ISSN 1392-3730/eISSN 1822-3605

JOURNAL of CIVIL ENGINEERING and MANAGEMENT

2026
Volume 32
Issue 1
Pages 84–101
https://doi.org/10.3846/jcem.2026.24922

AN INTEGRATED VISION-BASED DYNAMIC COLLISION RISK ASSESSMENT FRAMEWORK OF WORKERS AND MOBILE MACHINERY ON CONSTRUCTION SITE

Xingrong GAO [1], Wenjing KANG [2], Yunting LIU [1], Qian ZHANG [3], Ting-Kwei WANG [4]✉

[1] *School of Management Science and Real Estate, Chongqing University, Chongqing 40044, China*
[2] *Tianjin Binhai New Area Natural Resources Survey and Registration Center, Tianjin 300450, China*
[3] *School of Design and Built Environment, Curtin University, Perth, Australia*
[4] *Department of Civil Engineering, National Kaohsiung University of Science and Technology, Kaohsiung 82444, Taiwan*

Article History:
- received 11 April 2022
- accepted 10 June 2025

Abstract. The concepts of environmental impact index and spatial conflict degree have gained prominence in enhancing the controllability of collision accidents and mitigating the likelihood of collisions during construction processes. Nevertheless, prior studies predominantly focused on exploring collisions in terms of proximity and congestion between pairs of entities, thereby overlooking a comprehensive consideration of workers, construction machinery, environmental factors, and the spatial interaction of specific activities.To this end, this study aims to propose an integrated vision-based dynamic collision risk assessment framework by setting workers and mobile machinery as targeted research objectives and embedding a comprehensive risk assessment model in the proposed framework, thereby comprehensively assessing four types of risk factors (i.e., proximity, congestion, environmental impact index, and spatial conflict), and visualize the hierarchy of risk warnings. Firstly, a comprehensive risk assessment model was developed by using the fuzzy comprehensive evaluation method. This is followed by developing a dynamic risk assessment framework to extract the spatial information of the monitored objects by using computer vision as the underlying data of the risk factors. Finally, the proposed integrated framework was validated by an experimental study. This experiment's safety risk assessment results are consistent with expectations, which largely illustrates the effectiveness of the evaluation model constructed in this paper. For the vision module, the accuracy of classification and monitoring is more than 95%, and the speed of the object detection algorithm to process the video is about 10 frames per second, which shows the feasibility of this study.

Keywords: dynamic collision risk assessment, fuzzy comprehensive evaluation, environmental impact factor, space conflict, computer vision.

✉Corresponding author. E-mail: *tingkwei@nkust.edu.tw*

1. Introduction

Accident fatality and injury rates in the construction industry are two or three times higher than the average of other industries (Kang et al., 2017). Of all incidents, collisions are considered one of the four most life-threatening construction injuries, accounting for 17% of the total construction site mortality. Of these, 75% of accidents involved heavy equipment, such as trucks and cranes, and most of which were workspace collisions involving workers and equipment induced (Dagan & Isaac, 2015). The construction industry is typically labor-intensive. High-frequency movement of workers, materials, and construction machinery tends to lead to space congestion, and spatial collisions may not only reduce worker productivity but also result in safety incidents (Moon et al., 2014).

Traditional safety monitoring methods rely on manual inspections, which are time-consuming and labor-intensive (Chi & Caldas, 2012). Hence, it is crucial to automate construction operations to prevent incidents related to workers and equipment. Computer vision thus has become one of the most popular agendas in construction site safety management due to its advantages of user-friendliness, reusability, and high-level understanding (Martinez et al., 2019b). It not only can obtain information about the location of the monitored object but also can gain an understanding of the complex work patterns by digital images to effectively monitor the dynamic environment changes (Zhong et al., 2019).

Furthermore, previous studies on the collision risk of the construction site primarily focus on proximity and congestion by analyzing the relationship between pairs of entities (Kim et al., 2016). Existing studies (Kim et al., 2016; Son et al., 2019) have indicated that the incidence of collisions is not solely correlated with the proximity to construction machinery and the level of congestion at the construction site, but is also influenced by environmental factors and the extent of spatial conflict. However, those studies emphasize the relationships between paired entities (Zhang et al., 2020), with minimal attention given to the comprehensive effects of environmental factors and all surrounding entities in analyzing collision risks on construction sites. In addition, although prior research has focused on identifying risks and issuing warnings, few efforts have been made to investigate how these risks influence subsequent activities.

Therefore, it is essential for managers to understand the risk factors, how to quantify them, and how they impact subsequent activities. From management practices, this research proposes the following research questions:

RQ1. How can risk evaluation factors and spatial conflict metrics be systematically identified and quantified in environments involving mobile machinery and human workers?

RQ2. What are the effective methodologies to assess and model interactions between workers and dynamic entities, including their impact on safety-critical activities in evolving operational contexts?

RQ3. How can a dynamic, computer vision-based framework be developed to evaluate collision risks by integrating entity-specific characteristics and real-time spatial-temporal data?

To address the aforementioned research questions, the study has the following research objectives.

The research aims to develop a comprehensive, dynamic risk assessment framework that leverages computer vision to monitor and evaluate collision risks involving workers and mobile machinery. This involves identifying and validating key risk evaluation factors, establishing robust and quantifiable spatial conflict metrics, and modeling worker-entity interactions to understand their implications on safety and operational workflows. Ultimately, the framework will provide a proactive, real-time approach to enhancing safety in high-risk industrial environments.

The concept of spatial conflict is first proposed to transform the interaction between entities (workers and machinery) from distance to area. Then, the environmental impact index is developed to extend the research scope from the interaction of person-machine to person-machine-environment. Thereafter, the authors developed a comprehensive automated method for dynamic risk assessment of workers and mobile machinery based on diverse entity attributes.

2. Literature review

2.1. Risk assessment methods

Construction projects are inherently complex and risky, involving numerous uncertainties that can impact project outcomes. Risk assessment is a critical component of project management, ensuring that potential hazards are identified, evaluated, and mitigated to ensure the safety of workers, the integrity of the project, and adherence to timelines and budgets. Traditional risk assessment methods in construction often rely on static analyses, which identify potential risks during the planning phase and assess their likelihood and impact. For example, qualitative risk assessment involves identifying risks and categorizing them based on their likelihood and impact. Techniques such as risk matrices and SWOT analyses are commonly used (Raveendran et al., 2022). Quantitative risk assessment approach assigns numerical values to risks, often using probabilistic models to estimate the likelihood and impact of potential hazards. Techniques such as Fault Tree Analysis (FTA) and Event Tree Analysis (ETA) are examples of quantitative methods (Raveendran et al., 2022). Other methods include Monte Carlo simulation (Zhang et al., 2017), fuzzy theory (Seker & Zavadskas, 2017), and data fusion (Chen et al., 2019). Given the dynamic changes in construction sites, safety risk control procedures should be monitored through dynamic safety risk assessment and alerts to reduce accidents (Isaac & Edrei, 2016). The validity of current dynamic risk assessment methods heavily relies on the data that is primarily derived from feedforward signals, attempted events, and incident reports (Khakzad et al., 2014). Even though common static analysis models are often required, more dynamic risk analysis methods and automated processing software are still lacking (Paltrinieri et al., 2015).

2.2. Computer vision in construction safety

There are two types of hazards are generally found on construction sites (Golovina et al., 2016; Park et al., 2016): (1) Static hazards, which are inherent to the building's design and include elements such as temporary works, hazardous substance storage, site traffic control, and physical risks like floor openings for services or stairwells; and (2) Dynamic hazards, which involve the spatial-temporal movement of resources, such as workers, heavy equipment, and cranes being transported over active work areas. Construction site safety has long garnered significant attention from numerous researchers and stands among the most prolific domains for computer vision applications (Martinez et al., 2019a). Computer vision has revolutionized hazard detection in construction sites by enabling real-time identification of potential dangers. Research has primarily focused on how computer vision can be used to prevent people from entering work areas while heavy equipment is in operation (Kim et al., 2016, 2017). For example, Kim et al. (2016) integrated computer vision with a fuzzy inference method to monitor and evaluate a person's safety while working near heavy machinery. Notably, investigations in this realm primarily concentrate on ensuring the safety of construction workers, focusing on two principal dimensions: the individual worker and the interplay between workers and their surrounding environment (Zhang et al., 2020).

One avenue of research involves integrating ergonomics to examine workers' unsafe behaviors. Han and Lee (2013) proposed a vision-based motion capture and recognition framework to detect critical unsafe behaviors among construction workers on site. By extracting skeletal models from video for three-dimensional reconstruction, they demonstrated that predefined unsafe behaviors in the database can be detected (Han & Lee, 2013). Similarly, Guo et al. (2018) integrated computer vision techniques, construction safety knowledge, and ergonomic principles to develop a real-time skeleton-based recognition approach and constructed an unsafe-behavior database. By simplifying dynamic movements into static postures, they successfully identified unsafe behaviors (Guo et al., 2018).

For exploring the interplay between workers and their surrounding environment, vision-based distance measurement techniques can be broadly categorized into two approaches. The first approach relies on depth information to determine three-dimensional coordinates (Seo et al., 2015), while the second approach involves direct measurement using two-dimensional images (Kim et al., 2019). Chen et al. identified worker location and posture as two key quantitative features and proposed a location–posture integration principle for evaluating the safety risk of construction worker behaviors (Chen et al., 2019).

2.3. Spatial conflict

Previous studies highlighted that workers performing tasks near working equipment are the main causes of collisions (Luo et al., 2018). Therefore, it is necessary to decrease the possibility of collisions between people and mobile equipment at construction sites by effective collision monitoring. Computer vision thus can be used to automatically and constantly identify, classify, and monitor critical tasks and estimate motion, behaviors, and positions at construction sites. It is capable of providing rich information for different research purposes (Zhong et al., 2019). Existing studies mainly use it as an automated processing tool to assist target identification and tracking and systematically retrieve dynamic spatiotemporal data in order to assess the risk of physical collisions at construction sites (Zhang et al., 2020). Dynamic workspace monitoring can not only facilitate safety management in complex construction environments but also reinforce more accurate and proactive safety management.

Moreover, space occupancy is critical to the accuracy of conflict detection (Choi et al., 2014). Inter-object proximity and crowdedness were identified as major factors in collision research (Hu et al., 2020). The proximity between entities can directly cause safety incidents and congestion in the construction site may increase the possibility of collision accidents. Information on the location of workers and equipment, speed and direction of movement, crowdedness of the site environment, and workers' perceptions of risk were found to have significant impacts on crashes (Wang & Razavi, 2017). Hinze et al. (2005) thus classified causes of accidents into three categories: equipment-related cases, human factor-related cases, and environment-related cases.

Furthermore, the construction site environment is always noisy and crowded (Ballesteros et al., 2010). Large machinery is usually recognized as the main source of noise. Due to the labor-intensive nature of the construction industry, a large number of workers are exposed to noise hazards (Ballesteros et al., 2010). In a noisy environment, workers often concentrate on their work and ignore the dangers and potential collisions in the surrounding environments (Kim et al., 2016). The safety and productivity performance of workers in noisy and crowded worksite conditions are usually poor. There is a lack of effective ways to avoid work area congestion during construction processes (Zhang et al., 2015). Several previous studies included temporal-spatial conflict resolution as one of the steps in workspace planning and proposed qualitative and quantitative measures in this regard (Kassem et al., 2015). The conflict ratio, the ratio of the conflict volume to the required space volume, is usually used as a criterion of conflict severity (Akinci et al., 2002). Congestion is generally defined in collision studies as the number of entities within a certain range (Kim et al., 2016).

2.4. Research gaps

Both industry and academia have endeavored to integrate computer vision into construction safety compliance and hazard detection. Arrowsight Inc. (Europe Construction Tech Review, 2020) leverages machine learning (ML) and computer vision algorithms, in conjunction with Remote Video Auditing (RVA) services, to monitor and analyze risky behaviors, resulting in a remarkable reduction of unsafe worker behavior from approximately 50% to below 5%. Similarly, Viso Suite™ (Viso.ai, 2024) underscores the adoption of modern machine learning analysis to automate the detection of anomalous events, thereby improving safety measures on construction sites, including the identification of hazardous situations such as individuals in close proximity to heavy construction machinery. However, these industry approaches primarily focus on identifying potential hazards and may not adequately assess and quantify dynamic risks.

From an academic standpoint, numerous tools and models have been proposed in previous studies to address collision issues on construction sites by aiding in the allocation of site space for activities under specific conditions. However, commonly used simulation techniques still present limitations, as it is impractical to pre-plan the movement paths of all individuals, materials, and machinery before construction commences (Kim et al., 2016). Therefore, there is a need for a real-time dynamic early warning system. Previous studies have endeavored to develop dynamic safety assessment systems and proximity alert systems with the assistance of computer vision. However, the primary objectives of these studies were centered solely on workers and equipment, and the experimental results did not account for the distinctions among various entity types (Hu et al., 2020). Notably, existing discussions predominantly center on proximity and congestion risks (Zhang et al., 2020).

Scant research has been conducted to address safety issues by considering the individual attributes of workers and mobile machinery at various moments. Existing studies primarily focus on the relationship between paired entities (Zhang et al., 2020), while little study has considered a comprehensive impact of environmental factors and all surrounding entities on construction site collision risk analysis. Different from congestion, spatial conflict is characterized by the degree of spatial conflict between entities' activities and it is more reflective of the interaction between entities. Hitherto, the spatial conflict has not been concertized in safety-based proximity collision studies (Zhang et al., 2020), and dynamic workspaces have not yet been considered. Even though discussions have been found in identifying risks and then issuing warnings, little attempt has been made to explore the impact of risks on subsequent activities.

This paper aims to propose an integrated vision-based dynamic collision risk assessment framework with the consideration of the differences between workers and machines in terms of physical and work attributes and try to differentiate the assessment mechanisms. The concept of spatial conflict is introduced in adjacent collisions and environmental factors are included in the framework to achieve dynamic real-time collision risk assessment for multiple entity types with graded early warning.

3. Research methods

The research roadmap involves four major steps, as shown in Figure 1. First, data are collected via image, video, expert knowledge, and project schedule. Second, a comprehensive assessment model is established. This is followed by the development of the dynamic risk assessment framework. Last, an experimental study is conducted to validate the integrated vision-based dynamic collision risk assessment framework. These steps are detailed in the following sections.

For data collection, the computer vision data are collected via images and videos. The video data were processed in the form of video frames with continuous screenshots every 0.1 seconds for image processing as the base data for case validation. Expert knowledge was obtained using a semi-structured questionnaire, which used an intensive sampling method for expert data collection. Based on previous studies (Chen et al., 2019; Guo et al., 2018; Yu et al., 2017), we took 10 sample sizes. Ten questionnaires were distributed, and ten valid questionnaires were returned. The experts involved had an average of five years of experience in construction and 3 years of experience in building safety management. Around 80% of the respondents had experience in large construction projects. Experts need to determine the weights of the four factors in Figure 1 based on the hierarchical analysis (AHP), followed by the delineation of the spatial conflict degree and the environmental impact index within the given interval. Of note, we only adopted the data (reliability coefficient greater than or equal to 0.7) that had satisfied consistency. The project schedule was derived from the construction site of a real construction project in Chongqing, China.

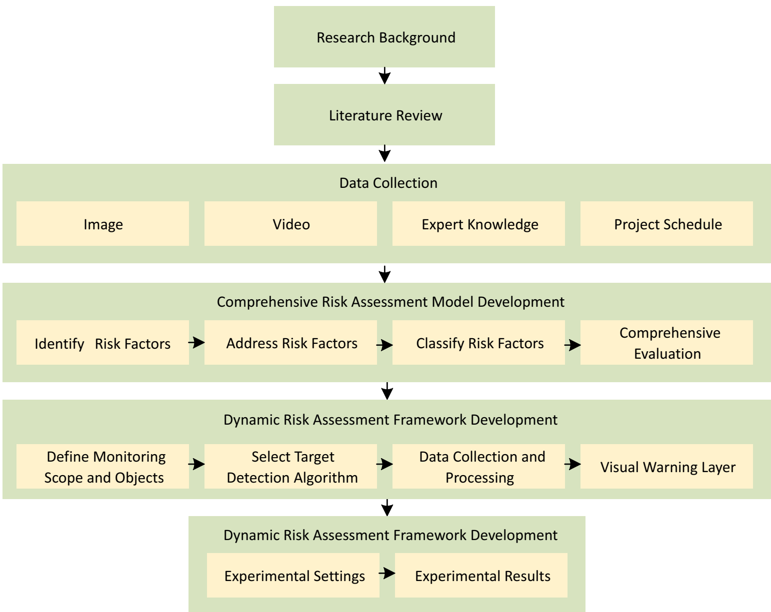

Figure 1. Research roadmap

4. Comprehensive risk assessment model

Fuzzy Comprehensive Evaluation (FCE) is a method based on fuzzy mathematics that is capable of quantifying and analyzing data that are not easily defined precisely (Liu et al., 2018). The application of the FCE method to develop risk assessment models enables quantitative analysis of fuzzy and uncertainty factors. By describing fuzzy and uncertain decision information, fuzzy sets are introduced. The fuzzy set is characterized by the membership function, which transforms qualitative evaluation into quantitative evaluation according to the fuzzy mathematical theory of membership degree (Holecek & Talašová, 2010). The comprehensive risk assessment model developed in this study involves three steps: (1) identify risk factors, (2) address risk factors, and (3) comprehensive evaluation.

4.1. Identify risk factors

The hazards associated with collision risks were first summarized through a literature review. Firstly, related literature was obtained from the Web of Science (WoS) Collection by searching the keywords "construction safety", "safety prediction", and "safety factors" and then classified into the following four areas: worker-related, machine-related, environment-related, and management-related risks.

Secondly, considering the scope of this study, the selection of risk factors needs to meet the following requirements: (1) visual monitorability at the construction site; (2) whether to increase the possibility of an accident; (3) related to a collision accident. Table 1 shows the screening results for risk factors.

Finally, according to Table 1, based on the screening and summary analysis, four risk factors meet the aforementioned three requirements: proximity, environmental impact index, congestion, and spatial conflict. The four selected factors meet the requirements and have been verified to have significant impacts on safety management.

4.2. Address risk factors

The fuzzy processing of risk factors includes two steps: (1) define risk factors and select appropriate methods to quantify the factors; (2) grade the factors and generate fuzzy sets to serve as basic data for the construction of the membership function.

Table 1. Screening table for risk factors

Areas	Risk factors	Visual monitorability at the construction site	Whether to increase the possibility of an accident	Related to a collision accident
Worker-related	Worker quality (work experience, safety awareness, etc.) (Nguyen et al., 2016; Sawicki & Szóstak, 2020)		Yes	
	Worker posture (Chen et al., 2019; Guo et al., 2018; Yu et al., 2017)	Yes	Yes	
	Proximity to hazard sources (Hu et al., 2020; Son et al., 2019; Zhang et al., 2017)	Yes	Yes	Yes
	Working height (Hallowell et al., 2017; Nguyen et al., 2016)	Yes	Yes	
	Personal protective equipment (PPE) (Alruqi & Hallowell, 2019; Fang et al., 2018)	Yes	Yes	
	Improper technical operation (Poh et al., 2018)		Yes	
Machine-related	Control of construction equipment and vehicles (Chen et al., 2020; Kanan et al., 2018; H. Kim et al., 2016)		Yes	Yes
	Condition of construction facilities (Zhu et al., 2016)		Yes	
Environment-related	Congestion of the workspace (Chen et al., 2019; Gheisari & Esmaeili, 2019; Kim et al., 2016; Ning et al., 2018)	Yes	Yes	Yes
	Construction environment (geohydrology, underground pipelines, etc.) (Esmaeili et al., 2015; Zhi-Qiang & Ya-mei, 2016)	Yes	Yes	
	Space conflict (Dashti et al., 2021; Zhang et al., 2015)	Yes	Yes	Yes
	Environmental impact index (Noise, dust, etc.) (Esmaeili et al., 2015; Zhi-Qiang & Ya-mei, 2016)	Yes (source of noise)	Yes	Yes
	Construction site safety facilities layout (Nguyen et al., 2016)	Yes	Yes	
Management-related	The proportion of workers without safety training (Alizadeh et al., 2015; Sawicki & Szóstak, 2020)		Yes	Yes
	Management structure and staffing (Guo & Yiu, 2016; Salas & Hallowell, 2016)		Yes	
	Construction Technology (Sawicki & Szóstak, 2020; Zhi-Qiang & Ya-mei, 2016)		Yes	

4.2.1. Define risk factors

(1) Proximity

Proximity embodies the distance relationship with the source of danger, which is defined as the smallest distance between machines or between machines and workers (Kim et al., 2016). In this research, proximity is defined as the minimum distance between machines or between machines and people, represented by $Mind_i$. Based on this, the diagram of proximity is shown in Figure 2a, where the proximity of the worker is $d_3\left(d_3 < d_2 < d_1\right)$ and the proximity of machinery is $d_1'\left(d_1' < d_2'\right)$. The video from the surveillance camera will be captured. The camera's frame rate will be 30 fps (i.e., frames per second) For real-time monitoring, the video from the surveillance camera will be captured. The camera's frame rate will be 30 fps (i.e., frames per second), which means that d_1, d_2, d_3, d_1' and d_2' are dynamically changing. The speed and direction of equipment can be reflected and calculated in terms of dynamically changing d_1, d_2, d_3, d_1' and d_2'.

(2) Environmental impact index

Frequently, risk variables include the environmental impact index (Esmaeili et al., 2015; Zhi-Qiang & Ya-mei, 2016). Among environmental impact indices, noise is not only as dangerous as any other hazardous conditions, but also more difficult to control (Lee et al., 2019). In addition, the impact of noise needs to be quantified. Moreover, adhering to the previously mentioned criteria for the selection of risk factors in this study, it is imperative for these factors to be subject to visual monitoring at the construction site. Given that the source of noise can be visually tracked, this study exclusively considers noise as the designated environmental index. Subsequently, noise is posited as the solitary environmental impact index, serving to characterize the influence of environmental risk sources on accident occurrences in this study. It is assumed that construction machinery is the only source of noise, and the noise will decrease along a straight line of divergence, regardless of the diffraction and reflection of noise. Therefore, the environmental impact index reflects the extent to which the noise hazard interferes with workers' awareness of the risk, characterized by the superimposed effect of the noise value that the average noise value of the mobile machinery decreases as the distance to the targeted entity (Lee et al., 2019).

For workers, the environmental impact index is the noise superimposed value of the average value of all mechanical noise according to the decreasing linear distance from the worker, denoted by $SV\left(dB_i\right)$. For machinery, it is the noise superimposed value of the average noise of all machinery (including research machinery) according to the decreasing distance from the targeted machinery, expressed by $SVdB_i'$.

The environmental impact index of the targeted entity can be obtained by Eqns (1)–(2) (Ning et al., 2018). In formula (1), $SV(dB)$ is the environmental impact index of the entity understudy, and i is the number of noise sources, where i =1, 2, ..., n; Le_i (dB) is the noise value of noise source decreasing by the distance to the entity. In Eqn (2), $Y(dB)$ is the attenuation value of noise, and d_i is the linear distance of the center of mass between the noise source and the targeted entity. As shown in Figure 2b, the environmental impact index of workers is $SV\left(dB_{10}, dB_{20}, dB_{30}\right)$, and the environmental impact index of machine A is $SV\left(dB_{21}', dB_{31}', dB_1\right)$.

$$SV = 10\log_{10}\sum_{i=1}^{n} 10^{0.1\times Le_i}; \quad (1)$$

$$Y = 5.548\ln\left(d_i\right) - 1.042. \quad (2)$$

(3) Congestion

Since lifting, hoisting, and moving materials are common activities on construction sites, workers are at risk of being struck by these moving objects in a congested work environment (Gheisari & Esmaeili, 2019). Spatial congestion has been defined in the literature (Kim et al., 2016; Zhang et al., 2020), and this paper defines congestion as the number of entities (workers versus machinery) (Kim et al., 2016).

(4) Space conflict

Space conflict reflects whether the spatial requirements of workers and machinery can be met, and represents the severity of the spatial conflict. Workers need to work in an environment with no obstacles within an average distance of 3.5 m to ensure their work safely and efficiently (Zhang et al., 2015). Hence, the space requirement of workers is defined as the physical space occupied by workers themselves extending 3.5 m outwards, which includes the working space of the workers and also includes the buffer space to ensure the productivity and safety of the workers.

Referring to the calculation of mechanical hazardous space occupancy, the average boundary extension width of 9.6 m occupied by the dangerous space of the three-common machinery (dump trucks, loaders, and excavators) is selected as the reference for the space occupation of construction machinery (Wang et al., 2019). Regarding the space conflict, Akinci et al. (2002) used the ratio of the space conflict volume to the required space volume as the standard of conflict severity. To facilitate data acquisition by computer vision, this study defines the space conflict as the ratio of the conflicting space area divided by the required space area, that is $C_i = \frac{S_C}{S_R}$, whereby C_i represents the space conflict degree of a monitoring object, and S_C indicates the conflict area between the monitoring object and the surrounding monitoring entities, S_R represents the working space required by the monitored object itself. As shown in Figure 2c, the space conflict of worker W_1 is $C_{W_1} = \frac{S_C^5}{S_R^{W_1}}$, whereby $S_R^{W_1}$ represents the space demand of worker W_1, and S_C^5 indicates space conflicts. Similarly, the space conflict of construction machinery B is calculated by: $C_B = \frac{\left(S_C^1 + S_C^3 + S_C^4\right)}{S_R^B}$.

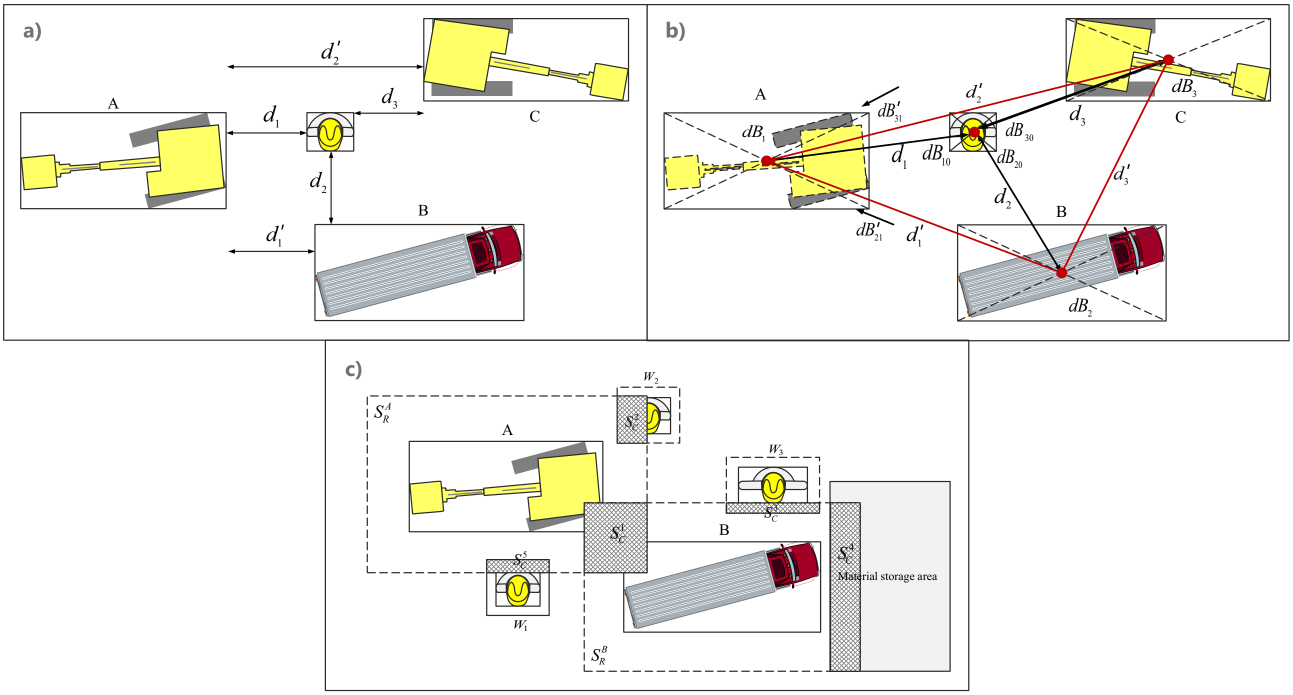

Figure 2. Schematic definition of proximity, environmental impact index and congestion

4.2.2. Classify risk factors

(1) Proximity

The level of proximity is described as close, normal, and farther, corresponding to high, medium, and low risk. Kim et al. (2016) mentioned that the risk classification of the proximity between human beings and moving machinery was obtained. Since heavy machinery generally occupies a large area, based on the proximity of people and mobile machinery, consideration was given to adding the length of a GENLYON C500 dump truck (about 10 meters) as the basis for grading the proximity of mobile machinery, as shown in Table 2.

(2) Environmental impact index

The environmental impact index conveys the risk of noise from machinery, described as large, medium, and small, corresponding to high, medium, and low risk. Due to the scarcity of criteria for grading the environmental impact indices, a semi-structured questionnaire was used to gain knowledge, and the boundaries of the three levels of classification were determined, as presented in Table 2.

(3) Congestion

The level of congestion is described as scattered, normal, and dense, corresponding to high, medium, and low risk. according to Kim et al. (2016), the risk classification of the congestion among workers and moving machinery was obtained as shown in Table 2.

(4) Space conflict

The degree of space conflict reflects the ratio of space demand occupied, and the levels are expressed as severe, moderate, and mild, corresponding to high, medium, and low risk. Due to the lack of criteria for rating the severity of space conflict, a semi-structured questionnaire was used in this paper to gain knowledge, and three levels of division boundaries are obtained, as shown in Table 2.

Table 2. Level of proximity, environmental impact index, congestion, space conflict

Risk factors	Entity type	Risk level		
		Low	Medium	High
Proximity	workers	Farther (≥10.5 m)	Normal (4~14 m)	Close (0~6.5 m)
	mobile machines	Farther (≥20.5 m)	Normal (10~27 m)	Close (0~16.5 m)
Environmental impact index	workers	Small (<81 dB)	Medium (78~99 dB)	Large (>94d B)
	mobile machines	Small (<85 dB)	Medium (82~107 dB)	Large (>99 dB)
Congestion	workers	Scattered (0~4 entities)	Normal (2~7 entities)	Dense (>5 entities)
	mobile machines	Scattered (0~4 entities)	Normal (2~7 entities)	Dense (>5 entities)
Space conflict	workers	Mild (0~0.4)	Moderate (0.31~0.74)	Severe (≥0.57)
	mobile machines	Mild (0~0.35)	Moderate (0.25~0.70)	Severe (≥0.49)

After the fuzzy expression of the risk factors, the factor set for fuzzy comprehensive evaluation can be expressed as U = (Proximity, Environmental Impact Index, Congestion, Space Conflict), $U_p = (f_1, f_2, f_3, f_4)$ for the factor set for workers, and $U_m = (f_1', f_2', f_3', f_4')$ for the factor set of machinery. The weight set indicates the relative significance of the four factors. Table 3 displays the weight ratio of the risk factors obtained from expert interview statistics.

The weight set for workers are
$W_p = (w_{f_1}, w_{f_2}, w_{f_3}, w_{f_4}) = (0.29, 0.15, 0.16, 0.40)$,
$W_m = (w_{f_1}', w_{f_2}', w_{f_3}', w_{f_4}') = (0.29, 0.20, 0.12, 0.39)$
for machinery, and the evaluation set as
$V = (\text{High, Medium, Low})$, recorded as $V = (v_1, v_2, v_3)$.

Table 3. Level of proximity, environmental impact index, congestion, space conflict

Proximity	Environmental Impact Index	Congestion	Space Conflict
The weight ratio of risk factors for workers			
29%	15%	16%	40%
The weight ratio of risk factors for mobile machinery			
29%	20%	12%	39%

4.3. Comprehensive evaluation

According to the classification description of risk factors and the results of obtaining expert knowledge in Section 4.2.2., the membership function proposed by Mardani et al. (2015) is performed in Python 3.7, as shown in Figures 3–6.

As such, the fuzzy relationship matrix R (see Eqn (3)) from factor set U to comment set V is obtained:

$$R = \begin{bmatrix} f_1(w_1) & f_1(w_2) & f_1(w_3) \\ f_2(w_1) & f_2(w_2) & f_2(w_3) \\ f_3(w_1) & f_3(w_2) & f_3(w_3) \\ f_4(w_1) & f_4(w_2) & f_4(w_3) \end{bmatrix}, \tag{3}$$

whereby, $f_1(w_i)$ represents the membership of the proximity of the monitoring entity, and $f_2(w_i)$ represents the membership of the congestion of the monitoring entity. The fuzzy comprehensive evaluation is obtained through the matrix synthesis operation:

$$B = W \times R. \tag{4}$$

B is a fuzzy distribution, showing the membership of each risk level of a monitored entity, namely the feature vector of the fuzzy level. A processing method of magnifying the fuzzy number by a certain multiple is adopted to calculate the final evaluation base (Liu & Tian, 2019). This paper constructs three comprehensive risk levels $(\text{High, Medium, and Low})$ and five Pre-Warning levels (I, II, III, IV, V). the median of Pre-Warning Level I, III, and V is selected. We used Eqn (5) to calculate the final risk value (V) using the fuzzy distribution method:

$$V = \text{median}_{V_1} \times B_1 + \text{median}_{V_3} \times B_2 + \text{median}_{V_5} \times B_3, \tag{5}$$

whereby V_1 represents the risk range for Pre-Warning Level I, V_3 denotes the risk range for Pre-Warning Level III, V_5 indicates the risk range for Pre-Warning Level V,

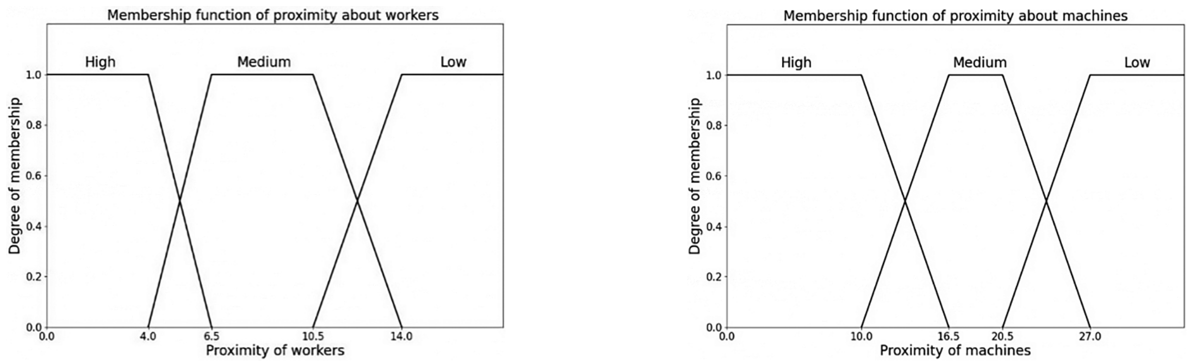

Figure 3. Membership function of proximity

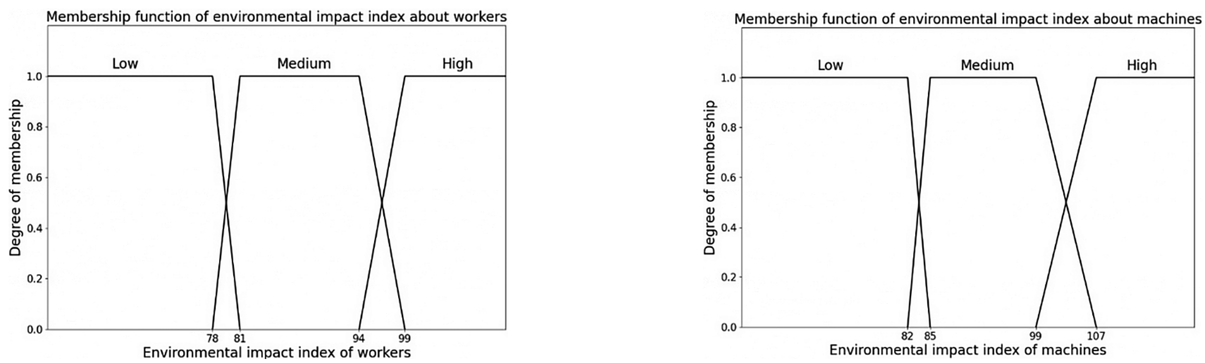

Figure 4. Membership function of environmental impact index

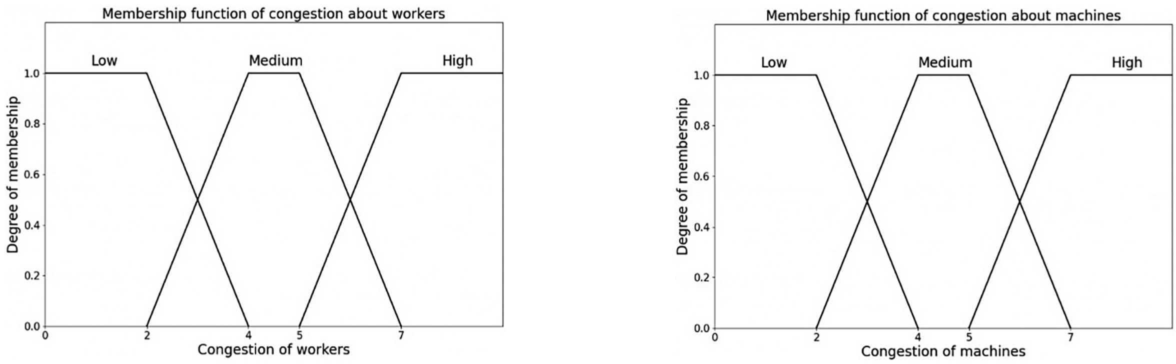

Figure 5. Membership function of congestion

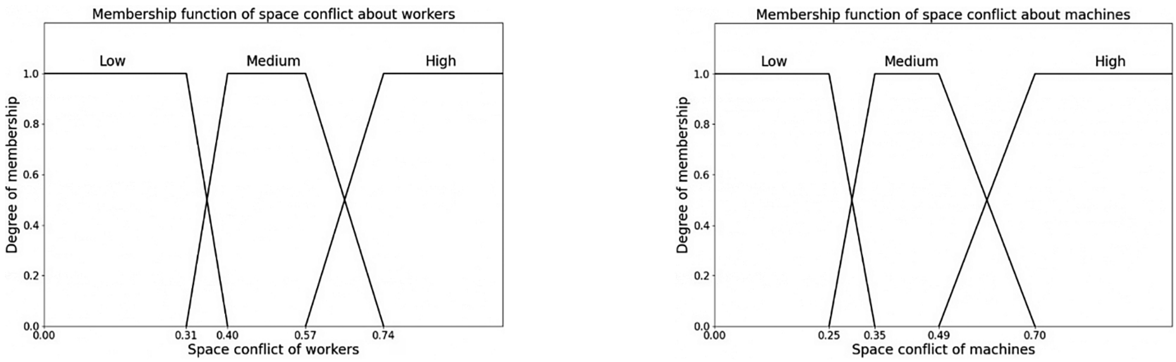

Figure 6. Membership function of space conflict

and V represents the ultimate risk value which is used as input to the hierarchical warning mechanism and vision expression.

5. Dynamic risk assessment framework

In this study, computer vision is used to solve the problem of dynamic collection and automated updating of safety information at construction sites. It is combined with the traditional dynamic risk analysis methods for risk modeling. The dynamic risk assessment framework includes four aspects: (1) define monitoring scope and objects; (2) select target detection algorithms; (3) data collection and processing; and (4) visual warning layer. These four aspects are shown in Figure 7.

5.1. Define monitoring scope and objects

Machinery in working condition is usually surrounded by an area of activity where workers are exposed to the risk of collision, and the area is referred to as a hazardous area. Take 30m as the range of the hazardous area centered on mobile machinery in this paper (Wang et al., 2019).

The key monitoring objects are workers and moving machinery. Besides, materials are also considered, because they will occupy the construction site and affect the interaction between workers and the surrounding environment, but the stacking of materials is relatively static.

5.2. Select target detection algorithm

The computer vision technology is indeed a deep learning algorithm. Among the deep learning algorithms, R-CNN (Lou & Cui, 2007), Faster R-CNN (Sun et al., 2018), YOLO (Li et al., 2018), and SSD (Shu et al., 2019) are the most widely used in object detection.

All existing object detection algorithms can achieve dynamic monitoring, but due to different performances and different speeds of processing pictures, there is a huge difference in the time interval. The safety of workers on the construction site must consider the value of time, and the image processing speed is relatively high. In addition, the monitoring range is relatively large, and the probability of small targets is relatively high. Since YOLOv3 excels in detection speed and small target capture, it was decided to use the YOLOv3 in object detection.

5.3. Data collection and processing

Training the target detection model requires a large amount of image data, there are high requirements for hardware equipment, and ordinary laptops usually cannot complete high-quality model training. Therefore, an open source YOLOv3 model (Redmon & Farhadi, 2018) trained on the public COCO dataset (Veit et al., 2016) was used. Some materials for the test were collected using the MAVIC MINI drone, which is only 249 grams and with a maximum rotation angle of 83 degrees and 12 megapixels. It can realize GPS hovering stably, which is powerful and convenient. Drones have been used in the field of construction for quality inspection, safety inspection, building inspection and measurement, progress monitoring, and on-site monitoring (Gheisari & Esmaeili, 2019). It is technically feasible to conduct safety monitoring on construction sites. In addition, mobile phones and cameras were also used to collect the data. In this experiment, the model configuration details are shown in Table 4.

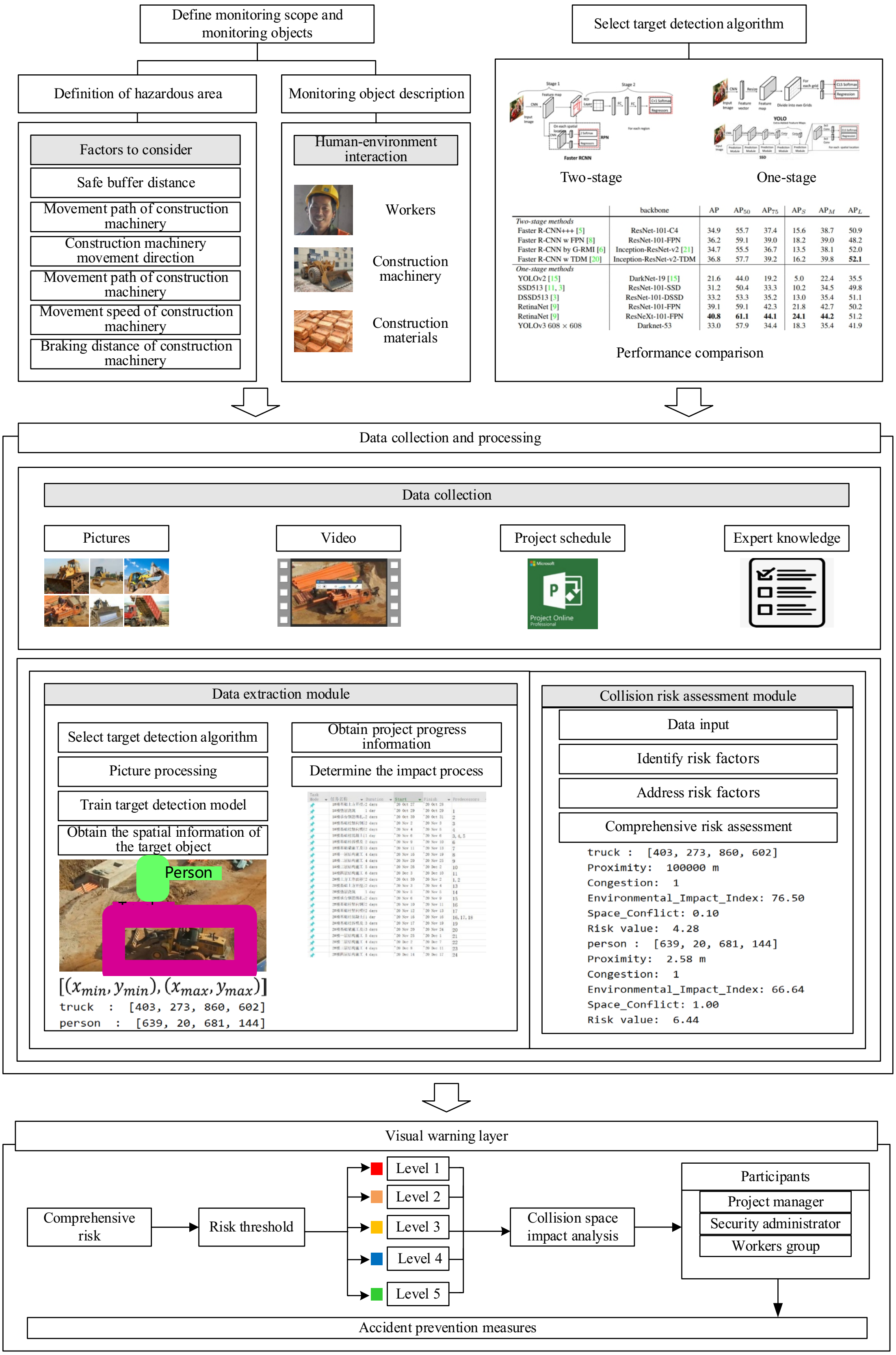

Figure 7. Dynamic risk assessment framework based on computer vision

Table 4. Model configuration details

Model Configuration	Details
Running Environment	TeslaGPU_P4_8G video memory single card_12 core CPU_40G memory
The version of the deep learning framework	Paddle Paddle 1.7.1
Programming language version	Python3.7
Deployment method	Public Cloud API
Object detection	YOLOv3

The object detection algorithm was performed using YOLOv3. The recorded dataset was processed by this algorithm. A total of 32 videos (each video is around 40 seconds) were acquired for training in this dataset. The frame rate of the camera used was 30 fps (i.e. frames per second). For the object detection algorithm, the extracted mAP is 91.3%. For the dataset, a simple cross-validation method is used in this paper to classify the training and test datasets to avoid overfitting. The train_test_split function of sklearn was used to divide the dataset into two subsets: training and test dataset. The ratio is set to 7:3. The learning rate corresponding to a single batch_size is set and will increase linearly according to the batch_size and the number of Graphic Processing Unit.

The data processing layer include two modules. The visual extraction module is based on the open-source python library named ImageAI, MySQL Database and the YOLOv3 target detection model. It enables to identify and obtain spatial information about the monitored objects, and the location information is the basis of safety data. Take the upper left corner of the image as the coordinate origin, adding to the right as the *x* axis, and adding downward as the *y* axis. The position coordinates of each monitoring object can be expressed as $\left(x_{\min}, y_{\min}, x_{\max}, y_{\max}\right)$, whereby $\left(x_{\min}, y_{\min}\right)$ are the coordinates of the upper left corner of the monitoring object, $\left(x_{\max}, y_{\max}\right)$ is the coordinates of the lower right corner of the monitoring object. Meanwhile, the project schedule was imported into the MySQL database and manually matched to the construction sections that are currently at risk, the successor of the current construction activity will be affected. The collision risk assessment module uses the attributes and spatial information of the monitored object in the proposed fuzzy comprehensive risk assessment model. The module outputs the attributes of the monitored object, spatial information, and the values of four risk factors, i.e., proximity, environmental impact index, congestion, and spatial conflict, and the comprehensive risk value.

5.4. Visual warning layer

The early warning module is important for implementing risk mitigation measures after a risk assessment and can be used as a basis for discovering potential hazards and guiding safety decision-making. In view of providing project managers with automated safety management tools for construction sites, the early warning mechanism in this paper is designed in the form of providing visual early warning information to rank risks in different colors. In addition, the design of the early warning mechanism can be determined according to the risk preference of the safety manager and the actual situation of the project. Table 5 provides a visual representation of the hierarchical alerting. The module is divided into five levels, namely I (extremely dangerous), II (strongly dangerous), III (moderately dangerous), IV (relatively dangerous), and V (slightly dangerous), referring to the warning signals of red, orange, yellow, blue, and green, respectively.

6. Experimental study

To operationalize the framework, a prototype is developed and serves as the practical implementation of the framework. The experimental study shows the feasibility of the developed prototype. The prototype in the experiment is carried out to (1) verify the usability of the selected target detection algorithm using data collected from real scenarios, and (2) assess whether the design of the assessment model is reasonable and effective based on the results of the comprehensive risk assessment.

6.1. Experimental settings

This study collected data from the experiment at the construction site of a construction project in Chongqing, China. Data was collected using mobile phones and drones. The entire experiment was conducted under Windows 10. Anaconda was installed to create a Python 3.7 environment, while TensorFlow-GPU1.15.0, keras2.3.1, CUDA10.0,

Table 5. Visualized warnings at different levels

Early warning level	Risk value	Early warning signal	Severity
Ⅰ	8~10	Red	Extremely Dangerous (ED)
Ⅱ	6~8	Orange	Strongly Dangerous (SD)
Ⅲ	4~6	Yellow	Moderately Dangerous (MD)
Ⅳ	2~4	Blue	Relatively Dangerous (RD)
Ⅴ	0~2	Green	Lightly Dangerous (LD)

and cuDNN7.6 were installed to complete the configuration of YOLOv3's operating environment. The algorithm was executed on a laptop configured with Intel(R) Core (TM) i5-10210U CPU @ 1.60GHz 2.11GHz and NVIDIA GeForce MX350. The specific experimental steps are as follows:

Step 1: Prepare the tools for picture and video data collection. There are four main requirements for the settings of camera equipment: first, the placement of the camera is critical to ensure that the desired construction elements are captured. The camera should be placed in a location that provides an unobstructed view of the construction site and the elements that need to be captured; Second, under the condition of satisfying the needs of obtaining required data, the camera settings should be adjusted to ensure that the captured images or videos are of high quality and that the desired construction elements are clearly visible; The third is to specify specifications for the camera equipment's durability and accuracy, i.e., that the image or videos must be sufficiently clear and have a resolution that satisfies the minimal requirements for video recognition, and the equipment should also be durable, able to resist bad weather such as sunshine, heavy rain, frost and cold on the construction site; Fourth, the data captured by the cameras should be processed in a timely manner to ensure that the desired construction elements are extracted accurately. This includes using computer vision algorithms to identify and extract the relevant elements from the images or videos captured by the camera. In this paper, MAVIC MINI drones and smartphones were mainly used. The higher the pixel of the acquisition tool, the better the recognition result. The drone aerial camera used in this paper had 12 megapixels and the smartphone had 40 megapixels.

Step 2: Select the scene where workers and construction machinery work together as the monitoring object. The monitoring range was estimated with the construction machinery as the core, starting from the outer boundary in the overhead view of the farthest reachable operation mode of the construction machinery and extending outward for 30 m, leaving clear markings at the boundary. The boundary was marked in the same way along the movement path of the construction machinery. To cover the farthest boundary mark on the screen, when fixing the visual equipment. Defining the monitoring scope helps identify workers who may appear in the hazardous area, discover potential collision risks, and reduce accidents. Once the monitoring range has been determined, the height of the visual equipment can be essentially fixed. Finally, we can adjust the visual device to select a better viewing angle and monitor position.

Step 3: Calculate the ratio between the actual distance and the pixels according to the fixed visual equipment. A 1 m long ruler was used as a marker in the monitoring screen. The placement position must be aligned with one side of the monitoring screen. The actual distance represented by each pixel was calculated by conversion. This ratio needs to be transformed into the risk assessment model to ensure the accuracy of the calculation.

Step 4: The distance pixel ratio was adjusted according to the filming data, and the video and picture data of the construction site were processed with YOLOv3 to output detection results. Using the pymysql library, the construction schedule is imported into the database, and subsequent construction activities that may be affected were inferred based on the risk of current construction activities. Collision risk was assessed only for workers and construction machinery, and visual warning signals were expressed in grades with colors and scores.

6.2. Experimental results

As part of the framework implementation, the prototype comprises the following components: a visual extraction module and a collision risk assessment module.

6.2.1. Visual extraction module

The performance of the visual processing module could affect the operation of the entire framework. As shown in Figure 8, scene 1a–1b is obtained by adjusting the parameter named minimum percentage probability of the target detection model from 25% to 40% for the same image. In scene 1b, the presence of machinery was not detected, and the safety level of workers working in the vicinity would be far lower than the actual situation.

Scene 1: The picture size is 1300 × 672 pixels, the range of picture capture is 37 m × 17 m, and each pixel is approximately 0.02 m. The experimental results are shown in Figure 9.

The visual extraction module classifies workers and construction machinery. In this paper, this module was simplified and the performance of the training set has not yet been evaluated. The validation of the vision module focused on the object detection performance of the new test data; the detection results of 600 images were analyzed by taking continuous screenshots of the videos collected at the construction site every 0.1 s. The classification accuracy almost reached 100%, with errors only showing up when the view was heavily obscured or the image was too blurred. Therefore, the validation of vision systems primarily focused on object detection performance rather than classification.

Precision and recall are widely used metrics to test detection performance. Precision is defined as the amount of true and positive detected objects (TP) divided by TP plus the number of false and positive detected objects (FP), recall is expressed as the sum of the TP divided by the TP and the number of false and negative objects (FN). Since high precision can guarantee monitoring performance while being less affected by low recall rate and the level of security of the monitored objects is related to the surrounding entities, the object's loss rate FN is a very important indicator of the accuracy of the test data is 94.58% and the loss rate is 11.6%.

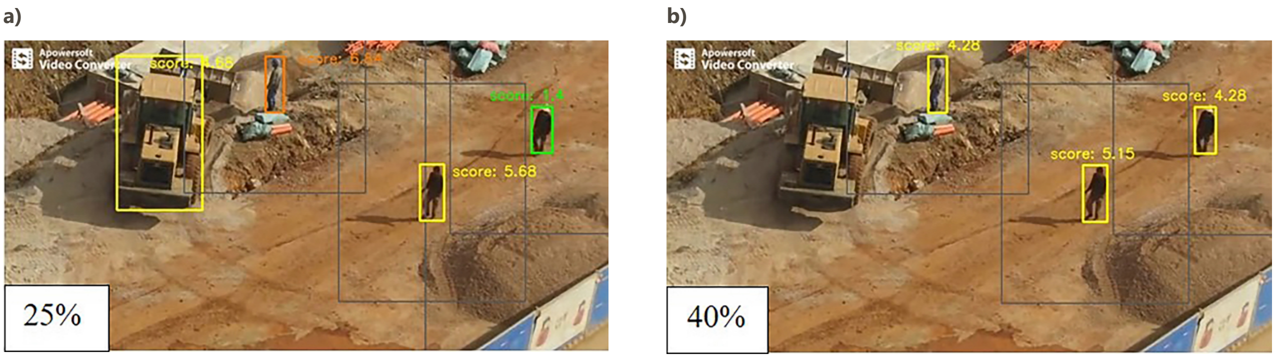

Figure 8. Experimental results of scene 1

Figure 9. Experimental results of scene 2 at the moment $t_1 - t_4$

6.2.2. Collision risk assessment module

Figure 9 shows the experimental results at different moments in a certain scenario. In Scene 2, the proximity of worker 2 at moment t_1 is 12.42 m, the environmental impact index is 62.24 dB, the congestion is 2, the spatial conflict is 0.08, and the risk level is V. In the process of approaching the construction machinery, the proximity of worker 2 gradually decreases, the environmental impact index increases as the distance decreases, the degree of congestion changes as worker 3 enters the scene, the spatial conflict gradually increases, and the risk level increases from V–IV–III. In addition, at moment t_4, the proximity of Worker 2 was 6.84 m, the environmental impact index was 64.13 dB, the degree of congestion was 3, the spatial conflict was 1.00, and the risk level was III.

Furthermore, the congestion of the loader at moment t_2 is 2 and the spatial conflict is 0.16. At moment t_4, due to more spatial conflicts with workers and the appearance of worker 3, the congestion becomes 3 and the spatial conflict turns to 0.23. When the index remains unchanged, the risk value of 4.28 changes to 4.68. Table 6 shows the specific security risk assessment data. Judging from the evaluation results, the fuzzy comprehensive evaluation model constructed in this paper can accurately assign security levels to monitoring entities.

Scene 2: The picture size is 1300 × 672 pixels, the range of picture capture is 37 m × 17 m, and each pixel is approximately 0.02 m. The experimental results are shown in Figure 9 and the monitoring data are shown in Table 6.

7. Discussion

The proposed dynamic collision risk assessment framework in this paper considers the environmental factors and the interaction between multiple entities, introduces the concept of spatial conflict degree, and outputs four risk factors and comprehensive risk values, allowing for a nuanced understanding of collision threats.

In Scenes 1 and 2, the spatial interactions between construction machinery and nearby workers were demonstrated, with proximity and congestion factors calculated respectively, followed by the output of corresponding risk values. The reliability of the risk assessment results depends, on one hand, on the rationality of the membership functions constructed within the comprehensive risk assessment model, and on the other hand, is closely tied to the performance of the object detection algorithm.

As previously noted, the comprehensive evaluation model is influenced by the design of membership functions. In constructing these functions, expert questionnaires were employed to gather knowledge and simulate human reasoning processes, thereby introducing an inevitable degree of subjectivity. For instance, in this study, a distance of less than 4 meters between a worker and moving machinery is regarded as a definite high-risk zone due to noise exposure. However, some safety managers argue that this threshold should not fall below 5 meters. At present, there is no unified standard for threshold design or classification of risk levels. As safety risk perceptions vary across construction sites, the anticipated levels of risk associated with monitored entities may differ. Consequently, risk management decisions largely depend on the judgment of safety managers. According to differing risk preferences, the internal parameters of the evaluation module must be adjusted to achieve the desired level of risk control.

The classification and monitoring accuracy within the vision module exceeds 95%, underscoring its high reliability in detecting potential hazards. Furthermore, the object detection algorithm operates at a speed of approximately 10 frames per second, proving the technology's feasibility for real-time safety monitoring. In actual construction settings, it is essential to schedule the entry of personnel, materials, and machinery in a logical order, while also optimizing site layouts and managing space effectively. These measures significantly enhance onsite safety management by reducing potential conflicts and facilitating more efficient supervision.

Table 6. Monitoring data for scene 2

Moment	Monitoring entity	Spatial information	Proximity (m)	Environmental Impact Index (dB)	Congestion	Space Conflict	Risk Value	Risk Level
t_1	Worker 1	[661, 33, 697, 152]	1.44	67.80	2	1.00	6.44	II
	Worker 2	[1210, 193, 1260, 309]	12.42	62.24	2	0.08	1.52	V
	truck	[305, 108, 589, 424]	∞	76.50	2	0.12	4.28	III
t_2	Worker 1	[686, 5, 722, 128]	2.88	67.57	2	1.00	6.44	II
	Worker 2	[1082, 223, 1137, 337]	10.80	63.09	2	0.28	2.06	IV
	truck	[332, 57, 542, 357]	∞	76.50	2	0.16	4.28	III
t_3	Worker 1	[627, 13, 664, 137]	1.94	68.39	3	1.00	6.84	II
	Worker 2	[984, 243, 1041, 355]	9.08	63.76	3	0.82	5.68	III
	Worker 3	[1215, 108, 1264, 205]	13.70	62.02	3	0.31	1.50	V
	truck	[311, 36, 530, 375]	∞	76.50	3	0.18	4.68	III
t_2	Worker 1	[552, 18, 590, 143]	1.08	68.63	3	1.00	6.84	II
	Worker 2	[840, 268, 902, 392]	6.84	64.13	3	1.00	5.68	III
	Worker 3	[1097, 139, 1147, 255]	11.98	62.25	3	0.42	3.63	IV
	truck	[174, 24, 498, 302]	∞	76.50	3	0.23	4.68	III

This study establishes an integrated risk assessment mechanism that synthesizes proximity and congestion factors, digitizing the evaluation process within the model. The internal parameters can be readily modified to accommodate various construction scenarios and risk preferences, demonstrating the model's adaptability and scalability. This renders it a valuable reference for more complex safety management tasks. The experimental risk assessment results align closely with expectations, thereby affirming the effectiveness of the proposed evaluation model.

To compare this study with existing literature, the integration of spatial conflict degree in this study quantifies both proximity and congestion, which marks a substantial improvement over previous models that typically treat these factors independently. For instance, Teizer and Cheng (2015) proposed a proximity-based hazard indicator focusing on near-miss incidents but did not consider spatial congestion as a quantifiable risk factor. The unified spatial risk quantification in this study stands out by blending these concepts into a single model that is both dynamic and sensitive to site variability.

In addition, this study goes beyond simple detection by integrating the results into a fuzzy logic-based evaluation engine, yielding interpretable risk scores. This makes the proposed system not only real-time but also operationally actionable, addressing a gap noted by Gan et al. (2024), who pointed out that many models fail to connect detection outputs with strategic site safety controls.

The proposed model's use of customizable membership functions derived from expert questionnaires introduces both a strength and a discussion point. While Kim et al. (2016) and Shin and Kim (2022) used fuzzy inference systems with fixed thresholds, they lacked adaptability. The proposed model's capacity to reconfigure risk perception based on project-specific expert input enhances practical applicability, particularly in environments where safety culture and spatial norms differ. This adaptability is critical, as safety perception is not universal – what one site considers a safe distance may be unacceptable elsewhere.

Finally, while previous research like that by Zhang et al. (2025) addressed collision risks using spatial modeling and trajectory prediction, their frameworks lacked the proposed framework's degree of customization and interoperability. The proposed framework has the ability to adjust internal parameters based on different construction site profiles or safety management styles gives the framework superior transferability across contexts, a feature rarely addressed in prior work.

8. Conclusions

The construction site environment is always noisy and crowded, and a large number of workers are exposed to the risk of collision with machinery. To solve this problem and reduce the possibility of accidents, a Fuzzy Comprehensive Evaluation (FCE) model based on computer vision is proposed in this study. It identifies four important factors that affect worker safety, namely proximity, environmental impact index, congestion, and spatial conflict. The integrated vision-based dynamic collision risk assessment framework can dynamically monitor workers and machinery by computer vision to obtain spatial information on monitored objects. Then, the collected spatial information was used as the basic data for safety risk assessment. Next, a fuzzy comprehensive evaluation model was utilized to integrate the four factors. Finally, we visualized the evaluation results to assist construction site management and verified the technical feasibility and model validity by undertaking an experimental study.

The findings of this study contribute to the existing body of knowledge of computer vision by (1) establishing a dynamic collision risk assessment framework based on the difference of entity attributes (workers and mobile machinery), which is scalable and modifiable; (2) proposing a spatial conflict and environmental impact index in collision studies and establishing a more comprehensive multi-index risk assessment model. This model considers the interaction between workers, machinery, and the environment; (3) assesses the influence between entities (workers and machinery) from distance interaction to area interaction; and (4) extrapolates the subsequent construction activities that may be affected based on the risk of current construction activities. Also, this study extends fuzzy logic principles to a dynamic construction environment by incorporating spatial conflict degrees and real-time object detection metrics, broadening the scope of fuzzy evaluation methodologies. In addition, this study employs expert elicitation for membership function design and combining it with advanced computer vision techniques bridges gaps between theoretical safety assessment models and practical construction-site realities. By demonstrating how dynamic interactions among workers, materials, and machinery can be systematically quantified and monitored, the study offers a new lens for understanding and managing collision risks, thus enriching theoretical discussions on proactive accident prevention.

Regulatory bodies and industry associations should work toward establishing standard safe-distance thresholds and risk-level classifications that reflect local contexts, thereby minimizing ambiguities in on-site safety management. At the same time, governments and professional organizations can encourage the adoption of risk assessment methods by offering research funding, tax incentives, or technology transfer programs, thereby fostering a broader culture of safety within the construction industry. Industry-wide certification and specialized training programs would further ensure that safety managers are well-versed in fuzzy logic principles and vision-based detection systems, empowering them to interpret risk analyses accurately and adjust parameters as needed.

For future studies, with the continuous optimization of deep learning algorithms, target detection can achieve faster speed, higher precision, and accuracy. With the improvement of the data set, future studies can focus on obtaining more details on-site management, such as (1) identifying the worker's identity information; (2) tracking the worker's work path; and (3) identifying different types of vehicles as well as transport material information.

Acknowledgements

This research was supported in part by the National Science and Technology Council (NSTC) of Taiwan under Grant NSTC 113-2221-E-992 -034 and 114-2221-E-992 -004 -MY2.

Author contributions

Xingrong Gao: Software, Validation, Formal analysis, Investigation, Resources. Wenjing Kang: Data Curation, Writing – Original Draft, Visualization, Writing – Editing. Yunting Liu: Data Curation, Writing – Original Draft, Visualization. Qian Zhang: Writing – Review & Editing. Ting-Kwei Wang: Conceptualization, Methodology, Supervision, Funding acquisition.

References

Akinci, B., Fischen, M., Levitt, R., & Carlson, R. (2002). Formalization and automation of time-space conflict analysis. *Journal of Computing in Civil Engineering*, *16*(2), 124–134. https://doi.org/10.1061/(ASCE)0887-3801(2002)16:2(124)

Alizadeh, S. S., Mortazavi, S. B., & Mehdi Sepehri, M. (2015). Assessment of accident severity in the construction industry using the Bayesian theorem. *International Journal of Occupational Safety and Ergonomics*, *21*(4), 551–557. https://doi.org/10.1080/10803548.2015.1095546

Alruqi, W. M., & Hallowell, M. R. (2019). Critical success factors for construction safety: Review and meta-analysis of safety leading indicators. *Journal of Construction Engineering and Management*, *145*(3), Article 04019005. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001626

Ballesteros, M. J., Fernández, M. D., Quintana, S., Ballesteros, J. A., & González, I. (2010). Noise emission evolution on construction sites. Measurement for controlling and assessing its impact on the people and on the environment. *Building and Environment*, *45*(3), 711–717. https://doi.org/10.1016/j.buildenv.2009.08.011

Chen, H., Luo, X., Zheng, Z., & Ke, J. (2019). A proactive workers' safety risk evaluation framework based on position and posture data fusion. *Automation in Construction*, *98*, 275–288. https://doi.org/10.1016/j.autcon.2018.11.026

Chen, C., Zhu, Z., & Hammad, A. (2020). Automated excavators activity recognition and productivity analysis from construction site surveillance videos. *Automation in Construction*, *110*, Article 103045. https://doi.org/10.1016/j.autcon.2019.103045

Chi, S., & Caldas, C. H. (2012). Image-based safety assessment: Automated spatial safety risk identification of earthmoving and surface mining activities. *Journal of Construction Engineering and Management*, *138*(3), 341–351. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000438

Choi, B., Lee, H.-S., Park, M., Cho, Y. K., & Kim, H. (2014). Framework for work-space planning using four-dimensional BIM in construction projects. *Journal of Construction Engineering and Management*, *140*(9), Article 04014041. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000885

Dagan, D., & Isaac, S. (2015). Planning safe distances between workers on construction sites. *Automation in Construction*, *50*, 64–71. https://doi.org/10.1016/j.autcon.2014.12.008

Dashti, M. S., RezaZadeh, M., Khanzadi, M., & Taghaddos, H. (2021). Integrated BIM-based simulation for automated time-space conflict management in construction projects. *Automation in Construction*, *132*, Article 103957. https://doi.org/10.1016/j.autcon.2021.103957

Esmaeili, B., Hallowell, M. R., & Rajagopalan, B. (2015). Attribute-based safety risk assessment. II: Predicting safety outcomes using generalized linear models. *Journal of Construction Engineering and Management*, *141*(8), Article 04015022. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000981

Europe Construction Tech Review. (2020). *Top 10 safety and compliance solution companies-2020*. https://safety-and-compliance.constructiontechreview.com/vendors/top-safety-compliance-companies-2020.html

Fang, W., Ding, L., Luo, H., & Love, P. E. (2018). Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction*, *91*, 53–61. https://doi.org/10.1016/j.autcon.2018.02.018

Gan, W., Gu, K., Geng, J., Qiu, C., Yang, R., Wang, H., & Hu, X. (2024). A novel three-stage collision-risk pre-warning model for construction vehicles and workers. *Buildings*, *14*(8), Article 2324. https://doi.org/10.3390/buildings14082324

Gheisari, M., & Esmaeili, B. (2019). Applications and requirements of unmanned aerial systems (UASs) for construction safety. *Safety Science*, *118*, 230–240. https://doi.org/10.1016/j.ssci.2019.05.015

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

Guo, B. H., & Yiu, T. W. (2016). Developing leading indicators to monitor the safety conditions of construction projects. *Journal of Management in Engineering*, *32*(1), Article 04015016. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000376

Guo, H., Yu, Y., Ding, Q., & Skitmore, M. (2018). Image-and-skeleton-based parameterized approach to real-time identification of construction workers' unsafe behaviors. *Journal of Construction Engineering and Management*, *144*(6), Article 04018042. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001497

Hallowell, M. R., Alexander, D., & Gambatese, J. A. (2017). Energy-based safety risk assessment: Does magnitude and intensity of energy predict injury severity?. *Construction Management and Economics*, *35*(1–2), 64–77. https://doi.org/10.1080/01446193.2016.1274418

Han, S., & Lee, S. (2013). A vision-based motion capture and recognition framework for behavior-based safety management. *Automation in Construction*, *35*, 131–141. https://doi.org/10.1016/j.autcon.2013.05.001

Hinze, J., Huang, X., & Terry, L. (2005). The nature of struck-by accidents. *Journal of Construction Engineering and Management*, *131*(2), 262–268. https://doi.org/10.1061/(ASCE)0733-9364(2005)131:2(262)

Holecek, P., & Talašová, J. (2010). FuzzME: a new software for multiple-criteria fuzzy evaluation. *Acta Universitatis Matthiae Belii ser. Mathematics*, *16*, 35–51.

Hu, Q., Bai, Y., He, L., Cai, Q., Tang, S., Ma, G., Tan, J., & Liang, B. (2020). Intelligent framework for worker-machine safety assessment. *Journal of Construction Engineering and Management*, *146*(5), Article 04020045. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001801

Isaac, S., & Edrei, T. (2016). A statistical model for dynamic safety risk control on construction sites. *Automation in Construction*, *63*, 66–78. https://doi.org/10.1016/j.autcon.2015.12.006

Kanan, R., Elhassan, O., & Bensalem, R. (2018). An IoT-based autonomous system for workers' safety in construction sites with real-time alarming, monitoring, and positioning strategies. *Automation in Construction*, *88*, 73–86. https://doi.org/10.1016/j.autcon.2017.12.033

Kang, Y., Siddiqui, S., Suk, S. J., Chi, S., & Kim, C. (2017). Trends of fall accidents in the US construction industry. *Journal of Construction Engineering and Management*, *143*(8), Article 04017043. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001332

Kassem, M., Dawood, N., & Chavada, R. (2015). Construction workspace management within an Industry Foundation Class-Compliant 4D tool. *Automation in Construction*, *52*, 42–58. https://doi.org/10.1016/j.autcon.2015.02.008

Khakzad, N., Khan, F., & Paltrinieri, N. (2014). On the application of near accident data to risk analysis of major accidents. *Reliability Engineering* & *System Safety*, *126*, 116–125. https://doi.org/10.1016/j.ress.2014.01.015

Kim, H., Kim, K., & Kim, H. (2016). Vision-based object-centric safety assessment using fuzzy inference: Monitoring struck-by accidents with moving objects. *Journal of Computing in Civil Engineering*, *30*(4), Article 04015075. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000562

Kim, K., Kim, H., & Kim, H. (2017). Image-based construction hazard avoidance system using augmented reality in wearable device. *Automation in Construction*, *83*, 390–403. https://doi.org/10.1016/j.autcon.2017.06.014

Kim, D., Liu, M., Lee, S., & Kamat, V. R. (2019). Remote proximity monitoring between mobile construction resources using camera-mounted UAVs. *Automation in Construction*, *99*, 168–182. https://doi.org/10.1016/j.autcon.2018.12.014

Lee, S. C., Kim, J. H., & Hong, J. Y. (2019). Characterizing perceived aspects of adverse impact of noise on construction managers on construction sites. *Building and Environment*, *152*, 17–27. https://doi.org/10.1016/j.buildenv.2019.02.005

Li, J., Su, Z., Geng, J., & Yin, Y. (2018). Real-time detection of steel strip surface defects based on improved yolo detection network. *IFAC-PapersOnLine*, *51*(21), 76–81. https://doi.org/10.1016/j.ifacol.2018.09.412

Liu, H., & Tian, G. (2019). Building engineering safety risk assessment and early warning mechanism construction based on distributed machine learning algorithm. *Safety Science*, *120*, 764–771. https://doi.org/10.1016/j.ssci.2019.08.022

Liu, J., Gong, E., Wang, D., & Teng, Y. (2018). Cloud model-based safety performance evaluation of prefabricated building project in China. *Wireless Personal Communications*, *102*, 3021–3039. https://doi.org/10.1007/s11277-018-5323-3

Lou, X., & Cui, B. (2007). Boundedness and exponential stability for nonautonomous RCNNs with distributed delays. *Computers* & *Mathematics with Applications*, *54*(4), 589–598. https://doi.org/10.1016/j.camwa.2007.02.007

Luo, X., Li, H., Cao, D., Yu, Y., Yang, X., & Huang, T. (2018). Towards efficient and objective work sampling: Recognizing workers' activities in site surveillance videos with two-stream convolutional networks. *Automation in Construction*, *94*, 360–370. https://doi.org/10.1016/j.autcon.2018.07.011

Mardani, A., Jusoh, A., & Zavadskas, E. K. (2015). Fuzzy multiple criteria decision-making techniques and applications–Two decades review from 1994 to 2014. *Expert Systems with Applications*, *42*(8), 4126–4148. https://doi.org/10.1016/j.eswa.2015.01.003

Martinez, P., Ahmad, R., & Al-Hussein, M. (2019a). A vision-based system for pre-inspection of steel frame manufacturing. *Automation in Construction*, *97*, 151–163. https://doi.org/10.1016/j.autcon.2018.10.021

Martinez, P., Al-Hussein, M., & Ahmad, R. (2019b). A scientometric analysis and critical review of computer vision applications for construction. *Automation in Construction*, *107*, Article 102947. https://doi.org/10.1016/j.autcon.2019.102947

Moon, H., Dawood, N., & Kang, L. (2014). Development of workspace conflict visualization system using 4D object of work schedule. *Advanced Engineering Informatics*, *28*(1), 50–65. https://doi.org/10.1016/j.aei.2013.12.001

Nguyen, L. D., Tran, D. Q., & Chandrawinata, M. P. (2016). Predicting safety risk of working at heights using Bayesian networks. *Journal of Construction Engineering and Management*, *142*(9), Article 04016041. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001154

Ning, X., Qi, J., & Wu, C. (2018). A quantitative safety risk assessment model for construction site layout planning. *Safety Science*, *104*, 246–259. https://doi.org/10.1016/j.ssci.2018.01.016

Paltrinieri, N., Khan, F., & Cozzani, V. (2015). Coupling of advanced techniques for dynamic risk management. *Journal of Risk Research*, *18*(7), 910–930. https://doi.org/10.1080/13669877.2014.919515

Park, J., Marks, E., Cho, Y. K., & Suryanto, W. (2016). Performance test of wireless technologies for personnel and equipment proximity sensing in work zones. *Journal of Construction Engineering and Management*, *142*(1), Article 04015049. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001031

Poh, C. Q., Ubeynarayana, C. U., & Goh, Y. M. (2018). Safety leading indicators for construction sites: A machine learning approach. *Automation in Construction*, *93*, 375–386. https://doi.org/10.1016/j.autcon.2018.03.022

Raveendran, A., Renjith, V., & Madhu, G. (2022). A comprehensive review on dynamic risk analysis methodologies. *Journal of Loss Prevention in the Process Industries*, *76*, Article 104734. https://doi.org/10.1016/j.jlp.2022.104734

Redmon, J., & Farhadi, A. (2018). *Yolov3*: *An incremental improvement*. arXiv. https://doi.org/10.48550/arXiv.1804.02767

Salas, R., & Hallowell, M. (2016). Predictive validity of safety leading indicators: Empirical assessment in the oil and gas sector. *Journal of Construction Engineering and Management*, *142*(10), Article 04016052. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001167

Sawicki, M., & Szóstak, M. (2020). Quantitative assessment of the state of threat of working on construction scaffolding. *International Journal of Environmental Research and Public Health*, *17*(16), Article 5773. https://doi.org/10.3390/ijerph17165773

Seker, S., & Zavadskas, E. K. (2017). Application of fuzzy DEMATEL method for analyzing occupational risks on construction sites. *Sustainability*, *9*(11), Article 2083. https://doi.org/10.3390/su9112083

Seo, J., Han, S., Lee, S., & Kim, H. (2015). Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics*, *29*(2), 239–251. https://doi.org/10.1016/j.aei.2015.02.001

Shin, Y.-S., & Kim, J. (2022). A vision-based collision monitoring system for proximity of construction workers to trucks enhanced by posture-dependent perception and truck bodies' occupied space. *Sustainability*, *14*(13), Article 7934. https://doi.org/10.3390/su14137934

Shu, Y., Duan, W., & Jiao, C. (2019). SSD evolution model in HF etching of fused silica optics. *Optik*, *181*, 372–377. https://doi.org/10.1016/j.ijleo.2018.12.067

Son, H., Seong, H., Choi, H., & Kim, C. (2019). Real-time vision-based warning system for prevention of collisions between workers and heavy equipment. *Journal of Computing in Civil Engineering*, *33*(5), Article 04019029. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000845

Sun, X., Wu, P., & Hoi, S. C. (2018). Face detection using deep learning: An improved faster RCNN approach. *Neurocomputing*, *299*, 42–50. https://doi.org/10.1016/j.neucom.2018.03.030

Teizer, J., & Cheng, T. (2015). Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazard areas. *Automation in Construction*, *60*, 58–73. https://doi.org/10.1016/j.autcon.2015.09.003

Veit, A., Matera, T., Neumann, L., Matas, J., & Belongie, S. (2016). Coco-text: Dataset and benchmark for text detection and recognition in natural images. *arXiv*. https://doi.org/10.48550/arXiv.1601.07140

Viso.ai. (2024). *Viso Suite*. https://viso.ai/

Wang, J., & Razavi, S. (2017). A comprehensive spatio-temporal network-based model for dynamic risk analysis on struck-by-equipment hazard. In *Computing in Civil Engineering 2017* (pp. 384–391). ASCE. https://doi.org/10.1061/9780784480847.048

Wang, Y., Liao, P.-C., Zhang, C., Ren, Y., Sun, X., & Tang, P. (2019). Crowdsourced reliable labeling of safety-rule violations on images of complex construction scenes for advanced vision-based workplace safety. *Advanced Engineering Informatics*, *42*, Article 101001. https://doi.org/10.1016/j.aei.2019.101001

Yu, Y., Guo, H., Ding, Q., Li, H., & Skitmore, M. (2017). An experimental study of real-time identification of construction workers' unsafe behaviors. *Automation in Construction*, *82*, 193–206. https://doi.org/10.1016/j.autcon.2017.05.002

Zhang, S., Teizer, J., Pradhananga, N., & Eastman, C. M. (2015). Workforce location tracking to model, visualize and analyze workspace requirements in building information models for construction safety planning. *Automation in Construction*, *60*, 74–86. https://doi.org/10.1016/j.autcon.2015.09.009

Zhang, L., Wu, X., Zhu, H., & AbouRizk, S. M. (2017). Perceiving safety risk of buildings adjacent to tunneling excavation: An information fusion approach. *Automation in Construction*, *73*, 88–101. https://doi.org/10.1016/j.autcon.2016.09.003

Zhang, M., Shi, R., & Yang, Z. (2020). A critical review of vision-based occupational health and safety monitoring of construction site workers. *Safety Science*, *126*, Article 104658. https://doi.org/10.1016/j.ssci.2020.104658

Zhang, Y., Guo, X., & Ma, Y. (2025). *A traffic conflict risk assessment model for highway construction zones based on trajectory data* (SAE Technical Paper 2025-01-7217). https://doi.org/10.4271/2025-01-7217

Zhi-Qiang, H., & Ya-mei, Z. (2016). Research on risk assessment technology of the major hazard in harbor engineering. *Procedia Engineering*, *137*, 843–848. https://doi.org/10.1016/j.proeng.2016.01.324

Zhong, B., Wu, H., Ding, L., Love, P. E., Li, H., Luo, H., & Jiao, L. (2019). Mapping computer vision research in construction: Developments, knowledge gaps and implications for research. *Automation in Construction*, *107*, Article 102919. https://doi.org/10.1016/j.autcon.2019.102919

Zhu, Z., Park, M.-W., Koch, C., Soltani, M., Hammad, A., & Davari, K. (2016). Predicting movements of onsite workers and mobile equipment for enhancing construction site safety. *Automation in Construction*, *68*, 95–101. https://doi.org/10.1016/j.autcon.2016.04.009