators in law enforcement in Tennessee revealed valuable insights about how crash investigations can be advanced using automated vehicle sensor data. The first research question asks what pertinent information is lacking in crash investigations. According to the survey, vehicle and occupant dynamics are the most requested information currently lacking in crash investigations, and this information can be provided by CAV sensors. The respondents also indicated a need for standardization in data retrieval processes, and many comments expressed a demand for state-funded training regarding CAV data. The second research question asks how crash investigators can best prepare to use CAV data in crash investigations. To this end, guidance on the appropriate training for law enforcement is provided, with the most pertinent topics being a comprehensive understanding of various CAV sensors and their uses and how to access this data from manufacturers using the necessary equipment. Furthermore, the results from factor analysis also emphasize the need for integrating digital data provided by CAV sensors, specialized and sophisticated training of crash investigative officers, and adopting standardized protocols for CAV crash investigation to improve its efficiency and effectiveness. Future research can include follow-up surveys that assess whether any of the suggested training has been implemented and if there is an improvement in CAV technology literacy among crash investigators.

ACKNOWLEDGMENTS

The authors express their sincerest gratitude to Mr. William “Buck” Campbell, Dr. Jerry Everett, and the Tennessee Highway Safety Patrol Office for his assistance with this project. This research was sponsored by the Collaborative Sciences Center for Road Safety (CSCRS), US Department of Transportation.

AUTHOR CONTRIBUTIONS

The authors confirm their contribution to the paper as follows: **Study conception and design:** A. Khattak, M. King; **Data collection:** A. Khattak, M. King; **Analysis and interpretation of results:** A. Khattak, M. King, S. Usman; **Draft manuscript preparation:** M. King, S. Usman, M. Adeel, A. Khattak. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

[1] Clamann, M., and A. J. Khattak. Advancing Crash Investigation With Connected and Automated Vehicle Data Collaborative Sciences Center for Road Safety (CSCRS). 2022.

[2] Scanlon, J. M., K. D. Kusano, T. Daniel, C. Alderson, A. Ogle, and T. Victor. Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain. *Accident Analysis & Prevention,* Vol. 163, 2021, p. 106454.

[3] Owings, R. P. *Submittal of the Final Report of the NHTSA R%D Event Data Recorder (EDR) Working Group to Docket*, 2001.

[4] Lee, J., and Y. Lee. Estimating vehicle speed by analyzing the acoustic frequency of dashboard camera sound. *Forensic Science International,* Vol. 338, 2022, p. 111384.

[5] Giovannini, E., A. Giorgetti, G. Pelletti, A. Giusti, M. Garagnani, J. P. Pascali, S. Pelotti, and P. Fais. Importance of dashboard camera (Dash Cam) analysis in fatal vehicle–pedestrian crash reconstruction. *Forensic Science, Medicine and Pathology,* Vol. 17, No. 3, 2021, pp. 379-387.

[6] Stanton, N. A., P. M. Salmon, G. H. Walker, and M. Stanton. Models and methods for collision analysis: A comparison study based on the Uber collision with a pedestrian. *Safety Science,* Vol. 120, 2019, pp. 117-128.

[7] Yakar, İ., E. Kanun, G. M. Oğuz, Ş. Bozduman, S. Bilgi, and K. Karataş. A Review on the Usability of Mobile Phone-based Close-Range Photogrammetry, Terrestrial Laser Scanning and UAVs in Traffic Accident Modeling. *Intercontinental Geoinformation Days,* Vol. 1, 2020, pp. 13-16.

[8] Catapang, A. N., and M. Ramos. Obstacle detection using a 2D LIDAR system for an Autonomous Vehicle.In *2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, 2016. pp. 441-445.

[9] Broggi, A., A. Fascioli, M. Carletti, T. Graf, and M. Meinecke. A multi-resolution approach for infrared vision-based pedestrian detection. In *IEEE Intelligent Vehicles Symposium, 2004*, 2004. pp. 7-12.

[10] Zolock, J., C. Senatore, R. Yee, R. Larson, and B. Curry. The use of stationary object radar sensor data from advanced driver assistance systems (ADAS) in accident reconstruction. 2016.

[11] Arvin, R., M. Kamrani, and A. Khattak. How instantaneous driving behavior contributes to crashes at intersections: Extracting useful information from connected vehicle message data. *Accident Analysis & Prevention,* Vol. 127, 2019, pp. 118-133.

[12] Chen, Q., Y. Xie, Y. Ao, T. Li, G. Chen, S. Ren, C. Wang, and S. Li. A deep neural network inverse solution to recover pre-crash impact data of car collisions. *Transportation Research Part C: Emerging Technologies,* Vol. 126, 2021, p. 103009.

[13] Cormier, J., J. Funk, G. Beauchamp, and D. Pentecost. Pycrash: An Open-Source Tool for Accident Reconstruction. *SAE Technical Paper*, 2021, pp. 01-0896.

[14] Muggenthaler, H., S. Drobnik, M. Hubig, M. Schönpflug, and G. Mall. Fall from a Balcony—Accidental or Homicidal? Reconstruction by Numerical Simulation. *Journal of Forensic Sciences,* Vol. 58, No. 4, 2013, pp. 1061-1064.