

Review

Deep Learning-Based Applications for Safety Management in the AEC Industry: A Review

Lei Hou ¹, Haosen Chen ¹, Guomin (Kevin) Zhang ¹ and Xiangyu Wang ^{2,3,*}

- ¹ School of Engineering, RMIT University, Melbourne, VIC 3000, Australia; lei.hou@rmit.edu.au (L.H.); s3523496@student.rmit.edu.au (H.C.); kevin.zhang@rmit.edu.au (G.Z.)
- ² School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330013, China
- ³ Australasian Joint Research Centre for Building Information Modelling, Curtin University, Perth, WA 6102, Australia
- * Correspondence: xiangyu.wang@curtin.edu.au; Tel.: +61-041-668-3029

Abstract: Safety is an essential topic to the architecture, engineering and construction (AEC) industry. However, traditional methods for structural health monitoring (SHM) and jobsite safety management (JSM) are not only inefficient, but also costly. In the past decade, scholars have developed a wide range of deep learning (DL) applications to address automated structure inspection and on-site safety monitoring, such as the identification of structural defects, deterioration patterns, unsafe workforce behaviors and latent risk factors. Although numerous studies have examined the effectiveness of the DL methodology, there has not been one comprehensive, systematic, evidence-based review of all individual articles that investigate the effectiveness of using DL in the SHM and JSM industry to date, nor has there been an examination of this body of evidence in regard to these methodological problems. Therefore, the objective of this paper is to disclose the state of the art of current research progress and determine the relevant gaps, challenges and future work. Methodically, CiteSpace was employed to summarize the research trends, advancements and frontiers of DL applications from 2010 to 2020. Next, an application-focused literature review was conducted, which led to a summary of research gaps, recommendations and future research directions. Overall, this review gains insight into SHM and JSM and aims to help researchers formulate more types of effective DL applications which have not been addressed sufficiently for the time being.

Keywords: machine learning; deep learning; jobsite safety management; structural health monitoring; workforce safety



Citation: Hou, L.; Chen, H.; Zhang, G.; Wang, X. Deep Learning-Based Applications for Safety Management in the AEC Industry: A Review. *Appl. Sci.* **2021**, *11*, 821. <https://doi.org/10.3390/app11020821>

Received: 1 December 2020
Accepted: 14 January 2021
Published: 16 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The architecture, engineering and construction (AEC) sector is a significant driver of economic activity around the world [1]. Structure- and workplace-related safety accidents have the potential to be life-threatening [2]. Unfortunately, these are always some of the most overlooked things in the sector. In the United States, around 40% of bridges are over 50 years old, and more than 9% of them are rated as structurally deficient, which would draw a total cost for bridge rehabilitation of around \$123 billion [3]. In addition to the need to design more robust structures under various loads [4,5], efficient structural monitoring is also important for aging infrastructure. Accurate structural health assessments are the basis for the decision-making of infrastructure maintenance, repair and rehabilitation. Typically, structure health monitoring (SHM) relates to different approaches, such as conducting regular visual inspections or relying on structural monitoring sensors [6]. Visual inspections require experienced inspectors to carry inspection instruments to reach the structure surface and conduct the inspection, and such a process can be labor-intensive, time-consuming and sometimes risky. Sensor-based monitoring can identify defects from both the structure surface and interior, and it is more reliable when the sensors are functional [7,8]. As time goes by, however, the accuracy may be compromised due to changing environments or

sensor aging problems. Under these circumstances, noise filtering approaches could be used to correct the data. However, this is also tedious and requires expertise.

Similarly, workforce safety issues in jobsite safety management (JSM) are also a challenge for the AEC industry [9]. For example, the US Occupational Health and Safety Agency (OSHA) recorded a surprisingly high death toll of 1008 construction worker fatalities in 2018 that were mainly caused by common on-site accidents, such as being struck by falling objects and falling from heights [10,11]. Traditionally, construction on-site safety monitoring relies on site patrols and surveillance [12]. However, the complex nature of site dynamics would make on-site safety monitoring more difficult and less proactive [13]. In addition, the fatigue level of workers cannot be accurately identified.

Over the past few years, researchers have been formulating various machine learning (ML) applications for various fields [14–17]. Prime ML applications in SHM and JSM include structure damage detection [18] and on-site worker safety monitoring [19]. The rapid evolution of graphics processing units (GPUs) has dramatically improved the computational capacity for processing ML algorithms, which has led to the advent of an increasing amount of deep learning (DL) applications that are underpinned by improved GPU performance [20]. In particular, the convolutional neural network (CNN), a DL algorithm, achieved extraordinary results in the ImageNET Large Scale Visual Recognition Challenge 2012 (ILSVRC2012), which is a benchmark in object classification and detection for thousands of object classes and millions of images [21]. Currently, DL has outperformed many advanced algorithms in numerous fields [20,22,23]. More and more, DL applications are being developed and deployed to address image classification, data augmentation and object detection problems [24–27]. Besides that, scholars have also made encouraging progress in integrating DL and natural language processing (NLP) for the text extraction of construction safety reports [28,29]. Through analyzing and classifying such reports, hidden dangers can be identified in time. Therefore, corresponding measures can be taken to avoid similar accidents in the future. It can be seen that ML and DL have great potential in image recognition and data analysis and are likely to be the best options to address the challenges of SHM and JSM.

To understand the research forefront in relation to applying ML and DL to address SHM and JSM, this paper provides a comprehensive literature review that covers ML- and DL-related research and development. This study gains insight into SHM and JSM and aims to help researchers formulate more types of effective DL applications which have not been addressed sufficiently for the time being. The remainder of the paper is organized as follows: Section 2 details the review methodology and scientometrics analysis; Section 3 demonstrates the SHM, JSM, DL and ML review; Section 4 denotes the limitations and future research directions; and Section 5 concludes this study.

2. Methodology

2.1. Literature Search

Over the past decade, researchers have progressively concentrated on the DL-based studies and have published an increasing amount of articles. Such an amount of research papers may challenge researchers to find the research focus and current states. To systematically identify and analyze the state-of-the-art SHM, JSM, DL and ML applications, this study used the Web of Science (WoS) as the data source for searching for articles. A variety of keywords, terms and variations were attempted, including TITLE-ABS-KEY (“deep learning” OR “machine learning” OR “convolutional neural network*” OR CNN* OR RNN OR “Recurrent neural network*”) AND TITLE-ABS-KEY (construction* OR site* OR civil* OR “AEC industr*”) AND TITLE-ABS-KEY (crack* OR “structu* health monitoring” OR SHM OR damage* OR defect* OR maintenance* OR inspection* OR behavi* OR safe* OR unsafe* OR fatigue* OR concrete* OR “computer vision*” OR “Natural Language Processing” OR NLP OR integration*). The time span of the search was from 2010 to 2020. The relevance of papers was ensured by reviewing each article’s title and abstract and excluding the irrelevant ones. Eventually, 527 papers remained.

2.2. Literature Analysis

In this study, CiteSpace [30] was used to provide a comprehensive understanding of the research hotspots and development trends for DL-based SHM and JSM. CiteSpace is a dynamic visual analysis tool which can draw knowledge maps diversely via clusters, network connectivity diagrams, nodes and so forth [31]. The uniqueness of this analysis is that emerging trends can be determined based on indicators derived by CiteSpace without domain experts' intervention or prior knowledge of the topic. Additionally, CiteSpace can expand a data set by collecting the most-cited references. This makes the data set more robust than defining the researched field with a list of predefined keywords. Figure 1 demonstrates a keyword co-occurrence network which consists of 367 keyword nodes and 1787 links generated from the literature database. The frequency of co-occurrence of the keyword is proportional to the font size of the keyword. Table 1 summarizes each keyword's occurrence. The trendy research themes are shown in Figure 2, as reflected in the keyword bursts.

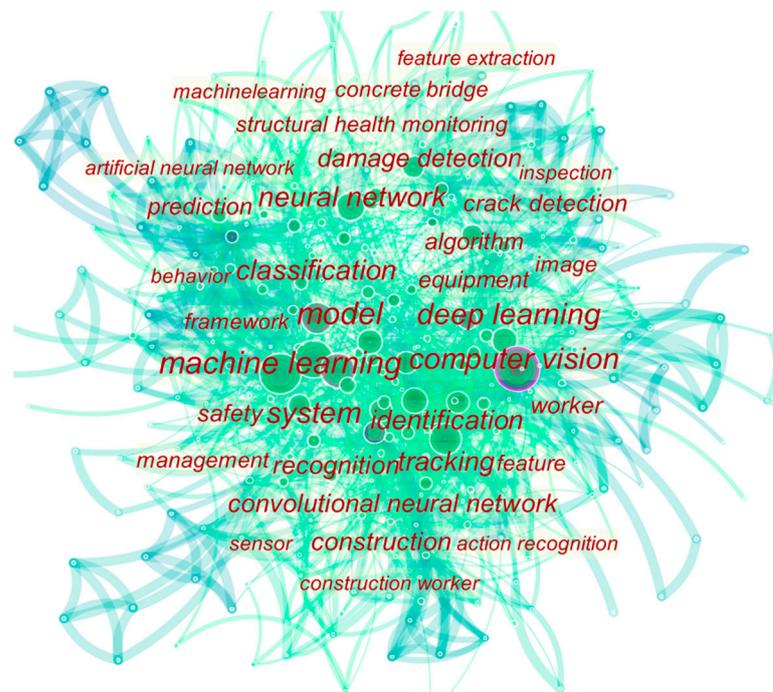


Figure 1. Co-occurrence network map.

Table 1. Keyword frequency.

Frequency	Centrality	Keyword
101	0.04	machine learning
92	0.11	model
89	0.09	system
88	0.24	computer vision
80	0.01	deep learning
60	0.15	classification
58	0.03	neural network
48	0.04	identification
44	0.05	tracking
42	0.05	recognition
39	0.04	convolutional neural network
39	0.04	damage detection
37	0.13	construction

Table 1. Cont.

Frequency	Centrality	Keyword
33	0.02	equipment
32	0.03	safety
32	0.01	algorithm
30	0.05	worker
28	0.03	crack detection
27	0.01	prediction
24	0.01	structural health monitoring

Top 6 Keywords with the Strongest Citation Bursts

Keywords	Year	Strength	Begin	End	2010 - 2020
computer vision	2010	5.34	2011	2015	
support vector machine	2010	3.11	2013	2017	
construction equipment	2010	2.72	2013	2016	
classification	2010	2.5	2014	2017	
tracking method	2010	2.49	2015	2016	
data mining	2010	2.52	2016	2017	

Figure 2. Strong citation burst keywords.

Co-citation analysis is an effective method for identifying the relationship of citations by annotating their citation and co-citation footmarks [32]. The literature co-citation network consisted of 492 nodes and 2075 links (Figure 3). Figure 4 identifies the citations with the strongest bursts. The references were sorted by starting time, and this order indicated the development of the research trend. Figure 3 shows the co-citation network clusters based on the log-likelihood ratio (LLR) test. The order number of clusters reflects the number of relevant papers published in an area. As can be seen from Figure 3, the major research results are mainly focused on cluster #0 body posture, cluster #1 DL-based roadway crack classification and cluster #2 infrastructure construction site. Therefore, the articles included in these clusters were analyzed primarily in order to identify emerging trends. Besides that, the latent semantic indexing (LSI) test and the mutual information (MI) test were also used for cluster labeling. The top-ranked clusters derived from these three algorithms are summarized in Table 2. Table 3 demonstrates the journals that were frequently cited by the acquired literature. It could be noted that Automation in Construction, Journal of Computing in Civil Engineering, Lecture Notes in Computer Science, Advanced Engineering Informatic and Computer-Aided Civil and Infrastructure Engineering were the most-cited journals in the ML, DL, SHM and JSM domains. Based on the derived clusters and keywords derived from CiteSpace, this study classified the collected papers into four categories using a systematic and manual process, namely vision-based damage detection, vibration-based damage detection, workers’ unsafe behavior detection and the analysis of construction safety documents. In the next section, the state-of-the-art ML- and DL-based technologies and applications are discussed in the corresponding categories.

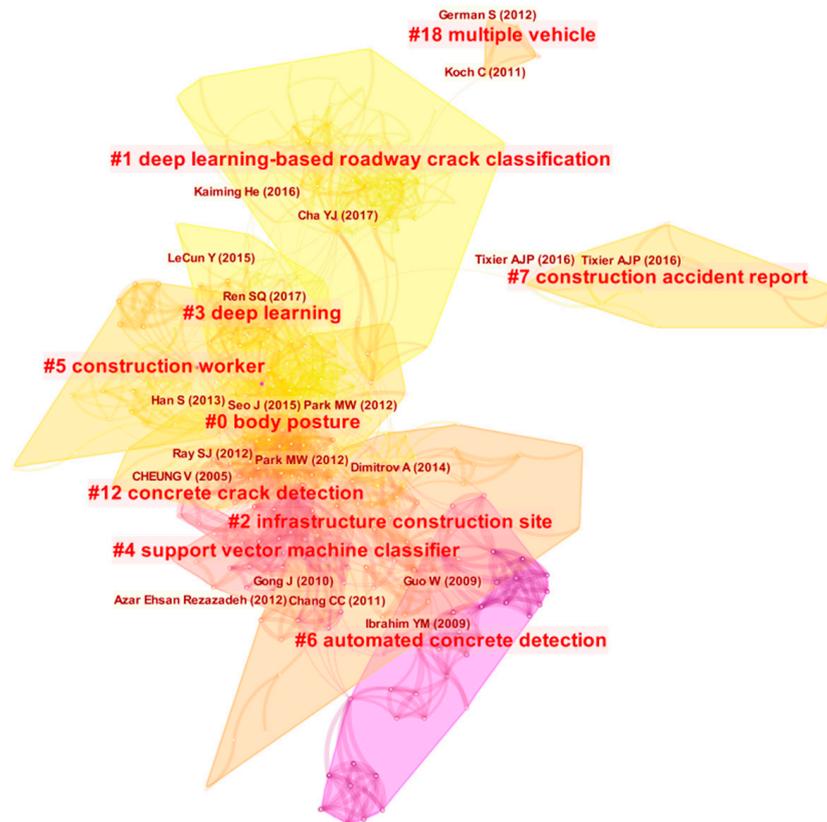


Figure 3. Cluster co-citation network.

Top 25 References with the Strongest Citation Bursts

References	Year	Strength	Begin	End	2010 - 2020
Gong J, 2010, J COMPUT CIVIL ENG, V24, P252, DOI	2010	6.08	2011	2018	██████████
Yang J, 2010, ADV ENG INFORM, V24, P428, DOI	2010	4.15	2011	2018	██████████
Brilakis I, 2011, ADV ENG INFORM, V25, P713, DOI	2011	5.92	2012	2017	██████████
Chi S, 2011, COMPUT-AIDED CIV INF, V26, P368, DOI	2011	5.48	2012	2016	██████████
Teizer J, 2009, ADV ENG INFORM, V23, P452, DOI	2009	5.44	2012	2017	██████████
Zou JH, 2007, J COMPUT CIVIL ENG, V21, P238, DOI	2007	4.13	2012	2015	██████████
Golparvar-Fard M, 2009, J INF TECHNOL CONSTR, V14, P129	2009	5.62	2013	2017	██████████
Azar ER, 2012, J COMPUT CIVIL ENG, V26, P769, DOI	2012	5.14	2013	2017	██████████
Gong J, 2011, AUTOMAT CONSTR, V20, P1211, DOI	2011	4.98	2013	2018	██████████
Park MW, 2012, J COMPUT CIVIL ENG, V26, P541, DOI	2012	4.85	2013	2017	██████████
Azar ER, 2012, AUTOMAT CONSTR, V24, P194, DOI	2012	4.3	2013	2018	██████████
Navon R, 2007, AUTOMAT CONSTR, V16, P474, DOI	2007	3.64	2013	2015	██████████
Bohn JS, 2010, J CONSTR ENG M, V136, P632, DOI	2010	3.06	2013	2017	██████████
Cheng T, 2011, AUTOMAT CONSTR, V20, P1173, DOI	2011	3.05	2013	2016	██████████
Bosche F, 2010, ADV ENG INFORM, V24, P107, DOI	2010	3.22	2014	2017	██████████
Azar ER, 2013, J CONSTR ENG M, V139, P785, DOI	2013	4.7	2015	2018	██████████
Golparvar-Fard M, 2009, J COMPUT CIVIL ENG, V23, P391, DOI	2009	4.37	2015	2017	██████████
Park MW, 2011, AUTOMAT CONSTR, V20, P905, DOI	2011	4.29	2015	2017	██████████
Memarzadeh M, 2013, AUTOMAT CONSTR, V32, P24, DOI	2013	3.79	2015	2016	██████████
Teizer J, 2015, ADV ENG INFORM, V29, P225, DOI	2015	3.52	2016	2018	██████████
Han S, 2013, J COMPUT CIVIL ENG, V27, P635, DOI	2013	3.02	2016	2020	██████████
Koch C, 2011, ADV ENG INFORM, V25, P507, DOI	2011	3.44	2017	2018	██████████
Shan Du, 2011, 2011 3rd International Conference on Computer Research and Development (ICCRD 2011), V0, P25, DOI	2011	3.3	2018	2020	██████████
Seo J, 2015, ADV ENG INFORM, V29, P239, DOI	2015	3.14	2018	2020	██████████
Ray SJ, 2012, ADV ENG INFORM, V26, P439, DOI	2012	3.09	2018	2020	██████████

Figure 4. Strong citation burst references.

Table 2. Top-ranked clusters derived from three algorithms.

Size	Silhouette	LSI	LLR	MI
82	0.844	fuzzy inference	body posture	jobsite video stream
66	0.977	crack detection	deep learning-based crack classification	rapid damage assessment
54	0.895	computer vision	infrastructure construction site	tunnel construction project
39	0.831	deep learning	deep learning-based crack classification	field surveillance video
38	0.911	construction workers	construction worker	biomechanical analysis
38	0.896	automated 2d detection	support vector machine	automated 2D detection
37	0.979	parameter optimization	automated concrete detection	deep learning
13	0.993	improved deep learning approach	construction accident report	deep learning
6	0.971	concrete crack detection	pixel-level bridge	deep learning
4	0.996	multiple vehicles	multiple vehicles	personal protective equipment

Table 3. Cited journal frequencies (Impact Factors (IFs) are based on the first cited year).

Co-Citation Count	First Cited Year	Cited Journals	Impact Factors (SCI)
271	2010	AUTOMAT CONSTR	3.13
229	2010	J COMPUT CIVIL ENG	1.42
187	2011	LECT NOTES COMPUT SC	1.17
184	2010	ADV ENG INFORM	3.10
161	2010	COMPUT-AIDED CIV INF	5.29
142	2012	J CONSTR ENG M	2.00
107	2016	SENSORS-BASEL	3.32
104	2011	INT J COMPUT VISION	6.56
97	2014	EXPERT SYST APPL	5.23
81	2018	IEEE ACCESS	4.64
77	2016	NATURE	40.13
76	2015	J MACH LEARN RES	3.13
74	2018	NEUROCOMPUTING	5.19
72	2015	COMMUN ACM	4.46
69	2016	ENG STRUCT	2.99

3. Review of DL and ML Safety Applications

3.1. SHM

3.1.1. Vision-Based Damage Detection

In recent years, researchers have used computer vision-based methods to conduct the visual inspection of surface defects and have attested considerable merits [33–35]. These methods are primarily based on image processing techniques (IPTs), such as histogram transformation, texture recognition and edge detection [1,36]. However, these methods are vulnerable to lighting condition changes and image distortion issues.

To enhance the performance of IPT-based approaches for defect detection, researchers have integrated ML algorithms [37–39]. Technically, ML algorithms can efficiently classify different damage features extracted from IPTs. ML-based methods mostly focus on identifying typical structural defects such as cracks [40–47], rusting [48,49], spalling [50,51] and loose bolts [52]. Nevertheless, these methods require defect features to be clearly defined and extracted using proper classifiers. Overall, these methods lack efficiency, feasibility and accuracy. Rapidly developing DL techniques are expected to solve the problems mentioned above. The CNN, as an end-to-end model, can improve the efficiency of defect detection and localization significantly because it can learn the defect features automatically from the labeled defects in the training samples. Normally, the process of using a CNN to determine defects in images is as follows: a fixed-size sliding window is used to scan and separate the image into small patches, and then a well-trained CNN is used to detect the defects on each small patch separately. Because the scales and shapes of defects may vary, it is difficult to find an appropriate window size to fit all kinds of them in practice.

To overcome the drawback above, a region-based CNN (R-CNN) [53] was proposed to replace the sliding windows method. The R-CNN is a two-stage detector. First, it employs a selective search approach [54,55] to generate region proposals. Then, the defect features can be extracted from the regions for classification and be highlighted by bounding boxes. However, the proposed region may be overlapped and, therefore, it may increase the computational burden [56].

To increase the efficiency of object detection, a Fast R-CNN [57] and Faster R-CNN [58] were proposed successively. Researchers have currently employed these two-stage detectors to detect the surface damages of structures. For instance, Cha et al. [56] used a Faster R-CNN to identify five classes of structural defects in both steel and concrete structures. A Faster R-CNN architecture modified by Li et al. [59] could determine the small concrete defects even in a complex image background. Although two-stage detectors can provide high accuracy, the generation of a region proposal hinders the detection speed. Therefore, two-stage detectors are hard to use for achieving real-time detection.

Due to the above limitations, single-stage detectors (e.g., the single shot Multibox detector (SSD) [60] and you only look once (YOLO) [61]) were proposed by combining object classification and localization into one single convolution network. The removal of the proposal generating step is the main feature of single-stage detectors. Besides that, they can predict multiple bounding boxes simultaneously. Hence, the detection speed of single-stage detectors can be improved significantly. Similarly, YOLOv3 was used by Zhang et al. [5] to detect the cracks, pop-out, delamination and exposed rebar of the bridge with relatively high accuracy.

However, using bounding boxes to identify the damage may not be suitable, as such rectangular boxes cannot determine the boundaries of the crack textures accurately. In this case, pixel-level semantic segmentation, which assigns classification labels to each image pixel instead of generating bounding boxes, allows for better defect localization and analysis [2]. A fully convolutional network (FCN) [19] is a pixel-by-pixel network for semantic segmentation. It predicts the class (crack or non-crack) of each pixel by using a deconvolution layer to upsample the last convolutional layer. Compared with the methods above, the pixel-level method can completely separate the damage feature from the background by highlighting each pixel of the defect. At present, the FCN method has already been exploited to identify concrete cracks [55,62] and other structure defects [63–65]. For example, Zhang et al. [66] proposed an FCN-based, pixel-level asphalt pavement crack detector. Similarly, Bang et al. [67] employed an FCN as a part of their developed network for road crack detection based on digital images produced by black box cameras. The results proved that an FCN is optimal for defect information segmentations. Except for segmenting the crack, it can also provide valuable damage information, such as crack widths for damage assessment. When annotating defects in an image sample, a small portion of non-cracked surfaces is likely to be labeled as defects. Such random errors are inevitable and not easy to measure. To reduce the impact of such uncertainty, Tong et al. [68] combined an FCN and a Gaussian-conditional random field for pavement defect detection. The developed framework can address the uncertainty of defect labeling.

3.1.2. Vibration-Based Damage Detection

Although a pixel-level representation of structural defects is beneficial for SHM, it can only identify the damage level on the structure surface and is not competent to infer the performance of internal structural components which may have been deteriorated in advance [69]. Vibration data is the main type of source of data utilized in SHM. Technically, any structural damage will change the stiffness and mass distributions of the structure and lead to differences in the natural frequencies and mode shapes [70,71]. Hence, vibration-based SHM methods have the potential to detect internal structural damages by analyzing the abnormal data acquired from the sensors (e.g., accelerometers). The previous research of vibration-based SHM mainly focused on setting up a real physical model to imitate the status of a real structure. Basically, this model-driven method employs mathematical

modeling and physical laws to represent the monitored structure [72]. Hence, the level and location of the damage can be determined accurately by analyzing and solving the model. Nevertheless, it is challenging to build and solve such a complicated model when the complexity of the monitored structure increases and the environmental factors are considered. Currently, model-driven methods have been progressively replaced by data-driven methods [73,74]. The most critical drawback of the model-driven approach is that modeling usually requires expertise and is time-consuming. Unlike the model-driven method, the data-driven method can identify the anomaly data directly by measuring the data collected from the sensors. Most of the data-driven method is based on the ML paradigm [75]. As the appropriate sensors' layouts can improve the efficiency and accuracy of data collection and transmission, ML algorithms, such as a genetic algorithm (GA) [76], have also been used for the determination of optimal sensor layouts. However, when applying vibration-based SHM methods in practice, the natural frequencies of the structure are easily affected by environmental factors (e.g., temperature) [77,78]. For example, if a structure has some small-scale damages, the changes in the natural frequency of the structure would possibly be suppressed by those environmental variables. Some scholars have conducted several analyses on the evolution of structural properties and their relationship with changes in environmental parameters [79,80]. Among them, the monitoring of the Z24 bridge is emblematic for addressing this issue [81]. Although significant efforts have been made in this regard, it requires comprehensive expertise and is time-consuming [82].

Currently, DL methods can potentially remedy the issue, as they can fully utilize the sensor data by automatically extracting the data features. Therefore, without expertise, even a delicate anomaly can be perceived. Recently, Ni et al. [83] presented a 1D CNN-based algorithm along with autoencoder data compression to identify anomaly data in a long-span suspension bridge. The results showed that the developed algorithm could achieve a precision of 97.53%. Similarly, Avci et al. [84] presented a 1D CNN-based method to detect the structural damages on a steel frame by using wireless sensor data. Azimi and Pekcan [85] introduced a novel CNN-based approach which could detect and locate the damage in a large-scale structure. Lin et al. [86] trained a deep CNN by feeding it unrefined sensor data and applied it for identifying the damage of a simulated beam structure. The results revealed that the trained CNN could detect structural damage with high accuracy, even if the test data were noisy. Zhang et al. [2] proposed a CNN to detect structural stiffness and mass changes. The developed CNN achieved good results on both the in-lab structure and the in-service bridge. Gulgec et al. [87] used a DL-based method for steel fatigue assessment. Compared to the traditional method, which is costly and laborious, their proposed method could achieve a high detection accuracy with a low cost.

Vibration-based methods integrated with DL have proven that they can perform well in damage detection, and the costs of these methods are relatively low. However, these methods still need much effort in data labeling, which is tedious. There is a novel application using an unsupervised learning method, which does not require data labeling. Guo et al. [88] presented an unsupervised learning method called a sparse coding algorithm for SHM. Sparse coding was employed to learn the feature representations from unstructured vibration data to improve the performance of damage detection. Through simulation, different degrees of damage were conducted, and the results showed that the proposed method outperformed other ML methods, such as logistic regression and decision trees, with a precision of 98%. However, this application was only validated on the simulated model. As the performance on real-world structures has yet to be verified, this may become a notable topic for future research.

3.2. JSM

3.2.1. Workers' Unsafe Behavior Detection

On-site surveillance videos or images have been used for automated unsafe behavior detection in recent years. Variables such as hard hats, safety vests and workers can be

detected by using certain computer vision techniques (e.g., a background subtraction algorithm [89], the histograms of oriented gradients (HOG) method [90], and the scale-invariant feature transform (SIFT) [91]). Nowadays, such methods which require much work for feature extraction are being replaced by DL gradually.

Mneymneh et al. [92] developed a CNN-based framework that could determine if workers (even they are moving) were wearing hard hats on the construction site. Xie et al. [93] modified a CNN to detect workers' hard hats, and the model produced excellent results in the mean average precision (mAP) performance metric. Similarly, the Faster R-CNN [94] and SSD methods [95] were also employed to detect hard hats.

Fang et al. [96] modified the Faster RCNN to identify if workers equipped harnesses properly. Kolar et al. [97] employed a VGG-16 model to detect if safety guardrails were installed correctly to prevent workers from falling from heights. Siddula et al. [98] integrated a Gaussian mixture model (GMM) with CNNs to detect roofers on roof construction sites. This research can alleviate roof site fall risks.

In the unsafe activities identification area, Ding et al. [19] coupled a long short-term memory (LSTM) model [99] with CNNs to identify if the worker would climb a ladder unsafely [19]. Kim et al. [100] developed an image-based risk prevention system to display the safety-related information of each construction worker on a wearable augmented reality (AR) device. Luo et al. [101] utilized a Faster R-CNN to determine workers' activities based on construction site images. Considering that temporal information is necessary for dynamic activities detection, Luo et al. [102] later improved the framework for video-based worker activity recognition by helping the temporal information emerge. Some researchers have also investigated construction vehicle detection using DL. Kim et al. [103] employed a region-based FCN to detect construction vehicles. Fang et al. [104] used a Faster R-CNN to identify the spatial relationship of workers and excavators on construction sites. This study provided a basic prototype of the site safety alert system, which can prevent workers from being hit by heavy equipment. Son et al. [105] used a Faster R-CNN to identify on-site workers in diverse poses against complex backgrounds.

3.2.2. Analysis of Construction Safety Documents

On-site managers can benefit from analyzing safety reports, as they can acquire details about the events and circumstances that result in safety accidents. Hence, corresponding actions can be taken to prevent similar accidents in the future [106]. With the development of safety management, scholars have proposed various text classification methods to classify and analyze accident causation [107]. More recently, researchers have tried to apply various ML algorithms to construction-related accident analysis. Tixier et al. [108] developed a predictive model which used random forest (RF) [109] and stochastic gradient tree boosting (SGTB) methods [110]. This model can forecast the different types of injuries recorded by on-site injury reports. Later, Tixier et al. [111] presented an improved NLP method to extract outcome variables and injury precursors from the unstructured injury-related text. This method can reduce the labor cost of text analysis. Chokor et al. [112] adopted an unsupervised learning approach for injury report classification based on K-means clustering. The text mining approach presented outstanding results regarding recall and precision. Goh et al. [113] applied six ML approaches to classify the near-miss accident reports. The results illustrated that ML algorithms performed better than traditional text classification methods. However, ML requires hand-crafted featuring, which limits the generalization of the classifier and also may affect the adaptability.

DL-based methods perform excellently, in regard to text mining and classification, when compared with traditional ML [114]. Word2Vec [115] is a popular DL method that adopts word embedding technology. Word2Vec can be trained by public semantic resources such as Wikipedia when utilizing it for text classification [116]. It avoids the manual featuring process and augments knowledge concurrently. However, there is relatively limited research focusing on this topic. Future research can be conducted by combining Word2Vec and computer vision in JSM. For example, CNN can be used to detect unsafe

events from the construction site images taken from surveillance cameras. Then, image captioning technology provided by Word2Vec can be used to explain the event in text formation. In this way, an automated safety report generation system based on construction images can be developed.

4. Challenges and Recommendations for Future Work

4.1. Limitations of the Data Set and Weak Generalization of DL Models

To achieve a high detection or classification accuracy, normally a massive labeled image is needed to train a deep neural network well. However, there are few civil engineering-related data sets available to the public, and it is not easy to obtain such data. To this end, it is common to pre-train a CNN model with an extensive common object dataset (e.g., ImageNet, MNIST and CIFAR-10). Normally, most of the popular pre-trained CNN models (e.g., VGG-16, ResNet50, Inceptionv3, and YOLOv3) are available online. Such models with pre-trained weight have already learned to extract the basic features of the images. Researchers can download these models as the backbone of their DL structures for conducting specific tasks. This process is called transfer learning [20,117], and many DL-based applications have employed it in practice. However, a pre-trained model is still required to learn the nuanced features of the specific targets. Given the lack of available datasets for training, researchers in construction are normally required to create their own by labeling images manually. This process is time-consuming, tedious and expensive. Since the specific database has to be created, it tends to be limited and, thus, can hinder the generalization ability of the model. To this end, data augmentation methods can be employed to upsize the dataset synthetically. In general, those methods include spatial flipping, cropping, bending and other deformations of the images [118]. Typically, the semantic meaning of the labels will not be changed by these augmentation approaches. Therefore, the newly generated images are highly associated with the original ones. In this way, the generalization of the DL model may not be improved effectively.

In order to tackle these issues, a generative adversarial network (GAN) may be a possible option. GANs were first introduced by Goodfellow et al. [119] and can be used as data generation models. A GAN consists of a generator and a discriminator. The principle of a GAN is to let the discriminator evaluate the new data produced by the generator. As the recurrent training progresses, the performance can be improved gradually during the process. This kind of unsupervised learning has been proven to be powerful in many applications. In recent years, scholars start using GANs to produce synthetic images. Compared to the data augmentation methods, the images generated by GANs are more diverse and distinct [5]. In addition, researchers have used the advantages of GANs to generate face images. For example, Jin et al. [120] applied a GAN to anime character creation. The GAN demonstrated stable performance in the innovative face creation of anime characters. Zhang et al. [5] designed a GAN which illustrated good outcomes in natural scene images. Besides those examples, Kitchen and Andy [121] leveraged a GAN in the health sector. They applied a deep GAN to produce synthetic prostate lesion images and utilized these images to upsize their training dataset. However, there are relatively few GAN-related studies conducted in the AEC safety industry. Implementing a GAN in practice might be a possible solution for tackling the construction-related data set challenge.

Similarly, there are few public vibration data sets available for researchers to train their classifiers. Hence, future research should focus on finding an alternative method to train those classifiers when sensor data is insufficient. One recommended method for building up training data sets is combining real-life data from an undamaged structure with simulated data for the damage scenarios collected from an in-lab model. Besides that, as more and more emerging algorithms come out, unsupervised learning that can run on a small amount of labeled and unlabeled sensor data may become a hot topic in the future.

4.2. Directions of Future Studies

With the advances of powerful cloud-based computing and the DL platform, the importance of data is more obvious than ever. However, there are still many types of data from SHM and JSM that have not been fully utilized by ML and DL technologies. Considering the development of emerging technologies and underutilized data types, the recent development direction of DL-based SHM and JSM research can be divided into the following categories.

4.2.1. DL-Based Seismic Vibration Control for SHM

Seismic vibration control is, in conjunction with SHM, another very important field for safety management [122–124]. In order to improve the stability of the structure when an earthquake occurs, the active controller and the passive base-isolated systems are generally integrated together [125]. Some scholars have started to use ML algorithms to evaluate the failure probability of base-isolated systems [122] and design active control schemes [123]. However, sensor malfunction and time delay are among the major concerns that impact a vibration control system's performance under seismic vibration [59,126,127]. The LSTM model has been proven to be effective in distinguishing long-term dynamical dependencies over sequential frames [99]. Therefore, it is possible to predict the response of structures to overcome the breakdown of signal and time delay issues by using LSTM networks, which results in developing robust and efficient control systems.

4.2.2. Visual-Based On-Site Fatigue Monitoring

In the field of JSM, fatigue monitoring is also important. Given the increasing complexity of on-site dynamics, the heavy equipment operators, as well as their operations and judgment, play an essential role in ensuring safety and productivity on the construction site. Nevertheless, as the work becomes more intense, their cognitive awareness may become more jeopardized, which may lead to a safety hazard. It is noteworthy that Tam and Fung [128] revealed that approximately 60.5% of crane operators would continue to work even feeling fatigued after long working hours under a tight construction schedule, and about 52.6% of the crane operators experienced a lack of breaks, as they found it inconvenient to frequently move in or out within the narrow workspace. Hence, automated fatigue monitoring and warnings can provide timely support for this cohort. As the fatigue condition of heavy equipment operators and workers is relatively easy to be detected via analyzing their facial expressions, the real-time monitoring of their faces through cameras can be an effective, feasible and non-invasive method to identify drowsiness and avoid accidents [129]. Currently, CNNs have been used in real-time fatigue monitoring of on-road driving [130–132]. Some researchers have employed CNNs to extract video-level features and then integrate them into LSTM models to analyze the temporal information for fatigue identification [133,134]. However, the potential of a CNN-LSTM model regarding fatigue monitoring of heavy equipment operators is far from being comprehensively explored. Therefore, these DL-based applications have the potential to be leveraged to detect driver fatigue in the construction field.

4.2.3. Possible Integrations with Other Digital Technologies

Future research can integrate DL with other novel technologies, such as AR and virtual reality (VR) [135,136], building information modeling (BIM) [137–140], 5G technology and the Internet of things (IoT) [64], into real-time SHM and JSM to extract information from the massive amounts of data continuously received from wireless sensor networks. Such research can further expand the scope of SHM and JSM through various advanced sensors to provide decision-making assistance accurately. Being capable of interpreting unstructured data in large volumes, DL technology can be employed to facilitate the entire integrated system cost-effectively and intuitively. Future research can propose such a new paradigm in relation to IoT-based sensor data collection, transformation and visualization as part of the DL-based SHM and JSM applications. Nowadays, unmanned aerial vehicles

(UAVs), mobile phones and AR devices (e.g., HTC VIVE Focus Plus, Microsoft Hololens and Google Glass) have become an integral part of human activities. These compact devices can normally collect digital data and have good computing and network technology. Moreover, most of these devices are programmable, having the potential to achieve real-time and cloud-based assessments. Currently, some popular CNNs (e.g., MobileNet-SSD and YOLOv3-Lite) have been proposed. These networks do not require a high computing power and can maintain relatively stable performance. Therefore, it is possible to conduct on-site safety inspection through these mobile devices. With the advent of the 5G era, low-latency collaboration and intelligent optimization become possible. For example, UAVs can be paired with mobile phones to obtain and process image data while embedding light-weight DL algorithms. In addition, AR technologies have been broadly exploited in construction through superimposing virtual elements onto the construction site [84,141,142]. By integrating these advanced digital devices and technologies with DL, it is possible to establish a real-time hazard alerting system. First, site images or videos are transmitted to the mobile phone continuously by using UAVs. Then, a mobile phone with DL algorithms installed can be used to detect hazardous activities based on these image-based data. When a worker enters the hazardous area (e.g., working too close to the construction vehicles), the AR device worn by the worker can provide timely alerts through popping up a warning in the worker's point of view. However, current lightweight DL algorithms are only qualified for detecting and localizing objects with apparent features in the images; things such as cracks and minor defects of structures with subtle features may not be applicable in this case. Therefore, an optimized version of the DL structure is still needed in the future.

5. Conclusions

DL-based applications in the AEC safety industry are becoming more and more widespread [143,144]. The purpose of this paper was to summarize the past decade of research in SHM- and JSM-based on ML and DL applications and offer possible solutions for current research challenges. First, this study started with scientometrics analysis by using CiteSpace to visualize the knowledge map of ML and DL applications in the AEC safety industry. Second, this study reviewed the related state-of-the-art literature and identified the main challenges of current research. Additionally, possible suggestions for future research directions were provided. It is believed that this comprehensive review can inspire researchers to develop more types of practical DL applications in the future.

Author Contributions: L.H., H.C., G.Z. and X.W. contribute to literature retrieval, literature classification, and generalization of major knowledge. L.H. and H.C. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: The authors would like to acknowledge KAEFER Integrated Services for their sponsorship of this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Koch, C.; Georgieva, K.; Kasireddy, V.; Akinci, B.; Fieguth, P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Inform.* **2015**, *29*, 196–210. [[CrossRef](#)]
2. Gholipour, G.; Zhang, C.; Mousavi, A.A. Nonlinear numerical analysis and progressive damage assessment of a cable-stayed bridge pier subjected to ship collision. *Mar. Struct.* **2020**, *69*, 102662. [[CrossRef](#)]
3. Spencer, B.F., Jr.; Hoskere, V.; Narazaki, Y. Advances in computer vision-based civil infrastructure inspection and monitoring. *Engineering* **2019**, *5*, 199–222. [[CrossRef](#)]

4. Sun, L.; Yang, Z.; Jin, Q.; Yan, W. Effect of Axial Compression Ratio on Seismic Behavior of GFRP Reinforced Concrete Columns. *Int. J. Struct. Stab. Dyn.* **2020**, *20*, 2040004. [[CrossRef](#)]
5. Abedini, M.; Mutalib, A.A.; Zhang, C.; Mehrmashhadi, J.; Raman, S.N.; Alipour, R.; Momeni, T.; Mussa, M.H. Large deflection behavior effect in reinforced concrete columns exposed to extreme dynamic loads. *Front. Struct. Civ. Eng.* **2020**, *14*, 532–553. [[CrossRef](#)]
6. Flah, M.; Nunez, I.; Chaabene, W.B.; Nehdi, M.L. Machine learning algorithms in civil structural health monitoring: A systematic review. *Arch. Comput. Methods Eng.* **2020**, 1–23.
7. Sun, L.; Li, C.; Zhang, C.; Liang, T.; Zhao, Z. The strain transfer mechanism of fiber bragg grating sensor for extra large strain monitoring. *Sensors* **2019**, *19*, 1851. [[CrossRef](#)]
8. Kordestani, H.; Zhang, C. Direct use of the savitzky–golay filter to develop an output-only trend line-based damage detection method. *Sensors* **2020**, *20*, 1983. [[CrossRef](#)]
9. Fang, W.; Ding, L.; Love, P.E.D.; Luo, H.; Li, H.; Peña-Mora, F.; Zhong, B.; Zhou, C. Computer vision applications in construction safety assurance. *Autom. Constr.* **2020**, *110*, 100980. [[CrossRef](#)]
10. Prasanna, P.; Dana, K.; Gucunski, N.; Basily, B. Computer-vision based crack detection and analysis. In Proceedings of the Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, San Diego, CA, USA, 12–15 March 2012; International Society for Optics and Photonics: Bellingham, WA, USA, 2012; p. 834542.
11. Sousa, V.; Almeida, N.M.; Dias, L.A. Risk-based management of occupational safety and health in the construction industry—Part 1: Background knowledge. *Saf. Sci.* **2014**, *66*, 75–86. [[CrossRef](#)]
12. Luo, X.; Li, H.; Yang, X.; Yu, Y.; Cao, D. Capturing and Understanding Workers’ Activities in Far-Field Surveillance Videos with Deep Action Recognition and Bayesian Nonparametric Learning. *Comput. Aided Civ. Infrastruct. Eng.* **2019**, *34*, 333–351. [[CrossRef](#)]
13. Wang, M.; Wong, P.; Luo, H.; Kumar, S.; Delhi, V.; Cheng, J. Predicting Safety Hazards Among Construction Workers and Equipment Using Computer Vision and Deep Learning Techniques. In Proceedings of the International Symposium on Automation and Robotics in Construction (ISARC), Banff Alberta, AB, Canada, 21–24 May 2019; IAARC Publications: Banff Alberta, AB, Canada, 2019; pp. 399–406.
14. Amezcua-Sanchez, J.P.; Park, H.S.; Adeli, H. A novel methodology for modal parameters identification of large smart structures using MUSIC, empirical wavelet transform, and Hilbert transform. *Eng. Struct.* **2017**, *147*, 148–159. [[CrossRef](#)]
15. Li, H.; Song, D.; Liu, Y.; Li, B. Automatic Pavement Crack Detection by Multi-Scale Image Fusion. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 2025–2036. [[CrossRef](#)]
16. Zhang, H.; Qiu, Z.; Cao, J.; Abdel-Aty, M.; Xiong, L. Event-Triggered Synchronization for Neutral-Type Semi-Markovian Neural Networks With Partial Mode-Dependent Time-Varying Delays. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *31*, 4437–4450. [[CrossRef](#)]
17. Yue, H.; Wang, H.; Chen, H.; Cai, K.; Jin, Y. Automatic detection of feather defects using Lie group and fuzzy Fisher criterion for shuttlecock production. *Mech. Syst. Signal Process.* **2020**, *141*, 106690. [[CrossRef](#)]
18. Oliveira, H.; Correia, P.L. Supervised strategies for cracks detection in images of road pavement flexible surfaces. In Proceedings of the IEEE 2008 16th European Signal Processing Conference, Lausanne, Switzerland, 25–29 August 2008; pp. 1–5.
19. Ding, L.; Fang, W.; Luo, H.; Love, P.E.D.; Zhong, B.; Ouyang, X. A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Autom. Constr.* **2018**, *86*, 118–124. [[CrossRef](#)]
20. Chen, H.; Chen, A.; Xu, L.; Xie, H.; Qiao, H.; Lin, Q.; Cai, K. A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. *Agric. Water Manag.* **2020**, *240*, 106303. [[CrossRef](#)]
21. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
22. Qian, J.; Feng, S.; Tao, T.; Hu, Y.; Li, Y.; Chen, Q.; Zuo, C. Deep-learning-enabled geometric constraints and phase unwrapping for single-shot absolute 3D shape measurement. *APL Photonics* **2020**, *5*, 046105. [[CrossRef](#)]
23. Qiu, T.; Shi, X.; Wang, J.; Li, Y.; Qu, S.; Cheng, Q.; Cui, T.; Sui, S. Deep Learning: A Rapid and Efficient Route to Automatic Metasurface Design. *Adv. Sci.* **2019**, *6*, 1900128. [[CrossRef](#)]
24. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)] [[PubMed](#)]
25. Koziarski, M.; Cyganek, B. Image recognition with deep neural networks in presence of noise—dealing with and taking advantage of distortions. *Integr. Comput. Aided Eng.* **2017**, *24*, 337–349. [[CrossRef](#)]
26. Molina-Cabello, M.A.; Luque-Baena, R.M.; Lopez-Rubio, E.; Thurnhofer-Hemsi, K. Vehicle type detection by ensembles of convolutional neural networks operating on super resolved images. *Integr. Comput. Aided Eng.* **2018**, *25*, 321–333. [[CrossRef](#)]
27. Wang, P.; Bai, X. Regional parallel structure based CNN for thermal infrared face identification. *Integr. Comput. Aided Eng.* **2018**, *25*, 247–260. [[CrossRef](#)]
28. Lukic, D.; Littlejohn, A.; Margaryan, A. A framework for learning from incidents in the workplace. *Saf. Sci.* **2012**, *50*, 950–957. [[CrossRef](#)]
29. Sanne, J.M. Incident reporting or storytelling? Competing schemes in a safety-critical and hazardous work setting. *Saf. Sci.* **2008**, *46*, 1205–1222. [[CrossRef](#)]

30. Chen, C.; Song, M. Visualizing a field of research: A methodology of systematic scientometric reviews. *PLoS ONE* **2019**, *14*, e0223994. [[CrossRef](#)]
31. Chen, C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 359–377. [[CrossRef](#)]
32. Chen, C. The citespace manual. *Coll. Comput. Inform.* **2014**, *1*, 1–84.
33. Rose, P.; Aaron, B.; Tamir, D.E.; Lu, L.; Hu, J.; Shi, H. Supervised computer-vision-based sensing of concrete bridges for crack-detection and assessment. In Proceedings of the Transportation Research Board 93rd Annual Meeting, Washington, DC, USA, 12–16 January 2014.
34. Nayyeri, F.; Hou, L.; Zhou, J.; Guan, H. Foreground–background separation technique for crack detection. *Comput. Aided Civ. Infrastruct. Eng.* **2019**, *34*, 457–470. [[CrossRef](#)]
35. Chan, B.; Guan, H.; Hou, L.; Jo, J.; Blumenstein, M.; Wang, J. Defining a conceptual framework for the integration of modelling and advanced imaging for improving the reliability and efficiency of bridge assessments. *J. Civ. Struct. Health Monit.* **2016**, *6*, 703–714. [[CrossRef](#)]
36. Jiang, Q.; Shao, F.; Gao, W.; Chen, Z.; Jiang, G.; Ho, Y. Unified No-Reference Quality Assessment of Singly and Multiply Distorted Stereoscopic Images. *IEEE Trans. Image Process.* **2019**, *28*, 1866–1881. [[CrossRef](#)] [[PubMed](#)]
37. Amezcua-Sanchez, J.P.; Adeli, H. Synchrosqueezed wavelet transform-fractality model for locating, detecting, and quantifying damage in smart highrise building structures. *Smart Mater. Struct.* **2015**, *24*, 065034. [[CrossRef](#)]
38. Butcher, J.B.; Day, C.R.; Austin, J.C.; Haycock, P.W.; Verstraeten, D.; Schrauwen, B. Defect detection in reinforced concrete using random neural architectures. *Comput. Aided Civ. Infrastruct. Eng.* **2014**, *29*, 191–207. [[CrossRef](#)]
39. Jiang, X.; Adeli, H. Pseudospectra, MUSIC, and dynamic wavelet neural network for damage detection of highrise buildings. *Int. J. Numer. Methods Eng.* **2007**, *71*, 606–629. [[CrossRef](#)]
40. Feng, D.; Feng, M.Q. Experimental validation of cost-effective vision-based structural health monitoring. *Mech. Syst. Signal Process.* **2017**, *88*, 199–211. [[CrossRef](#)]
41. Torok, M.M.; Golparvar-Fard, M.; Kochersberger, K.B. Image-based automated 3D crack detection for post-disaster building assessment. *J. Comput. Civ. Eng.* **2014**, *28*, A4014004. [[CrossRef](#)]
42. Yamaguchi, T.; Hashimoto, S. Fast crack detection method for large-size concrete surface images using percolation-based image processing. *Mach. Vis. Appl.* **2010**, *21*, 797–809. [[CrossRef](#)]
43. Yeum, C.M.; Dyke, S.J. Vision-based automated crack detection for bridge inspection. *Comput. Aided Civ. Infrastruct. Eng.* **2015**, *30*, 759–770. [[CrossRef](#)]
44. Yoon, H.; Elanwar, H.; Choi, H.; Golparvar-Fard, M.; Spencer, B.F., Jr. Target-free approach for vision-based structural system identification using consumer-grade cameras. *Struct. Control Health Monit.* **2016**, *23*, 1405–1416. [[CrossRef](#)]
45. O’Byrne, M.; Schoefs, F.; Ghosh, B.; Pakrashi, V. Texture analysis based damage detection of ageing infrastructural elements. *Comput. Aided Civ. Infrastruct. Eng.* **2013**, *28*, 162–177. [[CrossRef](#)]
46. Nishikawa, T.; Yoshida, J.; Sugiyama, T.; Fujino, Y. Concrete crack detection by multiple sequential image filtering. *Comput. Aided Civ. Infrastruct. Eng.* **2012**, *27*, 29–47. [[CrossRef](#)]
47. Zalama, E.; Gómez-García-Bermejo, J.; Medina, R.; Llamas, J. Road crack detection using visual features extracted by Gabor filters. *Comput. Aided Civ. Infrastruct. Eng.* **2014**, *29*, 342–358. [[CrossRef](#)]
48. Koch, C.; Paal, S.G.; Rashidi, A.; Zhu, Z.; König, M.; Brilakis, I. Achievements and challenges in machine vision-based inspection of large concrete structures. *Adv. Struct. Eng.* **2014**, *17*, 303–318. [[CrossRef](#)]
49. Chen, P.-H.; Shen, H.-K.; Lei, C.-Y.; Chang, L.-M. Support-vector-machine-based method for automated steel bridge rust assessment. *Autom. Constr.* **2012**, *23*, 9–19. [[CrossRef](#)]
50. Dawood, T.; Zhu, Z.; Zayed, T. Machine vision-based model for spalling detection and quantification in subway networks. *Autom. Constr.* **2017**, *81*, 149–160. [[CrossRef](#)]
51. German, S.; Brilakis, I.; DesRoches, R. Rapid entropy-based detection and properties measurement of concrete spalling with machine vision for post-earthquake safety assessments. *Adv. Eng. Inform.* **2012**, *26*, 846–858. [[CrossRef](#)]
52. Cha, Y.-J.; You, K.; Choi, W. Vision-based detection of loosened bolts using the Hough transform and support vector machines. *Autom. Constr.* **2016**, *71*, 181–188. [[CrossRef](#)]
53. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587.
54. Uijlings, J.R.; Van De Sande, K.E.; Gevers, T.; Smeulders, A.W. Selective search for object recognition. *Int. J. Comput. Vis.* **2013**, *104*, 154–171. [[CrossRef](#)]
55. An, Y.; Chatzi, E.; Sim, S.H.; Laflamme, S.; Blachowski, B.; Ou, J. Recent progress and future trends on damage identification methods for bridge structures. *Struct. Control Health Monit.* **2019**, *26*, e2416. [[CrossRef](#)]
56. Cha, Y.J.; Choi, W.; Suh, G.; Mahmoudkhani, S.; Büyüköztürk, O. Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 731–747. [[CrossRef](#)]
57. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.

58. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1137–1149. [[CrossRef](#)] [[PubMed](#)]
59. Li, R.; Yuan, Y.; Zhang, W.; Yuan, Y. Unified vision-based methodology for simultaneous concrete defect detection and geolocalization. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 527–544. [[CrossRef](#)]
60. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
61. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
62. Yang, X.; Li, H.; Yu, Y.; Luo, X.; Huang, T.; Yang, X. Automatic pixel-level crack detection and measurement using fully convolutional network. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1090–1109. [[CrossRef](#)]
63. Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
64. Cui, Z.; Ke, R.; Pu, Z.; Wang, Y. Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. *arXiv* **2018**, arXiv:1801.02143.
65. Xue, Y.; Li, Y. A fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 638–654. [[CrossRef](#)]
66. Jin, Y.; Zhang, J.; Li, M.; Tian, Y.; Zhu, H.; Fang, Z. Towards the automatic anime characters creation with generative adversarial networks. *arXiv* **2017**, arXiv:1708.05509.
67. Bang, S.; Park, S.; Kim, H.; Kim, H. Encoder–decoder network for pixel-level road crack detection in black-box images. *Comput. Aided Civ. Infrastruct. Eng.* **2019**, *34*, 713–727. [[CrossRef](#)]
68. Tong, Z.; Yuan, D.; Gao, J.; Wang, Z. Pavement defect detection with fully convolutional network and an uncertainty framework. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *35*, 832–849. [[CrossRef](#)]
69. Kashif Ur Rehman, S.; Ibrahim, Z.; Memon, S.A.; Jameel, M. Nondestructive test methods for concrete bridges: A review. *Constr. Build. Mater.* **2016**, *107*, 58–86. [[CrossRef](#)]
70. Huang, Z.; Zheng, H.; Guo, L.; Mo, D. Influence of the Position of Artificial Boundary on Computation Accuracy of Conjugated Infinite Element for a Finite Length Cylindrical Shell. *Acoust. Aust.* **2020**, *48*, 287–294. [[CrossRef](#)]
71. Adeli, H.; Jiang, X. *Intelligent Infrastructure: Neural Networks, Wavelets, and Chaos Theory for Intelligent Transportation Systems and Smart Structures*; CRC Press: Boca Raton, FL, USA, 2008.
72. Ofsthun, S.C.; Wilmering, T.J. Model-driven development of integrated health management architectures. In Proceedings of the 2004 IEEE Aerospace Conference Proceedings (IEEE Cat. No. 04TH8720), Big Sky, MT, USA, 6–13 March 2004; pp. 3692–3705.
73. Ji, S.; Sun, Y.; Shen, J. A method of data recovery based on compressive sensing in wireless structural health monitoring. *Math. Probl. Eng.* **2014**, *2014*, 546478. [[CrossRef](#)]
74. Wu, C.; Wu, P.; Wang, J.; Jiang, R.; Chen, M.; Wang, X. Critical review of data-driven decision-making in bridge operation and maintenance. *Struct. Infrastruct. Eng.* **2020**, 1–24. [[CrossRef](#)]
75. Mousavi, A.A.; Zhang, C.; Masri, S.F.; Gholipour, G. Structural Damage Localization and Quantification Based on a CEEMDAN Hilbert Transform Neural Network Approach: A Model Steel Truss Bridge Case Study. *Sensors* **2020**, *20*, 1271. [[CrossRef](#)] [[PubMed](#)]
76. Zhang, X.; Liang, D.; Zeng, J.; Asundi, A. Genetic algorithm-support vector regression for high reliability SHM system based on FBG sensor network. *Opt. Lasers Eng.* **2012**, *50*, 148–153. [[CrossRef](#)]
77. Gentile, C.; Guidobaldi, M.; Saisi, A. One-year dynamic monitoring of a historic tower: Damage detection under changing environment. *Meccanica* **2016**, *51*, 2873–2889. [[CrossRef](#)]
78. Kita, A.; Cavalagli, N.; Ubertini, F. Temperature effects on static and dynamic behavior of Consoli Palace in Gubbio, Italy. *Mech. Syst. Signal Process.* **2019**, *120*, 180–202. [[CrossRef](#)]
79. Zonno, G.; Aguilar, R.; Boroschek, R.; Lourenço, P.B. Environmental and Ambient Vibration Monitoring of Historical Adobe Buildings: Applications in Emblematic Andean Churches. *Int. J. Archit. Herit.* **2019**, 1–17. [[CrossRef](#)]
80. Moser, P.; Moaveni, B. Environmental effects on the identified natural frequencies of the Dowling Hall Footbridge. *Mech. Syst. Signal Process.* **2011**, *25*, 2336–2357. [[CrossRef](#)]
81. Peeters, B.; De Roeck, G. One-year monitoring of the Z24-Bridge: Environmental effects versus damage events. *Earthq. Eng. Struct. Dyn.* **2001**, *30*, 149–171. [[CrossRef](#)]
82. Zonno, G.; Aguilar, R.; Boroschek, R.; Lourenço, P.B. Analysis of the long and short-term effects of temperature and humidity on the structural properties of adobe buildings using continuous monitoring. *Eng. Struct.* **2019**, *196*, 109299. [[CrossRef](#)]
83. Ni, F.; Zhang, J.; Noori, M.N. Deep learning for data anomaly detection and data compression of a long-span suspension bridge. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *35*, 685–700. [[CrossRef](#)]
84. Avci, O.; Abdeljaber, O.; Kiranyaz, S.; Hussein, M.; Inman, D.J. Wireless and real-time structural damage detection: A novel decentralized method for wireless sensor networks. *J. Sound Vib.* **2018**, *424*, 158–172. [[CrossRef](#)]
85. Azimi, M.; Pekcan, G. Structural health monitoring using extremely compressed data through deep learning. *Comput. Aided Civ. Infrastruct. Eng.* **2020**, *35*, 597–614. [[CrossRef](#)]

86. Lin, Y.Z.; Nie, Z.H.; Ma, H.W. Structural damage detection with automatic feature-extraction through deep learning. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 1025–1046. [[CrossRef](#)]
87. Gulgec, N.S.; Takáč, M.; Pakzad, S.N. Structural sensing with deep learning: Strain estimation from acceleration data for fatigue assessment. *Comput. Aided Civ. Infrastruct. Eng.* **2020**. [[CrossRef](#)]
88. Guo, J.; Xie, X.; Bie, R.; Sun, L. Structural health monitoring by using a sparse coding-based deep learning algorithm with wireless sensor networks. *Pers. Ubiquitous Comput.* **2014**, *18*, 1977–1987. [[CrossRef](#)]
89. Chi, S.; Caldas, C.H. Automated object identification using optical video cameras on construction sites. *Comput. Aided Civ. Infrastruct. Eng.* **2011**, *26*, 368–380. [[CrossRef](#)]
90. Memarzadeh, M.; Golparvar-Fard, M.; Nibbles, J.C. Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors. *Autom. Constr.* **2013**, *32*, 24–37. [[CrossRef](#)]
91. Kim, H.; Kim, K.; Kim, H. Data-driven scene parsing method for recognizing construction site objects in the whole image. *Autom. Constr.* **2016**, *71*, 271–282. [[CrossRef](#)]
92. Mneymneh, B.E.; Abbas, M.; Khoury, H. Vision-based framework for intelligent monitoring of hardhat wearing on construction sites. *J. Comput. Civ. Eng.* **2019**, *33*, 04018066. [[CrossRef](#)]
93. Xie, Z.; Liu, H.; Li, Z.; He, Y. A convolutional neural network based approach towards real-time hard hat detection. In Proceedings of the 2018 IEEE International Conference on Progress in Informatics and Computing (PIC), Suzhou, China, 14–16 December 2018; pp. 430–434.
94. Fang, Q.; Li, H.; Luo, X.; Ding, L.; Luo, H.; Rose, T.M.; An, W. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos. *Autom. Constr.* **2018**, *85*, 1–9. [[CrossRef](#)]
95. Wu, J.; Cai, N.; Chen, W.; Wang, H.; Wang, G. Automatic detection of hardhats worn by construction personnel: A deep learning approach and benchmark dataset. *Autom. Constr.* **2019**, *106*, 102894. [[CrossRef](#)]
96. Fang, W.; Ding, L.; Luo, H.; Love, P.E.D. Falls from heights: A computer vision-based approach for safety harness detection. *Autom. Constr.* **2018**, *91*, 53–61. [[CrossRef](#)]
97. Kolar, Z.; Chen, H.; Luo, X. Transfer learning and deep convolutional neural networks for safety guardrail detection in 2D images. *Autom. Constr.* **2018**, *89*, 58–70. [[CrossRef](#)]
98. Siddula, M.; Dai, F.; Ye, Y.; Fan, J. Unsupervised feature learning for objects of interest detection in cluttered construction roof site images. *Procedia Eng.* **2016**, *145*, 428–435. [[CrossRef](#)]
99. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
100. Kim, K.; Kim, H.; Kim, H. Image-based construction hazard avoidance system using augmented reality in wearable device. *Autom. Constr.* **2017**, *83*, 390–403. [[CrossRef](#)]
101. Luo, X.; Li, H.; Cao, D.; Dai, F.; Seo, J.; Lee, S. Recognizing diverse construction activities in site images via relevance networks of construction-related objects detected by convolutional neural networks. *J. Comput. Civ. Eng.* **2018**, *32*, 04018012. [[CrossRef](#)]
102. Luo, X.; Li, H.; Cao, D.; Yu, Y.; Yang, X.; Huang, T. Towards efficient and objective work sampling: Recognizing workers' activities in site surveillance videos with two-stream convolutional networks. *Autom. Constr.* **2018**, *94*, 360–370. [[CrossRef](#)]
103. Kim, H.; Kim, H.; Hong, Y.W.; Byun, H. Detecting construction equipment using a region-based fully convolutional network and transfer learning. *J. Comput. Civ. Eng.* **2018**, *32*, 04017082. [[CrossRef](#)]
104. Fang, W.; Ding, L.; Zhong, B.; Love, P.E.D.; Luo, H. Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach. *Adv. Eng. Inform.* **2018**, *37*, 139–149. [[CrossRef](#)]
105. Son, H.; Choi, H.; Seong, H.; Kim, C. Detection of construction workers under varying poses and changing background in image sequences via very deep residual networks. *Autom. Constr.* **2019**, *99*, 27–38. [[CrossRef](#)]
106. Love, P.E.D.; Smith, J.; Teo, P. Putting into practice error management theory: Unlearning and learning to manage action errors in construction. *Appl. Ergon.* **2018**, *69*, 104–111. [[CrossRef](#)]
107. Chi, N.-W.; Lin, K.-Y.; Hsieh, S.-H. Using ontology-based text classification to assist Job Hazard Analysis. *Adv. Eng. Inform.* **2014**, *28*, 381–394. [[CrossRef](#)]
108. Tixier, A.J.P.; Hallowell, M.R.; Rajagopalan, B.; Bowman, D. Application of machine learning to construction injury prediction. *Autom. Constr.* **2016**, *69*, 102–114. [[CrossRef](#)]
109. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
110. Friedman, J.H. Stochastic gradient boosting. *Comput. Stat. Data Anal.* **2002**, *38*, 367–378. [[CrossRef](#)]
111. Tixier, A.J.P.; Hallowell, M.R.; Rajagopalan, B.; Bowman, D. Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports. *Autom. Constr.* **2016**, *62*, 45–56. [[CrossRef](#)]
112. Chokor, A.; Naganathan, H.; Chong, W.K.; El Asmar, M. Analyzing Arizona OSHA injury reports using unsupervised machine learning. *Procedia Eng.* **2016**, *145*, 1588–1593. [[CrossRef](#)]
113. Goh, Y.M.; Ubeynarayana, C.U. Construction accident narrative classification: An evaluation of text mining techniques. *Accid. Anal. Prev.* **2017**, *108*, 122–130. [[CrossRef](#)] [[PubMed](#)]
114. Stein, R.A.; Jaques, P.A.; Valiati, J.F. An analysis of hierarchical text classification using word embeddings. *Inf. Sci.* **2019**, *471*, 216–232. [[CrossRef](#)]
115. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.

116. Altinel, B.; Ganiz, M.C. Semantic text classification: A survey of past and recent advances. *Inf. Process. Manag.* **2018**, *54*, 1129–1153. [[CrossRef](#)]
117. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. *How Transferable Are Features in Deep Neural Networks?* Advances in Neural Information Processing Systems: Montreal, QC, Canada, 2014; pp. 3320–3328.
118. Montserrat, D.M.; Lin, Q.; Allebach, J.; Delp, E.J. Training object detection and recognition CNN models using data augmentation. *Electron. Imaging* **2017**, *2017*, 27–36. [[CrossRef](#)]
119. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the Advances in Neural Information Processing Systems 2, Montreal, QC, Canada; MIT Press: Cambridge, MA, USA, 2014; pp. 2672–2680.
120. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [[CrossRef](#)]
121. Kitchen, A.; Seah, J. Deep generative adversarial neural networks for realistic prostate lesion MRI synthesis. *arXiv* **2017**, arXiv:1708.00129.
122. Moeindarbari, H.; Taghikhany, T. Seismic reliability assessment of base-isolated structures using artificial neural network: Operation failure of sensitive equipment. *Earthq. Struct.* **2018**, *14*, 425–436.
123. Subasri, R.; Natarajan, A.M.; Sundaram, S.; Jianliang, W. Neural aided discrete PID active controller for non-linear hysteretic base-isolation building. In Proceedings of the 2013 9th Asian Control Conference (ASCC), Istanbul, Turkey, 23–26 June 2013; pp. 1–8.
124. Zheng, J.; Zhang, C.; Li, A. Experimental Investigation on the Mechanical Properties of Curved Metallic Plate Dampers. *Appl. Sci.* **2020**, *10*, 269. [[CrossRef](#)]
125. Zhang, C.; Wang, H. Robustness of the active rotary inertia driver system for structural swing vibration control subjected to multi-type hazard excitations. *Appl. Sci.* **2019**, *9*, 4391. [[CrossRef](#)]
126. Sugisawa, M.; Saeki, E.; Nakamura, H.; Hirabayashi, R.; Ichikawa, Y.; Ueki, M.; Hokari, M. Development of earthquake-resistant, vibration control, and base isolation technology for building structures. *Nippon Steel Tech. Rep.* **1995**, *66*, 37–46.
127. Zhang, X.; Wang, Y.; Chen, X.; Su, C.; Li, Z.; Wang, C.; Peng, Y. Decentralized Adaptive Neural Approximated Inverse Control for a Class of Large-Scale Nonlinear Hysteretic Systems With Time Delays. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *49*, 2424–2437. [[CrossRef](#)]
128. Tam, V.W.Y.; Fung, I.W.H. Tower crane safety in the construction industry: A Hong Kong study. *Saf. Sci.* **2011**, *49*, 208–215. [[CrossRef](#)]
129. Shi, S.-Y.; Tang, W.-Z.; Wang, Y.-Y. *A Review on Fatigue Driving Detection*; EDP Sciences: Essonne, France, 2017; p. 01019.
130. Zhang, F.; Su, J.; Geng, L.; Xiao, Z. Driver fatigue detection based on eye state recognition. In Proceedings of the 2017 International Conference on Machine Vision and Information Technology (CMVIT), Singapore, Singapore, 17–19 February 2017; pp. 105–110.
131. Dwivedi, K.; Biswaranjan, K.; Sethi, A. Drowsy driver detection using representation learning. In Proceedings of the 2014 IEEE International Advance Computing Conference (IACC), Gurgaon, India, 21–22 February 2014; pp. 995–999.
132. Zhang, W.; Murphey, Y.L.; Wang, T.; Xu, Q. Driver yawning detection based on deep convolutional neural learning and robust nose tracking. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–17 July 2015; pp. 1–8.
133. Guo, J.-M.; Markoni, H. Driver drowsiness detection using hybrid convolutional neural network and long short-term memory. *Multimed. Tools Appl.* **2019**, *78*, 29059–29087. [[CrossRef](#)]
134. Lyu, J.; Yuan, Z.; Chen, D. Long-term multi-granularity deep framework for driver drowsiness detection. *arxiv* **2018**, arXiv:1801.02325.
135. Hou, L.; Wu, S.; Zhang, G.K.; Tan, Y.; Wang, X. Literature Review of Digital Twins Applications in Construction Workforce Safety. *Appl. Sci.* **2021**, *11*, 339. [[CrossRef](#)]
136. Li, X.; Yi, W.; Chi, H.-L.; Wang, X.; Chan, A.P.C. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Autom. Constr.* **2018**, *86*, 150–162. [[CrossRef](#)]
137. Singh, V.; Gu, N.; Wang, X. A theoretical framework of a BIM-based multi-disciplinary collaboration platform. *Autom. Constr.* **2011**, *20*, 134–144. [[CrossRef](#)]
138. Wu, L.; Mokhtari, S.; Nazef, A.; Nam, B.; Yun, H.-B. Improvement of crack-detection accuracy using a novel crack defragmentation technique in image-based road assessment. *J. Comput. Civ. Eng.* **2016**, *30*, 04014118. [[CrossRef](#)]
139. Zhu, J.; Wang, X.; Wang, P.; Wu, Z.; Kim, M.J. Integration of BIM and GIS: Geometry from IFC to shapefile using open-source technology. *Autom. Constr.* **2019**, *102*, 105–119. [[CrossRef](#)]
140. Wu, P.; Wang, J.; Wang, X. A critical review of the use of 3D printing in the construction industry. *Autom. Constr.* **2016**, *68*, 21–31. [[CrossRef](#)]
141. Wang, X.; Truijens, M.; Hou, L.; Wang, Y.; Zhou, Y. Integrating Augmented Reality with Building Information Modeling: Onsite construction process controlling for liquefied natural gas industry. *Autom. Constr.* **2014**, *40*, 96–105. [[CrossRef](#)]
142. Hou, L.; Wang, X.; Bernold, L.; Love, P.E.D. Using animated augmented reality to cognitively guide assembly. *J. Comput. Civ. Eng.* **2013**, *27*, 439–451. [[CrossRef](#)]

-
143. Song, Y.; Wang, X.; Wright, G.; Thatcher, D.; Wu, P.; Felix, P. Traffic Volume Prediction With Segment-Based Regression Kriging and its Implementation in Assessing the Impact of Heavy Vehicles. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 232–243. [[CrossRef](#)]
 144. Qu, X.; Yu, Y.; Zhou, M.; Lin, C.-T.; Wang, X. Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: A reinforcement learning based approach. *Appl. Energy* **2020**, *257*, 114030. [[CrossRef](#)]