Direction of Arrival of Equipment Sound in the Construction Industry

K. Elelu[a], T. Le[a] and C. Le[b]

[a]Glenn Civil Engineering Department, Clemson University, USA.
[b]Department of Civil, Construction and Environmental Engineering, North Dakota State University, USA
E-mail: kelelu@g.clemson.edu, tuyenl@g.clemson.edu, chau.le@ndsu.edu

Abstract

The construction industry is often affected by unanticipated struck-by accidents, which often cause severe injuries and fatalities to the workers. Therefore, monitoring and tracking struck-by hazards in terms of the spatial relationship between a worker and a heavy vehicle is crucial to prevent such accidents. Current studies focus on using active sensors and implementing computer vision but not on the audibility of their safety signals. To address this issue, this paper utilizes sound, a ubiquitous data source present in every construction site, to track and separate equipment sound into different types and determine the direction of arrival (DOA) using the Open embeddeD Audition System (ODAS) framework. Each circular array performs DOA estimation independently using commercial software on two equipment sound sources, bulldozer (mobile) sound and hammer (stationary) sound. The DOAs are fed to a relational database, pre-processed, and used to perform the source tracking. This process provides a step towards monitoring the spatial relationship between workers and equipment with few labels of source location for calibration. The results of our study showed that this method was effective in identifying activities of multiple pieces of equipment in real-time in construction sites without the need for separating sound signals in advance. Future studies can focus on triangulating the exact location of the sound source with less computation power and monitoring how this helps improve workers' awareness of surrounding equipment.

Keywords –
Direction of Arrival; Machine Equipment; Struck-by Accident; Local coordinate

1 Introduction

For a long time, researchers have considered struck-by heavy vehicle accidents as one of the leading causes of death in the labor-intensive construction industry [1]. This is mostly due to the unpredictable nature of construction equipment and limited workspaces resulting in lifelong impairment, and fatalities [2]. Bureau of Labor Statistics in 2020 published that 15% of fatal occupational injuries resulted from hazardous contact with equipment and objects [3]. Also, according to Occupational Safety and Health Administration (OSHA) figures, struck-by-equipment hazards accounted for 58 % of struck-by incidents from 1995 to 2008 [4]. Therefore, contact collisions between construction workers on foot and equipment pose a significant risk to the safety and health of construction workers. However, the capacity to spot these collision accidents in advance on a site is vital to any construction project since preventing unanticipated catastrophes is always the best way to avoid them. Therefore, an automated safety monitoring of struck-by hazards has emerged as a potential option for effective safety management on a construction site [5]. In this paper, the authors propose the use of audio sound to extract important and useful information regarding equipment activities performed in the site. This is possible as audio-based activity identification is very easy to collect regardless of dynamic occlusions and different tools have signature distinguishable sounds which make it suitable for task identification [6].

2 Related Studies

Some past work has focused on the usage of visual sensors for localizing and characterizing the behavior of resources and has been extensively applied in the construction industry and achieved promising results by various methods [7]. Several active technologies have also been developed to provide proximity sensing and alerts for workers and equipment operators, such as image wearable devices [8], tactile-based wearable devices [9], Bluetooth-low energy [7], and global positioning system [11]. Some studies have used GPS data for long-distance detection of collision hazards. For example, an unsafe-proximity detection model focused on decreasing false alarms has been developed using a GPS-aided inertial navigation system sensor as the state tracking module [12]. Another GPS-related method for recording, identifying, and analyzing interactive hazardous near-miss situations between workers-on-foot and heavy construction equipment was presented using spatiotemporal data [13]. Shorter sensing devices like Bluetooth Low Energy (BLE)-based proximity sensing have been applied to address work-zone safety to allow understanding of dynamic spatial relationships among equipment, operator, workers, and a surrounding work environment [10]. Also, a spatiotemporal network-based model was developed at both entity and network levels to perform dynamic risk analysis on the struck-by-equipment hazard [14]. Sakhakarmi et al, develop a proximity sensor using a wearable tactile-based system for workers to help improve their hazards perception [9]. These smart and automated systems are effective but prone to false alarms. Wang et al. [12] presented two novel four-dimensional models, a time-sphere model and a time-cuboid model that effectively reduces the rate of false alarms; still, the data not wirelessly communicated made it challenging to apply in real-life situations. There is also a growing application of artificial intelligence in safety management such as using computer vision technology in image detection to monitor a comprehensive view of the area surrounding equipment via cameras installed on every side of the equipment body. This allows displaying the surrounding environment on the operator's monitor to protect the workers from potentially dangerous situations involving equipment operations [15]. Researchers also developed a digital twin using the Bayesian network that fed and updated real-time data from sensors that pro-actively forecast dangerous scenarios based on collision probability on affected workers-on-foot [16]. Most of these methods were implemented by installing sensors or electronic devices such as cameras directly on equipment or construction workers. Besides, those devices are relatively expensive; thus, they cannot be used on a wide scale in reality. Also, some base the alert detection on the information obtained from available hazard detection records, as more factors can cause dangers to workers that are not factored into the study. To address the gaps, there is a need to investigate the surveillance approach that is less expensive and adaptable to rigorous construction effect like during the excessive dust in visual equipment monitoring. Auditory surveillance using the sound collected from construction sites could address these issues. However, there is a lack of such a method of processing audio signals for equipment monitoring and safety against collision hazards in the construction field.

A few efforts have been made in other sectors' research in auditory surveillance for collision hazard detection. A real-time framework was created to detect multi-vessel collisions using a spatial clustering process to detect clusters of encounter vessels within each cluster from the vast number of monitored vessels in a surveyed sea area [17]. The framework effectively and efficiently detects encounter vessels and ranks collision risk indexes within each cluster. Another method for detecting collision hazards amongst motorcycles via accelerometer measures was proposed using a machine learning model [18]. The system was designed using data from an instrumented vehicle and validated in simulation. Nonetheless, research on auditory surveillance focusing on detecting collision hazards in construction job sites has not gained much attention in the research community.

Acoustic emissions from construction activities were used to calculate the working hour to allow field managers to know workers' work progress and productivity [19]. Success from this study shows how using auditory surveillance can prove positively in detecting equipment sound type. Another using surveillance technology in the construction industry helps support the construction industry's safety performance since the lack of sufficient visibility is the principal factor leading to fatalities [20]. As stated earlier, most sound-based surveillance technologies were only focused on monitoring construction work activities and equipment operations. For instance, a hybrid system was proposed for recognizing multiple construction equipment activities [18]. A supervised machine learning-based sound identification algorithm was implemented to enhance monitoring and performance of construction site activity [21]. A few studies attempted to develop new approaches for conducting an audio-based event detection system for safety. Experimental trials were designed to deploy sensing technology to provide alerts to proximity detection when heavy construction equipment and workers are in close proximity [22]. Nonetheless, the devices were installed on construction equipment only, not equipped on construction workers. Another approach using a machine learning algorithm can categorize sound events and make construction workers aware of possible safety risks and hazards [23]. Still, the sound data relating to collision hazards were only collected from a particular worksite. Such an approach is restrained because the sounds emitted by the equipment from various construction sites may differ and contribute a different amount of noise. Studies have been done on determining the optimum position for construction noise barrier location, and a comparison was made during each stage of the construction process. All in all, there is a need to develop an audio-based localization of stationary and mobile equipment framework. In doing so, this paper particularly focuses on extracting the local coordinate of the sound source and process this data to locate the position of construction equipment.

3 Methodology

This study introduces a framework for audio base activity detection to prevent the struck-by construction equipment hazards, which is illustrated in Figure. 1. This framework consists of three phases that include equipment sound extraction, data processing for DOA estimation, and worker's danger notification signal based on distance computed. First, the extraction of the sound record is done in real-time, which is possible with the microphone array mounted on the raspberry pi. The second phase has these audio data separated into different equipment sounds, containing each sound's identification with their local coordinate value uploaded into the database for data cleaning and processing. This is done by separating the mobile and stationary equipment with their unique sound identification and entering each sound type into a separate relational database where the DOA of the sound from the workers is estimated. Lastly, based on the DOA, the location and distance from workers are calculated as shown in equation 7 to notify the workers based on the danger zone specified. Each of these steps is described in further detail below.

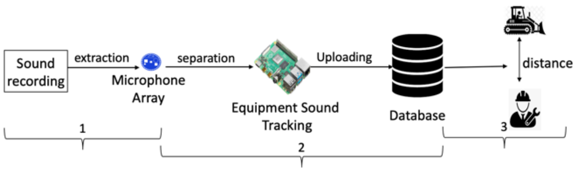


Figure 1: The overall framework for audio-based activity identification

3.1 Sound Processing Pipeline

We employ a ReSpeaker Mic Array v2.0 device with four high-performance digital microphones and twelve configurable RGB LED indications (see Figure 2) coupled to a Raspberry Pi 3 processor. This RP3 runs its own instance of the ODAS framework, which outputs direction of arrival (DOA) represented by a 3D unit vector in the array's local coordinate system, as shown in Figure 3. The DOA of the arrays lies on a virtual unit hemisphere with the z-axis facing up and lies on the positive side and is defined in equation 1 [24].

$$e = [e_x\, e_y\, e_z]^T \tag{1}$$

where $e_x^{\,2} + e_y^{\,2} + e_z^{\,2} = 1$, and $e_z > 0$.

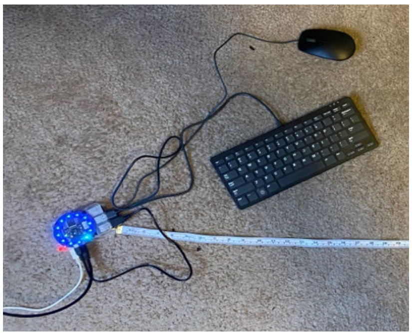


Figure 2: Experiment Setup

To maximize DOA localization on the microphone array, MacBook Pro and Android smartphone speakers are utilized to play a bulldozer (mobile) sound and a hammer (stationary) sound, respectively. The loudness of the bulldozer and hammer sound using a decibel-meter is in the range of 75dB to 79dB. The ODAS sound source tracking module is designed to detect both static and moving sources. Although it is capable of tracking up to 4 sources at this time, we only make use of two sound sources for this experiment. A single microphone array is installed in a controlled space (see Figure 4 for calibration locations). This array is positioned in the top-left middle part of the room, while the mic array is oriented with the z-axis facing the ceiling and the x-axis facing the northeast direction. It should be noted that the processing is independent and unaffected by the array's positions and orientations. These relations are useful for interpreting the findings and are not employed to locate the sound source. Microphone arrays are usually omnidirectional, which make their orientation relevant [23], and the multi-channel raw audio is sampled at 44,100 Hz from the ReSpeaker array, resampled by ODAS at 16,000 Hz, which then returns an updated DOA estimation [24]

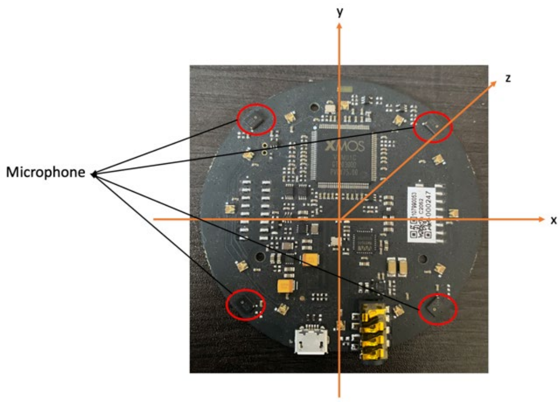


Figure 3: ReSepaker Microphone Array

We also estimate the angle in degrees to determine the direction in a cartesian plane with the x and y local coordinate collected using equation 2.

$$\theta = \tan\beta - 1\ (\frac{y}{x}) * (\frac{180}{pi}) \quad (2)$$

3.2 Experimental Set-up

This section describes how we set up our experiment as well as the technique we used to determine the DOA of construction equipment sound. Firstly, a bulldozer sound device is placed sequentially at four different locations away from the microphone array, as shown in Figure 4; after every 30sec each, the microphone array is moved to these new points to record a 3-component DOA of approximately 130 records estimated every 1s. The device is moved from position 1 to a new position 2, and this process is repeated for positions 3 and 4 while keeping the hammer sound at a single position 4 for the entire duration of the experiment. The procedures are repeated five times, and the measurements were performed in nearly a horizontal plane. After the experiment's conclusion in a controlled environment free from noise and environmental factors, a DOA local coordinate record between the range of 15,100-15,200 for both the stationary and mobile sound was recorded. A timer on an iPhone device is synchronized with the time on RP3. When left running continuously, each calibration point is estimated to record 3700-4000 for the 30s. The timestamp, id, energy, and local coordinate of DOA estimates are saved locally as a text file which is pushed to the cloud and loaded to a relational database for data cleaning and pre-processing. The data is then sorted and joined based on the time stamp. The result is a chronological table that contains the DOAs from all arrays synchronized using the timestamps. This table can be queried to get the full information for any period of time. ODAS tracks the loudest sound source and records the DOAs with energy greater than zero as dynamic. Once the experiment can verify the direction of arrival of construction equipment, which was done in this study, the experiment will be further expanded to detect collision hazard using two microphone array set-ups on the construction site. The first mic-array is a fixed at a point, and the second mic-array is a mobile mounted on the construction workers. To pinpoint the location and distance of the equipment from the worker, we considered several scenarios to present the derivation of the distance to the worker. One of the case scenarios is when a bulldozer and hammer equipment sound are played simultaneously, the bulldozer changes location during the period of operation, if a construction worker is in the danger zone, the framework should be able to notify the workers of the current status and vice versa. The performance will be tested both in a control environment and a construction site to examine the amount of false alarm generated from the notification which will help to determine wide range application on the construction site.

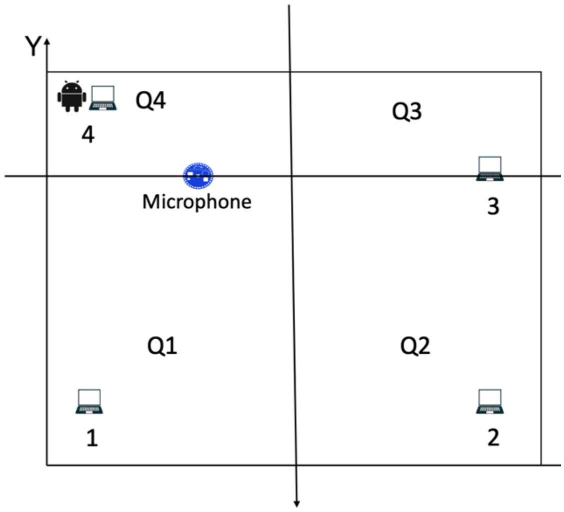


Figure 4: Position 1 to 4 and the location of microphone array.

4 Result and Discussion

As mentioned in Section 3, two sound devices were used for five different scenarios considering equipment like bulldozer and hammer. Figure 5 present the performance analysis result for all the scenarios, and its shows the bulldozer sounds being moved manually from calibration point one to four. The first position of the mobile device is at the positive x and y coordinate in Q1 with a value above the zero line. At position two, the coordinate enters the second quadrant Q2, with the x-axis changing to a negative value slightly below the zero point, as shown in Figure 4. At position three, even though there is a drop close to the zero line, it shows a positive y-axis and still slightly occupying the second quadrant's space. The fourth position shows a positive x-axis and a negative y-axis with values above zero and below zero, respectively. Lastly, the DOAs at experiment four at some point shows some roughness due to slight environmental disturbance and other issues. This is due to a substantial disturbance of the sound source at and near point three from the audio being played, confirmed by re-playing the audio sound along with that duration. A slightly straight line is observed for the stationary hammer sound as this maintained position four without any location changes.

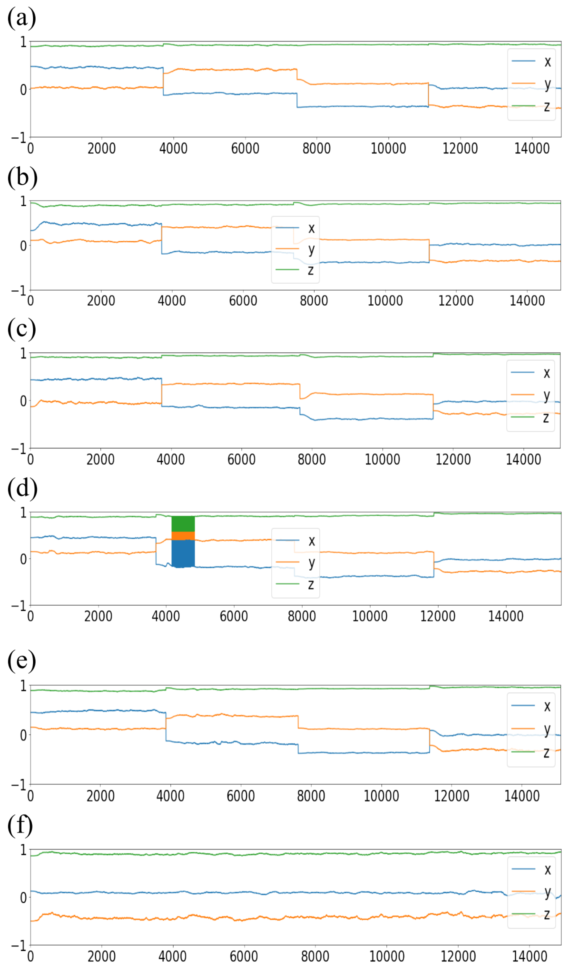
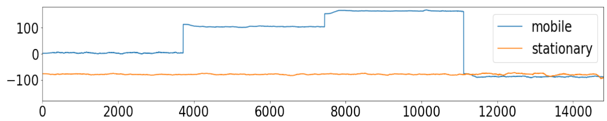


Figure 5: (a-e) Mobile sound DOAs measured from the five arrays and shown measurement from point 1 to 4 and (f) Stationary sound DOAs measured at a single location.

4.1 Angle of Arrival of Sound

This experiment shows we considered four planar wavefronts at direction-of-arrival of 1.47*, 106*, 163*, and 268* for points one, two, three, and four, respectively. The fourth point for mobile and the stationary device were coherent, as they were positioned at the same location in the experiment space.



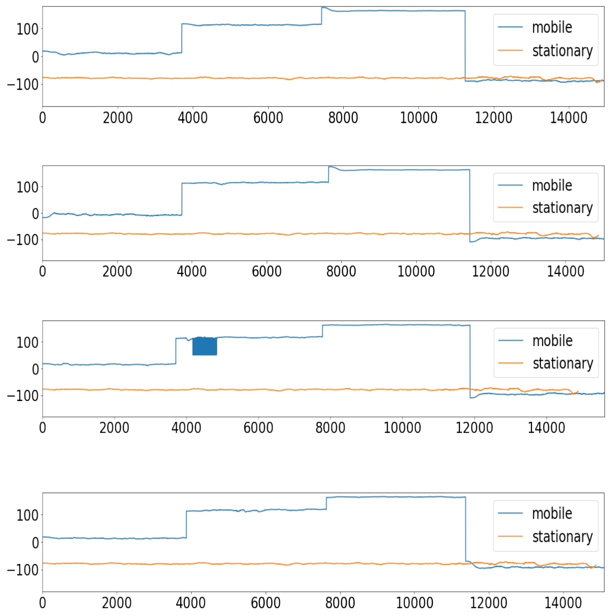
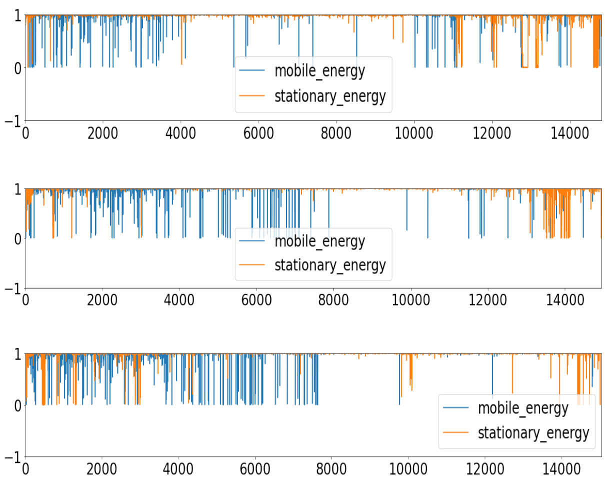


Figure 6: Mobile and stationary sound angle measurement in degree for the five arrays and shown measurement along the point 1 to 4

4.2 Energy Timeseries

The framework gives additional information about the loudness of the sound in a normalized form, with a scale from 0 to 1 for both the mobile and stationary sounds. Figure 7 shows the high stationary energy value, which is attributed to the uniqueness in the sound type, with is a loud intermittent sound that is not affected by sand or concrete noise. The end position of the experiment has a relatively lower energy value for the hammer sound; playing back the audio to confirm the information, we notice a decrease in sound while in operation.



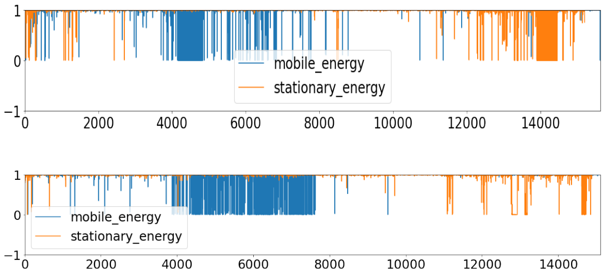


Figure 7: Energy propagation for experiment 1 to 5

4.3 Local coordinate mapping

In each of the DOA local coordinates for the x and y-axis, the projected shape, and the movement of bulldozer sound along the point indicated are recognized, allowing for determination of the movement with few labels on the initial source location. The magnitude of the vector's entries shows a deeper presence at these four calibration points as the device maintains some 30seconds—also, a path of the movement of the equipment sound. Since the experiment has entries for the five sets of measurements, their mapping is different and color-coded, as shown in Figure 8.

4.4 Discussion

The ODAS framework exploits the directivity of microphones to compute generalized cross-correlation phase transform (GCC-PHAT) between pairs of microphones [25] and compute time delay of arrival (TDOA) based on the microphones that are close to each other, which indirectly impacts the performance as an array with less microphone and evidently reduce the accuracy of detection [24]. Equipment sound like a hammer that has strong intermittent sound is easily picked up by the microphone and produces high energy value compared to continues prolong sound like a bulldozer which level of sound can be highly impacted by the materials it is working on and the condition of the machine. The fast, optimized processing strategies make it possible for this framework to perform all processing on low-cost hardware like Raspberry Pi 3, which is cheaper and more economical than the already available sound detection device. Lastly, due to the scope of this study, one particular limitation of this research is that the system does not account for noise filtering on the reSpeaker v2, and further work is needed to be done to eliminate background noise disrupting the sound capture.

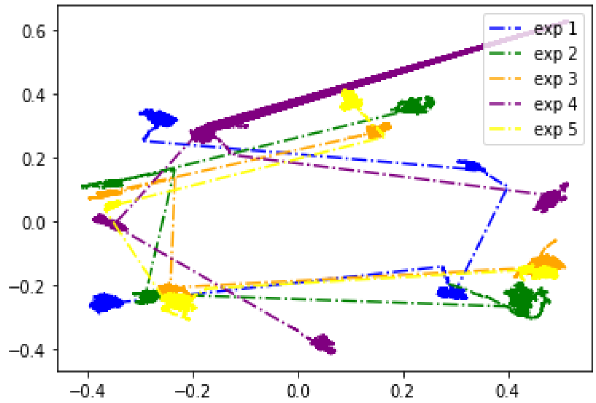


Fig 8: Equipment mapping of positions 1 to 4 showing the movement of equipment sound along the line.

5 Conclusion

This research focused on pro-active struck-by hazard detection in construction and introduced an approach to simulate a sound source localized from a series of local coordinate system arrays. Each computed array performs the DOA estimation independently and feeds its measurements to a data center where the DOA from all arrays are timestamped and preprocessed for data exploration and analysis. The method was demonstrated with a single circular microphone array in a controlled environment. The methods demonstrated here provide a step towards monitoring activities in construction sites with no training effort, as the device has an inbuilt algorithm to separate four sound sources. To make the device more stream-lined, we plan to design a flexible device to allow to be able to mount it on construction workers and will ultimately contribute to promoting a safer working environment for construction workers.

References

[1] W. Wu, H. Yang, Q. Li, and D. Chew, "An integrated information management model for proactive prevention of struck-by-falling-object accidents on construction sites," *Automation in Construction*, vol. 34, pp. 67–74, 2013, doi: 10.1016/j.autcon.2012.10.010.

[2] B. Shao, Z. Hu, Q. Liu, S. Chen, and W. He, "Fatal accident patterns of building construction activities in China," *Safety Science*, vol. 111, pp. 253–263, Jan. 2019, doi: 10.1016/j.ssci.2018.07.019.

[3] Bureau of Labor Statistics, "NATIONAL CENSUS OF FATAL OCCUPATIONAL INJURIES IN 2020," 2020. [Online]. Available:

www.bls.gov/iif/oshwc/cfoi/cfoi_rates_2020hb.xlsx.

[4] W. Wu, H. Yang, D. A. S. Chew, S. hua Yang, A. G. F. Gibb, and Q. Li, "Towards an autonomous real-time tracking system of near-miss accidents on construction sites," *Automation in Construction*, vol. 19, no. 2, pp. 134–141, Mar. 2010, doi: 10.1016/j.autcon.2009.11.017.

[5] J. Park, K. Kim, and Y. K. Cho, "Framework of Automated Construction-Safety Monitoring Using Cloud-Enabled BIM and BLE Mobile Tracking Sensors," *Journal of Construction Engineering and Management*, vol. 143, no. 2, p. 05016019, Feb. 2017, doi: 10.1061/(asce)co.1943-7862.0001223.

[6] K. M. Rashid and J. Louis, "Activity identification in modular construction using audio signals and machine learning," *Automation in Construction*, vol. 119, Nov. 2020, doi: 10.1016/j.autcon.2020.103361.

[7] D. Roberts and M. Golparvar-Fard, "End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level," *Automation in Construction*, vol. 105, Sep. 2019, doi: 10.1016/j.autcon.2019.04.006.

[8] 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, Nov. 2017, doi: 10.1016/j.autcon.2017.06.014.

[9] S. Sakhakarmi, J. W. Park, and A. Singh, "Tactile-based wearable system for improved hazard perception of worker and equipment collision," *Automation in Construction*, vol. 125, May 2021, doi: 10.1016/j.autcon.2021.103613.

[10] J. Park, X. Yang, Y. K. Cho, and J. Seo, "Improving dynamic proximity sensing and processing for smart work-zone safety," *Automation in Construction*, vol. 84, pp. 111–120, 2017, doi: https://doi.org/10.1016/j.autcon.2017.08.025.

[11] N. Pradhananga and J. Teizer, "Automatic spatio-temporal analysis of construction site equipment operations using GPS data," *Automation in Construction*, vol. 29, pp. 107–122, 2013, doi: https://doi.org/10.1016/j.autcon.2012.09.004.

[12] J. Wang and S. N. Razavi, "Low False Alarm Rate Model for Unsafe-Proximity Detection in Construction," *Journal of Computing in Civil Engineering*, vol. 30, no. 2, p. 04015005, Mar. 2016, doi: 10.1061/(asce)cp.1943-5487.0000470.

[13] O. Golovina, J. Teizer, and N. Pradhananga, "Heat map generation for predictive safety planning: Preventing struck-by and near miss interactions between workers-on-foot and construction equipment," *Automation in Construction*, vol. 71, pp. 99–115, 2016, doi: https://doi.org/10.1016/j.autcon.2016.03.008.

[14] J. Wang and S. Razavi, "Spatiotemporal Network-Based Model for Dynamic Risk Analysis on Struck-by-Equipment Hazard," *Journal of Computing in Civil Engineering*, vol. 32, no. 2, p. 04017089, Mar. 2018, doi: 10.1061/(asce)cp.1943-5487.0000732.

[15] H. Son, H. Seong, H. Choi, and C. Kim, "Real-time vision-based warning system for prevention of collisions between workers and heavy equipment," *Journal of Computing in Civil Engineering*, vol. 33, no. 5, p. 4019029, 2019.

[16] M. Leonardo, N. Berardo, C. Alessandro, R. Luigi, and D. G. Giuseppe Martino, "Development of a Digital Twin Model for Real-Time Assessment of Collision Hazards," 2020, pp. 14–19. doi: 10.3311/ccc2020-003.

[17] R. Zhen, M. Riveiro, and Y. Jin, "A novel analytic framework of real-time multi-vessel collision risk assessment for maritime traffic surveillance," *Ocean Engineering*, vol. 145, pp. 492–501, 2017, doi: https://doi.org/10.1016/j.oceaneng.2017.09.015.

[18] D. Selmanaj, M. Corno, and S. M. Savaresi, "Accelerometer-based data-driven hazard detection and classification for motorcycles," in *2014 European Control Conference (ECC)*, 2014, pp. 1687–1692. doi: 10.1109/ECC.2014.6862549.

[19] C. Cho, Y.-C. Lee, T. Zhang, and P. D. Student, "Sound Recognition Techniques for Multi-Layered Construction Activities and Events."

[20] J. W. Hinze and J. Teizer, "Visibility-related fatalities related to construction equipment," *Safety Science*, vol. 49, no. 5, pp. 709–718, Jun. 2011, doi: 10.1016/j.ssci.2011.01.007.

[21] T. Zhang, Y.-C. Lee, M. Scarpiniti, and A. Uncini, "A supervised machine learning-based sound identification for construction activity monitoring and performance evaluation," in *Construction Research Congress 2018*, 2018, pp. 358–366.

[22] E. D. Marks and J. Teizer, "Method for testing proximity detection and alert technology for safe construction equipment operation," *Construction Management and Economics*, vol. 31, no. 6, pp.

636–646, Jun. 2013, doi: 10.1080/01446193.2013.783705.

[23] Y. C. Lee, M. Shariatfar, A. Rashidi, and H. W. Lee, "Evidence-driven sound detection for prenotification and identification of construction safety hazards and accidents," *Automation in Construction*, vol. 113, May 2020, doi: 10.1016/j.autcon.2020.103127.

[24] P. Gerstoft *et al.*, "Audio scene monitoring using redundant ad-hoc microphone array networks," *IEEE Internet of Things Journal*, pp. 1–1, Aug. 2021, doi: 10.1109/jiot.2021.3103523.

[25] F. Grondin *et al.*, "ODAS: Open embeddeD Audition System," Mar. 2021, [Online]. Available: http://arxiv.org/abs/2103.03954