

A FRAMEWORK OF ON-SITE CONSTRUCTION SAFETY MANAGEMENT USING COMPUTER VISION AND REAL-TIME LOCATION SYSTEM

J. Zhang¹, D. Zhang^{2,3*}, X. Liu¹, R. Liu¹ and G. Zhong³

¹Department of Construction Management, Tianjin University, Tianjin, China

²School of Architecture, Tianjin University, Tianjin, China

³Tianjin University Research Institute of Architectural Design & Urban Planning, Tianjin University, Tianjin, China

³IBM China Development Lab, IBM, Beijing, China

* Corresponding author

ABSTRACT Construction safety-related research is either management-driven or technology-driven. Many emerging technologies have been investigated by researchers worldwide in different construction management applications. Although there are some limitations, recent advances in deep learning technology have made computer vision (CV) a very active research topic in the field of construction safety management. However, most CV-based studies focus on detecting unsafe behaviour without extending the work to include identifying, locating, and notifying the people involved. In this paper, a framework is proposed for an on-site construction safety management system using Fast R-Convolution Neural Network (CNN)-based computer vision and Bluetooth Low Energy (BLE)-based real-time location system (RTLS). CV can detect on-site entities, understand their spatial relations in a semantic way, and recognize actions of workers and construction equipment. Moreover, the trajectories of moving objects can be tracked, and the next location can be predicted. Combined with the low-cost and easy-to-implement BLE-based RTLS, workers involved in a potential construction hazard will be warned through a mobile application on their smartphones, via loud sounds and vibrations. Experimental studies were conducted at two construction sites to test the CV-based method and the RTLS-based worker identification and warning.

1. Introduction

Hazards at construction sites originate from many sources. First, construction is dynamic, temporary, and decentralized in nature (Li *et al.*, 2015). Second, hazards can arise from the unsafe behaviour of on-site workers resulting from poor attitudes towards safety (Choudhry and Fang, 2008).

Research on construction safety is mainly conducted from two perspectives: it is either management-driven or technology-driven (Zhou *et al.*, 2015). In general, it is assumed that enhanced construction safety management can effectively improve on-site safety performance and reduce the number of accidents. Research from the management perspective usually includes either safety management processes such as safety education and training or focuses on individual/organizational characteristics such as workers' attitude towards safety. Unfortunately, human errors do exist and cannot be completely eliminated through safety education and training; thus, various innovative technologies have been investigated to assist and improve on existing management-driven safety management practices. These technical approaches aimed to enhance rather than replace the management efforts (Teizer *et al.*, 2010).

The construction industry is not very keen to adopt innovative Information and Communication Technologies (ICT) mainly due to the unstructured and dynamic working environment (Gonzalez de Santos *et al.*, 2008). Along with advances in ICT,

researchers have gradually acknowledged the significance of new technologies to many construction management practices, including safety management (Zhou *et al.*, 2013). Some important innovative technologies that have been incorporated in previous safety-related research include Building Information Modelling (BIM) (e.g. Shen *et al.*, 2016), Radio Frequency Identification (RFID) (e.g. Li *et al.*, 2018), Real-Time Location System (RTLS) (e.g. Li *et al.*, 2016), Geographic Information System (GIS) (e.g. Jia *et al.*, 2012), laser scanning (e.g. Marks *et al.*, 2013), Virtual Reality (VR) (e.g. Sacks *et al.*, 2014), Augmented Reality (AR) (e.g. Chi *et al.*, 2013), Knowledge Management (KM) (e.g. Le *et al.*, 2014), Computer Vision (CV) (e.g. Ding *et al.*, 2018), robot (e.g. You *et al.*, 2018), Unmanned Aerial Vehicle (UAV) and remote sensing (e.g. Alizadehsalehi *et al.*, 2018).

Of the abovementioned technologies, CV has drawn attention for solving issues related to construction safety because of its capability to automate tasks, which the human visual system is not able to do (Seo *et al.*, 2015). CV can detect objects and their special relations, track the movement of objects, and understand actions of people and equipment at a construction site (Yang *et al.*, 2016). Although the study of CV in the field of construction management began more than a decade ago, research became more active only in the past few years along with advances in deep learning technology such as convolution neural network (CNN) and its adaptations such as Fast R-CNN,

which makes the understanding of captured images and video much faster and more accurate.

In most CV-based studies, unsafe behaviour detection is the major research objective. For example, Ding *et al.* (2018) combines CNN and long short-term memory to create a hybrid deep learning model to detect unsafe behaviours. There is a need to extend the study from detection of unsafe behaviour to cautioning and eliciting a reaction. In order to warn the worker who is in danger, it is necessary to determine who the worker is. RTLS is a practical solution for locating members of on-site construction crews, and a variety of RTLS technologies have been investigated for this purpose in the last decade (e.g. Li *et al.*, 2016).

This paper proposes an approach that combines CV and RTLS for proactive on-site safety management, where unsafe behaviours are detected by a CV system and people in danger are located by a RTLS and warned via several methods. The next two sections present a review of the development of CV and RTLS and their use in construction safety management. The framework of the proposed CV+RTLS approach will next be described, followed by experimental case studies. In the final section, conclusions will be presented and future work is discussed.

2. Computer Vision

2.1 Capabilities of a CV system

CV has the potential to understand images or videos obtained at complex and dynamic construction sites by. Inexpensive 2D video cameras and surveillance systems are now standard equipment at most construction sites for a number of reasons, including documentation of construction activity (Everett *et al.*, 1998) and security (Yang *et al.*, 2010). As hardware prices decrease, 3D data sensing devices, such as 3D laser scanning (Starbuck *et al.*, 2014) and RGB-D cameras (Ray and Teizer 2012) become more affordable, and these devices are able to provide greater accuracy than 2D data-based approaches. The imagery data processed from these devices enable three main capabilities in construction safety and health monitoring: object detection, object tracking, and action recognition (Seo *et al.*, 2015).

Object detection refers to the capability of recognizing a variety of entities at a construction site such as workers, equipment, and materials. Object detection is the foundation for all other CV approaches such as object tracking and action recognition – after all, without detecting an object, there is no way to track it or recognize its actions. In the last few decades, many 2D and 3D object detection algorithms have been developed, but their performance largely depends on the objects to be detected and the surrounding environment (e.g., available light, vantage point, and obstacles) (Seo *et al.*, 2015) and there is no one-size-fits-all solution.

CV achieves object tracking by assigning consistent labels to detected objects and tracking them in a sequence of still images from a video feed (Yilmaz *et al.*, 2006). Object tracking is able to accommodate the dynamic nature of construction sites, and

a CV-based approach has an obvious advantage in that no sensors need to be attached to the object. The major challenge in CV-based object tracking is the occlusions caused by interacting trajectories (Yang *et al.*, 2010). Aside from the different tracking algorithms reviewed by Seo *et al.* (2015), the accuracy is affected by the distance between the camera and the object being tracked.

Action recognition requires the CV to understand what a worker or a piece of equipment is doing even if it remains in one place (such as a person on a ladder reaching far to one side). Poppe (2010) reviewed human action recognition methods studied in last few decades and summarized a two-step approach to action recognition: (1) image representation to extract action features from the image, and (2) action classification to match the features with the known features of classified actions. Action recognition for equipment could follow a similar approach (Golparvar-Fard *et al.*, 2013).

2.2 Challenges of using CV in construction settings

While the study of CV in construction management began over a decade ago, few cases of actual application in the industry exist. The challenge for researchers is two-fold: (1) technical limitations of CV limit its use in the construction industry, and (2) the inefficient integration of CV with other on-site technology systems.

First, the accuracy and speed of object detection need to improve greatly to make CV more practical to apply in construction for example, proactive safety management. Rapid development of graphics processing unit (GPU) technology facilitates the CNN-based approach and largely improves the performance of object detection; otherwise, any CV application will be merely a castle in the sky. Second, most CV-related research is limited to the detection of some specific entities (workers and some personal protection equipment (PPE), for example) and the understanding of a certain contextual relationship (e.g., a worker not wearing PPE). It is necessary to have a systematic and semantic representation and an understanding of construction scenes for more comprehensive and practical applications. Third, 2D-based object tracking is not reliable, and 3D-based systems are not practical to use at construction sites, especially with the many obstacles present at these sites. Lastly, action recognition remains a challenge in construction, especially for construction equipment. The movements of construction equipment are complex and they vary widely across different types.

As mentioned before, CV has the potential to interpret site situations and identify unsafe behaviours, but it needs to interact with ICT systems that can initiate actions to prevent accidents. At least two systems are required: a location system and a warning system. In order to warn a worker who is in danger (or a worker who could potentially cause an accident, such as an excavator operator), the worker must be identified and located. In most cases, any equipment involved must also be located and its operator identified. After being identified, workers must be notified via an appropriate warning system: for example, speakers integrated into hardhats or LED lights on safety vests.

3. Real-Time Location System

3.1 Safety management using RTLS

A RTLS can be used to identify/track the location of workers, materials, and equipment at a construction site in real time (Li *et al.*, 2016). Thus, it can be used to address a number of on-site issues, including construction safety management.

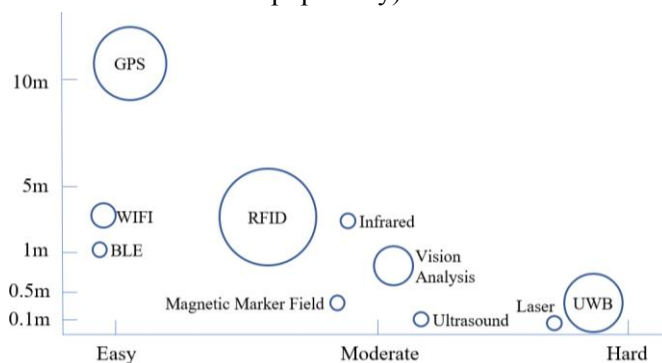
The most obvious and desirable goal of using a RTLS is to protect the lives of construction workers. Lee *et al.* (2012) developed an RFID-based RTLS for on-site personnel tracking that uses assistant tags to overcome signal attenuation. Similarly, Ding *et al.* (2013) integrated a Fiber Bragg Grating (FBG) sensor system and an RFID-based RTLS for a tunnel project in Wuhan, China. Wu *et al.* (2013) investigated the use of RTLS to proactively prevent accidents where workers are struck by falling objects. Teizer *et al.* (2010) studied a system that was supported by a RTLS to allow workers on foot to communicate with the operators of heavy equipment.

Another notable area of research using a RTLS is real-time safety management of equipment. Hwang (2012) investigated the use of an ultra-wideband (UWB) real-time system to prevent tower crane collisions. Wu *et al.* (2010) employed an ultrasound-based RTLS to track people, equipment and materials in order to predict near-miss accidents. A GPS-enabled RTLS was adopted in a study by Garcia *et al.* (2006) in an outdoor environment where the real-time traffic data near a construction site were captured.

3.2 Analysis of RTLS technologies

A RTLS can be used in many process management practices, such as safety management, production control, quality assurance, material management and logistics, and facility management (Li *et al.*, 2016). As such, many RTLS technologies have been extensively studied in the construction industry in the previous decade.

Figure 1 Comparison of RTLS technologies (where circle size reflects the popularity)



RFID has been reported to be used in most RTLS solutions because of its cost-effective and flexible approach for identifying individual objects. RFID has an accuracy range of 1 to 3 meters and is typically used indoors, as compared to a Global Positioning System (GPS), which has an accuracy range of 5 to 10 meters and is generally used in outdoor environments.

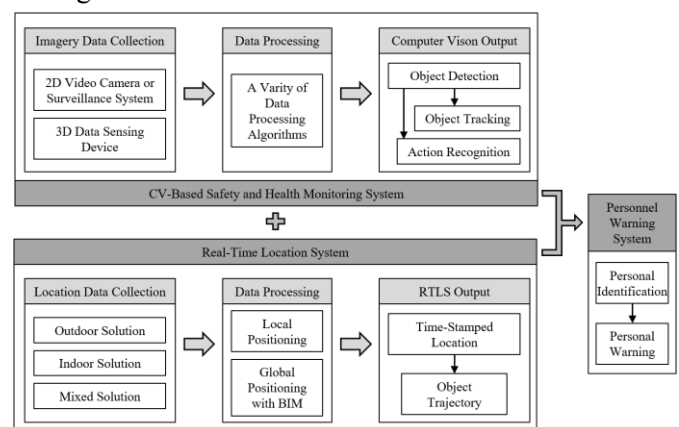
UWB technology is very accurate (with an accuracy range from 0.1 to 0.3 meters) and has a short response time, and it can be applied for both indoor and outdoor sites. However, a UWB system relies on a local area network (LAN) for deployment, and LAN is rarely available at construction sites. CV is another option; however, it requires access to extensive labelled training data sets before implementation, and the inefficiency of vision analysis in dark or dusty environments limits its use at construction sites. Wang and Razavi (2016) summarized the advantages and limitations of other technologies, including Bluetooth low energy (BLE), Wireless LAN, ultrasound, lasers, radar, infrared, and magnetic marker fields. Figure 1 shows the range of accuracy and the ease of implementation (both technically and economically) of the major RTLS technologies.

4. Methods

4.1 Framework

This paper presents a general framework for on-site safety management which could be adapted to fit many specific applications. An example scenario is the case of a worker who approaches a working excavator, where there is a need to warn both the worker and the excavator operator to eliminate the hazard. The following modules, described in Figure 2, would be required to work together in order to achieve this result.

Figure 2 A general framework of on-site safety management based on CV and RTLS



CV-based safety and health monitoring: In this module, imagery data collected using either 2D or 3D camera systems are processed by a variety of algorithms to realize object detection, on the basis of which the object tracking and action recognition are accomplished if needed. In the example scenario, the worker and the excavator must first be detected from a piece of video footage as objects at the construction site. Successful detection is not enough to conclude that the worker is in a dangerous situation. If the distance between worker and excavator is sufficient and the worker is stationary, or if the worker is moving but the worker is protected by a fencing system and the excavator is operating at a safe distance, or if the worker is moving in an area with no fencing system but the excavator is not being operated, then the worker is safe. However, the worker and the excavator operator need to be warned if the worker moves towards a working excavator and the area

around the excavator is not fenced. In this scenario, the CV needs to track the movement of the worker and predict his or her trajectory in following few seconds, understand the surrounding context, and recognize the action of the excavator.

RTLS: The RTLS module in this framework locates and tracks the movement of workers, materials, and equipment with an acceptable accuracy. Although CV is capable of tracking objects, the actual performance is insufficient to support industry practices. As such, a RTLS is used to assist the CV approach, for example, to calculate the distance between a worker and an excavator, track the movement of the worker, and predict the worker's trajectory. Through outdoor RTLS (e.g., GPS), indoor RTLS (e.g., RFID-based system), or mixed RTLS (e.g., UWB system), local coordinates of an object being tracked are calculated by the Received Signal Strength (RSS) method. Next, when local coordinates are combined with the BIM model and the location of signal receiving devices such as Bluetooth gateways, the global coordinates can be determined. The output of the RTLS is a dataset of time-stamped locations and a calculated object trajectory.

Personnel warning system: Once it has been recognised by CV that a person is in danger and the locations of the relevant people, materials, and equipment, are determined, a personnel warning system is needed to take action to notify key people and eliminate the danger. By combining the CV and RTLS, the people to be notified (the worker and the excavator operator in the example) can be identified. Methods to notify on-site workers include speakers built into hardhats, mobile phones, smart wrist bands, LED-equipped safety vests, and so on.

As an illustration, a CNN-based CV approach, a BLE-based RTLS, and a Bluetooth-based warning system are presented as best practices in their respective modules. These technologies are also used in the case studies in this research.

4.2 CNN-based CV

CNN is a CV approach for object detection and classification which can automatically extract a large number of image features and use those features for image classification (Kolar *et al.*, 2018). Treating an image as a matrix of pixels where each pixel represents one feature, LeCun *et al.* (1998) developed a CNN-based CV to process large images. However, CNN-based approach requires sufficient computing power to process large images, especially when dealing with a dataset that includes millions of images.

Hardware improvement in GPUs has enabled effective training of large CNN networks by stacking multiple convolutional and pooling layers to recognize features not only from static images but also from video footage (Zhang *et al.*, 2016). Girshick *et al.* (2014) proposed a method where, instead of selecting a huge number of regions, the R-CNN algorithm uses a selective search to extract just 2000 regions from the image. The same authors further improved the R-CNN algorithm by feeding the input image to the CNN to generate a convolutional feature map. This Fast R-CNN approach is able to more accurately identify objects than any other deep learning method that has

been proposed (Fang *et al.*, 2018), and it makes many CV applications possible in practice.

4.3 BLE-based RTLS

BLE technology can communicate contextual information among multiple connected devices (e.g., BLE sensors, mobile devices, and online computers) with minimal infrastructure in terms of size and cost (Park *et al.*, 2016). The low power required for operation gives the tracking device a very long working life. Because of its small size, a BLE device is suitable for use as a wearable device (when attached to a hardhat or a safety vest). Despite its potential, BLE-based RTLS has not received enough attention from researchers and professionals in the construction industry.

The proposed BLE-based RTLS used in this framework consists of four components: BLE devices (gateways and beacons), a communication channel, a cloud-based platform, and an application domain (Zhao *et al.*, 2017). Gateways are placed in designated locations, which are determined by the unique layout of the construction site. Combined with a BIM model, the global coordinates of those gateways are obtained. Beacons are attached to workers, materials, and equipment to track their locations in real time. Based on the RSS methods, the Bluetooth signals transmitted from a certain beacon and captured by three or more gateways are used to determine the location of the beacon. All data are relayed to a cloud-based system via a wireless communication device. A third-party application can use the data for function-specific applications (in an application domain) via an application programming interface (API).

4.4 Bluetooth-based warning system

The Bluetooth beacons used in the RTLS are carried by workers and are paired with the workers' mobile phones. Once a worker is identified as a person in danger, the system will send a command to the beacon which can trigger a loud sound and a vibration on the worker's mobile phone.

5. Experimental Cases

5.1 Imagery data collection and analysis

A high-rise apartment building project located in Tangshan, China, was selected to test the CV approach. A standard 2.0 Megapixel 2D camera was fixed on the crane tower at 8 meters above the ground. Most workers on-site were performing rebar binding work.

Figure 3 shows the result of object detection from Fast R-CNN analysis. As shown in Figure 3(a), hardhats are successfully detected even in a large-scale view. Other larger entities such as workers and equipment are also correctly detected without any issue. Figure 3(b) indicates the identification of several large holes, which are the potential hazard areas.

Figure 4 shows the result of object tracking. A selected worker was labelled, and his trajectory was traced and superimposed on the image. Due to the lack of 3D capability of the camera, some positions look a bit odd on the 2D view (e.g., in the red circle, it appears that a worker had already fallen into a hole).

Figure 3(a) Hardhat detection, (b) Hole detection

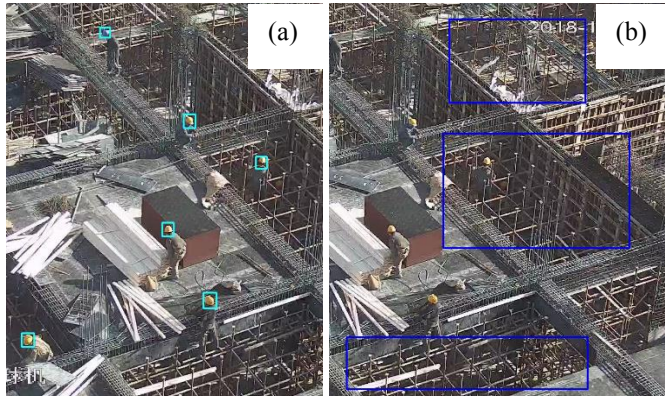
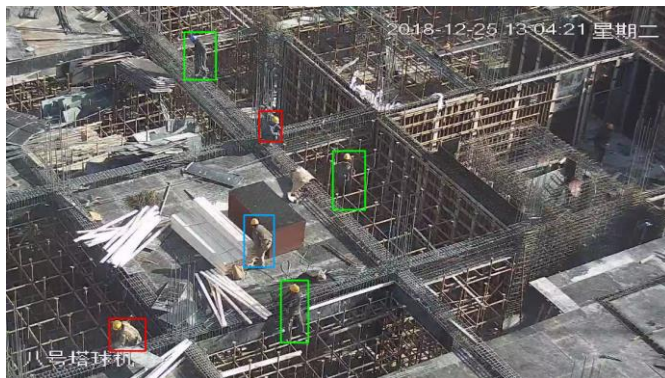


Figure 4 Object tracking of a selected worker



Figure 5 shows the result of action recognition. By matching features extracted from detected workers to a set of known action classification features, three types of action were recognized: standing (in green frames), stooping (in blue frames), and squatting (in red frames).

Figure 5 Action recognition of workers



5.2 RTLS deployment and testing

One recent experimental system of the proposed RTLS was deployed at a commercial building in the city of ShiJiaZhuang, China, as part of a mechanical, electrical, and plumbing (MEP) installation project.

The system at the ShiJiaZhuang project was set up on the fourth floor of the building, which is laid out in a rectangular shape (50 meters by 30 meters). The five stars shown in the layout in Figure 6 are Raspberry Pi gateways (pictured in Figure 7a) which were installed at a height of 2.5 meters. In the ShiJiaZhuang project, one beacon (Figure 7b) was used to track each of the five workers, and one additional beacon was used to track a welding tool. Huawei E3372h-510 USB sticks (Figure 7c) were used to send data to the cloud-based system via 4G/LTE mobile data communication. Calculated location information is displayed on a mobile app (Figure 7d).

Figure 6 Layout of gateways in Shijiazhuang Project

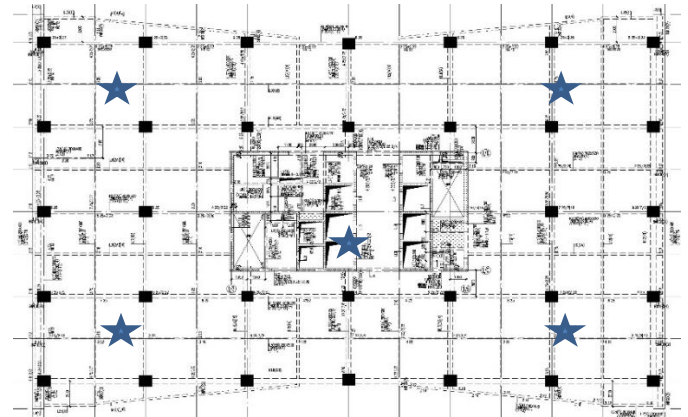
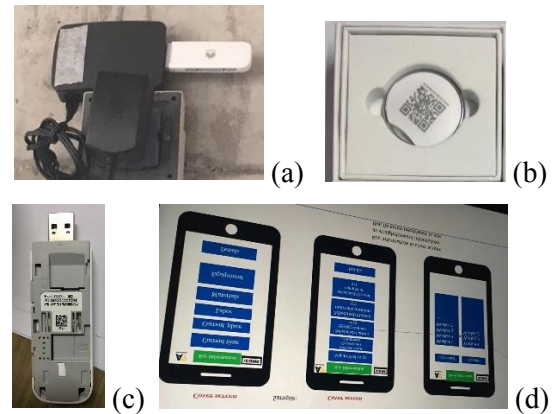


Figure 7 Photos of (a) Gateway, (b) Beacon, (c) Wireless Internet adapter, (d) Mobile phone interface



The RTLS implemented in the ShiJiaZhuang project demonstrated an average accuracy of 1 to 2 meters, which provides sufficient certainty to be useful for proactive safety management systems. In addition, Bluetooth communication devices do not require a line of sight (LOS) environment, which means they are more flexible for on-site deployment.

Another demonstrated advantage of the BLE-based RTLS is cost efficiency. The system used in the ShiJiaZhuang project was inexpensive: \$70 per gateway and \$10 per beacon, plus a fee of \$15 per month for the 4G/LTE mobile data plan. Providing a power supply to gateways was an issue at the beginning of the project, but the issue was resolved through the use of power banks. A single power bank costs about \$20; it supplies power during the workday and is recharged overnight.

6. Conclusion and Future Work

On-site safety management has always been a focus for research in the construction industry. As technology advances, new opportunities emerge to better manage safety at construction sites. This paper proposed a framework for an on-site safety management system using computer vision and a real-time location system. Using the Fast R-CNN method to analyze imagery data obtained from on-site cameras, the situation of the construction scene is semantically understood: the entities on site and their spatial relations, the trajectories of moving objects and the trends of their movement, and the actions of workers and construction equipment. When image analysis is combined with BLE-based RTLS, the locations of people involved in a potential safety hazard can be identified. Through the connection between the Bluetooth beacon used for RTLS and the workers' mobile phones, a simple warning method is designed to notify the relevant workers.

Some experimental studies were conducted separately. A construction site in Tangshan, China, was fitted with cameras to collect on-site video footage. Workers and large construction equipment were correctly detected, but action recognition was not easy because of the large field of view at the construction site. BLE-based RTLS was deployed and tested in a MEP installation site in Shijiazhuang, China. The accuracy obtained was about 2.5 meters, which is good enough for a low-cost, easy-to-implement location system. A prototype mobile application was developed to warn workers using sounds and vibrations.

The authors are planning to implement an integrated CV and RTLS solution in a new construction project in March 2019 to investigate the information coordination between the CV system and the RTLS. Also, considering that sound and vibration warnings from a mobile phone could be too weak to be noticed in a noisy construction environment with excavation activities and heavy vehicle traffic, some alternatives will be tested, for example, a Bluetooth enabled LED light strip on a worker's safety vest.

References

- Alizadehsalehi S *et al.* (2018) The effectiveness of an integrated BIM/UAV model in managing safety on construction sites. *International Journal of Occupational Safety and Ergonomics* 2018(9): 1-16.
- Chi H *et al.* (2013) Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. *Automation in Construction* 33: 116-122.
- Ding LY *et al.* (2013) Real-time safety early warning system for cross passage construction in Yangtze Riverbed Metro Tunnel based on the internet of things. *Automation in Construction* 36: 25-37.
- Everett JG *et al.* (1998) Time-lapse video applications for construction project management. *Journal of Construction Engineering and Management* 124(3): 204-209.
- Fang W *et al.* (2018) Falls from heights: A computer vision-based approach for safety harness detection. *Automation in Construction* 91: 53-61.
- Garcia C *et al.* (2006) Using GPS to measure the impact of construction activities on rural interstates. *Journal of construction engineering and management* 132(5): 508-515.
- Girshick RB *et al.* (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. *IEEE Conference on Computer Vision and Pattern Recognition*: 580-587.
- Golparvar-Fard M *et al.* (2013) Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Advanced Engineering Informatics* 27(4): 652-663.
- Gonzalez de Santos *et al.* (2008) Power assist devices for installing plaster panels in construction. *Automation in Construction* 17(4): 459-466.
- Hwang S (2012) Ultra-wide band technology experiments for real-time prevention of tower crane collisions. *Automation in Construction* 22: 545-553.
- Jia N *et al.* (2012) Development and Implementation of a GIS-Based safety Monitoring System for Hydropower Station Construction. *Journal of Computing in Civil Engineering* 26(1): 44-53.
- Kolar Z *et al.* (2018) Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Automation in Construction* 89: 58-70.
- Le QT *et al.* (2014) A social network system for sharing construction safety and health knowledge. *Automation in Construction* 46: 30-37.
- LeCun Y *et al.* (1998) Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11): 2278-2324.
- Lee HS *et al.* (2012) RFID-based real-time locating system for construction safety management. *Journal of Computing in Civil Engineering* 26(3): 366-377.
- Li H *et al.* (2015) Proactive behavior-based safety management for construction safety improvement. *Safety science* 75: 107-117.
- Li H *et al.* (2016) Real-time locating systems applications in construction. *Automation in Construction* 63: 37-47.
- Marks ED *et al.* (2013) Laser Scanning for Safe Equipment Design That Increases Operator Visibility by Measuring Blind Spots. *Journal of Construction Engineering and Management* 139(8): 1006-1014.
- Park J *et al.* (2016) Framework of automated construction-safety monitoring using cloud-enabled BIM and BLE mobile tracking sensors. *Journal of Construction Engineering and Management* 143(2): 05016019.
- Poppe R (2010) A survey on vision-based human action recognition. *Image Vision Computing* 28(6): 976-990.
- Ray SJ and Teizer J (2012) Real-time construction worker posture analysis for ergonomics training. *Advanced Engineering Informatics* 26(2): 439-455.
- Sacks R *et al.* (2014) Safety by design: dialogues between designers and builders using virtual reality. *Construction Management and Economics* 31(1): 55-72.

- Seo J *et al.* (2015) Computer vision techniques for construction safety and health monitoring. *Advanced Engineering Informatics* 29(2): 239-251.
- Shen X *et al.* (2016) Near-Miss Information Visualization Tool in BIM for Construction Safety. *Journal of Construction Engineering and Management* 142(4): 04015100.
- Starbuck R *et al.* (2014) A stereo vision-based approach to marker-less motion capture for on-site kinematic modeling of construction worker tasks. *Proceedings of the 15th International Conference on Computing in Civil and Building Engineering (ICCCBE)*, Orlando, FL, 2014.
- Teizer J *et al.* (2010) Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system. *Automation in Construction* 19(5): 630-640.
- Wu W *et al.* (2010) Towards an autonomous real-time tracking system of near-miss accidents on construction sites. *Automation in Construction* 19(2): 134-141.
- Wu W *et al.* (2013) An integrated information management model for proactive prevention of struck-by-falling-object accidents on construction sites. *Automation in Construction* 34: 67-74.
- Yang J *et al.* (2010) Tracking multiple workers on construction sites using video cameras. *Advanced Engineering Informatics* 24 (4): 428-434.
- Yang J *et al.* (2016) Vision-based action recognition of construction workers using dense trajectories. *Advanced Engineering Informatics* 30(3): 327-336.
- Yilmaz A *et al.* (2006) Object tracking: a survey. *ACM Computing Surveys(CSUR)* 38(4): 13.
- You S *et al.* (2018) Enhancing perceived safety in human–robot collaborative construction using immersive virtual environments. *Automation in Construction* 96: 161-170.
- Yu Y (2014) Integrated application of BIM and RFID in construction safety management. *Journal of Engineering Management* 2014(4): 87-92.
- Zhang X *et al.* (2016) Deep fusion of multiple semantic cues for complex event recognition. *IEEE Trans. Image Process* 25(3): 1033-1046
- Zhao J *et al.* (2017) Data analysis on applying real time tracking in production control of construction. In *Proceedings of Industrial Engineering and Engineering Management (IEEM)*: 573-577.
- Zhou Z *et al.* (2013) Applying advanced technology to improve safety management in the construction industry: a literature review. *Construction Management and Economics* 31(6): 606-622.
- Zhou Z *et al.* (2015) Overview and analysis of safety management studies in the construction industry. *Safety science* 75: 337-350.