Automation in Construction 148 (2023) 104748

Contents lists available at ScienceDirect

Automation in Construction

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

Monitoring safety behaviors of scaffolding workers using Gramian angular field convolution neural network based on IMU sensing data

Sungkook Hong [a], Jihwan Yoon [a], Youngjib Ham [b], Bogyeong Lee [c], Hyunsoo Kim [a,*]

[a] Department of Architectural Engineering, Dankook University 152 Jukjeon-ro, Suji-gu, Yongin-si, Gyeonggi-do, 16890; Republic of Korea
[b] Department of Construction Science, Texas A&M University, 3337 TAMU, College Station, TX 77843, United States of America
[c] Department of Architectural Engineering, Dankook University, Republic of Korea

ARTICLE INFO

Keywords:
Safety behavior
Safety compliance
Inertial measurement unit (IMU)
Convolution neural network (CNN)
Gramian angular fields (GAF)
Safety regulations

ABSTRACT

Monitoring workers' safety compliance is critical to the construction accident prevention. However, current practices mainly rely on the safety mangers' manual observation. To improve safety management, this study examines the feasibility of the automated framework to classify pre-defined safety behaviors of scaffolding workers and identify whether each safety behavior was carried out. To examine the fidelity of the framework, 35 scaffold workers were studied by using five inertial measurement unit (IMU) sensors and implementing the five safety regulations. To classify compliance with safety regulations, this study builds on the Gramian angular fields (GAFs) based convolution neural network (CNN). The proposed model is compared to five other classification algorithms. The performance evaluation results show that the proposed method is feasible to identify whether workers comply with safety regulations automatically. This outcome has the potential to improve safety management through personalized worker training or notification during work.

1. Introduction

In jobsites, workers may not occasionally follow safety regulations because they are excessively familiar with their work or lack accidental experience. Especially, such safety insensitivity linked to the feature that workers frequently work in a high place and with heavy equipment incurs fatal accidents. Most construction procedures have the associated safety regulations and rules to place safety managers at the sites to prevent accidents. However, it is inadequate to supervise every worker due to the limited number of safety managers on a construction site. Therefore, a majority of efforts have been made to prevent accidents by monitoring construction workers' behavioral tendencies and patterns [1–4]. Recently, studies for detecting hazard zones and changes in body stability have been mainly conducted [2,4–6]. Previous studies focused on finding hazards, hazard zones, obstacles, and body stability, and they have extracted unsafe behaviors based on deep-learning algorithms. Yu et al. [7] validated a method that detects unsafe behavior of workers using an image-skeleton-based parameterized approach in construction sites. Although several studies have tried to extract unsafe behaviors, unsafe behavior appears to have different forms and features by individual workers. Extracting such individual unsafe behavior is challenging to categorize various unsafe behaviors since the behaviors vary from individuality under diverse situations. Here, safety regulations for preventing describe how a worker should act in a specific situation, and thus monitoring workers' safety compliance with them could be effective in evaluating an individual's level of safety compliance. For example, if one worker complies with the safety regulations at 90% and another at 70%, safety managers could focus more on the latter (i.e., workers who comply less with the safety regulations) for safety accident prevention.

Most fatal accidents are fall accidents on construction sites [1,8]. In particular, scaffolding-related accidents result in approximately 60 deaths and 4500 injuries every year, according to the United States Bureau of Labor Statistics (BLS) [9]. These high risks of accidents are closely related to the nature of the scaffolding works. Scaffolding is installed in various ways in places where accessibility to workers is poor, such as the facade of the upper floor of the building. Therefore, the scaffolding worker needs to keep safety regulations strictly than other process workers and must check their safety equipment (e.g., safety helmet, safety hook, etc.) and working environments (e.g., whether the scaffold is connected, etc.). However, workers' carelessness (non-compliance with safety regulations) is one of the causes of accidents

* Corresponding author.
E-mail addresses: dk_hsk22@dankook.ac.kr (S. Hong), jihwan@dankook.ac.kr (J. Yoon), yham@tamu.edu (Y. Ham), bglee_@dankook.ac.kr (B. Lee), hkim13@dankook.ac.kr (H. Kim).

https://doi.org/10.1016/j.autcon.2023.104748
Received 27 September 2022; Received in revised form 19 December 2022; Accepted 6 January 2023
Available online 13 January 2023
0926-5805/© 2023 Elsevier B.V. All rights reserved.

[10]. For example, when workers work on scaffolding, their working behaviors without fastening the safety hook corresponds to the carelessness of safety [11]. To prevent the safety accident of scaffolding workers, an effective monitoring system of the compliance with safety regulations is required.

Recently, the shortage of management personnel has been addressed by monitoring the overall situation in the workplace through several closed-circuit televisions (CCTV). In this context, studies have been conducted to classify worker behaviors through computer vision techniques [12]. Despite the benefits, it is difficult to expect CCTV systems to cover all construction sites without blind spots. Given that covering the entire site corresponds to the fundamental of safety that there should be no place to be excluded, relying on a vision-based approach to establish an extensive monitoring system is still challenging. Although the safety monitoring system must be continuous and cover extensive sight, vision sensors (e.g., CCTV, camera, LiDAR, Radar, etc.) are expensive to build monitoring systems without any blind spots. Given the nature of temporary structures such as scaffoldings that repeat assembling and disassembling frequently, vision sensors attached have limitations in monitoring work continuously.

On the other hand, motion sensors such as IMU has an advantage in that it is relatively cheaper than vision sensors and can monitor worker behaviors for a long time, even with lower power usage. In addition, it is a small-sized sensor that does not interfere with worker performance and can collect motion data continuously regardless of the location. In this context, this study aims to examine the feasibility of a framework for monitoring workers' safety behaviors regarding safety compliance. The data collected through the IMU sensor attached to the worker body is used to analyze whether the worker's behavior follows the safety regulation or not. This study builds upon the following hypotheses: (1) distinguishable characteristics of motion data will appear depending on each safety regulation, and (2) IMU data can be used to identify whether a scaffolding worker complies with the safety regulation properly or not. This study builds on the GAF-CNN algorithm for classifying and identifying safety behaviors. For collecting data, experiments were conducted with 35 scaffolding workers. The collected data is used to train and test the proposed model. After training and testing the model, the performance evaluation results are compared with other algorithms.

This paper is organized in the following order. First, the approaches used in the existing safety behavior analysis and related state-of-the-art techniques are reviewed. Then, the overall methodology including the selection process of safety regulations through the survey of experienced practitioners are described. The rest of this paper discusses the results of experiments designed for validation.

2. Research background

Occupational health and safety have long been critical issues in the construction industry. To build a safe environment, extensive efforts have been made. The Center for Construction Research and Training (CPWR) and the Construction Industry Institute (CII) have established training programs for construction workers and explored the latest technologies for applying to work sites [13,14]. There have been studies to automate safety management by monitoring construction sites using various approaches (e.g., sensor-based technology, computer vision) [15–18]. For example, Yan et al. [19] extracted ergonomically hazardous activities using IMU sensors. Another approach that detects collapsing of body balance has been used to identify potential accidents [20]. In this regard, Dzeng et al. [21] detected the balance collapsing of tiling workers by using the built-in smartphone accelerometer. In the previous study by Wu et al. [22], detection of balance loss was used to examine the potential risk of accidents by using a radio frequency identification system and ultrasonic transceiver. Because 80–90% of construction accidents are caused by unsafe activities and behaviors of construction workers [23], prior research on safety monitoring, including the above, were geared toward extracting unsafe activities occurring in potential accidents [24,25].

Kim et al. [26] showed that the abnormality of subjects differed by their capability that respond to the hazard. In this regard, categorizing unsafe behaviors may be difficult since various unsafe behaviors occur differently on individuality and characteristics of situations. Meanwhile, safety behaviors that are strictly defined by safety regulations are likely standardized, and thus these behaviors are less affected by individuality and situations. Given that most construction works are supposed to follow safety regulations, monitoring the safety compliance of individual workers can provide a chance to improve jobsite safety. Therefore, this paper aims to (1) classify various safety behaviors and (2) identify whether workers properly conduct safety behaviors determined by safety regulations.

With the advancement of deep-learning algorithms and a graphic processing unit, vision-based behavior recognition have been applied to extract unsafe behavior in construction [27,28]. Han et al. [29] detected unsafe actions of construction workers using vision-based depth sensors. Detecting the personal protection equipment (PPE) such as safety harnesses using a vision-based approach was also validated by Fang et al. [28]. Besides, motion sensors such an IMU has showed high suitability that is highly durable, available in low-visibility conditions, and low-power usage [30–32]. With these advantages, IMU is one of the widely used sensors in the construction industry [24]. Joshua and Varghese [33] showed that IMU data can classify characteristics of activities. Yoon et al. [34] attempted to examine the gait stability of ironworkers based on IMU data. These studies validated the feasibility of using IMU sensors to extract subtle characteristics of the activities of workers. Although these studies focused on classifying each activity by IMU data and extracting unsafe behaviors in their task sequence, it is also essential to identify whether a worker's behavior follows the safety regulation.

Previous studies monitored potential accidents by extracting unsafe behaviors using various sensors. They focused on defining and extracting unsafe behaviors of workers using various sensing approaches. They concentrated on the attitude at a specific point in time or on the relationship with external elements (mainly equipment) rather than on the worker's continuous behavior. For example, they detected a lack of necessary equipment, or poor fastening of safety equipment. However, this can be a missing link in terms of compliance with safety regulations, which requires proper implementation of safety activities within continuous behaviors. Therefore, there is a knowledge gap in measuring the continuous behavior and analyze the activities of the workers therein to identify whether the safety regulations have been properly implemented. This study used IMU sensors to collect workers' activities corresponding to the experimental conditions to analyze their bodily movement. The following section describes the experimental design and the collection and analysis of workers' behavioral data.

3. Methodology

This study hypothesized that each safety regulation has distinguishable characteristics and that each safety behavior that workers conduct will be assessed to identify whether it meets the criteria. To test the feasibility, the motion data of 35 scaffold workers were collected through IMU sensors reflecting whole body movements (IMUs are attached to the head, both wrists, and both ankles). The worker performed five safety behaviors while walking on the scaffold. The experimental setup includes five zones, and a worker should conduct a safety behavior in each zone. For example, when the workers enter the first zone, scaffold workers conduct the first safety behavior (checking personal protection equipment (PPE)) determined by the safety regulations. The collected IMU data were compared with the simultaneously recorded video to confirm how the time series data corresponding to the actual situation. Then, collected data are analyzed through three machine learning and three deep learning algorithms, respectively. In this procedure, two-way approaches are used to analyze IMU data. One is activity classification, which includes the total dataset (5250 datasets). From the total dataset, the authors have attempted to classify five regulations. The other one is to identify whether each type of safety behavior (1050 dataset for each safety behavior) properly follows the safety regulation. Fig. 1 illustrates the process; (1) experiment and data collection; (2) comparing the actual movement to IMU data and labeling; (3) time series data preprocessing by the signal vector magnitude (SVM) and Gramian angular field (GAF); (4) applying machine learning and deep learning algorithms; (5) classification performance reports.

3.1. Experimental design

Scaffolding is essential in construction projects for easy installation, quick dismantling, and securing ample workspace. Since various tasks are performed simultaneously, safety accidents such as falls and falling objects are the most common in scaffolding works. Therefore, the experiment of this study is designed and conducted to test whether the workers' safety behaviors follow the safety regulations correctly through wearable sensing.

A total of 35 subjects were clinically healthy. The minimum age is 22, and the maximum age is 37. The experiment was conducted on subjects with at least three years of experience in scaffolding. Their average age was 30.53 years old. Their fatigue state did not affect the experiment through sufficient rest for at least 8 h before the experiment. Subject details are described in Table 1.

Each subject was equipped with a total of five IMU sensors. One sensor was attached to a worker's head, two sensors were attached to both wrists, and two sensors were attached to both ankles. IMU sensors collect accelerometer and gyroscope signals with a 128 Hz sampling rate. In the experiment, all subjects wore a safety helmet, a safety vest, safety gloves, a safety harness, safety shoes, and gaiters. Fig. 2 (a) shows the location of the attached IMU sensors. Subjects were asked to follow five safety regulations. The length of each zone is 1000 mm (the length of each step board). For the safety, scaffolding was installed on the flat asphalt with a sufficient bearing capacity of scaffoldings, and the position of the step board was installed at the height of 150 mm from the ground.

Safety regulations are selected by interviews with skilled safety managers of construction sites (a total of six safety managers having more than ten years site experience). They responded to the interview question that requests what important tasks during scaffold are. By combining the answers of the safety managers, five safety regulations are selected. The first safety regulation is set as the inspection of personal protective equipment (PPE). The regulation for the first zone (Zone 1: PPE inspection) is that workers inspect all their equipment by tapping them with both hands at least twice. Tapping the safety helmet multiple times ensures that the helmet has not been peeled off owing to any external impact. Visual inspection can be difficult to verify that the helmet is worn correctly. In the second zone (Zone 2: Responding to warning alarm), a worker is asked to respond to the warning alarm. When the alarm starts, a worker is asked to look around twice. In the third zone (Zone 3: Joint inspection), a worker inspects whether the joint of scaffoldings is appropriately installed. To complete the zone, a worker inspects and hammers two marked joints in Zone 3 at least twice. In the fourth zone (Zone 4: Step board inspection), workers inspect whether the step board of the scaffold is assembled adequately by pounding the step board with their own foot at least twice. Lastly, Zone 5 is set to hooking the safety hook. In this zone, a worker hooks the safety hook and inspects it by pulling it at least twice. In the following, the five classified safety activities are denoted as (Z1) PPE, (Z2) Warning Alarm, (Z3) Joint, (Z4) Step board, and (Z5) Safety hook. The selected safety regulations were based on the safety guidelines of the Occupational Safety and Health Administration (OSHA) and Korea Occupational Safety and Health Administration (KOSHA). Fig. 2 (c) represents the safety codes of OSHA and KOSHA guidelines, which serve as base for the safety regulations used in this study. All subjects performed 30 sets of experiments, and a total of 5250 safety behavior data collected. A description of each zone and the experiment setup is shown in Fig. 2 (b). The entire experiment procedure was simultaneously video recorded to identify activities. Additionally, before the experiments, every subject was trained by safety managers for all safety regulations. The subjects were also instructed to thoroughly obey the safety regulations with accurate movement. Even though they had tried to comply with all regulations properly, a few activities are classified as cases of non-compliance with safety regulations.

3.2. Data preprocessing and segmentation

Human motion data from multi-sensors increases the likelihood of producing meaningful results in safety and health risk assessments [24]. In this study, the locations of five IMU sensors are 1) head, 2) both wrists, and 3) both ankles. Each acceleration and gyroscopic graph have three axes (x, y, and z, respectively).

Since workers' activities are measured continuously, they are segmented into small data forms for analyzing data. Segmentation techniques are primarily categorized into activity-defined windows, event-define widows, and sliding windows [35]. Because fused data are manually annotated by referencing the judgment of a group of six safety experts for supervised learning, sliding window approaches that simplify data preprocessing are not suitable for this study. In addition, event-defined approaches that extract specific events can skip the safety regulations or worker's movement. Therefore, activity-defined window approaches that detect activity changes are applied to extract safety regulations. In the labeling process, the activity classification results by experts show that every activity concludes within five seconds of an implementation. Therefore, selecting a five-second window size can include every activity without omitting data points of safety regulation.

While machine learning algorithms and LSTM utilize raw data, data for training convolution neural network (CNN) algorithms based on images need to be transformed into simplified images. The signal vector magnitude (SVM*) of the IMU data is used to represent the intensity and severity of movement and reduce the size of the data by one-third. The SVM formula is used to reduce data noise and secure the calculation process's efficiency [36,37]. In this study, 3-axis accelerometer and 3-axis gyroscope data from 5 different body parts (Head-acc, Head-gyro, Right Wrist-acc, Right Wrist-gyro, Left Wrist-acc, Left Wrist-gyro, Right Ankle-acc, Right Ankle-gyro, Left Ankle-acc, and Left Ankle-gyro) is calculated through the SVM shown in Eq. (1). A total of 30 graphs shrinks into 10 graphs, and they are aligned in parallel and transformed into images by each activity of safety regulation.

$$SVM_{ij} = \left[\sum\nolimits_{k=1}^{n} \sqrt{x_k^2 + y_k^2 + z_k^2} \right] \tag{1}$$

In the study by Wang et al. [38], the framework that encodes time series as images using the Gramian Angular Field (GAF) was proposed. GAF has advantages that it can preserve temporal dependency. GAF image can compress extensive IMU graphs transformed into polar coordinates into one image. Generally, the Gramian matrix has a problem that has a large size of n×n, n is the length of the raw time series. However, the data preprocessing segments raw time series data by 5 s of window size. Consequently, GAF images have 640×640 size, which is affordable for training CNN. For transforming IMU data to GAF image, raw IMU data is transformed into polar coordinates. Before that, given a time series $X = \{x_1, x_2, \ldots, x_n\}$ of n real-time value, X need to be rescaled for making all value that they into the interval [−1,1] or [0,1] by:

$$x_{-1}^{\sim i} = \frac{(x_i - max(X) + (x_i - min(X)\,)}{max(X) - min(X)} \tag{2}$$

or

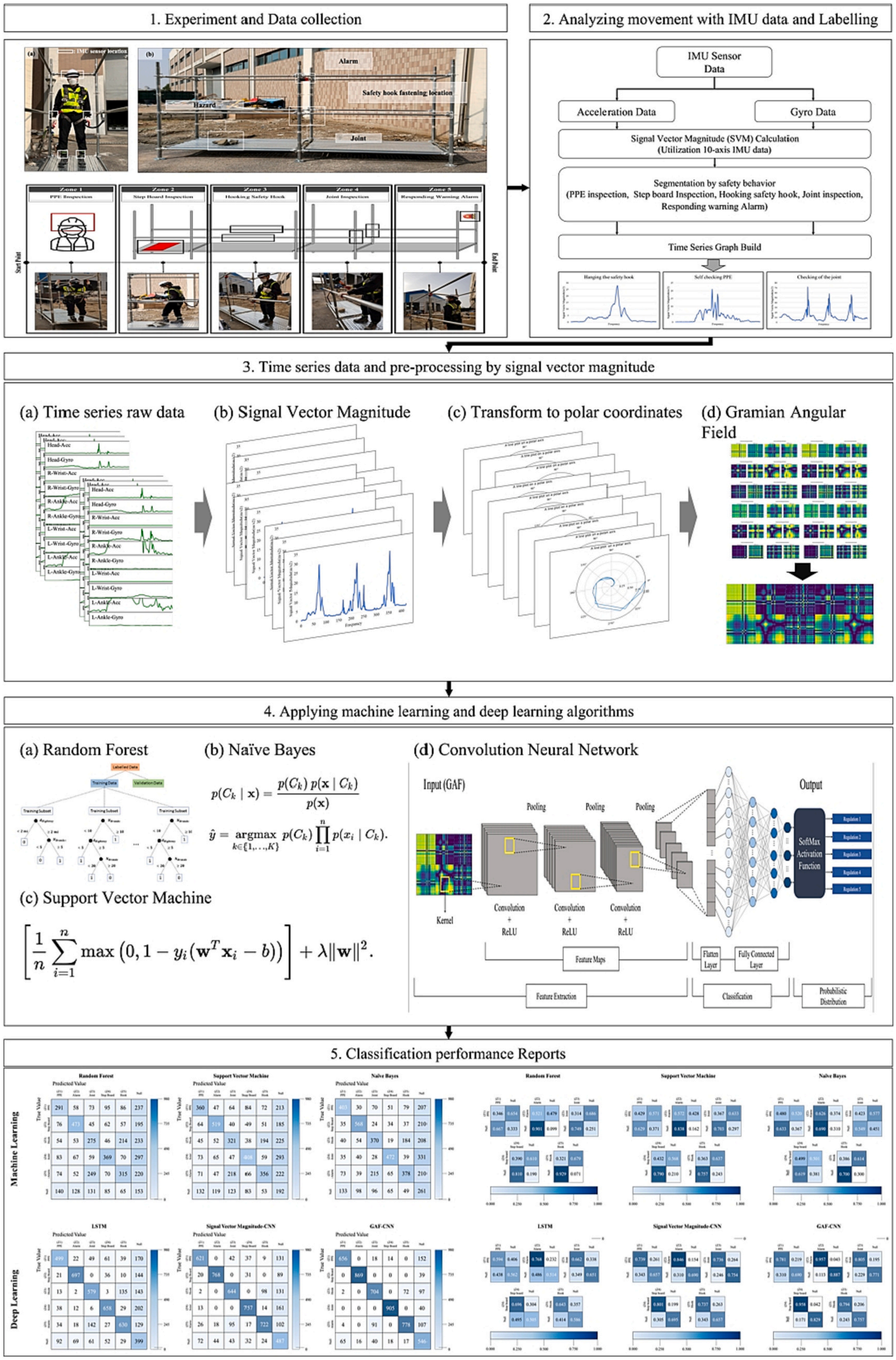

Fig. 1. Framework for safety behavior monitoring in scaffolding work.

Table 1
Subject Information.

Statistical parameters	Height (cm)	Weight (kg)	Age (years)
Mean	171.72	72.56	30.53
Median	171.13	70.53	32
Standard Deviation	6.68	8.71	4.34
Minimum Value	161.55	58.51	22
Maximum Value	183.13	87.86	37

$$x_0^{\sim i} = \frac{(x_i - min(X))}{max(X) - min(X)} \tag{3}$$

Furthermore, rescaled time series $\widetilde{X}$ is transformed into polar coordinates by encoding the value as angular cosine and the time stamp with the equation as follows:

$$\begin{cases} \varnothing = arccos\left(x_i^{\sim}\right), -1 \leq x_i^{\sim} \leq 1, x_i^{\sim} \in \widetilde{X} \\ r = \dfrac{t_i}{N}, t_i \in N \end{cases} \tag{4}$$

In Eq. (4), t_i refers time stamp and N is regularized span of the polar coordinates which is constant factor [39]. The transformed polar coordinates of IMU raw data can exploit the angular perspective to identify temporal correlation. The Gramian summation angular field is defined in Eq. (5).

$$GASF = \left[cos(\varnothing_i - \varnothing_j)\right] = \widetilde{X}' \bullet \widetilde{X} - \sqrt{I - \widetilde{X}^2} \bullet \sqrt{I - \widetilde{X}^2} \tag{5}$$

Fig. 3 (a) represents the visual changes in the transformation process from raw time series data to GAF image. Fig. 3 (b) describes the featured time series and GAF imange of each safety regulation.

3.3. Data analysis

In this study, the proposed GAF-CNN algorithm is compared with five other algorithms. Three machine learning and two deep learning algorithms are applied to compare the classification and identification performances. The five different algorithms are Random Forest (RF) [40], Support Vector Machine (SVM**) [41,42], Naïve Bayes (NB) [43], Long-Short Term Memory (LSTM) [44,45], and SVM*-CNN. This study was conducted to verify the frameworks for classifying each activity and identifying whether workers follow the safety regulations correctly. For examining the feasibility, a two-way approach is used for model training. Firstly, model training for activity classification proceeds. 75% of the total dataset was used for training each algorithm model, and 25% of the total dataset, which is non-overlapped, was used to verify the performance of the model. Then, data was segregated by behaviors to identify behaviors correctly following safety regulations. In each safety regulation, 75% of behaviors correctly following safety regulation data and 75% of behaviors incorrectly following safety regulation data were applied to model training. The other 25% of each data not used to train the model was used for model verification. Data segregation was performed randomly. Tables 2 and 3 represent the parameters of machine learning models in the order of RF and SVM**. Table 4 describes the parameters of deep learning-based models—LSTM and CNN (including SVM*-CNN and GAF-CNN). To investigate the feasibility of the algorithm for this study, both classification and identification models were verified based on the F-1 score as well as precision, recall, and accuracy score.

4. Results

A total of 5250 safety regulations were collected for each activity. In all experiment processes, workers' activities were collected simultaneously through videos. A total of 5250 collected data from workers were analyzed and classified by experts into behaviors following safety regulations correctly and behaviors not following safety regulations. The results are as follows. (Z1) PPE: 840 compliance, 210 non-compliance; (Z2) Warning Alarm: 908 compliance, 142 non-compliance; (Z3) Joint Inspection: 875 compliance, 175 non-compliance; (Z4) Step Board Inspection: 945 compliance, 105 non-compliance; and (Z5) Safety Hook: 980 compliance, 70 non-compliance.

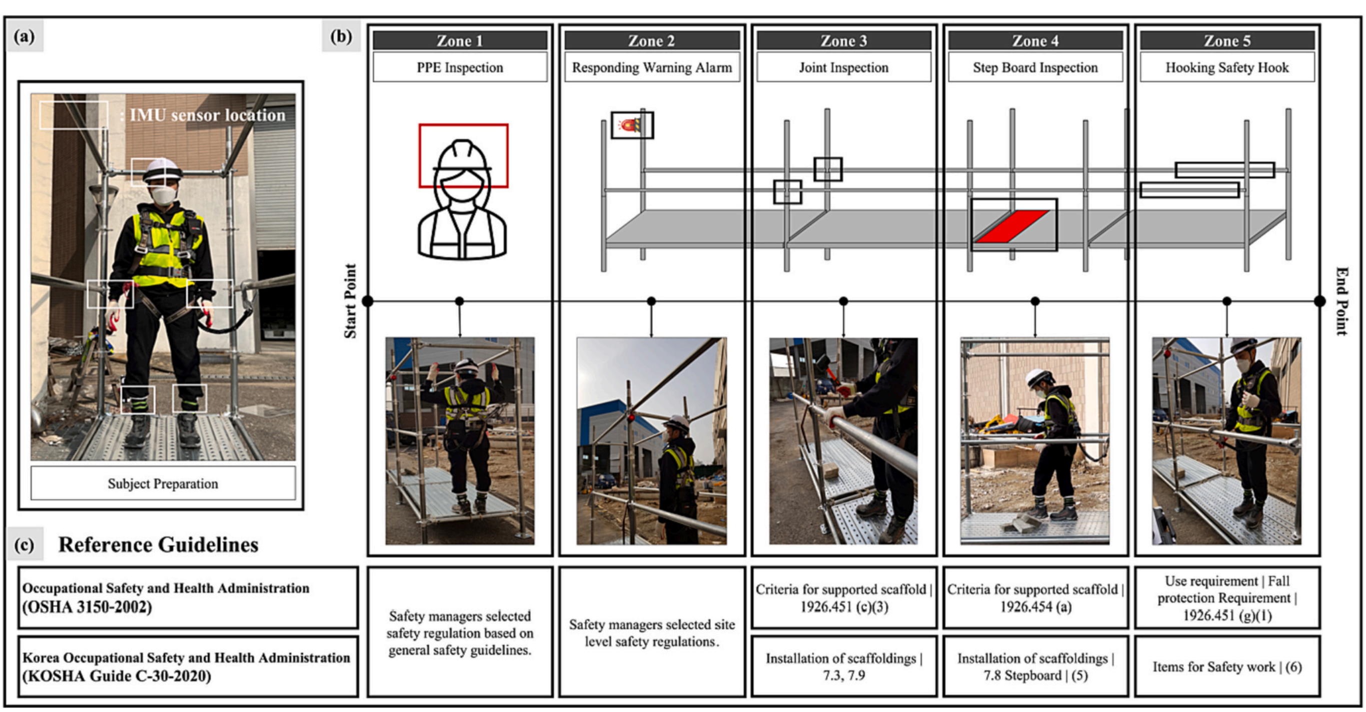

Fig. 2. Overview of the experiment; (a) subject preparation, (b) order of safety regulations, and (c) reference guidelines for each safety regulation.

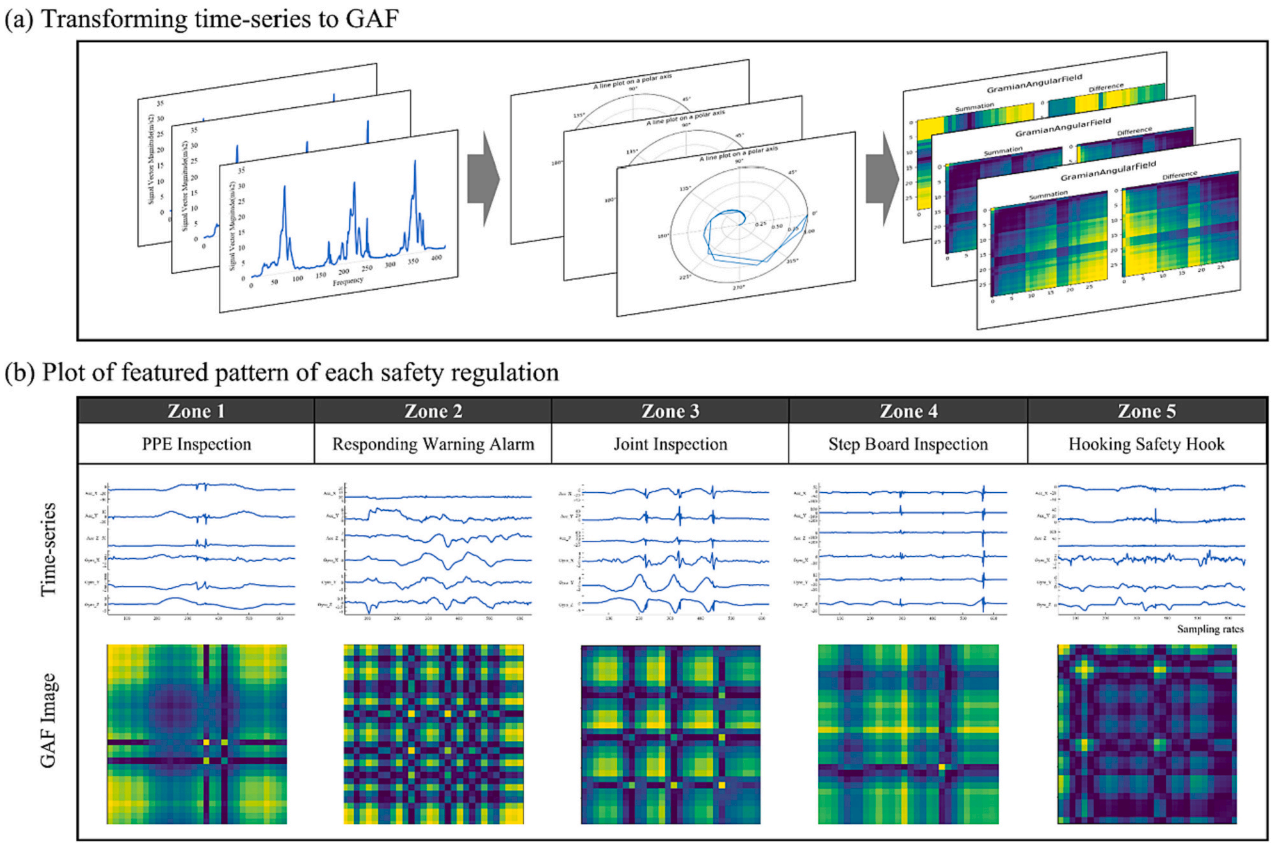

Fig. 3. Time series data and transformed GAF image: (a) transforming process of time-series data, and (b) plot of featured data pattern of each data form.

Table 2
Parameter table of RF.

Parameter name	Value
n_estimators	65
max_depth	15
max_features	auto
min_samples_split	15
min_samples_leaf	15

Table 3
Parameter table of SVM**.

Parameter name	Value
C	1.0
gamma	auto
degree	3
Coef0	0.0
kernal	rbf

Table 4
Hyperparameter table of LSTM and CNN.

Hyperparameter	LSTM	CNN
Layers	3 LSTM layer	2 Conv2D
Filter (# of filters)	–	5*5 (20, 50)
Epoch	250	200
Optimizer	ADAM	ADAM
Learning Rate	0.001	0.001
Batch size	64	16
Activation function	ReLU / Softmax	ReLU / Softmax

4.1. Activity classification results

A total of 35 workers implemented five safety regulations 30 times, respectively. In the procedure of activity classification, the authors have attempted to validate that the model can classify each activity by supervised machine learning and deep learning algorithms. Therefore, model learning was conducted by assigning a label to the data for each activity, and all behaviors incorrectly following regulations were given a Null label. Accuracy in the three machine learning models was derived from 78.58% at RF, 80.36% at SVM**, and 82.23% at NB, respectively. Precision and recall values were derived significantly lower than the accuracy because the performance decrease derived from error caused by the inclusion of non-compliance regulation as Null labels, which is relatively difficult to distinguish. Precision (0.115 at RF, 0.144 at SVM, and 0.183 at NB) and recall (0.218 at RF, 0.274 at SVM, and 0.372 at NB) value of Null is lower than average precision (0.357 at RF, 0.411 at SVM, and 0.467 at NB) and recall (0.357 at RF, 0.411 at SVM, and 0.467 at NB). Therefore, compared to the model's accuracy, the F1-score was relatively low. Each F1-score is 0.357, 0.411, and 0.467, in the order of RF, SVM, and NB.

On the other hand, the overall indicators derived from the deep learning model are higher than the machine learning model. Training accuracy from deep learning models is 88.65%, 92.06%, and 94.97% in order of LSTM, SVM*-CNN, and GAF-CNN. GAF-CNN has the highest performance in activity recognition which shows 0.849 in F1-score. After that, SVM*-CNN shows the second highest of F1-score at 0.762. LSTM derives the lowest representing 0.659 of F1-score within deep learning training results.

(Z2) Warning Alarm has the highest indicators across all algorithms except in recall of GAF-CNN. In the GAF-CNN algorithm, recall of (Z4) Step Board Inspection has the highest indicator as 0.958. Overall, (Z2) shows the highest classification performance, followed by (Z4). Although (Z1), (Z3), and (Z5) are mainly implemented by workers' hands, (Z2) and (Z4) are performed by physical movement of other parts of the body such as the head or foot. Consequently, characteristic signals generated from body attachment sensors other than hands show higher classification performance than activities mainly using hands. Activity classification results are represented in Table 5.

Fig. 4 presents confusion matrices of activity classification. In (Z3), a high acceleration signal is applied to the subject's body by hitting the structure, and in (Z5), a high acceleration signal is also applied to the body in the process of confirming the fastening of the hook. Therefore, similarity in the signal pattern of a specific movement occurs in the two activities. Therefore, (Z3) and (Z5) have the most errors and exhibit a high error tendency in the confusion matrix. On the other hand, (Z2) and (Z4) were generally distinguishable with high probability.

4.2. Activity identification results

It may be more difficult to identify whether a safety behavior undertaken by a worker meets certain criteria (correctly following the safety regulation) than to classify different activities. The supervised learning was conducted for each activity to see if the proposed GAF-CNN algorithm can identify whether safety behaviors correctly follow safety regulations. The algorithm adequacy through model evaluation indicators, including F1-score and accuracy, was derived in the order of RF, SVM**, NB, LSTM, SVM*-CNN, and GAF-CNN. In all algorithms, (Z2) has highest indicators in accuracy and F1-score, which is 0.464, 0.516, 0.583, 0.733, 0.825, and 0.948 in accuracy and 0.627, 0.671, 0.722, 0.833, 0.893, and 0.969 in F1-score. (Z4) follows (Z2) in every indicator except accuracy using the SVM** algorithm (0.410 of accuracy at SVM** in (Z4)). In this case, (Z1) has the second highest value in accuracy, which is 0.417. On the GAF-CNN algorithm, every indicator in every safety regulation outperforms the other algorithms. All indicators in (Z2) and (Z4) show high performance above 0.945. This tendency is similar to the activity classification results. Contrary to the activity classification results, identification results tend to have a higher F1-score than accuracy.

When switching model learning from a machine learning algorithm to a deep learning algorithm, average accuracy and average F1-score increase significantly. The difference between the average accuracy and the average F-1score is 0.054%p, 0.055%p in RF to SVM**, and 0.056%p and 0.051%p when changing from SVM** to NB, respectively. On the other hand, when the machine learning algorithm changes from NB to the deep learning algorithm LSTM, it shows a difference in an average accuracy of 0.192%p and an average F1-score difference of 0.165%p. The detailed indicators are presented in Table 6.

Fig. 5 shows confusion matrices derived based on the results of each algorithm. Learning results derived through machine learning do not yield diagonal results, and the form of an inverse diagonal also occurs. The confusion matrix, which shows the reverse diagonal trend, changes the diagonal direction with darker colors when the deep learning algorithm is used. In the case of using the deep learning algorithm, the best results are derived in identifying the activity of (Z2) and (Z4).

5. Discussions

5.1. Effect of movement characteristics of body parts on results

As shown in the classification and identification results, the performance indicators of (Z2) and (Z4) outperform other safety behaviors (Z1, Z3, and Z5). It may be notable to investigate why these two behaviors show higher performance indicator values. The suggested method (i.e., GAF-CNN) includes the procedures that finds features that are used for facilitating classification. In the case of (Z2), the action

Table 5
Accuracy of training data classifying each activity by algorithm (%).

	Machine Learning Result																	
	Random Forest						Support Vector Machine						Naive Bayes					
Training Accuracy	78.58%						80.36%						82.23%					
Precision	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null
	0.405	0.569	0.331	0.508	0.390	0.115	0.483	0.611	0.395	0.560	0.454	0.144	0.561	0.685	0.461	0.669	0.493	0.183
Recall	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null
	0.346	0.521	0.314	0.390	0.321	0.218	0.429	0.572	0.367	0.432	0.363	0.274	0.480	0.626	0.423	0.499	0.386	0.372
F1-Score	0.357						0.411						0.467					
Deep Learning Result																		
	LSTM						Signal Vector Magnitude-CNN						GAF-CNN					
Training Accuracy	88.65%						92.06%						94.97%					
Precision	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null
	0.716	0.850	0.692	0.786	0.722	0.336	0.824	0.925	0.782	0.866	0.833	0.442	0.902	0.982	0.825	0.966	0.897	0.557
Recall	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null	(Z1)	(Z2)	(Z3)	(Z4)	(Z5)	Null
	0.594	0.768	0.662	0.696	0.643	0.568	0.739	0.846	0.736	0.801	0.737	0.694	0.781	0.957	0.805	0.958	0.794	0.778
F1-Score	0.659						0.762						0.849					

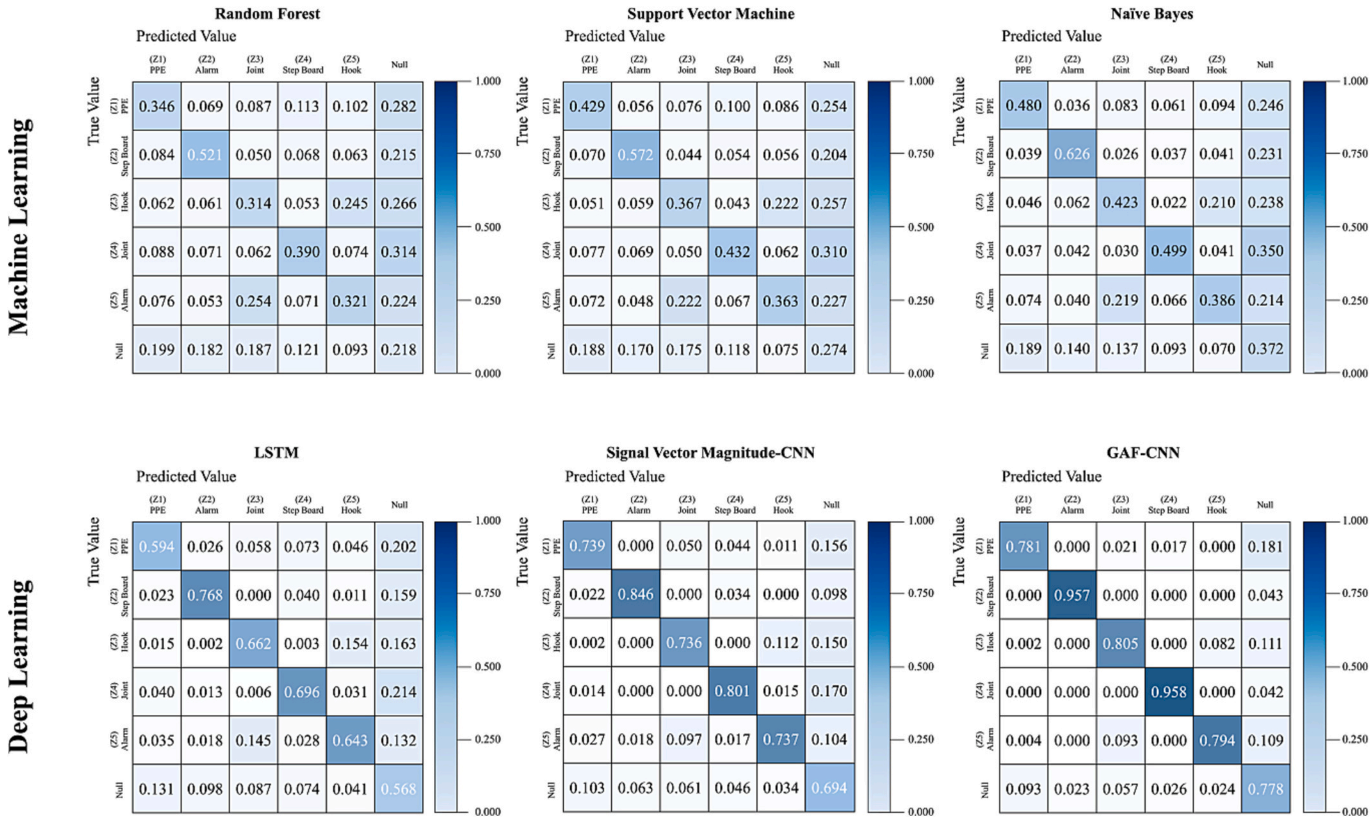

Fig. 4. Results of confusion matrix each classifier.

Table 6
Accuracy and F1-score of identifying standardized activity by algorithms (%).

Zone	Approved Regulation	Accuracy	of approved regulation by each classifier				
		F1-Score					
	Disapproved Regulation	RF	SVM**	NB	LSTM	SVM* -CNN	GAF -CNN
(Z1)	840	0.344	0.417	0.457	0.588	0.723	0.763
PPE	210	0.458	0.541	0.586	0.697	0.810	0.840
(Z2)	908	0.464	0.516	0.583	0.733	0.825	0.948
Alarm	142	0.627	0.671	0.722	0.833	0.893	0.969
(Z3)	875	0.304	0.355	0.428	0.660	0.739	0.799
Joint	175	0.429	0.487	0.552	0.764	0.825	0.870
(Z4)	945	0.370	0.410	0.488	0.677	0.790	0.945
Step Board	105	0.528	0.568	0.637	0.795	0.873	0.969
(Z5)	980	0.305	0.355	0.380	0.639	0.731	0.791
Hook	70	0.463	0.513	0.537	0.769	0.837	0.877

SVM* - Signal Vector Magnitude.
SVM** - Support Vector Machine.

"looking around" means rotating a head. The main features that classify activities seems to be judged by signals from rotating head movement, which is (Z2), within various signals from subjects' movement. Because (Z1), (Z3), (Z4) and (Z5) doesn't include movement that looking around the sites few times, while (Z2) have extraordinary movement (i.e., rotating head).

In addition, (Z4) has a featured characteristic, such as stamping a step board twice. In general, this stamping action is unique in scaffolding work. Although there are steps during the experiment, the signals from steps and stampings are different from each other. During the general scaffolding work situations, walking and stamping patterns can be distinguishable in the IMU data.

On the other hand, the safety regulations (Z1, Z3, and Z5) focus on hand and arm movement. Among these three behaviors, (Z1) shows the lowest performance indicators in the deep learning algorithms, including the GAF-CNN algorithm. By analyzing the miss-classified instances, the authors found an interesting point. In the (Z1), workers were asked to tap their safety helmet twice with both hands. However, there are some instances when workers tap their helmet just once. During the labeling process, tapping a safety helmet twice with two hands is determined as correct behavior. Most workers tapped a safety helmet twice with both hands (right and left hands touching the helmet simultaneously) that is represented in video recording of Fig. 6 (A). However, one worker tapped a safety helmet in both hands (right-hand taps the helmet first, then left taps the helmet) that is represented in video recording of Fig. 6 (B). This worker just tapped once (one right-hand tap and one left-hand tap) in every 30 experiment trials. Although this action was labeled as Null, the algorithm classified it as

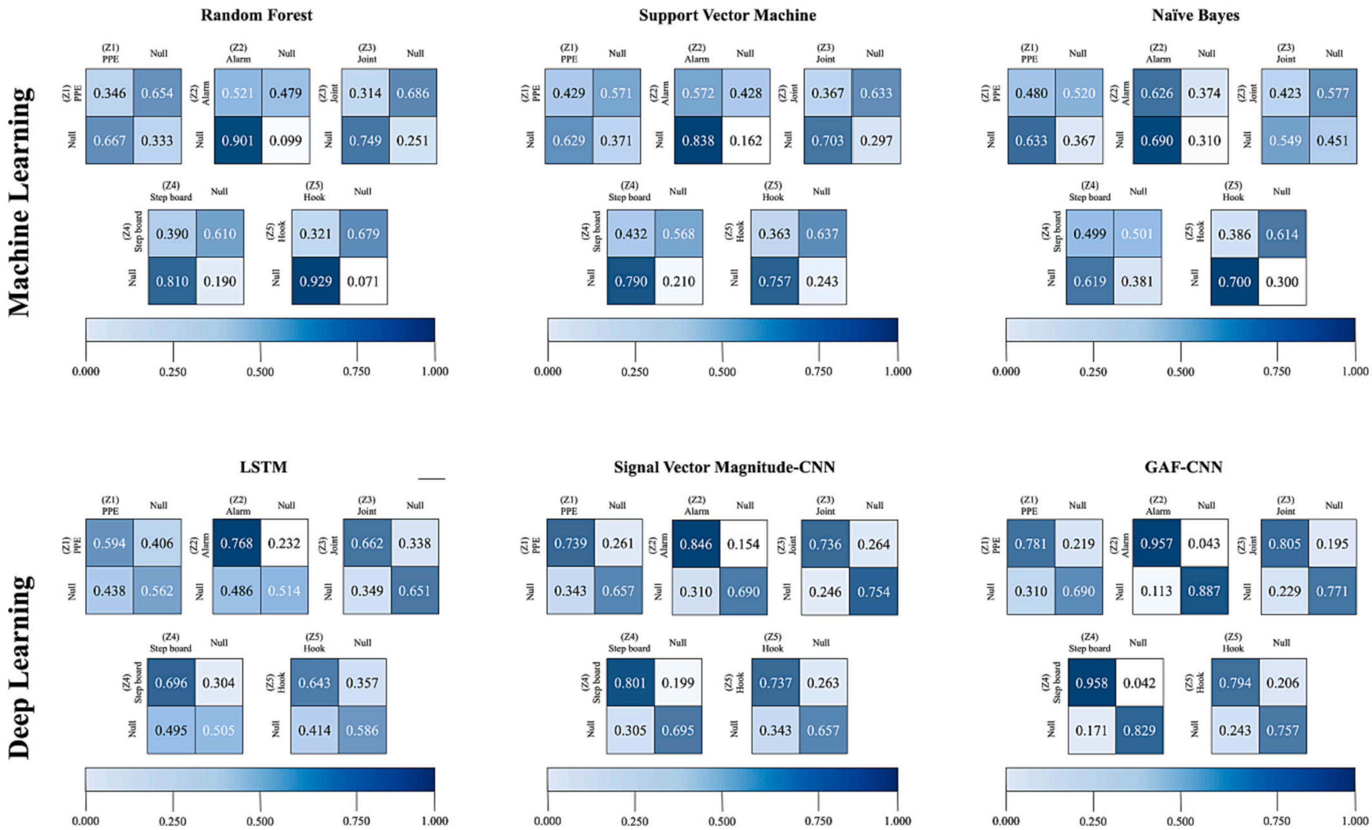

Fig. 5. Results of confusion matrix each classifier.

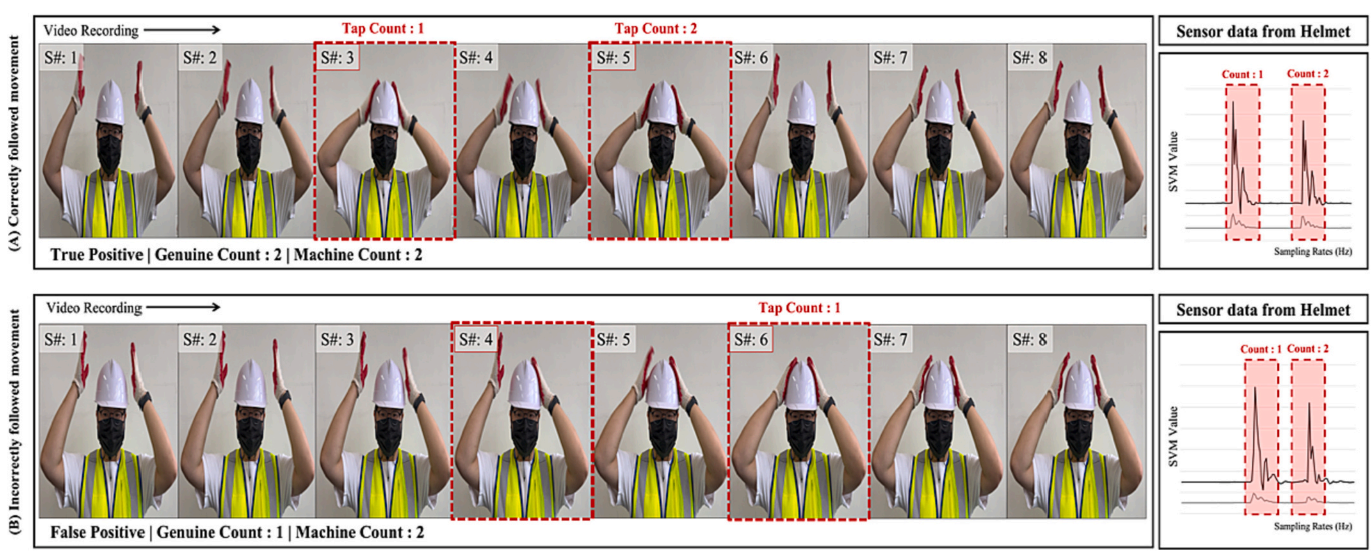

Fig. 6. Experiment to reproduce the cause of false negatives.

correctly following the regulation since tapping a safety helmet in both hands once makes similar signals to tapping a safety helmet twice (correctly following the regulation). This is because collected data from both cases have similar pattern of signals. This confused action may decrease the evaluation performance.

Moreover, the hands of the worker are used extensively for various tasks. For example, workers do not let their hands rest, such as using various tools and cell phones. Therefore, the movement of the hands of the workers acts as a noise of identifying safety regulations, making it difficult to determine whether to implement safety regulations correctly.

5.2. *Expected benefits and potential applications*

This study focuses on the IMU sensors to classify workers' activities and automatically makes a judgment on whether movements have been performed in accordance with a specific criterion. In addition, through the analysis of individual workers, it is possible to provide a personalized safety measure to individual workers (e.g., notifications, corrections, warnings, etc.). Furthermore, if the proposed method is operated in real-time, it is possible to provide a notification about the violation of individual safety regulations to all workers at any location, improving safety management.

Furthermore, although the experiments conducted in this study are limited to scaffolding work, the authors envision that the contents of this study include the possibility of detecting safety regulations in other works. Because all algorithms used in this study can be used to analyze data collected using IMU sensors, they are applicable to other construction tasks other than scaffolding.

5.3. Limitations and future research

A large amount of learning data is required to examine whether safety regulations have been appropriately complied with. Labeling (compliance or non-compliance) should be performed depending on the workers' compliance with safety regulations. In this study, two types of data (image-based and time-series based) were manually labeled. Although the results of this study confirmed that the model learned by the labeled data shows high accuracy, considering the characteristics of construction sites where many workers work simultaneously, automatically performing labeling is considered necessary. In other fields (especially in the smart factory field), if the collected data shows a certain level of deviation, it is classified as anomaly data to minimize the resources required for labeling (unsupervised learning-based classification). To apply labeling methods in other fields, analysis of the threshold level of behavior data (which is for classification as anomaly data) must be conducted in advance. In addition, research is required to be conducted on a method that can automatically label workers' behaviors in consideration of the characteristics of the construction site through various approaches. Similarly, because various situations may occur owing to the specificity of the construction site, applying the safety regulations adopted in this study directly to the site is challenging. Selecting regulations to prevent safety accidents that may occur under various field conditions should be studied in advance. Further research is required on the process of analyzing the validity of the proposed framework according to the safety behavior of selected regulations.

However, the result of learning by converting time-series data into a GAF image secured >90% in this study. Nevertheless, it is considered necessary to apply and analyze various encoding methods developed. Various types of behavioral characteristics occurring in the construction site can affect classification accuracy according to the encoding method. Therefore, comparing and analyzing various encoding techniques are necessary to examine the suitability of techniques for each behavior or task in the future studies.

6. Conclusion

This study examined the feasibility of a framework that automatically identifies whether workers' behavior is in accordance with the safety regulations. The authors hypothesized that scaffolding activities have distinguishable characteristics and that data collected from IMU would be able to classify activities. In this study, a system scaffold was installed, and six safety managers selected five pre-defined safety regulations. During the experiment, 35 healthy subjects conducted five types of safety behaviors related to scaffolding work. When the subjects were conducting the experiment, the video records collected during the experiment were used to determine whether the worker's behavior conforms to the safety regulations. The workers' movements were labeled for each zone and safety regulations and trained to three machine learning algorithms, and three deep learning algorithms. Each model derives the performance of classifying behaviors and the performance of identifying behaviors that follow the safety regulations correctly. Safety behavior classification results presented that RF produced the lowest accuracy and F1-score, and GAF-CNN showed the highest performance. In addition, within the classification results, (Z2) and (Z4) achieved the highest accuracy and F1-score. Similarly, safety behavior identification results showed that the highest performance was achieved with GAF-CNN and that identification results of (Z2) and (Z4) were more evident than the other regulations. Because (Z2) and (Z4) were mainly performed using body parts (i.e., feet or head) other than hands, it was assumed that they have easier conditions to predict activity between (Z2) and (Z4) due to the IMU data characteristics where sensor values were measured for each body part.

The main contribution of this study is that it represents the feasibility of automatically detecting compliance with safety regulations using IMU sensor's. Given that safety behaviors defined by safety regulations reflect less individuality of workers than other behaviors, this study found that the safety behavior of workers in jobsites can be classified and identified with IMU sensors. This finding can be the basis for advanced safety management by continuous activity monitoring approaches that can decrease human error and unstable behaviors under a situation with safety regulations. Moreover, safety management can be improved through personalized safety management training and countermeasures based on individual characteristics and behavioral patterns (safety compliance level) of workers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgment

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) and the Ministry of Education (NRF-2022R1F1A1072450, 2021R1|1A1A01052305).

References

[1] Y. Kang, S. Siddiqui, S.J. Suk, S. Chi, C. Kim, Trends of fall accidents in the U.S. construction industry, J. Constr. Eng. Manag. 143 (2017) 04017043, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001332.

[2] M.F. Antwi-Afari, H. Li, Fall risk assessment of construction workers based on biomechanical gait stability parameters using wearable insole pressure system, Adv. Eng. Inform. 38 (2018) 683–694, https://doi.org/10.1016/j.aei.2018.10.002.

[3] W. Chen, Z. Jiang, H. Guo, X. Ni, Fall detection based on key points of human-skeleton using OpenPose, Symmetry. 12 (2020) 744, https://doi.org/10.3390/sym12050744.

[4] J. Yoon, B. Lee, J. Chun, B. Son, H. Kim, Investigation of the relationship between Ironworker's gait stability and different types of load carrying using wearable sensors, Adv. Eng. Inform. 51 (2022), 101521, https://doi.org/10.1016/j.aei.2021.101521.

[5] M. Antwi-Afari, H. Li, S. Anwer, S. Yevu, Z. Wu, P. Antwi-Afari, I. Kim, Quantifying workers' gait patterns to identify safety hazards in construction using a wearable insole pressure system, Saf. Sci. 129 (2020) 104855, https://doi.org/10.1016/j.ssci.2020.104855.

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

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

[8] Fatal and nonfatal falls, slips, and trips in the construction industry: The Economics Daily: U.S. Bureau of Labor Statistics, (n.d.). https://www.bls.gov/opub/ted/2021/fatal-and-nonfatal-falls-slips-and-trips-in-the-construction-industry.htm (accessed August 13, 2021).

[9] Scaffolding - Overview | Occupational Safety and Health Administration, (n.d.). https://www.osha.gov/scaffolding (accessed January 18, 2022).

[10] N. Chau, G.C. Gauchard, D. Dehaene, L. Benamghar, C. Touron, P.P. Perrin, J.-M. Mur, Contributions of occupational hazards and human factors in occupational injuries and their associations with job, age and type of injuries in railway workers, Int. Arch. Occup. Environ. Health 80 (2007) 517–525, https://doi.org/10.1007/s00420-006-0158-8.

[11] M. Szóstak, B. Hoła, P. Bogusławski, Identification of accident scenarios involving scaffolding, Autom. Constr. 126 (2021), 103690, https://doi.org/10.1016/j.autcon.2021.103690.

[12] Y. Yu, W. Umer, X. Yang, M.F. Antwi-Afari, Posture-related data collection methods for construction workers: a review, Autom. Constr. 124 (2021), 103538, https://doi.org/10.1016/j.autcon.2020.103538.

[13] CII - Education, (n.d.). https://www.construction-institute.org/resources/education (accessed July 25, 2022).

[14] T. Simpson, CPWR | Disaster Response, CPWR |. (n.d.). https://www.cpwr.com/training/training-programs/disaster-response/ (accessed July 25, 2022).

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

[16] H. Li, G. Chan, M. Skitmore, Visualizing safety assessment by integrating the use of game technology, Autom. Constr. 22 (2012) 498–505, https://doi.org/10.1016/j.autcon.2011.11.009.

[17] J. Park, K. Kim, Y.K. Cho, Framework of automated construction-safety monitoring using cloud-enabled BIM and BLE mobile tracking sensors, J. Constr. Eng. Manag. 143 (2017) 05016019, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001223.

[18] R. Sacks, A. Perlman, R. Barak, Construction safety training using immersive virtual reality, Constr. Manag. Econ. 31 (2013) 1005–1017, https://doi.org/10.1080/01446193.2013.828844.

[19] X. Yan, H. Li, A.R. Li, H. Zhang, Wearable IMU-based real-time motion warning system for construction workers' musculoskeletal disorders prevention, Autom. Constr. 74 (2017) 2–11, https://doi.org/10.1016/j.autcon.2016.11.007.

[20] Y.-C. Fang, R.-J. Dzeng, Accelerometer-based fall-portent detection algorithm for construction tiling operation, Autom. Constr. 84 (2017) 214–230, https://doi.org/10.1016/j.autcon.2017.09.015.

[21] R.-J. Dzeng, Y.-C. Fang, I.-C. Chen, A feasibility study of using smartphone built-in accelerometers to detect fall portents, Autom. Constr. 38 (2014) 74–86, https://doi.org/10.1016/j.autcon.2013.11.004.

[22] W. Wu, H. Yang, D.A.S. Chew, S. Yang, A.G.F. Gibb, Q. Li, Towards an autonomous real-time tracking system of near-miss accidents on construction sites, Autom. Constr. 19 (2010) 134–141, https://doi.org/10.1016/j.autcon.2009.11.017.

[23] H. Lingard, S. Rowlinson, Occupational Health and Safety in Construction Project Management, n.d. https://www.routledge.com/Occupational-Health-and-Safety-in-Construction-Project-Management/Lingard-Rowlinson/p/book/9781138103191 (accessed July 25, 2022).

[24] C.R. Ahn, S. Lee, C. Sun, H. Jebelli, K. Yang, B. Choi, Wearable sensing technology applications in construction safety and health, J. Constr. Eng. Manag. 145 (2019) 03119007, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001708.

[25] W. Fang, L. Ding, P.E.D. Love, H. Luo, H. Li, F. Peña-Mora, B. Zhong, C. Zhou, Computer vision applications in construction safety assurance, Autom. Constr. 110 (2020), 103013, https://doi.org/10.1016/j.autcon.2019.103013.

[26] H. Kim, C.R. Ahn, K. Yang, Identifying safety hazards using collective bodily responses of workers, J. Constr. Eng. Manag. 143 (2017) 04016090, https://doi.org/10.1061/(ASCE)CO.1943-7862.0001220.

[27] W. Fang, B. Zhong, N. Zhao, P.E.D. Love, H. Luo, J. Xue, S. Xu, A deep learning-based approach for mitigating falls from height with computer vision: convolutional neural network, Adv. Eng. Inform. 39 (2019) 170–177, https://doi.org/10.1016/j.aei.2018.12.005.

[28] W. Fang, L. Ding, H. Luo, P.E.D. Love, Falls from heights: a computer vision-based approach for safety harness detection, Autom. Constr. 91 (2018) 53–61, https://doi.org/10.1016/j.autcon.2018.02.018.

[29] S. Han, S. Lee, F. Peña-Mora, Vision-based detection of unsafe actions of a construction worker: case study of ladder climbing, J. Comput. Civ. Eng. 27 (2013) 635–644, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000279.

[30] L. Sanhudo, D. Calvetti, J.P. Martins, N.M.M. Ramos, P. Mêda, M.C. Gonçalves, H. Sousa, Activity classification using accelerometers and machine learning for complex construction worker activities, J. Build. Eng. 35 (2021), 102001, https://doi.org/10.1016/j.jobe.2020.102001.

[31] R. Akhavian, A.H. Behzadan, Smartphone-based construction workers' activity recognition and classification, Autom. Constr. 71 (2016) 198–209, https://doi.org/10.1016/j.autcon.2016.08.015.

[32] Y. Tian, J. Zhang, L. Chen, Y. Geng, X. Wang, Single wearable accelerometer-based human activity recognition via kernel discriminant analysis and QPSO-KELM classifier, IEEE Access. 7 (2019) 109216–109227, https://doi.org/10.1109/ACCESS.2019.2933852.

[33] L. Joshua, K. Varghese, Accelerometer-based activity recognition in construction, J. Comput. Civ. Eng. 25 (2011) 370–379, https://doi.org/10.1061/(ASCE)CP.1943-5487.0000097.

[34] J. Yoon, B. Lee, J. Chun, B. Son, H. Kim, Investigation of the relationship between Ironworker's gait stability and different types of load carrying using wearable sensors, Adv. Eng. Inform. 51 (2022), 101521, https://doi.org/10.1016/j.aei.2021.101521.

[35] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, I. Rojas, Window size impact in human activity recognition, Sensors. 14 (2014) 6474–6499, https://doi.org/10.3390/s140406474.

[36] D.M. Karantonis, M.R. Narayanan, M. Mathie, N.H. Lovell, B.G. Celler, Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring, IEEE Trans. Inf. Technol. Biomed. 10 (2006) 156–167, https://doi.org/10.1109/TITB.2005.856864.

[37] N. Bidargaddi, A. Sarela, L. Klingbeil, M. Karunanithi, Detecting walking activity in cardiac rehabilitation by using accelerometer, in: Sensor Networks and Information 2007 3rd International Conference on Intelligent Sensors, 2007, pp. 555–560, https://doi.org/10.1109/ISSNIP.2007.4496903.

[38] Z. Wang, T. Oates, Imaging time-series to improve classification and imputation, in: Proceedings of the 24th International Conference on Artificial Intelligence, Association for the Advancement of Artificial Intelligence Press, Buenos Aires, Argentina, 2015, pp. 3939–3945, https://doi.org/10.48550/arXiv.1506.00327.

[39] N. Hatami, Y. Gavet, J. Debayle, Classification of time-series images using deep convolutional neural networks, in: A. Verikas, P. Radeva, D.P. Nikolaev, J. Zhou (Eds.), 2017 the 10th International Conference on Machine Vision (ICMV 2017), SPIE, Vienne, Austria, 2017, https://doi.org/10.1117/12.2309486.

[40] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[41] D.A. Pisner, D.M. Schnyer, Chapter 6 - support vector machine, in: A. Mechelli, S. Vieira (Eds.), Machine Learning, Academic Press, 2020, pp. 101–121, https://doi.org/10.1016/B978-0-12-815739-8.00006-7.

[42] S. Suthaharan, Support vector machine, in: S. Suthaharan (Ed.), Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning, Springer US, Boston, MA, 2016, pp. 207–235, https://doi.org/10.1007/978-1-4899-7641-3_9.

[43] S. Chen, G.I. Webb, L. Liu, X. Ma, A novel selective naïve Bayes algorithm, Knowl.-Based Syst. 192 (2020), 105361, https://doi.org/10.1016/j.knosys.2019.105361.

[44] F.A. Gers, D. Eck, J. Schmidhuber, Applying LSTM to time series predictable through time-window approaches, in: R. Tagliaferri, M. Marinaro (Eds.), Neural Nets WIRN Vietri-01, Springer, London, 2002, pp. 193–200, https://doi.org/10.1007/978-1-4471-0219-9_20.

[45] M. Ullah, H. Ullah, S.D. Khan, F.A. Cheikh, Stacked lstm network for human activity recognition using smartphone data, in: 2019 8th European Workshop on Visual Information Processing (EUVIP), 2019, pp. 175–180, https://doi.org/10.1109/EUVIP47703.2019.8946180.