Enhanced Machine Learning Classification Accuracy for Scaffolding Safety Using Increased Features

Sayan Sakhakarmi[1]; JeeWoong Park, A.M.ASCE[2]; and Chunhee Cho[3]

Abstract: Despite regular safety inspections and safety planning, numerous fatal accidents related to scaffold take place at construction sites. Current practices relying on human inspection are not only impractical but also ineffective due to dynamic construction activities. Furthermore, a scaffold typically consists of multiple bays and stories, which leads to complexity in its structural behaviors with various modes of failure. However, previous studies considered only a limited number of failure cases for a simple one-bay scaffold while exploring machine-learning (ML) approaches to predict safety conditions. Thus, the authors have proposed an approach to monitor a complicated scaffolding structure in real time. This study explored a method of classifying scaffolding failure cases and reliably predicting safety conditions based on strain data sets from scaffolding columns. Furthermore, the research team successfully enhanced the predicting accuracy of ML classification by the proposed self-multiplication method to increase the number of features such as strain data sets. Implementation of the proposed methodology is expected to enable the monitoring of a large, complex system at construction sites. **DOI: 10.1061/(ASCE)CO.1943-7862.0001601.** © *2018 American Society of Civil Engineers.*

Author keywords: Scaffold; Construction site; Safety; Machine learning; Support vector machine (SVM).

Introduction

The construction industry is one of the major contributors to fatal accidents every year, despite efforts and precautions taken to prevent accidents in construction sites. Statistics show that 21.1% of fatalities in the private industry occurred in construction (OSHA 2015). Among these fatalities, fall protection and scaffolding regulation are ranked higher in the top 10 most frequently cited Occupational Safety and Health Administration (OSHA) standards violation (OSHA 2015). Although safety of the construction labors is perceived as the topmost priority among contractors (Wu et al. 2015), the conventional visual method of safety inspection is not effective to identify potential safety hazards in the dynamic construction site. Thus, researchers have implemented various sensing technologies in the construction site to improve workplace safety.

Research interests on sensing technologies and following applications have spread across various aspects of construction and resulted in new tools for ensuring and improving safety planning and management in the construction site. Researchers have utilized building information modeling (BIM) technology (Collins et al. 2014; Kim et al. 2016b; Sulankivi and Kähkönen 2010; Zhang et al. 2015) and the construction schedule integrated with information on planned site activities (Navon and Kolton 2006, 2007) to improve safety management of the construction site before actual construction. Also, various researchers have implemented wireless sensing devices to identify hazardous conditions by collecting and processing real-time site data. Motion sensors (Chen et al. 2017; Jebelli et al. 2014; Kim et al. 2017; Liu et al. 2012; Yang et al. 2014, 2016, 2017) are another type of technology that has been used to identify hazard conditions on the basis of workers' motion data. Similarly, researchers have tested other wireless sensing technologies such as radio frequency identification (Pradhan et al. 2009) integrated with BIM for real-time tracking and locating workers in the construction site (Costin et al. 2014; Fang et al. 2016). Also, laser technology (Golparvar-Fard et al. 2011), Bluetooth sensors (Cho et al. 2017; Li et al. 2016; Park et al. 2017a, b), and infrared and ultrasonic sensors (Lee et al. 2009) have been tested to improve real-time safety of construction sites. Furthermore, image processing (Fang et al. 2018; Han and Lee 2013; Jung et al. 2018; Seo et al. 2015) has also been implemented as a method to detect hazard conditions.

Along with sensing technologies, machine-learning techniques have been continuously used in various studies related to construction safety (Chen et al. 2017; Park and Brilakis 2012, 2016; Seo et al. 2015; Tixier et al. 2016; Yang et al. 2014, 2016, 2017). Despite the significant progress accomplished by such data-based algorithms, several researchers have indicated relatively small sizes of data used in research as a limitation associated with the reliability of using such approaches (Kim et al. 2017; Liu et al. 2012; Yang et al. 2017). Hence, there is a need for a reliable real-time assessment of construction safety to prevent hazardous situations. Therefore, our research team focused on an automated method of real-time condition assessment of the construction site. However, as the construction site is large and involved with many disparate activities, equipment, and materials, it will be complex to develop such a system to monitor the entire construction safety. Thus, the team investigated scaffolding structures, as they are one of the most used temporary structures in the construction industry.

Relevant Literature for Scaffolding

One of early automated efforts with BIM and scaffolding structures by Kim and Teizer (2014) introduced the automation of design and

[1]Ph.D. Student, Dept. of Civil and Environmental Engineering and Construction, Univ. of Nevada, Las Vegas, NV 89154. Email: sayan.sakhakarmi@unlv.edu

[2]Assistant Professor, Dept. of Civil and Environmental Engineering and Construction, Univ. of Nevada, Las Vegas, NV 89154. Email: jee.park@unlv.edu

[3]Assistant Professor, Dept. of Civil and Environmental Engineering, Univ. of Hawaii at Manoa, Honolulu, HI 96822 (corresponding author). Email: chunhee@hawaii.edu

Note. This manuscript was submitted on May 1, 2018; approved on August 3, 2018; published online on November 30, 2018. Discussion period open until April 30, 2019; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Construction Engineering and Management***, © ASCE, ISSN 0733-9364.

planning of scaffolds based on the sequences of structural elements being constructed in the site. Cho et al. (2018b) proposed a scaffolding structure modeling technique to precisely analyze and predict scaffolding structures. They accounted for the changes in boundary conditions and material parameters to resemble actual scaffolding structures under various loading combinations. Vision-based scaffolding research also included a video image processing technique (Jung 2014) and series of images (Jung et al. 2018) to identify potential structural failures of temporary structures.

To identify potential hazard situations during construction phases early in the project planning stage, Collins et al. (2014) proposed a method to incorporate safety risk factors for activities involving scaffolds in a BIM model integrated with construction schedules. A similar study was conducted by Kim et al. (2016b), which integrated scaffolding structures and construction schedules in a BIM model. Information on detailed work plans of construction crews was incorporated in the scaffolding-BIM integration to automatically identify potential safety hazards. Their system successfully identified potential hazard conditions in the early planning stage, and construction personnel could prerecognize the hazard information in planning, designing, and inspection of the scaffolds. Although they showed considerable progress by integrating various resources in one platform to realize automated safety planning, and such efforts are useful for planning, their applicability is limited for use during the construction stage.

With an effort to assess the real-time safety of temporary structures, Yuan et al. (2016) proposed a cyberphysical system to identify and communicate potential hazards situation. The study considered a single-bay scaffolding platform system of dimensions $1.524 \times 1.524 \times 2.134$ m ($5 \times 5 \times 7$ ft), equipped with various sensors to measure information such as load, strain, acceleration, and displacement. After extracting measured information, potential hazard situations were identified through structural performance analysis of the virtual structure, and the information was communicated to the workers and site managers through mobile application. However, the study was limited to the identification of potential hazard resulting from base settlement of a single-bay scaffold only. The conventional analysis is possible for a small system attempted in their research, but when the system becomes large and complex with many different failure modes, the real-time performance of such a system is unknown as the conventional structural analysis may take time. For a fully automated solution, the system is required to (1) classify failure modes considering both local and global failures and (2) analyze a scaffolding structure without consuming too much time during the implementation stage.

As a solution of the real-time monitoring of scaffolding structures, our research team developed the boundary condition and model updating techniques (Cho et al. 2018b) and demonstrated the feasibility of a machine learning (ML)–based approach to predict the safety conditions of a scaffolding structure (Cho et al. 2018a). The studies showed methods to account for the changes in the boundary conditions of scaffolds and an ML approach to circumvent the complex rigorous structural analysis for safety condition protection by investigating strain patterns from a simple scaffolding structure—a single-bay, 2-story scaffold—with four classification cases. Despite the successful demonstration of real-time safety monitoring, the studies were limited to fully prove the extent to which an ML approach can perform reliably on a large scaffolding structure that will inherently have a large number of failure cases, including local and global failures.

In reality, scaffolding structures are built on a dynamic construction, which typically generates more complicated conditions to be considered. As a result, the dynamic nature of site activities leads to a large number of potential local failure cases for the scaffolds, apart from the failure cases considered in previous studies (Cho et al. 2018a; Yuan et al. 2016). As local failures potentially cause critical structural issues, it is important to categorize cases of local failures while assessing the structural safety condition of a structure. However, until now, there has been little attention to adequately study different failure cases associated with a complex scaffold and demonstrate the capability of an ML approach in assessing the safety condition of such scaffolding structures. In this paper, our research team thus classified structural failure cases consisting of the subcategories of potential local and global failure cases of a complex scaffold structure by using an ML approach. As an ML algorithm, the support vector machine (SVM) has demonstrated the ability to classify a large number of failure cases.

The proposed methodology used a database consisting of a large number of strain data sets, generated from extensive structural analysis, for different safety cases for a four-bay, 3-story scaffolding structure; this database was to eliminate the step of rigorous real-time analysis. Randomly generated load combinations were applied to a scaffold model for producing strain data sets. Investigating the strain data sets corresponding to various failure cases, SVM classifiers were trained to predict the scaffolding safety conditions. For testing and training purposes, various proportions of data sets (i.e., break data rate) were used to study the effect of the data size on ML prediction accuracy. The study team believed that the prediction accuracy would significantly decrease with the introduction of a large number of structural safety classification cases for a complex scaffolding structure adopted in this study. Moreover, the strain features of the data sets generated from an FEM model would not be sufficient to accurately classify the larger number of failure cases. Thus, the research team hypothesized that an increased number of features would better reflect the complexity resulting from many different modes of failures in different combinations of strain data, thus resulting in better ML classification accuracy. Therefore, this study pertains to an investigation of refined failure modes, their impact on prediction accuracy, and a resolution by refining the feature vectors in SVM.

Objective and Scope

The objective of this research was to enable real-time monitoring of large and complex scaffolds at construction sites in an attempt to prevent construction workers from tragic scaffolding collapses. The local and global failure cases of a complex scaffolding structure were precategorized, which are criteria for estimating the status of safety conditions. Then, this research investigated strain patterns of the categorized structural conditions and capability of ML to correctly predict the structural state of the scaffold. In this investigation, the authors demonstrated a challenge that is inherent to a system with a small number of components in a strain vector called features. To overcome this challenge, the proposed method combined strain components by self-multiplication to increase the number of features including nonlinear characteristics.

The actual construction site uses a complex scaffold structure. However, conducting an actual test for assessing the safety condition of a complex scaffolding structure is too risky and complex as it would involve the applications of complex load combinations. Along with an increased risk from a complex scaffolding structure, testing with such a complex structure is expensive as well. To overcome this limitation, this research adopted a scaffolding modeling technique (Cho et al. 2018b) that demonstrated a material and geometric parameters modeling method under various loading cases. This technique showed improved accuracy of a scaffold modeling, compared with actual scaffolding structural behaviors. The study adopted complex four-bay 3-story scaffolding structure in a finite-element model for further analysis for collecting ML training data sets.

Methodology

The proposed methodology was composed of three basic steps as shown in Fig. 1. The first step was to identify the safety conditions for a four-bay 3-story scaffold structure based on the global and local failure cases, which required construction of a numerical scaffold model. Then, the next step pertained to the construction of a database consisting of a large number of strain data sets by randomly applying various loading combinations in widely varying ranges; this technique is important because of its direct association with the challenges imposed by complex systems. Once 1,000 strain data sets were obtained for each failure case, 80% of the strain data sets for each loading case were used to train with SVM, and the remaining 20% strain data sets were used to test prediction accuracy. These steps are further explained in the following sections.

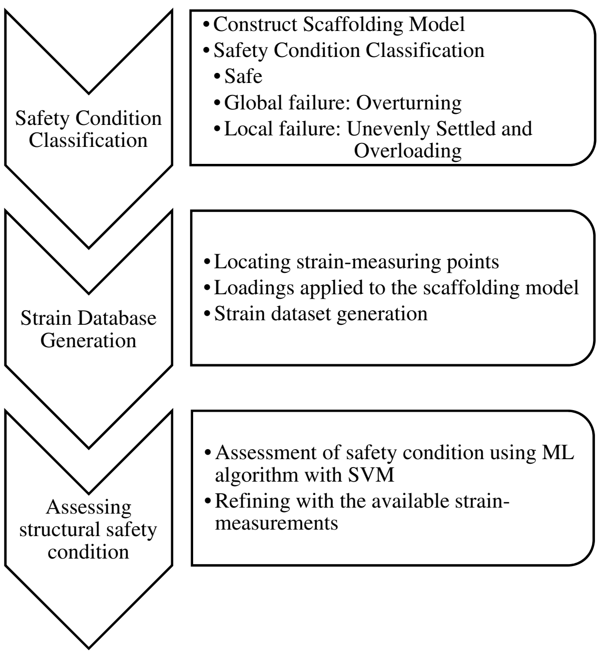

Fig. 1. General approach of the study.

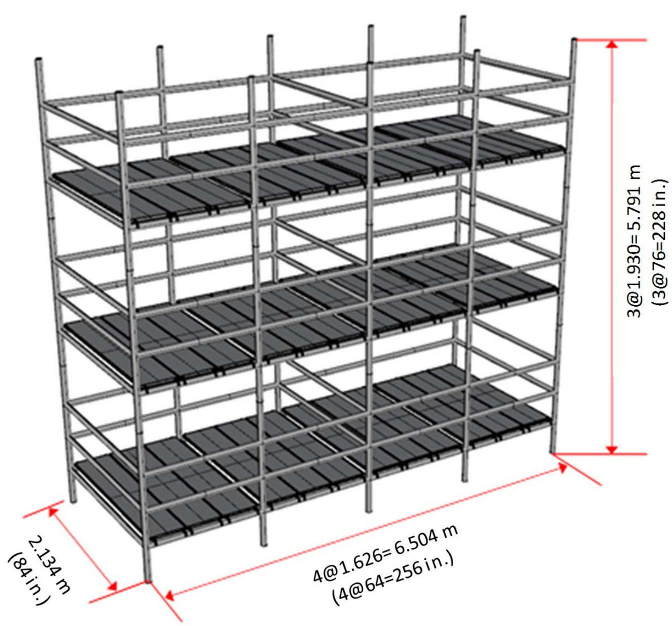

Fig. 2. Numerical scaffolding model.

Safety Condition Classification for a Scaffolding Structure

Numerical Scaffolding Model

To reflect the complex nature of a scaffolding in the construction site, a four-bay 3-story scaffolding structure was adopted in this study. Our previous study (Cho et al. 2018b) demonstrated that a scaffolding structure could be numerically modeled by using a finite-element method to reflect the actual structural behaviors. Therefore, our research team implemented the same principle to model a four-bay 3-story scaffolding structure by using the commercial software package COMSOL Multiphysics. The numerical scaffold model is shown in Fig. 2. Pipes and planks of the scaffold

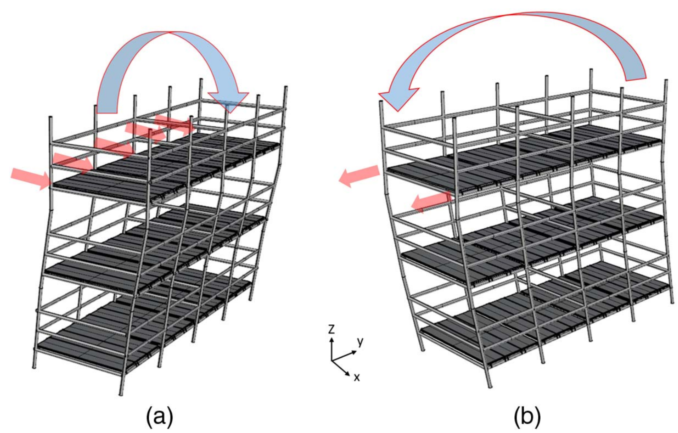

Fig. 3. Overturning failure: (a) side sway in X-direction; and (b) side sway in Y-direction.

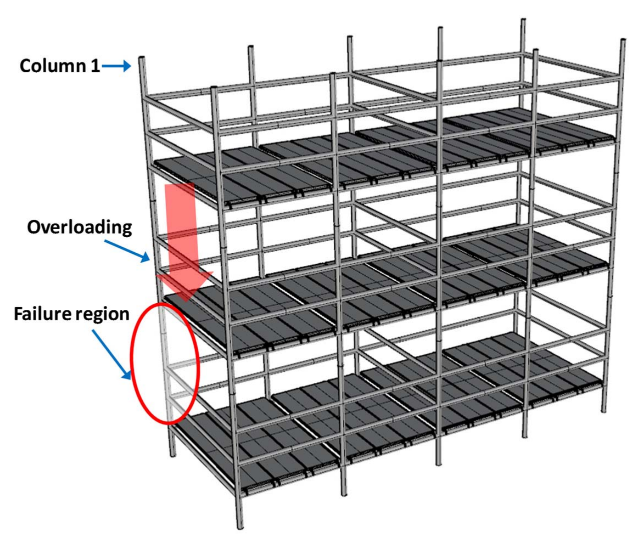

Fig. 4. Overloading failure.

were modeled with beam-column elements and shell elements, respectively. The transformable boundary conditions (Cho et al. 2018b) were assigned to present more realistic conditions between the columns and foundations of the scaffold model.

Table 1. Safety condition cases

Structural safety condition	Level of failure	Type of failure	Number of subcases
Safe	—	—	1
Unsafe	Global failure	Overturning	2
	Local failure	Unevenly settled	10
		Overloading	10

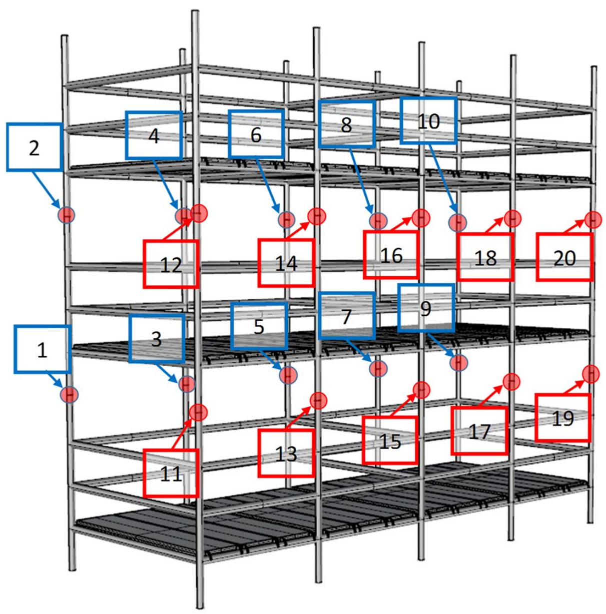

Fig. 5. Scaffolding model with strain-measuring points.

The scaffolding design follows scaffold specifications (OSHA 2017). The dimension of the unit scaffold is 1.626 m (64 in.) in length, 2.134 m (84 in.) in width, and 1.930 m (76 in.) in height, and the overall dimension of the scaffolding was 6.502 m (256 in.) in length, 2.134 m (84 in.) in width, and 5.791 m (228 in.) in height.

Safety Condition Classification

The structural conditions were categorized as either safe or unsafe in a large scale. However, the unsafe category could be further classified into various subcases based on the nature of failure modes, either global- or local-level failures on different external loading conditions imposed on the scaffolding structure. Thus, the subcategories of the unsafe category included global and local failures.

Global Failure. The overturning failure results from the lateral movement of the structure in any direction. In this study, the overturning effects resulting from two side-sway movements in the X- and Y-directions were considered. The subcases for overturning failure are shown in Fig. 3.

Local Failure. According to unsafe conditions resulting from local-level failures of the scaffolding structure, they were further classified into unevenly settled and overloading failures. In the unevenly settled failure case, some of scaffolding foundations could be down by a certain degree due to the unsettling ground condition or combinations of loads and the ground condition. Therefore, scaffolding columns on the uneven settlement may not support loadings properly and burden adjacent columns by imposing parts of its loads on them. Depending on the ground condition and combinations of loads and the ground condition, settlements might occur in one or more of the vertical members of the scaffold. However, although multiple local failures occur in columns, the first failure identification is more important than any others as it suffices for an indication of a safety problem. Therefore, this study focused on identifying the first failure of vertical members. As the scaffold model considered in this study had 10 columns, the uneven settled condition was further categorized into 10 subcases, each of which was a case of unevenly settled failures.

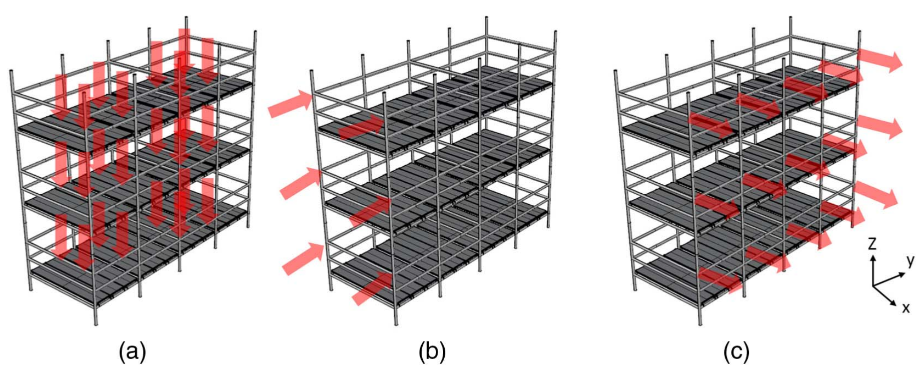

Fig. 6. Types of loadings applied to the scaffolding model: (a) gravity force; (b) lateral force (Y-direction); and (c) lateral force (X-direction).

Table 2. Ranges of randomly applied loads

Range of each load	Safe	Side sway X	Side sway Y	Uneven settlement	Overloading
Gravity force (uniform load) (N/m^2)	0 to −1,400	0 to −1,400	0 to −1,400	0 to −1,400	−1,400 to −2,000
Lateral Y force (point load) (N)	−500 to +500	+5,000 to +10,000 −5,000 to −10,000	−500 to +500	−500 to +500	−500 to +500
Lateral X force (point load)	−100 to +100	−100 to +100	+8,000 to +15,000 −8,000 to −15,000	−100 to +100	−100 to +100

Table 3. Sample strain data sets generated from random load application

Safety condition	Strain measurements ($\mu\varepsilon$) at different strain-measuring locations																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Safe	−11.2	−25.5	−16.2	−9.1	−4.6	−11.3	−20.4	−22.8	−15.8	−8.8	−5.3	−12.7	−6.6	−3.5	−0.9	−3.3	−8.8	−12.4	−8.5	−4.6
	−9.6	−18.0	−20.4	−8.2	−9.0	−11.7	−18.5	−23.1	−10.2	−13.1	−4.7	−8.0	−6.3	−1.7	−6.5	−4.7	−10.0	−11.9	−6.4	−6.5
	−14.6	−22.7	−25.1	−7.5	−4.3	−12.0	−21.0	−27.0	−14.6	−2.8	−8.5	−12.2	−9.4	−1.4	−4.3	−6.6	−13.0	−11.6	−6.9	−1.8
	−14.7	−17.3	−17.6	−14.7	−5.4	−16.2	−23.6	−27.1	−14.1	−7.6	−9.1	−8.8	−6.9	−7.6	−2.8	−9.1	−9.9	−13.1	−5.3	−2.1
	−11.0	−18.7	−21.8	−10.3	−4.0	−11.1	−22.9	−21.3	−16.0	−9.7	−7.8	−9.2	−9.8	−7.6	−3.5	−8.5	−10.5	−10.7	−9.5	−7.1
Overturning	33.1	22.9	19.8	31.1	28.1	−37.6	−51.6	−52.7	−57.1	−36.5	18.1	15.9	12.4	20.9	13.4	−21.2	−28.5	−28.9	−37.5	−17.2
	31.9	23.2	20.2	31.1	28.9	−37.9	−53.3	−54.4	−55.3	−36.2	18.0	15.6	12.4	20.9	13.5	−21.2	−28.6	−29.3	−36.2	−17.5
	−8.1	−24.5	−27.5	−15.9	0.7	−9.7	−25.9	−28.5	−16.4	0.2	−3.1	−4.6	−9.0	−8.8	0.4	−3.5	−5.5	−8.9	−9.2	−0.2
	−15.6	−28.0	−17.3	−28.9	−5.5	−16.3	−28.3	−18.5	−29.5	−6.9	−8.0	−15.4	−11.7	−21.0	−6.0	−8.0	−15.5	−12.2	−21.6	−6.8
	−8.0	−23.2	−24.3	−20.9	−4.1	−8.4	−24.1	−25.1	−22.0	−4.5	−4.0	−5.5	−4.9	−13.0	−3.7	−4.1	−6.3	−5.7	−13.1	−3.9
Unevenly settled	−2.2	−24.9	−23.1	−13.6	−7.1	−5.5	−2.1	−36.3	−22.9	1.8	−2.3	−18.3	−0.1	−0.3	−12.5	−3.5	−17.0	−5.4	−11.2	0.6
	−1.2	−10.5	−0.3	0.1	−5.4	−3.0	−11.3	−15.0	−3.4	−6.5	−1.0	−2.2	−3.2	−3.6	−2.6	−1.5	−2.1	−11.8	−9.0	−1.8
	−17.1	−3.8	−20.5	−0.5	−8.0	1.3	−3.3	−17.2	−7.4	−11.2	−2.5	−0.8	−8.6	−4.5	−6.0	1.0	−4.8	−5.3	−2.0	−3.2
	0.7	−10.8	−14.0	−13.3	−1.3	−4.1	−5.9	−15.4	−29.3	−4.8	2.6	−2.2	−7.6	−0.9	−0.3	−1.6	−1.1	−0.5	−16.3	−2.4
	6.2	−2.3	−22.7	−12.8	−0.1	−7.6	−3.5	−19.6	−32.9	−2.6	1.2	−0.9	−4.1	−10.3	−1.2	−4.3	−0.4	−10.9	−0.7	−0.4
Overloading	−103.6	−51.8	−6.8	−19.5	−43.6	−25.6	−52.6	−15.8	−41.6	−45.9	−31.8	−33.6	−3.7	−18.9	−24.2	−15.3	−42.9	−13.5	−31.6	−28.5
	−2.7	−79.0	−62.3	−20.6	−55.3	−64.4	−8.9	−58.7	−44.8	−62.0	−2.5	−41.1	−53.7	−17.0	−33.7	−36.7	−5.8	−29.4	−26.3	−46.0
	−28.6	−12.3	−21.5	−28.0	−63.2	−30.4	−53.2	−57.3	−91.7	−23.0	−22.6	−6.4	−19.3	−18.8	−38.6	−27.0	−42.4	−38.4	−6.4	−11.8
	−0.5	−17.0	−10.2	−38.9	−24.7	−20.0	−50.7	−31.4	−34.5	−94.0	−0.3	−13.2	−9.5	−31.0	−15.9	−15.4	−48.3	−29.0	−34.2	−32.4
	−41.5	−43.2	−38.1	−34.2	−60.4	−6.5	−30.0	−23.8	−23.6	−77.6	−21.5	−34.3	−24.1	−21.0	−55.2	−5.4	−28.1	−23.1	−12.0	−21.6

Overloading is another unsafe failure. Fig. 4 illustrates an example of the overloading failure case resulting from the local failure of an individual member of the scaffold structure. In the simulation, the uncertainty as to where loading would occur was taken into account by applying random amounts of loads on random combinations of bays and floors to the scaffold.

Although multiple member failures might occur over time, it is unlikely that they occur exactly at the same time. Thus, timely detection of the potential local failure of a single member as explained in the unevenly settled case would suffice for an indication of a safety hazard. Hence, this study used 10 subcases of the unsafe overloading category. Table 1 summarizes the safety condition cases classified into 23 subcases.

Data Generation through Structural Analysis

Locating Strain-Measuring Points

For the generation of required strain data sets, the first step was to locate strain-measuring points in the scaffolding model. As shown in Fig. 2, the scaffold model consisted of 10 vertical members. Two strain-measuring points were evenly located on each of the vertical members with 20 strain-measuring points as indicated in Fig. 5.

Loadings Applied to the Scaffolding Model

After locating the strain-measuring points in the scaffolding model, the next step was to generate random loadings that corresponded to the 23 different failure cases as discussed in the Safety Condition Classification section. For this purpose, the design of loading types follows ASCE 7 (ASCE 2016), and loading amplitudes of gravity

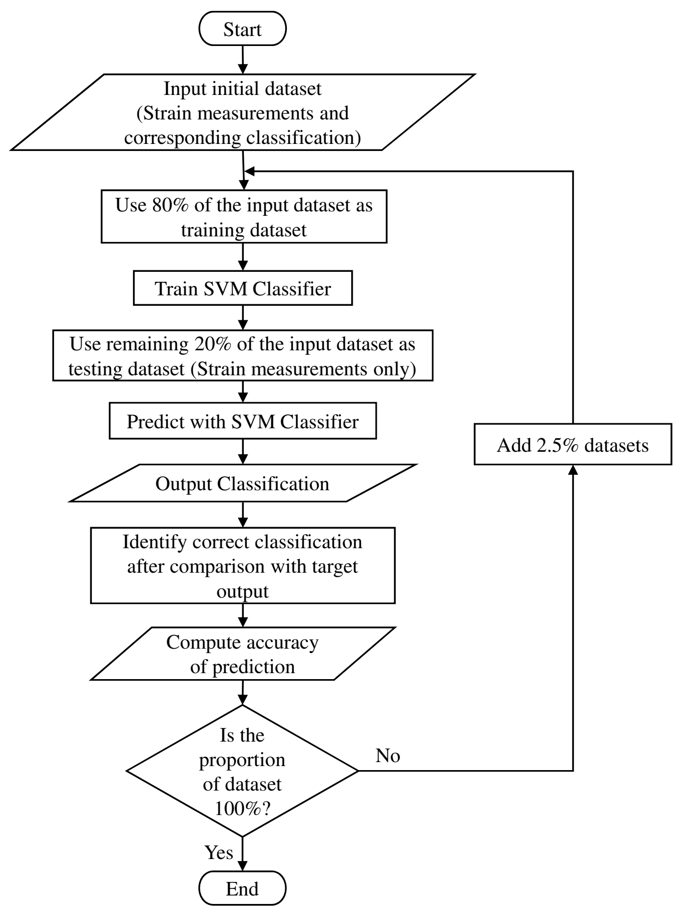

Fig. 7. Safety condition assessment.

J. Constr. Eng. Manage., 2019, 145(2): 04018133

forces and lateral forces were randomly applied to the scaffolding model as shown in Fig. 6. The verification of such a method must entail complex loads and corresponding data generation. This is not possible to be conducted in an actual scenario, which is why this study adopted computer simulation together with an ML approach to extensively verify its prediction capability in various loading cases.

In the case of the gravity force, uniform loads were applied to random unit scaffold as shown in Fig. 6(a). In the case of the lateral force in the Y-direction, two-point loads were randomly applied to each floor as shown in Fig. 6(b). Similarly, in the case of the lateral forces in the X-direction, five-point loads were randomly applied to each floor as shown in Fig. 6(c). Table 2 shows the ranges of these three types of random loadings applied to the model for simulating various safety conditions.

Strain Data Set Generation

Using a number of combinations of the three types of the random loadings, strain-measurement data sets were generated through structural analysis. For each of the 23 different failure cases, 1,000 strain data sets were generated, each composed of 20 strain measurement features corresponding to the 20 strain-measuring points. Hence, a database was formed with 23,000 strain data sets, each consisting of 20 strain features. Table 3 shows sample sets of the strain data sets generated by this simulation. As seen in Table 3, strain data that fall in the same category of the safety condition show significant differences in values because of the random applications of the loads.

Assessing Structural Safety Condition of the Scaffold

Assessment of Safety Condition

Previous studies using SVM for applications using sensor data (Chen et al. 2017; Yang et al. 2014, 2016) demonstrated its reliability in solving classification problems. The SVM is a supervised ML technique that involves training of the binary SVM classifiers with available data sets and then using the SVM classifiers to predict results. As each structural failure case of a scaffold has a unique set of strain data, the binary classification performed by SVM on each of the classified cases gives a reliable evaluation of accuracy on each classification. Therefore, an SVM algorithm was implemented for the assessment of the safety conditions of the scaffold based on strain measurements.

Fig. 7 illustrates the analysis procedure, which is specifically designed to investigate the effect of the number of available data sets on prediction accuracy. The first trial used 575 data sets, which equals 2.5% of the entire 23,000 data sets in training and predicting the results. This study used 80% of these data sets for training of the SVM classifiers. Then, with the remaining 20% strain data set, the trained SVM classifiers were used to determine the accuracy of successfully predicting the safety conditions of the scaffolding structure. This process was repeated until all data sets were covered for the accuracy test; that is, the data sets were incremented by 2.5% in the subsequent trials followed by training and prediction. This incremental analysis conducted 40 tests using various proportions of the data sets. While computing the accuracy of prediction using SVM for each test data set, the predicted outputs for each case were first concatenated and then compared with the target output. Then, the prediction accuracy was determined by dividing the total number of correct results by the total number of the test data sets.

Refining with the Available Strain Measurements

In the proposed methodology, the strain-measurement data sets formed the basis for predicting the structural safety condition of

Table 4. Sample calculation of a data set for creating new strain features

Parameter	Value
Original data set	
X1	−6.6
X2	−18.3
X3	−14
X4	−17.1
X5	−13.5
X6	−14.7
X7	−20.4
X8	−19.2
X9	−20.6
X10	−17.1
X11	−4.52
X12	−10.2
X13	−8.8
X14	−9
X15	−4.1
X16	−6.8
X17	−13.5
X18	−11.4
X19	−7.6
X20	−7.7
New strain features (190 for each data set)	
X1 · X2	121.3
X1 · X3	92.6
X1 · X4	113.4
X1 · X5	89.2
X1 · X6	97.2
X1 · X7	134.8
X1 · X8	126.8
X1 · X9	136.1
X1 · X10	113.2
X1 · X11	29.8
X1 · X12	67.6
X1 · X13	58.6
X1 · X14	59.6
X1 · X15	27.1
X1 · X16	45.2
X1 · X17	89.1
X1 · X18	75.8
X1 · X19	50.5
X1 · X20	50.8
X2 · X3	258.2
X2 · X4	316.2
X2 · X5	248.6
X2 · X6	270.8
X2 · X7	375.6
X2 · X8	353.4
X2 · X9	379.4
X2 · X10	315.4
X2 · X11	83.1
X2 · X12	188.4
X2 · X13	163.4
X2 · X14	166.2
X2 · X15	75.4
X2 · X16	126.1
X2 · X17	248.5
X2 · X18	211.3
X2 · X19	140.7
X2 · X20	141.5
X18 · X19	87.9
X18 · X20	88.4
X19 · X20	58.9

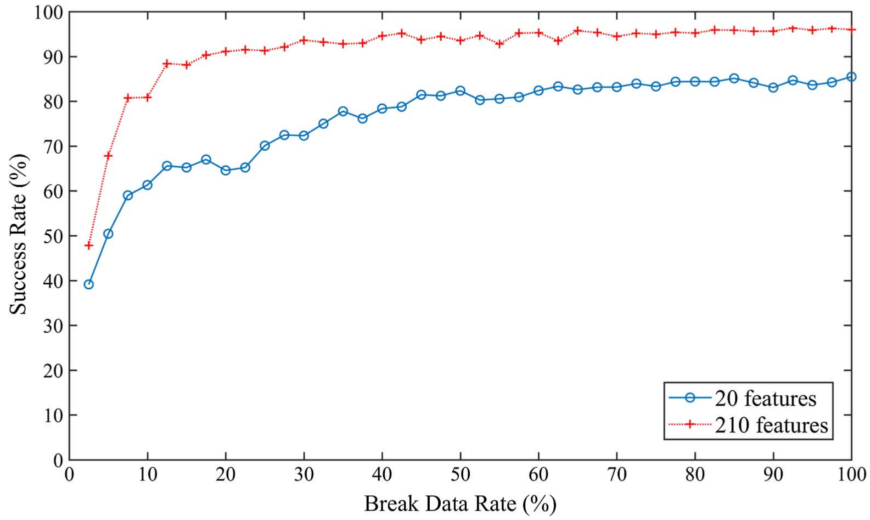

Fig. 8. Performance comparison with different numbers of strain data features.

the scaffold. Thus, the more strain features were used for assessing safety condition of the scaffold, the more accurate prediction was achieved. However, it is not practical to measure a large number of strain values from an actual scaffolding structure. Therefore, self-multiplication was applied to the feature vectors to get new features in an effort to increase the number of feature data. By introducing the second-degree polynomial features as result of self-multiplication, 190 strain features were added to the original 20 features, which resulted in 210 strain features for each of the 23,000 data sets. Table 4 shows a sample calculation for a strain data set in which the new strain features are obtained by multiplying two original strain features. The SVM algorithm was applied to both data sets—original data set with 20 features and modified data set with 210 features—for the safety analysis of the scaffolding structure. The steps elaborated in the "Assessment of Safety Condition" section were applied in this analysis.

Table 5. Comparison of prediction results

Proportion of data set (%)	Total number of data set	Success rate of prediction with SVM (%) 20 features (A)	210 features (B)	% improvement in prediction[a]
2.5	575	39.13	47.83	22.22
10.0	2,300	61.30	80.87	31.91
20.0	4,600	64.57	91.09	41.08
30.0	6,900	72.32	93.62	29.46
40.0	9,200	78.37	94.57	20.67
50.0	11,500	82.35	93.57	13.62
60.0	13,800	82.39	95.29	15.66
70.0	16,100	83.17	94.47	13.59
80.0	18,400	84.40	95.22	12.81
90.0	20,700	83.04	95.65	15.18
100.0	23,000	85.48	96.00	12.31

[a]$I = (B - A)/A \times 100$.

Result and Discussion

The accuracy of predicting the structural safety conditions of the four-bay 3-story scaffold model was determined using SVM for varying numbers of data sets. Fig. 8 plots the results of prediction accuracy with respect to the number of used data sets and the success rates for the 20 feature and 210 feature cases. The break data rate indicates the percentage of the data used from the entire 23,000 data sets.

For the data set with 20 strain features, the success rates of predicting the structural safety conditions of the scaffold ranged from 39.13% to 85.48% for the 40 tests conducted with varying numbers of strain data sets ranging from 575 (2.5%) to 23,000 data sets (100%). For both plots, a rapid increase in the prediction accuracy was observed for the beginning range of the break data rates. However, in the case with 20 feature vectors, the improvement in accuracy was flattened out around 84%, reaching about 85% only when the entire data sets were used for the test; no significant improvement was observed after 50% of the break data rate. A success rate of 85%, that is a failure rate of 15, may not be considered as a reliable solution to be used for safety applications. A similar trend was observed in the case of 210 strain features, but the accuracy was higher than the original 20 strain feature cases as it reached 96%. The implementation of SVM with the increased strain features significantly improved prediction accuracy.

Table 5 shows the results of improvement in prediction accuracy for selected data sets. It is evident that the rate of prediction improved with an increase in the size of the data sets used in training for both data sets. In the case with 20 strain features, the prediction accuracy increased from 39.13% to 82.35% while the data set sizes changed from 575 to 11,500. However, in the case with 210 strain features, it did not require a larger data set for such improvement in prediction accuracy as the success rate increased to 91.09% for 4,600 strain data sets from the prediction accuracy of 47.83% for 575 strain data sets. The results presented clear evidence of improvement; a faster learning rate was observed with higher improvement rates in the early data sets, and a higher prediction rate was observed for all data sets.

Conclusion

Although a considerable number of construction workers work on scaffolds and their safety is vital, a limited number of studies have assessed the real-time structural safety condition of complex scaffolding structures to improve worker safety. Therefore, the study team explored an ML approach by investigating the effect of feature vectors for classifying potential structural failure cases of a complex scaffold structure. The failure cases included 23 modes of local and global failures with 20 original strain features per data set and 210 increased strain features per data set over 23,000 data sets.

From the results of the simulation, it was concluded that 20 strain features were insufficient to accurately represent the failure cases generated by complex loading combinations applied to each of the data sets. While the maximum accuracy with the 20 strain feature case was 85.48%, that with the 210 strain feature case was 96%. The comparison of the prediction accuracy between the two cases exposed the effect of increased strain feature vectors for exactly the same loading cases. Since the 210 strain features generated by the self-multiplication contained higher-order polynomial factors, the relationships identified by nonlinearity were more accurately captured by the SVM algorithm in producing more accurate classifiers. This study confirmed the applicability of SVM for conducting real-time analysis of scaffolding structures with a feed of continuous strain data. Therefore, its implementation in the construction site is expected to improve worker safety, especially working on scaffolds.

This study used SVM to generate additional features and explore their effects on prediction accuracy with respect to varying numbers of data sets for the randomly generated 23,000 simulations. The study used strain data collected from the columns of a scaffold because they are critical factors associated with collapse failures. Minor failures that may occur to planks were not considered because these do not seriously affect structural collapse. The structural failure cases considered in this study were categorized on the basis of the assumption that the timely identification of a single-member failure in a scaffold is sufficient to prevent its potential collapse. However, incorporation of more failure cases that can represent multiple local failures would be more beneficial for engineers and site managers to manage failed scaffolding structures. Future studies on this subject would require the generation of an algorithm to automatically identify all possible failure cases. In this case, SVM may consume considerable resources in the prediction stage as it conducts binary classification for each of the failure modes individually. Thus, future studies may require a more advanced technique such as deep learning if time constraints exist in the prediction phase.

Data Availability Statement

Data generated or analyzed during the study are available from the corresponding author by request. Information about the *Journal*'s data-sharing policy can be found here: https://ascelibrary.org/doi/10.1061/(ASCE)CO.1943-7862.0001263.

References

ASCE. 2016. *Minimum design loads for buildings and other structures*. ASCE 7. Reston, VA: ASCE.

Chen, J., J. Qiu, and C. Ahn. 2017. "Construction worker's awkward posture recognition through supervised motion tensor decomposition." *Autom. Constr.* 77: 67–81. https://doi.org/10.1016/j.autcon.2017.01.020.

Cho, C., J. Park, K. Kim, and S. Sakhakarmi. 2018a. "Machine learning for assessing real-time safety conditions of scaffolds." In *Proc., 35th Int. Symp. on Automation and Robotics in Construction (ISARC 2018)*. Waterloo, Canada: International Association for Automation and Robotics in Construction.

Cho, C., S. Sakhakarmi, K. Kim, and J. Park. 2018b. "Scaffolding modelling for real-time monitoring using a strain sensing approach." In *Proc., 35th Int. Symp. on Automation and Robotics in Construction (ISARC 2018)*. Waterloo, Canada: International Association for Automation and Robotics in Construction.

Cho, C., S. Sakhakarmi, J. Park, and P. P. Shrestha. 2017. "Automated RSSI-based tracking sensor deployment using electromagnetic simulation." In *Proc., Seoul Int. Conf. on Applied Science and Engineering*, 142–151. Hong Kong: Asia-Pacific Education Research Association.

Collins, R., S. Zhang, K. Kim, and J. Teizer. 2014. "Integration of safety risk factors in BIM for scaffolding construction." In *Proc., Int. Conf. on Computing in Civil and Building Engineering*, edited by R. R. Issa, and I. Flood, 307–314. Reston, VA: ASCE.

Costin, A., N. Pradhananga, and J. Teizer. 2014. "Passive RFID and BIM for real-time visualization and location tracking." In *Proc., Construction Research Congress*, 169–178.

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

Fang, Y., Y. K. Cho, S. Zhang, and E. Perez. 2016. "Case study of BIM and cloud-enabled real-time RFID indoor localization for construction management applications." *J. Constr. Eng. Manage.* 142 (7): 05016003. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001125.

Golparvar-Fard, M., J. Bohn, J. Teizer, S. Savarese, and F. Peña-Mora. 2011. "Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques." *Autom. Constr.* 20 (8): 1143–1155. https://doi.org/10.1016/j.autcon.2011.04.016.

Han, S., and S. Lee. 2013. "A vision-based motion capture and recognition framework for behavior-based safety management." *Autom. Constr.* 35: 131–141. https://doi.org/10.1016/j.autcon.2013.05.001.

Jebelli, H., C. R. Ahn, and T. L. Stentz. 2014. "The validation of gait-stability metrics to assess construction workers' fall risk." In *Proc., Computing in Civil and Building Engineering*, 997–1004. Reston, VA: ASCE.

Jung, Y. 2014. "An Approach to automated detection of failure in temporary structures using image processing." *J. Eng. Archit.* 2 (1): 49–61.

Jung, Y., H. Oh, and M. M. Jeong. 2018. "An approach to automated detection of structural failure using chronological image analysis in temporary structures." *Int. J. Constr. Manage.* 1–8, in press. https://doi.org/10.1080/15623599.2017.1411457.

Kim, H., C. R. Ahn, and K. Yang. 2017. "Identifying safety hazards using collective bodily responses of workers." *J. Constr. Eng. Manage.* 143 (2): 4016090. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001220.

Kim, K., Y. K. Cho, and S. Zhang. 2016. "Integrating work sequences and temporary structures into safety planning: Automated scaffolding-related safety hazard identification and prevention in BIM." *Autom. Constr.* 70: 128–142. https://doi.org/10.1016/j.autcon.2016.06.012.

Kim, K., and J. Teizer. 2014. "Automatic design and planning of scaffolding systems using building information modeling." *Adv. Eng. Inf.* 28 (1): 66–80. https://doi.org/10.1016/j.aei.2013.12.002.

Lee, U.-K., J.-H. Kim, H. Cho, and K.-I. Kang. 2009. "Development of a mobile safety monitoring system for construction sites." *Autom. Constr.* 18 (3): 258–264. https://doi.org/10.1016/j.autcon.2008.08.002.

Li, H., X. Yang, F. Wang, T. Rose, G. Chan, and S. Dong. 2016. "Stochastic state sequence model to predict construction site safety states through real-time location systems." *Saf. Sci.* 84: 78–87. https://doi.org/10.1016/j.ssci.2015.11.025.

Liu, J., X. Zhang, and T. E. Lockhart. 2012. "Fall risk assessments based on postural and dynamic stability using inertial measurement unit." *Saf. Health Work* 3 (3): 192–198. https://doi.org/10.5491/SHAW.2012.3.3.192.

Navon, R., and O. Kolton. 2006. "Model for automated monitoring of fall hazards in building construction." *J. Constr. Eng. Manage.* 132 (7): 733–740. https://doi.org/10.1061/(ASCE)0733-9364(2006)132:7(733).

Navon, R., and O. Kolton. 2007. "Algorithms for automated monitoring and control of fall hazards." *J. Comput. Civ. Eng.* 21 (1): 21–28. https://doi.org/10.1061/(ASCE)0887-3801(2007)21:1(21).

OSHA (Occupational Safety and Health Administration). 2015. "Commonly used statistics." Accessed August 29, 2017. https://www.osha.gov/oshstats/commonstats.html.

OSHA (Occupational Safety and Health Administration). 2017. "Scaffold specifications: 1926 Subpart L App A." Accessed January 9, 2017. https://www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10917.

Park, J., J. Chen, and Y. K. Cho. 2017a. "Self-corrective knowledge-based hybrid tracking system using BIM and multimodal sensors." *Adv. Eng. Inf.* 32: 126–138. https://doi.org/10.1016/j.aei.2017.02.001.

Park, J., X. Yang, Y. K. Cho, and J. Seo. 2017b. "Improving dynamic proximity sensing and processing for smart work-zone safety." *Autom. Constr.* 84: 111–120. https://doi.org/10.1016/j.autcon.2017.08.025.

Park, M.-W., and I. Brilakis. 2012. "Construction worker detection in video frames for initializing vision trackers." *Autom. Constr.* 28: 15–25. https://doi.org/10.1016/j.autcon.2012.06.001.

Park, M.-W., and I. Brilakis. 2016. "Continuous localization of construction workers via integration of detection and tracking." *Autom. Constr.* 72: 129–142. https://doi.org/10.1016/j.autcon.2016.08.039.

Pradhan, A., E. Ergen, and B. Akinci. 2009. "Technological assessment of radio frequency identification technology for indoor localization." *J. Comput. Civ. Eng.* 23 (4): 230–238. https://doi.org/10.1061/(ASCE)0887-3801(2009)23:4(230).

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

Sulankivi, K., and K. Kähkönen. 2010. "4D-BIM for construction safety planning." In *Proc., W099-Special Track 18th CIB World Building Congress*, 117–128. Delft, Netherlands: CIB World.

Tixier, A. J. P., M. R. Hallowell, B. Rajagopalan, and D. Bowman. 2016. "Application of machine learning to construction injury prediction." *Autom. Constr.* 69: 102–114. https://doi.org/10.1016/j.autcon.2016.05.016.

Wu, C.-L., D.-P. Fang, P.-C. Liao, J.-W. Xue, Y. Li, and T. Wang. 2015. "Perception of corporate social responsibility: The case of Chinese international contractors." *J. Cleaner Prod.* 107: 185–194. https://doi.org/10.1016/j.jclepro.2015.04.143.

Yang, K., C. R. Ahn, M. C. Vuran, and S. S. Aria. 2016. "Semi-supervised near-miss fall detection for ironworkers with a wearable inertial measurement unit." *Autom. Constr.* 68: 194–202. https://doi.org/10.1016/j.autcon.2016.04.007.

Yang, K., C. R. Ahn, M. C. Vuran, and H. Kim. 2017. "Collective sensing of workers' gait patterns to identify fall hazards in construction." *Autom. Constr.* 82: 166–178. https://doi.org/10.1016/j.autcon.2017.04.010.

Yang, K., S. Aria, C. Ahn, and T. Stentz. 2014. "Automated detection of near-miss fall incidents in iron workers using inertial measurement units." In *Proc., Construction Research Congress 2014*, 935–944. Reston, VA: ASCE.

Yuan, X., C. J. Anumba, and M. K. Parfitt. 2016. "Cyber-physical systems for temporary structure monitoring." *Autom. Constr.* 66: 1–14. https://doi.org/10.1016/j.autcon.2016.02.005.

Zhang, S., K. Sulankivi, M. Kiviniemi, I. Romo, C. M. Eastman, and J. Teizer. 2015. "BIM-based fall hazard identification and prevention in construction safety planning." *Saf. Sci* 72: 31–45. https://doi.org/10.1016/j.ssci.2014.08.001.