Automation in Construction 110 (2020) 103013

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

Review

Computer vision applications in construction safety assurance

Weili Fang[a,b,c,d], Lieyun Ding[a,b,*], Peter E.D. Love[d], Hanbin Luo[a,b], Heng Li[e], Feniosky Peña-Mora[c], Botao Zhong[a,b], Cheng Zhou[a,b]

[a] Dept. of Construction Management, Huazhong University of Science and Technology, Wuhan 430074, China
[b] Hubei Engineering Research Center for Virtual, Safe and Automated Construction, Wuhan 430074, China
[c] Dept. of Civil Engineering and Engineering Mechanics, Columbia University in the City of New York, New York, NY 10027, United States of America
[d] School of Civil and Mechanical Engineering, Curtin University, Perth, WA 6845, Australia
[e] Dept. of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong, SAR, China

ARTICLE INFO

Keywords:
Computer vision
Deep learning
Digital technology
Safety

ABSTRACT

Advancements in the development of deep learning and computer vision-based approaches have the potential to provide managers and engineers with the ability to improve the safety performance of their construction operations on-site. In practice, however, the application of deep learning and computer vision has been limited due to an array of technical (e.g., accuracy and reliability) and managerial challenges. These challenges are a product of the dynamic and complex nature of construction and the difficulties associated with acquiring video surveillance data. In this paper, we design and develop a deep learning and computer vision-based framework for safety in construction by integrating an array of digital technologies with multiple aspects of data fusion. Then, we review existing studies that have focused on identifying unsafe behavior and work conditions and develop a computer-vision enabled framework that: (1) considers current progress on computer vision and deep learning for safety; (2) identifies the research challenges that can materialize with using deep learning to identify unsafe behavior and work conditions; and (3) can provide a signpost for future research in the emergent and fertile area of deep-learning within the context of safety.

1. Introduction

Worldwide construction is one of the most dangerous industries as people are susceptible to workplace accidents, injuries, and even fatalities. Approximately 7% of the world's workforce is employed in construction; however, the industry accounts for 30–40% of workplace fatalities [86]. According to the Occupation Safety and Health Administration (OSHA) in the United States, for example, a total of 991 people have been killed while working on construction sites since 2006 [1]. Across the Atlantic Ocean in the United Kingdom, fatalities are also a problem in construction accounting for a total of 29.86% of all workplace accidents [2,3]. In China, a total of 3843 fatal injuries were recorded in 2017 on construction sites, which has resulted in the industry being identified as the most hazardous in the country [4]. Notably, more than 90% of accidents are due to unsafe behavior and work conditions. It, therefore, follows that if we can moderate people's unsafe behavior and improve work conditions, then safety performance will naturally improve.

Technological developments aided by computer vision have been identified as a robust approach to automatically identify and recognize unsafe behavior and conditions [5–11]. As a result, a rich collection of images of people's actions and the work conditions that contribute to unsafe events have been accumulated [9–11]. From an engineering perspective, computer vision aims to automate tasks that the human visual system is unable to perform. The ability to automate tasks has been enhanced by deep learning (also known as deep structured learning or hierarchical learning). In particular, Convolutional Neural Networks (CNN), a class of deep learning networks, have been used for analyzing visual imagery (e.g., processing images and video) and overcoming the issues associated with the manual observation and recording of hazards (i.e., potential sources of harm) on construction sites.

While attention has been placed on the use of deep learning and computer vision to monitor the safety behavior and identify unsafe conditions on construction sites, there has been no state-of-the-art review that has examined its development and potential use in the future. A review is needed to determine the current limitations of deep learning and computer vision in construction. Furthermore, there is a need to identify the recurrent problems that researchers are confronted with during its implementation to manage safety. The upshot of performing

* Corresponding author at: Dept. of Construction Management, Huazhong University of Science and Technology, Wuhan 430074, China.
E-mail address: dly@hust.edu.cn (L. Ding).

https://doi.org/10.1016/j.autcon.2019.103013
Received 7 February 2019; Received in revised form 27 September 2019; Accepted 12 November 2019

such a review is to provide a pathway to ensure that future research provides a robust theoretical underpinning to be developed, and has relevance to practice.

Our review draws on developments that have been made in Artificial Intelligence (AI) in construction and other domains (e.g., autonomous vehicles). The review is used to design and develop a deep learning and computer vision-based framework. Our framework integrates digital technologies and multiple aspects of data fusion, which can be used to improve safety performance. The technologies that we have selected to develop our framework complement each other as evident in their widespread use to support the implementation of Industry 4.0.

We commence our paper by providing a setting for deep learning in computer vision and the nature of unsafe behavior and conditions that manifest on construction sites. Then, we review existing studies that have focused on identifying unsafe behavior and conditions in accordance with our enabling framework by placing emphasis on: (1) the existing progress on computer vision and deep learning for safety; (2) identifying the research challenges that can materialize when using deep learning to identify unsafe behavior and work conditions; and (3) providing a signpost for future research in the emergent and fertile area of deep-learning within the context of safety.

2. Understanding computer vision and deep learning

Within the field of computer science, machine learning, which is a subset of AI [12], has been applied widely to areas such as speech recognition [13], natural language processing [14], robot control [15] and computer vision [16]. Conventional machine learning approaches are limited in their ability to process data in their raw form [17]. The inability to process data arises because a considerable amount of engineering and domain knowledge is required to design a feature extractor [17: p.446].

Deep learning is a representation method that can be used to extract sophisticated features at high levels of abstraction automatically. This method can also learn from data with multiple levels of end-to-end representations [17]. By combining deep learning methods (e.g., neural networks) with computer vision, image features can be automatically extracted and used to learn from training data. One particular type of deep learning method that has been widely used is the CNN, as it has been able to accurately and reliably outperform other deep neural networks in areas such as image classification [18], object detection [19], and segmentation [20]. Deep learning has enabled developments in computer vision-based applications to thrive [21–23], for example, (1) autonomous vehicles [21]; and (2) automatic diagnosis of breast or skin-cancer [22]. In light of such developments, we suggest that deep learning and computer vision approaches can potentially provide us with the much-needed visual insights to accurately and reliably understand the nuances of tasks that are performed during the construction process. As a result, we will be able to identify the unsafe behavior and conditions that manifest by comparing information that is acquired with our existing hazard knowledge.

3. An Enabling Framework of Computer Vision-based for Safety

3.1. Unsafe behavior and condition monitoring tasks

Within the context of safety, we put the spotlight on unsafe behavior and the conditions that pervade practice on construction sites concerning standards and regulations in China (Fig. 1, source from Baidu). It has been observed that 88% of the accidents that occur on construction sites are due to unsafe behaviors and with 10% being attributable to unsafe work conditions [24]. The most common types of unsafe behavior identified on construction sites are [5,25–30]: (1) failure of personal protective equipment (PPE); (2) exposure to a hazardous area; and (3) failure to follow safety procedures.

Drawing on accident statistics produced by bodies such as the China State Administration of Work Safety, we can observe that unsafe work conditions are primarily attributable to: (1) coming into contact with plant; and (2) structural defects. In China, for example, there were 88, and 130 accidents related to crane collapses in 2014 and 2013, respectively [24,25]. Also, a considerable number of accidents were attributable to insufficient inspection and condition assessments of physical structures (i.e., detection of the defects and damage such cracking, spalling, defective joints, and corrosion) [31]. In Fig. 1, we present a series of examples identifying unsafe behavior and conditions.

3.2. Data sources for safety monitoring

To mitigate safety risks in construction, both manual inspection and digital technologies (i.e., deep learning and computer vision), have been used to identify and monitor hazards. As a consequence, data is recorded in various formats (e.g., safety reports, video, and photographs), which researchers have used for monitoring safety. For example, Robinson et al. [32] utilized machine learning to map the causal factors contributing to hazards from safety reports. While the extraction of such factors provides us with knowledge (i.e., discrete facts) about safety issues its assimilation to aid the understanding that is needed to reduce their occurrence is absent as a no context is provided. If we are to make any headway toward effectively embracing AI and realizing its benefits in construction, we need to provide a context, so that knowledge and understanding are coupled. In this instance, the challenge centers around the fusion of data with images/videos. Drawing on developments in AI and computer science that have fused different types of data (e.g., text) with image/video to better understand the nature of problem [33,34], we suggest that the use of safety reports, non-visual sensors in conjunction with deep learning and computer vision can provide an effective means for monitoring safety on-site.

3.3. Framework of computer vision-based for safety

To build an automatic computer vision-based safety monitoring system, we need to establish a data resource and then be able to analyze it to identify patterns of behavior and emerging trends. As digital technologies mature, they can be applied to create a robust learning environment and combined with deep learning and computer vision to aid the analysis of safety data to provide us with the understanding needed to improve safety performance. In sum, embracing and fusing digital technologies with deep learning and computer vision provides a mechanism to create new solutions to automatically identify unsafe behavior and conditions [9–11,35,36].

Considering the review presented above, we present in Fig. 2 a framework to enable computer vision to be used to monitor safety. Our framework combines both digital technologies and multiple aspects of data fusion. In our proposed framework, the scope of digital technologies encompasses machine learning, building information models (from now on referred to as model), ontology, AR/VR, and IoT.

4. Deep learning and computer vision for safety in construction

Our review of previous computer-vision research within the context of safety aligns with our developed framework (Tables 1 and 2). In this section, we provide an overview of the different deep learning and computer vision approaches that have been developed and their reported benefits.

4.1. Computer vision-based unsafe behavior recognition

With developments in deep learning, a plethora of approaches have been developed and applied to examine unsafe behavior in construction with computer vision [10,11,37]. As we noted above, deep learning and computer vision research have focused on the three categories of unsafe

Fig. 1. Examples of (A) unsafe behavior; (B) unsafe plant; (C) structural defects.

behavior. Such research has tended to leverage algorithms (i.e., object detection algorithms, object tracking, image classification, and activity recognition) that have achieved good levels of performance from the domain of computer science to extract knowledge and then are compared with rules, guidelines, or expert experiences to identify unsafe behavior. A notable study, for example, is the work of Fang et al. [11] who combined computer vision with a Mask R-CNN to identify individuals and structural supports. This knowledge was used to identify the unsafe behavior of individuals who had traversed structural supports and then determine the relationship between these objects.

We present in Table 1 a detailed summary of prior works on deep learning and computer vision used to identify unsafe behavior and their limitations. Despite the significant progress being made to identify unsafe behavior using computer vision, several problems remain unresolved as we highlight in Table 1.

4.2. Identification of unsafe conditions using computer vision

Research focusing on the recognition of unsafe conditions using computer vision has focused on the identification of a plant's location as well as status and structural defects (Table 2). Several deep learning and computer vision approaches have been developed to determine the surface quality of external structures (e.g., cracks). For example, Cha et al. [46] integrated computer vision with a CNN to detect cracks, though without computing the defect features from two-dimensional (2D) images. However, Cha et al.'s [46] approach achieved a detection

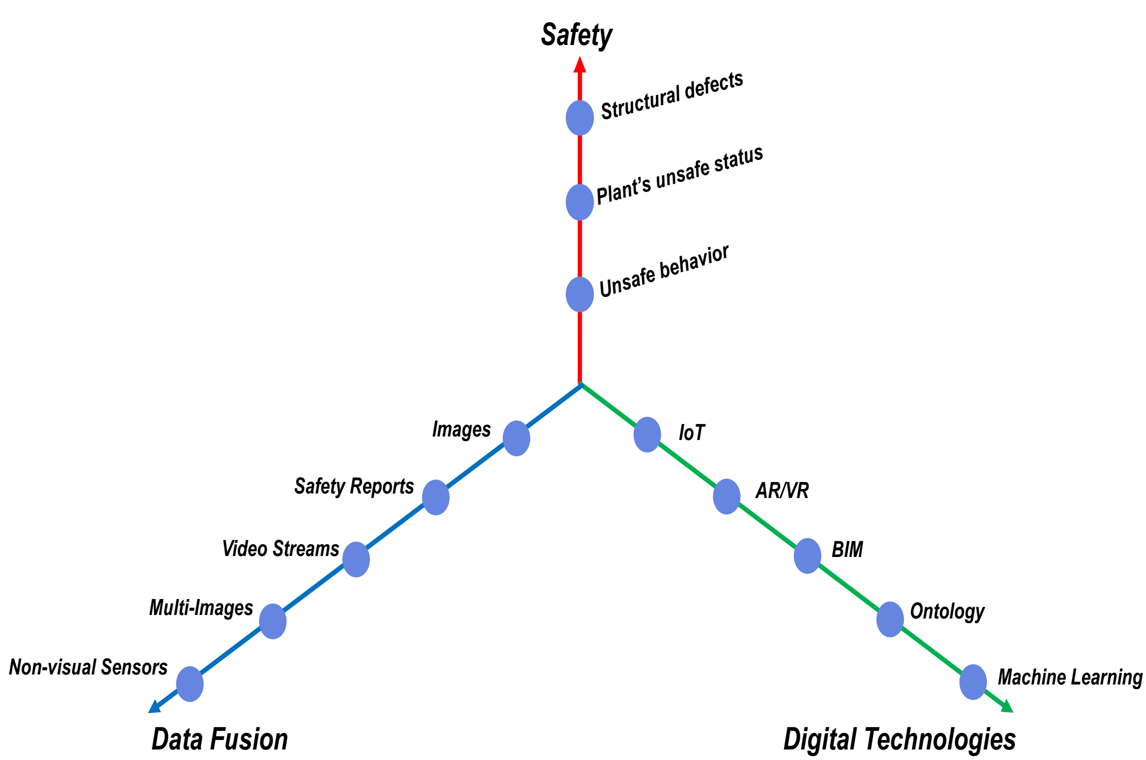

Fig. 2. Integration of data fusion and digital technologies with safety.

Table 1
Examples of unsafe behavior and prior works on deep learning and computer vision-based unsafe behavior recognition.

Types of hazards	Categories of problem	Types of data resource	Digital technologies	Descriptions	Limitations	References
Failure to use PPE	Entering working sites without wearing safety hardhat	2D Images	Deep learning	Applied Faster R-CNN to detect worker not wearing safety hardhat	Failure to detect worker; 'only part of head visible'	[37]
	Not wearing safety harness working in height	2D images	Deep learning	Applied Faster R-CNN to detect worker, and then used a classification network to identify if a worker wears a safety harness	The training database was small, and only a selected number of activities working at height were examined.	[10]
	Safety harness hanging does not meet requirements (e.g. hook to fixed object)	3D image	Deep learning	Deep learning-based 3D used for reconstruction approach build a 3D scene model, and then use 3D object detection to classify objects.		[38,39]☆
Entering hazardous area	Approaching or entering to unprotected sides (the edge of the floor, roof, balcony, roof hatch, foundation pit)	Video	Deep learning and ontology	Using semantic segmentation approach to identify scene (i.e., foundation pit, roof hatch) and ontology to reason if there has a safety barrier in a range of sides		[40,41] ☆
	Traversing structural support without any safety barrier	2D image	Deep learning	Applied an automatic computer-vision approach that utilizes Mask R-CNN to detect individuals traversing structural supports during the construction of a project	The model fails to detect this unsafe behavior due to the presence of occlusions	[11]
	Not taking warning during lifting unwanted workers get into the dangerous areas	2D image	Deep learning for object detection, activity detection approach	Use of deep learning-based object detection to detect people, warning signs, cranes, and then use an activity detection approach to identify crane's activities		[42,43] ☆
	Struck or be close to bulldozer/excavator that is backing up	UAV images	Deep learning	Applied a Yolo-v3 approach to localize objects, and use an image rectification method to measure the distance among objects	The mean absolute distance error for the distance measurement was less than 0.9 m, and the mean absolute percentage errors were approximately 4%.	[8]
	Approaching or entering to stair/elevator well	3D Image	Deep learning	Use of deep learning to identify specific areas (e.g., stairwell) and people. Then apply tracking to identify the person entering an area		[40,44] ☆
Failure to follow safety procedures	Transporting works by hanging objects	3D Images	Deep Learning and ontologies	Use of deep learning to detect heavy equipment and people,		[41,42] ☆
	Lifting objects underground without verifying the situation	3D images and videos	Deep Learning for object detection approach, 3D reconstruction approach, and 3D segmentation approach	The application of a 3D recognition approach to building a 3D model in real-time. Then use object detection to identify objects from the as-built 3D model. Attributes are extracted from the as-built 3D models. The spatial collision can be identified during operation in real time.		[38–40] ☆
	The main and auxiliary hook work simultaneously when operating a crane	Video	Deep learning	Use of deep learning to identify a cranes' activities to determine the main or auxiliary hook.		[43,45] ☆
	Falling from ladder and scaffold (i.e., do not reach too far to either said or to rear, not carrying tools or materials up the ladder)	Video	Deep learning	Development of a hybrid deep learning approach by integrating CNN and LSTM to identify unsafe behavior (i.e., abnormal climb ladder)	The database is small, which leads to error detection	[42]

Note: The ☆ denotes potential solutions for identifying unsafe behavior.

Table 2
Example of unsafe conditions and prior works on computer vision-based unsafe condition recognition.

Categories of problem			Types of data Resource	Digital technologies	Descriptions	Limitations	References
Unsafe plant	Overload	When the working range is increased, or weight exceeds the corresponding load, the overturning moment may be exceeded	Sensors, videos	Deep learning, IoT, and BIM	Using deep learning and sensors to collect data (locations, activities), and the mechanical parameters determined. Such data can be transmitted to a model using IoT. This will enable hazards to be visualized in the model in accordance with pre-determined safety rules.		[42,43,48,82] ☆
	Weak foundation	The foundation of the tower crane is subject to failure	Not achieved yet				
		Use of crane before the concrete foundation has cured and ready for use (e.g., hoisting)	Not achieved yet				
	Inclined lifting	Using inclined lifting in a complex environment in the presence of high winds	Sensors, videos	IoT, deep learning	Sensors are used to collect environmental data (e.g., wind speed), and computer vision is used to identify crane operation activities		[43,48,82] ☆
Structural defects	External surface quality	Pavement crack detection	3D image	Deep learning	An improved CrackNet is developed to detect pavement crack, which achieved precision, recall, and F-measure with 90.20%, 89.06%, and 89.62%, respectively.	The proposed approach has limited capability of acquiring a global context, which makes it a challenge to classify typical noise patterns	[49]
		Component type identification, spall conditions check, damage level evaluation, damage type determination	2D image	Deep learning	A deep transfer learning based on VGGNet is introduced to identify structure damages.	The deep TL brings new directions for the detection and evaluation of structural damage from images with a small database, which can achieve satisfying performances in basic component and damage recognition tasks.	[50]
	Internal structural quality	Grouting is not dense Grouting thickness is not enough Grouting density is insufficient	Radar image	Deep learning	Collecting internal structural data by using radar, and then generate an image by using these data. Deep learning can identify defects		[51] ☆

Note: The ☆ denotes potential solutions for unsafe condition recognition.

accuracy of 98% from the database that was created. Likewise, Xue and Li [47] proposed a fully convolutional network (FCN) model to detect cracks and leakages from 2D images for a shield tunnel and achieved a detection accuracy of 95%.

While there has been a significant body of work undertaken to determine the surface quality of structures (i.e., cracks), issues associated with the identifying defective joints (e.g., grouting), for example, remain unresolved. Similarly, in the case of plant, there has been a paucity of research that has focused on its location and status. This is due to the difficulties associated with the identification of their mechanical parameters and the extraction of attributes.

5. Challenges of computer vision in construction

While there has been an array of attempts to automate the identification of hazards on construction sites, no fully automatic computer vision-based system has yet been developed. We have however been able to identify people not wearing their PPE, but unable to determine: (1) the identity of a person not adhering to this requirement; and (2) if PPE is appropriately used (e.g., the hook of safety harness not being fixed to rail). Deep learning, however, can potentially provide us with data analytics to be able to identify hazards in real-time automatically. We next discuss the potential application and technical challenges of deep learning algorithms, particularly CNNs, for safety management.

5.1. Applying computer vision in practice

Having an extensive and high-quality database of images of varying types to engender a CNN's capacity to learn is a pre-requisite for identifying hazards (i.e., miss detected, or low detection accuracy) and ensuring the successful application of computer vision. However, the absence of an adequately sized database is a significant obstacle that stymies the use of computer vision for engendering effective safety monitoring. In comparison with publicly available datasets in computer science such as the ImageNet and Microsoft® Common Objects in Context (COCO), those required for construction possess unique characteristics that need to consider spatial conflicts, cluttered backgrounds, occlusions, various poses and scales and the dynamic and changing nature of its environment. In this case, many potential hazards are out of a person's sight (i.e., they may be struck-by a plant or equipment) and a structure's quality is not easy identified due to the limitations of deep learning models in being able to predict previously unobservable objects and extract concealed information using computer vision.

Due to the limited availability of datasets, researchers have had to use relatively small samples of images to undertake their experimental works to identify hazards. The use of small datasets has resulted in reported evaluation metrics such as precision, recall, and accuracy being problematic to compare and contrast with other approaches. Due to the variability in the quality of datasets used for training and testing, it is difficult to determine the validity and reliability of results that have been posted in the extant literature. Thus, there is a need for robust and objective evaluation criteria that we can use to compare and contrast various computer vision approaches that are promulgated for managing safety.

5.2. Technical challenges

As we note above, there has been a tendency of computer vision studies to rely on the use of small databases and then use supervised approaches to identify unsafe behavior. The corollary being weak generalizations due to: (a) the assumption that training and testing databases belong to the same distribution; and (b) machine learning being used to train small databases, which limits inter and intra-class variability. Consequently, this hinders their ability to accurately recognize unsafe behavior and enable generalizations to be made to different datasets [52]. The techniques of transfer learning and data augmentations (i.e., crop, flips, and random rotation) however can be applied to overcome the issues associated with the accuracy and reliability with using small training database.

Deep learning models learn correlations between input and output features but are unable to characterize causality. Despite this, we need to understand the interactions between people's behaviors and their corresponding working situations to be able to contextualize the information that often surrounds hazards that materialize on construction sites. For example, behavior-based safety (BBS) has been used to observe and identify people's unsafe actions. Then, feedback is directly provided to those who have committed an unsafe act with the intent of modifying their future behavior [36,53–55].

While the use of deep learning is capable of recognizing hazards, it is essential to realize that these approaches focus on addressing specific tasks related to safety. As a result, this poses a significant problem, as no single approach can be used to identify a range of unsafe behaviors, which renders it a costly and time-consuming process to implement computer vision in practice. We, therefore, need to develop new algorithms and train them to be able to detect a wide range of common unsafe behaviors and conditions that materialize while work is being performed on-site.

Deep learning models are akin to being a 'black box' and thus are opaque [57,65,77]. Strides have been made to address this issue by visualizing the contributions of individual nodes in a complex network using more than a million parameters. Ensuring the transparency of deep learning remains unsolved [65]. As we are unable to determine the exact features that have been extracted and learned from nodes, then understanding how detection was made and identifying the parameters that need to be adjusted to accurately detect hazards are issues that need be addressed to justify the adoption of deep learning [65].

6. Areas for future research in deep learning and computer vision

To address the above challenges and ensure computer vision can be effectively and efficiently applied to monitor safety, we propose potential areas for future research in accordance with our designed and developed framework presented above.

6.1. Combining deep learning and computer vision with digital technologies

The computer vision approaches that have been developed in construction tend to have low levels of information utilization and therefore require higher levels of accuracy to detect hazards. As safety regulations become more complicated, interdependent, and more stringent due statutory requirements, prevailing computer vision approaches will be unable to identify a range of unsafe behavior and conditions. If such approaches are not able to accommodate changing regulatory requirements and the nuances of construction, then they will become redundant.

6.1.1. Ontology and computer vision

To accommodate safety regulations, we suggest that ontology should be integrated with computer vision approaches that are designed and developed. Ontology is a formal conceptualization of knowledge, which is a simplified view of a domain that describes objects, concepts, and relationships between them [60]. The purpose of an ontology is to enable computer applications to represent and reason knowledge efficiently. When combined with computer vision, objects can be automatically detected and attributes extracted from images (i.e., classes and geometry). With this in mind, a semantic computer vision-based framework can be developed that comprises four procedures: (1) ontological model of hazards (e.g., unsafe behavior and the status of plant); (2) entity and attributes detection with computer vision; (3) extraction of spatial and temporal semantic-relationship from videos; and (4) reasoning data for hazard identification.

We suggest that a semantic model that integrates ontology and computer vision can be used for hazard identification with deep learning model, even when data is scarce. The combining of ontology and computer vision not only relies on accurately detecting objects, but also the use of the spatial-temporal relationship between them to reason hazards.

Several studies have demonstrated that existing computer-vision based approaches can satisfactorily detect a variety of objects [56], which thus renders the proposed semantic approach to be potentially beneficial without having a specific database for training.

6.1.2. Group with as-built visual data, as-planned model and IoT

We suggest that there is a need to develop a deep learning visual analytics system for project performance based on an as-built three dimensional (3D) semantic reconstruction model [58,59] and as-planned' model using computer vision. In doing so, the system would be able to provide an automatic and scalable method of producing quality 3D 'as-built' models from large amounts of images and video data from various sources so that it can align with an 'as-planned' model. It would also provide the ability to visualize safety related to situational awareness and facilitate claim analysis and accident investigations. This system would be able to use field snapshots or videos aligned with an 'as-planned' model for annotation, reporting, documentation, and communication. In this system, computer vision would not only be used to extract and identify objects and their attributes (i.e., distance and classes) but also construct an 'as-built' 3D semantic model in real-time. Simultaneously, sensors can be used to extract information (i.e., locations) from installed components' enabling data from the computer vision detection results to be transformed via the IoT. Thus, data can be stored and integrated within a 3D model of a constructed asset, which can be updated continuously in real-time. In doing so, potential hazards will be able to be identified in the model, with potential structural defects or failures being able to be recognized.

6.1.3. As-built visual data, AR/VR, and building information models

AR applications have been extensively used in construction and used to supplement either virtual or the real world and have been developed to [61]:

- retrieve information both during construction and facility management for safety;
- visualize underground utilities, improve visual perception for excavation safety and subsurface utility inspection; and
- obtain real-time 3D operational instructions that are overlaid on the actual site to assist assembly and other complex operations.

Prevailing AR-capabilities and mobile devices (e.g., Apple's ARKit for iOS and Google's ARCore for Android) are sophisticated enough to support our visualization applications identified [62]. There is a need to enable a resource constraint mobile device to support the deep learning powered analytics engine. A cloud-mobile hybrid or a pure mobile method may be a solution to address the problem. An optimized deep-learning algorithm, however, is needed to reduce the computing power so that real-time information retrieval can be enabled within a mobile/cloud computing environment.

6.2. Insight from multiple data fusion

Multiple data fusion is the process of integrating data from several sources to produce more consistent, accurate, and useful information than that can be provided by an individual supplier [63]. During construction safety, data can be generated from a wide range of sources (e.g., safety reports and non-visual sensors).

6.2.1. Utilizing video streams and multi-model fusion

Deep learning has generally been undertaken within a supervised context where data is labeled, but there has been a distinct shift toward using unsupervised models to improve the detection speed of objects as well their accuracy and reliability. Based on developments that have been made within the field of computer science for deep learning and computer vision [64,66–68], we suggest that a self-taught deep learning-based unsupervised learning approach can be used to identify hazards using video streaming. Here video streams would act as inputs into a deep learning model, which has a self-taught learning mechanism with adjustable parameters to enable continuous self-training video streams that form output frames for testing. This process would, therefore, improve the ability to generalize a model's output. Systems that are trained with videos can use each successive frame as a training database, whereby the goal is to predict the next frame. For example, if the goal is to detect hazards within a frame tn, then between t1 and tn-1 would be used for training model without the need for any human labeling. This process addresses not only the issues associated with limited training data but also the problem of assuming that the distribution of training and testing database are identical.

In light of the numerous remote sensing image data resources that are readily available, and have been utilized in various applications [69–73], we suggest that these technologies can be fused to identify hazards in construction as well automatically. Such sensing image data includes, for example, by fusing thermal and 2D images, a person's unsafe behavior can be identified such as smoking in construction as the temperature of the cigarette is higher than its immediate environment. Thus, it is easy to differentiate between cigarette and non-pothole using the thermal imaging technique. Similarly, image channels (e.g., optical images) fusion can improve the identification of hazards and provide complementary visual information that is useful for deep learning models to extract detailed information [74].

6.2.2. Alignment between computer vision and text reports

We can combine text and image data to enable a deep learning model to reason and understand the nature of risk. Here, the combination of text reports and image data has two avenues for research:

1. *Leveraging reports to improve the accuracy to identify unsafe behavior*: To enable computer vision to accurately identify hazards from images, we can combine deep learning and computer vision to extract and encode images from feature representations (i.e., important regions), and then use Natural Language Processing techniques to obtain feature presentations (e.g., words and semantic relations). Finally, hazards can be retrieved and identified by using image-sentence similarity; and
2. *Automatic generation of safety reports from images*: Prevailing on-site safety inspections predominately are dependent on pen and paper to record hazards that engineers observe, and then these handwritten records are transferred into a computer system to generate safety report. This manual process of transferring data, however, is a time-consuming and error-prone process. With advances being made in image caption algorithms [75,76], we can develop semantic image captions that can enable hazard information to be automatically described. This approach can assist site managers to automatically generate risk reports rather than having to undertake site walks to identify potential hazards.

6.2.3. Alignment between computer vision and non-visual sensor data

Data from numerous sensors positioned on a construction site can be used to detect hazards. Different types of sensors have been used on construction sites to collate safety data [27,78,79]. For example, location sensors (i.e., Radio Frequency Identification and Global Positioning Systems) can be used to identify those individuals entering dangerous work areas [27]. Thus, we suggest that by fusing images obtained from multiple non-visual sensors we can extend the range of hazards that deep learning-based computer vision can detect, which include identifying:

- *Unsafe behavior*: Sensors (i.e., locations sensors, identity sensors) can be used to acquire a person's coordinates and their identity [80,81]. Then, computer vision can be used to extract information (i.e., objects classes, activity, attributes) and 3D coordinates of objects [56,74,83]. Two types of information are synchronously integrated according to the coordinates obtained. An individual's unsafe behavior record can be recorded such as their actions, frequency of events, and location, once this information is obtained, it can be used for safety training;
- *Unsafe plant*: Sensors can be integrated to obtain mechanical parameters (e.g., bending moment and angle) of a crane during hoisting. Then, using computer vision, we can identify the context and plant that contribute to the presence of unsafe conditions; and
- *Structural defects*: Several researchers have used sensor data for structural health monitoring determine the cause of defects [84,85]. Research has also combined computer vision and deep learning to recognize defects from images. However, they have been unable to determine the quality of the internal composition of a piece of civil infrastructure (i.e., tunnel and bridge). Ground-penetrating radar (GPR) is a geophysical approach that uses electromagnetic radiation in the microwave band (UHF/VHF frequencies) of the radio spectrum and detects the reflected signals from subsurface structures. Here, we suggest that a computed tomography (CT) system can be developed by integrating radar images and deep learning-based computer vision to extract information and diagnose the quality of the internal structure.

7. Conclusions

Computer vision combined with deep learning provides the capability to automatically identify unsafe behavior and conditions on construction sites and therefore can be used to improve safety performance. Nonetheless, there remain several challenges that need to be addressed before construction can directly benefit from technological developments being made within the field of computer vision. In this paper, a review that examines the use of computer vision and deep learning for monitoring of unsafe behavior and conditions is conducted to identify these challenges, which are a product of the dynamic and complex nature of construction and the difficulties associated with acquiring video surveillance data. More specifically, the dearth of databases that can be used for training and testing deep learning models to identify unsafe actions and conditions requires development to put in place a foundation for the benefits of computer vision come to the fore. Notwithstanding this limitation, we have proposed a robust enabling framework for utilizing computer vision to improve safety performance in construction. By being able to integrate state-of-the-art digital technologies and unify multiple data resources our robust computer vision-based framework acts as a signpost for engendering future research in the emergent and fertile area of deep-learning within the context of safety.

Declaration of competing interest

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgment

The authors would like to thank Editor and the five anonymous reviewers for the constructive and insightful comments that have helped improve the quality of this manuscript. We also acknowledge the financial support provided by the National Natural Science Foundation of China (Grant No. 71732001, No. 51978302, No. 51678265, No. 51878311, No. 71821001) and China Scholarship Council (CSC).

References

[1] US. Occupational Safety and Health Administration, Fatal occupational injuries counts and rates by selected industries, https://www.bls.gov/news.release/cfoi.t04.htm, (2015).
[2] UK. Office for National Statistics, Employees and Self-Employed by Industry, https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/employeesandselfemployedbyindustryemp14, (August 13, 2019).
[3] U.K. Health and Safety Executive, Fatal Injuries Arising From Accidents at Work in Great Britain, http://www.hse.gov.uk/statistics/pdf/fatalinjuries.pdf, (2017).
[4] China state administration of work safety (work safety summary in 2017), http://www.chinasafety.gov.cn/newpage/aqfx/aqfx.htm.
[5] S. Han, S. Lee, A vision-based motion capture and recognition framework for behavior-based safety management, Autom. Constr., 35 (2013) 131–141. https://doi.org/https://doi.org/10.1016/j.autcon.2013.05.001.
[6] Q. Fang, H. Li, X. Luo, L. Ding, T.M. Rose, W. An, Y. Yu, A deep learning-based method for detecting non-certified work on construction sites, Adv. Eng. Inform. 35 (2018) 56–68 https://doi.org/10.1016/j.aei.2018.01.001.
[7] Y. Lin, Z. Nie, H. Ma, Structural damage detection with automatic feature-extraction through deep learning, Computer-Aided Civil and Infrastructure Engineering 32 (2017) 1025–1046 https://doi.org/10.1111/mice.12313.
[8] D. Kim, M. Liu, S. Lee, V.R. Kamat, Remote proximity monitoring between mobile construction resources using camera-mounted UAVs, Autom. Constr. 99 (2019) 168–182, https://doi.org/10.1016/j.autcon.2018.12.014.
[9] L. Ding, W. Fang, H. Luo, P.E.D. Love, B. Zhong, X. Ouyang, A deep hybrid learning model to detect unsafe behavior: integrating convolution neural networks and long short-term memory, Autom. Constr. 86 (2018) 118–124 https://doi.org/10.1016/j.autcon.2017.11.002.
[10] W. Fang, L. Ding, H. Luo, P.E.D. Love, Falls from heights: a computer vision-based approach for safety harness detection, Autom. Constr. 91 (2018) 53–61 https://doi.org/10.1016/j.autcon.2018.02.018.
[11] W. Fang, B. Zhong, N. Zhao, P.E.D. Love, H. Luo, J. Xue, S. Xu, A deep learning-based approach for mitigating falls from height with computer vision: convolutional neural network, Adv. Eng. Inform. 39 (2019) 170–177 https://doi.org/10.1016/j.aei.2018.12.005.
[12] M.I. Jordan, T.M. Mitchell, Machine learning: trends, perspectives, and prospects, Science 349 (2015) 255 https://doi.org/10.1126/science.aaa8415.
[13] S. Agarwalla, K.K. Sarma, Machine learning based sample extraction for automatic speech recognition using dialectal Assamese speech, Neural Netw. 78 (2016) 97–111 https://doi.org/10.1016/j.neunet.2015.12.010.
[14] C.V. Guimaraes, R. Grzeszczuk, G.S. Bisset, L.F. Donnelly, Comparison between manual auditing and a natural language process with machine learning algorithm to evaluate faculty use of standardized reports in radiology, J. Am. Coll. Radiol. 15 (2018) 550–553 https://doi.org/10.1016/j.jacr.2017.10.042.
[15] X. Chen, J. Guhl, Industrial robot control with object recognition based on deep learning, Procedia CIRP 76 (2018) 149–154 https://doi.org/10.1016/j.procir.2018.01.021.
[16] A. Brunetti, D. Buongiorno, G.F. Trotta, V. Bevilacqua, Computer vision and deep learning techniques for pedestrian detection and tracking: a survey, Neurocomputing 300 (2018) 17–33 https://doi.org/10.1016/j.neucom.2018.01.092.
[17] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436 https://doi.org/10.1038/nature14539.
[18] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM 60 (2017) 84–90 https://doi.org/10.1145/3065386.
[19] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation, 2014 IEEE Conference on Computer Vision and Pattern Recogn., 2014, pp. 580–587. https://doi.org/https://doi.org/10.1109/cvpr.2014.81.
[20] Y. Guo, Y. Liu, T. Georgiou, M.S. Lew, A review of semantic segmentation using deep neural networks, International Journal of Multimedia Information Retrieval 7 (2018) 87–93 https://doi.org/10.1007/s13735-017-0141-z.
[21] H.E.B. Russell, L.K. Harbott, I. Nisky, S. Pan, A.M. Okamura, J.C. Gerdes, Motor learning affects car-to-driver handover in automated vehicles, Science Robotics, 1 (2016) eaah5682. https://doi.org/https://doi.org/10.1126/scirobotics.aah5682.
[22] P. Mehta, B. Shah, Review on techniques and steps of computer-aided skin cancer diagnosis, Procedia Computer Science 85 (2016) 309–316 https://doi.org/10.1016/j.procs.2016.05.238.
[23] K.R. Forouzan Dehbashi Sharif, Machine translation and automatic post-editing in the translation of business letters and contracts, International Journal of Scientific Research and Management 4 (2016), https://doi.org/10.18535/ijsrm/v4i10.08.
[24] http://www.tadiao168.com/News/trade-259.aspx.
[25] Q. Yu, B. Sun, H. Luo, R. Zhou, Y. Li, Statistical analysis of tower crane accidents, Construction Safety 11 (2015) 10–13 (In Chinese).
[26] S. Chi, S. Han, Analyses of systems theory for construction accident prevention with specific reference to OSHA accident reports, Int. J. Proj. Manag. 31 (2013) 1027–1041 https://doi.org/10.1016/j.ijproman.2012.12.004.

[27] H. Li, M. Lu, S.-C. Hsu, M. Gray, T. Huang, Proactive behavior-based safety management for construction safety improvement, Saf. Sci. 75 (2015) 107–117 https://doi.org/10.1016/j.ssci.2015.01.013.

[28] S.Y. Guo, L.Y. Ding, H.B. Luo, X.Y. Jiang, A big-data-based platform of workers' behavior: observations from the field, Accid. Anal. Prev. 93 (2016) 299–309 https://doi.org/10.1016/j.aap.2015.09.024.

[29] C.B. Pettinger, Improving Occupational Safety and Health Interventions: A Comparison of Safety Self-efficacy and Safety Stages of Change, Ph.D. Dissertation Virginia Polytechnic Institute and State University, 2000, https://vtechworks.lib.vt.edu/bitstream/handle/10919/27395/dissertation_new.pdf?sequence=1.

[30] H.S. Lotlikar, A Behavior-Based Safety Approach for Construction Companies, MS dissertation University of Cincinnati, 2005, https://etd.ohiolink.edu/!etd.send_file?accession=ucin1116210023&disposition=inline.

[31] C. Koch, K. Georgieva, V. Kasireddy, B. Akinci, P. Fieguth, A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, Adv. Eng. Inform. 29 (2015) 196–210 https://doi.org/10.1016/j.aei.2015.01.008.

[32] S.D. Robinson, W.J. Irwin, T.K. Kelly, X.O. Wu, Application of machine learning to mapping primary causal factors in self-reported safety narratives, Saf. Sci. 75 (2015) 118–129 https://doi.org/10.1016/j.ssci.2015.02.003.

[33] F. Zhu, Z. Ma, X. Li, G. Chen, J.-T. Chien, J.-H. Xue, J. Guo, Image-text dual neural network with decision strategy for small-sample image classification, Neurocomputing 328 (2019) 182–188 https://doi.org/10.1016/j.neucom.2018.02.099.

[34] J. Ma, Y. Ma, C. Li, Infrared and visible image fusion methods and applications: a survey, Information Fusion 45 (2019) 153–178 https://doi.org/10.1016/j.inffus.2018.02.004.

[35] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, X. Yang, Automatic pixel-level crack detection and measurement using fully convolutional network, Computer-Aided Civil and Infrastructure Engineering 33 (2018) 1090–1109 https://doi.org/10.1111/mice.12412.

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

[37] Q. Fang, H. Li, X. Luo, L. Ding, H. Luo, T.M. Rose, W. An, Detecting non-hardhat-use by a deep learning method from far-field surveillance videos, Autom. Constr. 85 (2018) 1–9 https://doi.org/10.1016/j.autcon.2017.09.018.

[38] H. Fan, H. Su, L. Guibas, A point set generation network for 3D object reconstruction from a single image, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2463–2471. https://doi.org/https://doi.org/10.1109/cvpr.2017.264.

[39] J. Zhu, J. Zhu, X. Wan, C. Wu, C. Xu, Object detection and localization in 3D environment by fusing raw fisheye image and attitude data, J. Vis. Commun. Image Represent. 59 (2019) 128–139 https://doi.org/10.1016/j.jvcir.2019.01.005.

[40] F. Lateef, Y. Ruichek, Survey on semantic segmentation using deep learning techniques, Neurocomputing 338 (2019) 321–348 https://doi.org/10.1016/j.neucom.2019.02.003.

[41] L. Qiao, Y. Qie, Z. Zhu, Y. Zhu, U.K.u. Zaman, N. Anwer, An ontology-based modeling and reasoning framework for assembly sequence planning, Int. J. Adv. Manuf. Technol., 94 (2018) 4187–4197. https://doi.org/https://doi.org/10.1007/s00170-017-1077-4.

[42] K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask R-CNN, 2017 IEEE International Conference on Computer Vision (ICCV), (2017), pp. 2980–2988, https://doi.org/10.1109/iccv.2017.322.

[43] V.S. Murahari, T. Plotz, On attention models for human activity recognition, Proceedings of the 2018 ACM International Symposium on Wearable Computers, ACM, Singapore, Singapore, 2018, pp. 100–103. https://doi.org/https://doi.org/10.1145/3267242.3267287.

[44] S. Scheidegger, J. Benjaminsson, E. Rosenberg, A. Krishnan, K. Granström, Mono-camera 3D multi-object tracking using deep learning detections and PMBM filtering, 2018 IEEE Intelligent Vehicles Symposium (IV), 2018, pp. 433–440, , https://doi.org/10.1109/ivs.2018.8500454.

[45] J. Zhu, H. Zeng, S. Liao, Z. Lei, C. Cai, L. Zheng, Deep hybrid similarity learning for person re-identification, IEEE Transactions on Circuits and Systems for Video Technology 28 (2018) 3183–3193 https://doi.org/10.1109/tcsvt.2017.2734740.

[46] Y.-J. Cha, W. Choi, O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, Computer-Aided Civil and Infrastructure Engineering 32 (2017) 361–378 https://doi.org/10.1111/mice.12263.

[47] Y. Xue, Y. Li, A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects, Computer-Aided Civil and Infrastructure Engineering 33 (2018) 638–654 https://doi.org/10.1111/mice.12367.

[48] Y. Fang, Y.K. Cho, J. Chen, A framework for real-time pro-active safety assistance for mobile crane lifting operations, Autom. Constr. 72 (2016) 367–379, https://doi.org/10.1016/j.autcon.2016.08.025.

[49] A. Zhang, K.C.P. Wang, Y. Fei, Y. Liu, S. Tao, C. Chen, J.Q. Li, B. Li, Based fully automated pavement crack detection on 3D asphalt surfaces with an improved CrackNet, J. Comput. Civ. Eng., 32 (2018) 04018041. https://doi.org/https://doi.org/10.1061/(ASCE)cp.1943-5487.0000775.

[50] Y. Gao, K.M. Mosalam, Deep transfer learning for image-based structural damage recognition, Computer-Aided Civil and Infrastructure Engineering 33 (2018) 748–768 https://doi.org/10.1111/mice.12363.

[51] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, Q. Sun, Deep learning for image-based cancer detection and diagnosis − a survey, Pattern Recogn. 83 (2018) 134–149 https://doi.org/10.1016/j.patcog.2018.05.014.

[52] J. Tompson, R. Goroshin, A. Jain, Y. LeCun, C. Bregler, Efficient object localization using convolutional networks, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 648–656, , https://doi.org/10.1109/cvpr.2015.7298664.

[53] R.M. Choudhry, Behavior-based safety on construction sites: a case study, Accid. Anal. Prev. 70 (2014) 14–23 https://doi.org/10.1016/j.aap.2014.03.007.

[54] H. Luo, W. Fang, P.E.D. Love, L. Ding, Computer vision for behaviour-based safety in construction: A review and future directions, Adv. Eng. Inform. 43 (2020) 100980, https://doi.org/10.1016/j.aei.2019.100980.

[55] S. Guo, P. Zhang, L. Ding, Time-statistical laws of workers' unsafe behavior in the construction industry: a case study, Physica A: Statistical Mechanics and Its Applications 515 (2019) 419–429 https://doi.org/10.1016/j.physa.2018.09.091.

[56] W. Fang, L. Ding, B. Zhong, P.E.D. Love, H. Luo, Automated detection of workers and heavy equipment on construction sites: a convolutional neural network approach, Adv. Eng. Inform. 37 (2018) 139–149 https://doi.org/10.1016/j.aei.2018.05.003.

[57] W. Samek, T. Wiegand, K.-R. Müller, Detection of construction workers under varying poses and changing the background in image sequences via very deep residual networks, arXiv Preprint, 99 (2017) https://arxiv.org/pdf/1708.08296.pdf.

[58] P. Tang, D. Huber, B. Akinci, R. Lipman, A. Lytle, Automatic reconstruction of as-built building information models from laser-scanned point clouds: a review of related techniques, Autom. Constr. 19 (2010) 829–843 https://doi.org/10.1016/j.autcon.2010.06.007.

[59] R. Lu, I. Brilakis, Digital twinning of existing reinforced concrete bridges from labeled point clusters, Autom. Constr. 105 (2019) 102837https://doi.org/10.1016/j.autcon.2019.102837.

[60] T. Gruber, A translation appraoch to portable ontology specifications, Knowl. Acquis. 5 (2) (1993) 199–220, https://doi.org/10.1006/knac.1993.1008.

[61] M. Chu, J. Matthews, P.E.D. Love, Integrating mobile building information modeling and augmented reality systems: An experimental study, Autom. Constr. 85 (2018) 305–316 https://doi.org/10.1016/j.autcon.2017.10.032.

[62] M. J. Kim, X. Wang, P.E.D. Love, H. Li, S. C. Kang, Virtual reality for the built environment: a critical review of recent advances, Journal of Information Technology in Construction. 18 (2), 279–305. http://www.itcon.org/2013/14.

[63] J. Llinas, D.L. Hall, An introduction to multisensor data fusion, in The Proceedings of the IEEE (1997), pp. 6–23. https://doi.org/https://doi.org/10.1109/5.554205.

[64] H. Hu, N. Phan, J. Geller, H. Vo, B. Manasi, X. Huang, S. Di Lorio, T. Dinh, S.A. Chun, Deep self-taught learning for detecting drug abuse risk behavior in tweets, Computational Data and Social Networks, Springer International Publishing, Cham, 2018, pp. 330–342, , https://doi.org/10.1007/978-3-030-04648-4_28.

[65] Marcus, G. Deep learning: A critical appraisal. (2018) arXiv preprint arXiv:1801.00631.

[66] P. Kang, D. Kim, S. Cho, Semi-supervised support vector regression based on self-training with label uncertainty: An application to virtual metrology in semiconductor manufacturing, Expert Syst. Appl. 51 (2016) 85–106 https://doi.org/10.1016/j.eswa.2015.12.027.

[67] X. Tang, F. Guo, J. Shen, T. Du, Facial landmark detection by semi-supervised deep learning, Neurocomputing 297 (2018) 22–32 https://doi.org/10.1016/j.neucom.2018.01.080.

[68] P.K. Rhee, E. Erdenee, S.D. Kyun, M.U. Ahmed, S. Jin, Active and semi-supervised learning for object detection with imperfect data, Cogn. Syst. Res. 45 (2017) 109–123 https://doi.org/10.1016/j.cogsys.2017.05.006.

[69] Aparna, Y. Bhatia, R. Rai, V. Gupta, N. Aggarwal, A. Akula, Convolutional neural networks based potholes detection using thermal imaging, Journal of King Saud University - Computer and Information Sciences, (2019). https://doi.org/https://doi.org/10.1016/j.jksuci.2019.02.004.

[70] E. Shi, Q. Li, D. Gu, Z. Zhao, A Method of Weather Radar Echo Extrapolation Based on Convolutional Neural Networks, MultiMedia Modeling, Springer International Publishing, Cham, 2018, pp. 16–28. https://doi.org/https://doi.org/10.1007/978-3-319-73603-7_2.

[71] P. Ding, Y. Zhang, W.-J. Deng, P. Jia, A. Kuijper, A light and faster regional convolutional neural network for object detection in optical remote sensing images, ISPRS J. Photogramm. Remote Sens. 141 (2018) 208–218 https://doi.org/10.1016/j.isprsjprs.2018.05.005.

[72] N. Qin, X. Hu, H. Dai, Deep fusion of multi-view and multimodal representation of ALS point cloud for 3D terrain scene recognition, ISPRS J. Photogramm. Remote Sens. 143 (2018) 205–212 https://doi.org/10.1016/j.isprsjprs.2018.03.011.

[73] V.N. Nguyen, R. Jenssen, D. Roverso, Automatic autonomous vision-based power line inspection: a review of current status and the potential role of deep learning, Int. J. Electr. Power Energy Syst. 99 (2018) 107–120 https://doi.org/10.1016/j.ijepes.2017.12.016.

[74] H. Luo, C. Xiong, W. Fang, P.E.D. Love, B. Zhang, X. Ouyang, Convolutional neural networks: computer vision-based workforce activity assessment in construction, Autom. Constr. 94 (2018) 282–289 https://doi.org/10.1016/j.autcon.2018.06.007.

[75] M.Z. Hossain, F. Sohel, M.F. Shiratuddin, H. Laga, A comprehensive survey of deep learning for image captioning, ACM Comput. Surv. 51 (2019) 1–36, https://doi.org/10.1145/3295748.

[76] U. Nadeem, S.A.A. Shah, F. Sohel, R. Togneri, M. Bennamoun, Deep learning for scene understanding, Handbook of Deep Learning Applications, Springer International Publishing, Cham, 2019, pp. 21–51, , https://doi.org/10.1016/b978-0-12-811318-9.00020-x.

[77] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?": Explaining the predictions of any classifier, (2016), https://doi.org/10.18653/v1/n16-3020 Conference of the North American Chapter of the Associate for Computational Linguistics: Demonstractions.

[78] W. Lee, K.-Y. Lin, E. Seto, G.C. Migliaccio, Wearable sensors for monitoring on-duty and off-duty worker physiological status and activities in construction, Autom.

Constr. 83 (2017) 341–353 https://doi.org/10.1016/j.autcon.2017.06.012.
[79] S. Dong, H. Li, Q. Yin, Building information modeling in combination with real-time location systems and sensors for safety performance enhancement, Saf. Sci. 102 (2018) 226–237 https://doi.org/10.1016/j.ssci.2017.10.011.
[80] T. Cheng, J. Teizer, G.C. Migliaccio, U.C. Gatti, Automated task-level activity analysis through fusion of real-time location sensors and worker's thoracic posture data, Autom. Constr. 29 (2013) 24–39 https://doi.org/10.1016/j.autcon.2012.08.003.
[81] A. Kelm, L. Laußat, A. Meins-Becker, D. Platz, M.J. Khazaee, A.M. Costin, M. Helmus, J. Teizer, Mobile passive radio frequency identification (RFID) portal for automated and rapid control of personal protective equipment (PPE) on construction sites, Autom. Constr. 36 (2013) 38–52 https://doi.org/10.1016/j.autcon.2013.08.009.
[82] U.H. Govindarajan, A.J.C. Trappey, C.V. Trappey, A.C. Chuang, J.J. Sun, A review of essential standards and patent landscapes for the Internet of Things: A key enabler for Industry 4.0, Adv. Eng. Inform. 33 (2017) 208–229, https://doi.org/10.1016/j.aei.2016.11.007.
[83] X. Luo, H. Li, D. Cao, Y. Yu, X. Yang, T. Huang, Towards efficient and objective work sampling: recognizing workers' activities in site surveillance videos with two-stream convolutional networks, Autom. Constr. 94 (2018) 360–370 https://doi.org/10.1016/j.autcon.2018.07.011.
[84] D. Yang, J. Wang, D. Li, K.S.C. Kuang, Fatigue crack monitoring using plastic optical fibre sensor, Procedia Structural Integrity 5 (2017) 1168–1175 https://doi.org/10.1016/j.prostr.2017.07.029.
[85] X. Jin, S. Yuan, J. Chen, On crack propagation monitoring by using reflection spectra of AFBG and UFBG sensors, Sensors Actuators A Phys. 285 (2019) 491–500 https://doi.org/10.1016/j.sna.2018.11.052.
[86] R.Y. Sunindijo, P.X.W. Zou, Political skill for developing construction safety climate, ASCE Journal of Construction Engineering and Management 138 (5) (2012) 605–612 https://doi.org/10.1061/(asce)co.1943-7862.0000482.