https://doi.org/10.1016/j.jobe.2022.104448

Wearable Acceleration-based Action Recognition for Long-term and Continuous Activity Analysis in Construction Site

Yue Gong[1], Kanghyeok Yang[2], JoonOh Seo[3], Jin Gang Lee[4]

[1] Ph.D. Student, Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong. E-mail: yue.gong@connect.polyu.hk

[2] Assistant Professor, School of Architecture, Chonnam National University, Gwangju, Republic of Korea. E-mail: kyang@jnu.ac.kr

[3] Associate Professor, Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong. E-mail: joonoh.seo@polyu.edu.hk

[4] Postdoctoral Fellow, Department of Building and Real Estate, The Hong Kong Polytechnic University, Kowloon, Hong Kong. E-mail: jingang.lee@polyu.edu.hk

ABSTRACT

As construction is labor intensive, improvement in labor productivity is essential for achieving better project performance. Activity analysis, a widely adopted approach to improve labor productivity, measures the time spent on specific activities and can identify the root causes of low productivity. The use of automated action recognition using machine learning-based classification based on data (e.g., accelerations) collected from wearable sensors, which addresses the limitations of observation-based activity analysis, has been introduced as an effective means for monitoring and measuring activities. Despite the potential of acceleration-based action recognition, some challenges still need to be addressed from a practical perspective. For example, action categories defined in previous studies tend to be based on either body movements (e.g., walking, lifting, sitting, and standing) or work contexts (e.g., spreading mortar and laying a concrete block), thereby hindering the comprehensive understanding of the diverse nature of activities in construction. The approach needs to be further tested by noisy and continuous acceleration data collected from construction sites to validate its applicability and practicality in actual use. This research proposes a comprehensive hierarchical activity taxonomy (from Level 1 to Level 3) for acceleration-based action recognition by explicitly categorizing diverse construction activities in accordance with body movements and work contexts to address these issues. The proposed taxonomy was tested by using acceleration data collected from 18 construction workers, including formwork and rebar workers, at two construction sites in Hong Kong. Different machine-learning algorithms were implemented on the basis of hierarchically defined construction activities. Testing results indicate a competitive classification performance on Level 1 activities with 98% accuracy on the identification of work and idling. The prediction accuracy of Level 2 classification is also acceptable, with 90.6% and 86.6% classification accuracy for formwork and rebar work, respectively. Level 3 classification, which reaches an accuracy of 77.1% (formwork) and 74.9% (rebar work), requires further improvement before it can be applied in the construction field. The results of this study shall provide practical insights into the application of acceleration-based automated activity analysis for productivity monitoring.

KEYWORDS

Accelerometer; Action Recognition; Activity Taxonomy; Automation; Productivity; Wearable sensor

1. INTRODUCTION

The construction industry is one of the most labor-intensive industries, and a large portion of construction tasks still relies on the manual workforce (Ng & Tang, 2010). High dependency on manual workforce has been recognized as one of the fundamental causes of low productivity in construction (Jarkas, 2010). In practice, activity analysis, a work sampling method, has been widely used to improve labor productivity by continuously monitoring and measuring construction activities to eliminate the root causes of low productivity (Gouett et al., 2011). In particular, activity analysis quantifies the time spent on specific types of activities that are categorized as productive or nonproductive and then identifies any existing barriers to minimize nonproductive activities. However, current activity measurement mainly relies on time-consuming human observation, which may hinder the application of activity analysis in practice.

Recently, automated action recognition techniques using machine learning-based classification have been introduced, and in particular, the use of acceleration-based action recognition has shown its potential to replace human observers with wearable sensors and algorithms for continuous activity measurement without interfering with ongoing work (Hwang & Lee, 2017). Diverse construction activities involve specific body movements of construction workers, and these movements create unique acceleration signals. Acceleration-based action recognition tries to automatically capture these unique patterns from the signals by using machine learning algorithms and classify diverse construction activities. As action recognition is performed on the basis of a set of time-series acceleration data, the classification results can be used to automatically measure the time spent on specific activities in any construction tasks. Several researchers in construction have examined the reliability and validity of automated activity recognition by using acceleration data collected in laboratory settings or construction sites and demonstrated its great potential for activity analysis (Akhavian & Behzadan, 2016; Bangaru et al., 2021b; Cheng et al., 2013; Joshua & Varghese, 2014; Kwapisz et al., 2011; Sanhudo et al., 2021; Weiss et al., 2016).

Despite the usefulness of acceleration-based action recognition, a few challenges have been identified concerning its practical implementation in ongoing construction tasks. As machine learning algorithms deal with multiclass classification problems, their performance will be affected by how activities are defined. In the construction domain, the action categories tend to be determined on the basis of representative activities of construction work that are the most repeatedly performed. However, confusion among different activities frequently occurs because of the lack of consideration of body movements that will directly affect the pattern of acceleration signals from body-attached sensors. Considering the nonstandardized nature of field operations, the action recognition algorithms frequently suffer from noisy actions (e.g., actions that are unclearly predefined and labeled, or transitional actions). These issues will be more remarkable in acceleration data that are continuously collected in unstructured settings, such as actual construction sites.

This study proposes an acceleration-based action recognition approach by applying a new hierarchical work taxonomy that considers movement and work contexts. This taxonomy aims to extract useful information for activity analysis and reduce classification errors from action recognition algorithms that are based on acceleration data. A comprehensive and universally applicable work taxonomy for construction tasks is proposed by considering 1) whether activities will contribute to productivity, and 2) whether activities will involve unique body movements that can create distinguishable acceleration signals using machine learning algorithms. The proposed taxonomy is validated by using traditional feature-based machine learning and deep learning algorithms for acceleration-based action recognition. In particular, acceleration data are collected from 18 construction workers from two construction sites in an uncontrolled manner by using an inertial measurement unit (IMU) embedded in a smartwatch (i.e., Apple Watch) during concrete work (e.g., formwork and rebar installation) for two months. The collected data are labeled in accordance with the proposed work taxonomy to evaluate the validity of the taxonomy and the classification performance by applying various machine learning algorithms. On the basis of the action classification results, the usefulness of the proposed work taxonomy and its appropriate level of detail are discussed. Future research directions to enhance the practicability of automated activity recognition and activity analysis in a construction workplace are explored.

2. LITERATURE REVIEW ON AUTOMATED ACTION RECOGNITION FOR ACTIVITY ANALYSIS IN CONSTRUCTION

Although the definition of activity analysis can vary depending on the field of application, it commonly refers to a continuous process for improving productivity by workforce assessment through work sampling techniques and then identifying and eliminating factors that inhibit productivity in construction (CII 2010). Workforce assessment measures the time spent on specific activities and calculates direct-work rates as an indicator of productivity (Gouett et al., 2011). Work sampling based on observations is one of the widely used techniques for workforce assessment. Different types of predefined activities are recorded by an observer using a data collection form in a given time interval, and the recorded data can be used to calculate the direct-work rate. However, manual observation has been criticized because of the high cost of hiring observers, the possibility of interfering with ongoing work during observation, and the potential human errors when recording activities based on the observer's judgment (Khosrowpour et al., 2014).

Automated approaches for collecting activity data at construction sites by using sensors, including 1) location sensor-based, 2) visual sensor-based, and 3) wearable sensor-based approaches, have been proposed to address these issues (Khosrowpour et al., 2014). All these approaches have demonstrated their feasibility and applicability for efficiently tracking worker activities, and the wearable sensor-based approach has been recognized as the suitable method for long-term and comprehensive activity analysis during construction tasks (Chen et al., 2012; Wang et al., 2019). Location data that can be detected by using various location tracking sensors (e.g., UWB, RFID, GPS, and Bluetooth beacons) may provide useful information to determine idling or nonidling work, but they cannot differentiate nonidling work at a fixed position (e.g., hammering while standing). Vision-based approaches that can classify activities by analyzing consecutive images from a camera may provide the most accurate and reliable data for activity analysis. However, cameras installed at construction sites can cover only limited areas of the sites. The identification of that worker of interest from different video streams is required, a job that is relatively challenging, to continuously monitor a specific worker by using different cameras. Identity switches may frequently occur when multiple workers are found in the scene. The existence of blind spots is another problem of the vision-based approaches. Thus, vision-based approaches will be the best for activity analysis only at a designated area for a relatively short duration while the workers to be monitored will stay within the camera view. Compared with the two other approaches, the wearable sensor-based approach has comparative advantages in the continuous monitoring of multiple workers. Wearable sensors are attached to workers to collect data associated with construction activities, and the identity of diverse workers can be easily recognized. The body-attached sensor can collect data continuously throughout the construction site, and the collected data can be stored and transferred for further analysis by connecting it to a smartphone. Recently, small, and lightweight wearable sensors, such as wristbands and smart helmets, have become available, thereby minimizing discomfort during ongoing work.

Wearable sensors have drawn much attention and demonstrated their feasibility in acceleration data collection and action recognition (Akhavian & Behzadan, 2016). Most of commercial-grade wearable devices, including fitness trackers or smart watches, have an accelerometer or an IMU with an accelerometer, enabling to collect real-time acceleration signals that represent body movements. Acceleration-based action recognition aims to classify predefined activities by using machine learning algorithms based on the assumption that each action will create its own unique acceleration signals that are specific enough to differentiate diverse activities (Bao & Intille, 2004; Yang et al., 2008). These algorithms have been tested and validated for various applications, such as daily activities, sports, and construction activities. Joshua and Varghese (2011) used wired accelerometers to collect acceleration signals from masonry workers' waists. The classification accuracy for identifying fetching and spreading mortar, fetching, and laying bricks and filling joints is reported as 80%, showing a great potential of acceleration-based action recognition for activity analysis. In a study conducted by Joshua and Varghese (2011), this approach showed relatively good classification performance in identifying effective, contributory, and ineffective activities of ironworkers and carpenters. Akhavian and Behzadan (2016) used a smartphone to collect acceleration data on the participant's upper arm while simulating construction tasks. The machine learning model that was trained using the experimental data achieves over 90% accuracy in differentiating between idling and sawing. The classification accuracy for identifying loading, hauling, unloading, and returning ranges from 70% to 80%. A classification of laying block, adjusting block, removing mortar, and spreading mortar is reported to reach an overall accuracy of 88% in the masonry tests conducted by workers, who wore a single wristband on their dominant hand to collect acceleration information.

Despite the potential of acceleration-based action recognition for continuous and automated activity analysis, some challenges that need to be overcome before it can be put into practice. Previous studies mainly tested the feasibility of this approach in laboratory settings where participants were asked to conduct certain instructed activities (Weiss et al., 2016). However, the activities of interest in these studies, unlike tasks in real-life situations, have higher repetitiveness and fewer transition patterns between activities. For the field validation, Joshua and Varghese (2014) collected acceleration data from construction workers during actual construction tasks (e.g., ironwork and carpentry) and focused on classifying three categories of activities (i.e., effective, contributory, and ineffective tasks). Although the testing results showed approximately 90% and 78% accuracy for ironwork and carpentry, respectively, such results are achieved from a relatively short duration of data sampling (e.g., 30–40 min). This fact implies the need for further validations using long-term continuous data that include noisier signals. Previous studies mainly focused on the validation of the accuracy of action recognition rather than designing a better activity taxonomy that can offer a better understanding of the ongoing construction activities. In acceleration-based action recognition algorithms, the defining of actions will affect the classification performance and the utilization of the classification results. As shown in Table 1, acceleration-based action recognition studies tend to label acceleration signals in accordance with 1) movement-oriented activities, such as standing, sitting, walking, hammering, and screwing or 2) work context-oriented activities, such as spreading mortar, fetching and laying bricks, and filling joints for masonry work. The defining of action categories in the algorithm based on the characteristics of body movements will have an advantage of more accurate classification of activities because acceleration signals from the activities will create more distinguishable signals. However, the additional judgment of the types of activities may be needed to derive knowledge for activity analysis when applying such categories. The use of work context-oriented activities as labels for action recognition can provide more intuitive knowledge on measuring work expenditures during activity analysis, thereby helping evaluate work efficiency and expose delay issues. The work context-oriented activity recognition can result in poor classification performance, especially when the classified activities include similar body movements. From a work context perspective, activities for formwork include assembling and stripping of formwork, both of which involve hammering. Thus, acceleration-based action recognition algorithms cannot correctly classify such activities due to the similar patterns of acceleration signals. The activity taxonomy for action classification needs to be defined to include movement and work context to achieve high-performance classification results with rich information on construction activities.

Taxonomy criteria	**Activity category[1]**	**Classification accuracy[2]**	**Data collection method[3]**	**Research**
Motion	Basic task: Connecting, covering, cutting, digging,	-	Observation	Everett and Slocum (1994)

	finishing, inspecting, measuring, placing, planning, positioning, spraying, spreading			
Motion	Walking, tying rebar guiding crane, between activities	-	Automation (camera)	Buchholz et al. (2003)
Motion	Loading, pushing, unloading, returning, idling	87% to 97% (user-dependent) and 62% to 96% (user-independent)	Automation (smartphone)	Akhavian and Behzadan (2016)
Context	Work, material, travel, and idle	-	Automation (location sensor and accelerometer)	Cheng et al. (2013)
Context	Direct work, tools and materials, instructions and drawings, crane deliveries, minor contributory work, travel, idle, unexplained, waiting, no contact	-	Observation	Thomas and Daily (1983)
Context	Effective work, essential contributory work, ineffective work	90.1% (iron work) and 77.7% (carpentry)	Automation (IMU)	Joshua (2014)
Context	Spreading mortar, laying blocks, adjusting blocks, removing mortar	88.1%	Automation (IMU)	(Ryu et al., 2019)
Motion	Sitting, lying down, walking, walking upstairs, walking downstairs, stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, lie-to-stand	89.6%	Automation (accelerometer)	Hassan et al. (2018)
Motion	Jogging, walking, upstairs, downstairs, sitting, standing	97.6%	Automation (accelerometer)	(Ignatov, 2018)
Motion	Run, walk, still	92.7%	Automation (accelerometer)	(Lee et al., 2017)
Motion	Biological Motion Library (BML): knocking, lifting, throwing, walking. Multimodal Human Action Database (MHAD): jumping, jumping jacks, bending, punching, waving (two hands), waving (one hand), clapping, throwing, sit-down/stand-up, sit-down, stand-up	99% (BML) and 99% (MHAD)	Automation (magnetic induction sensor)	(Golestani & Moghaddam, 2020)
Motion and context	PAMAP2 Dataset: lie, sit, stand, walk, run,	94.5%	Automation (accelerometer)	(Xu et al., 2019)

	cycle, Nordic walk, iron, vacuum clean, rope jump, ascend and descend stairs, watch TV, computer work, drive car, fold laundry, clean house, play soccer			
Motion	MHEALTH Dataset: standing still, sitting and relaxing, lying down, walking, climbing stairs, waist bends forward, frontal elevation of arms, knees bending, cycling, jogging, running, jump front & back	99.6%	Automation (accelerometer and ECG)	(Gumaei et al., 2019)
Motion	Standing, bending-up, bending, bending-down, squatting-up, squatting, squatting-down, walking, twisting, working overhead, kneeling-up, kneeling, kneeling-down, and using stairs	94.7%	Automation (IMU)	(Kim & Cho, 2020)
Motion	Adjusting levelling jacks, carrying crossbars, carrying levelling jacks, and carrying scaffold plank, carrying scaffold frame, dragging scaffold plank, hammering, inserting jacks into scaffold frame, lifting scaffold plank from elbow to overhead, walking, wrenching, climb, downstairs, climb with tool bag, downstairs with tool bag	93.3%	Automation (accelerometer and ECG)	(Bangaru et al., 2021a)

Table 1. Activity taxonomy used in activity recognition research

3. METHODOLOGY

This research proposes a comprehensive activity taxonomy considering the characteristics of workers' movements and the work context that will serve as action labels for acceleration-based recognition algorithms and investigates the validity of the algorithms in practice by using continuously collected field data. Figure 1 illustrates the overall research framework. A comprehensive activity taxonomy aiming to effectively measure activities required for identifying productivity issues while minimizing possible confusion in action classification was proposed. For field validation, two local construction sites in Hong Kong were recruited, and continuous acceleration data during construction works (e.g., rebar and formwork) were collected by using an IMU-embedded smartwatch, and videos were simultaneously recorded by using a chest-mounted portable video camera for labeling activities. Machine learning-based classification algorithms were applied to the collected acceleration data for automatically classifying diverse activities that were defined on the basis of the proposed activity taxonomy. The validity of the proposed activity taxonomy for action recognition and the applicability for workforce assessment were examined on the basis of the classification performance.

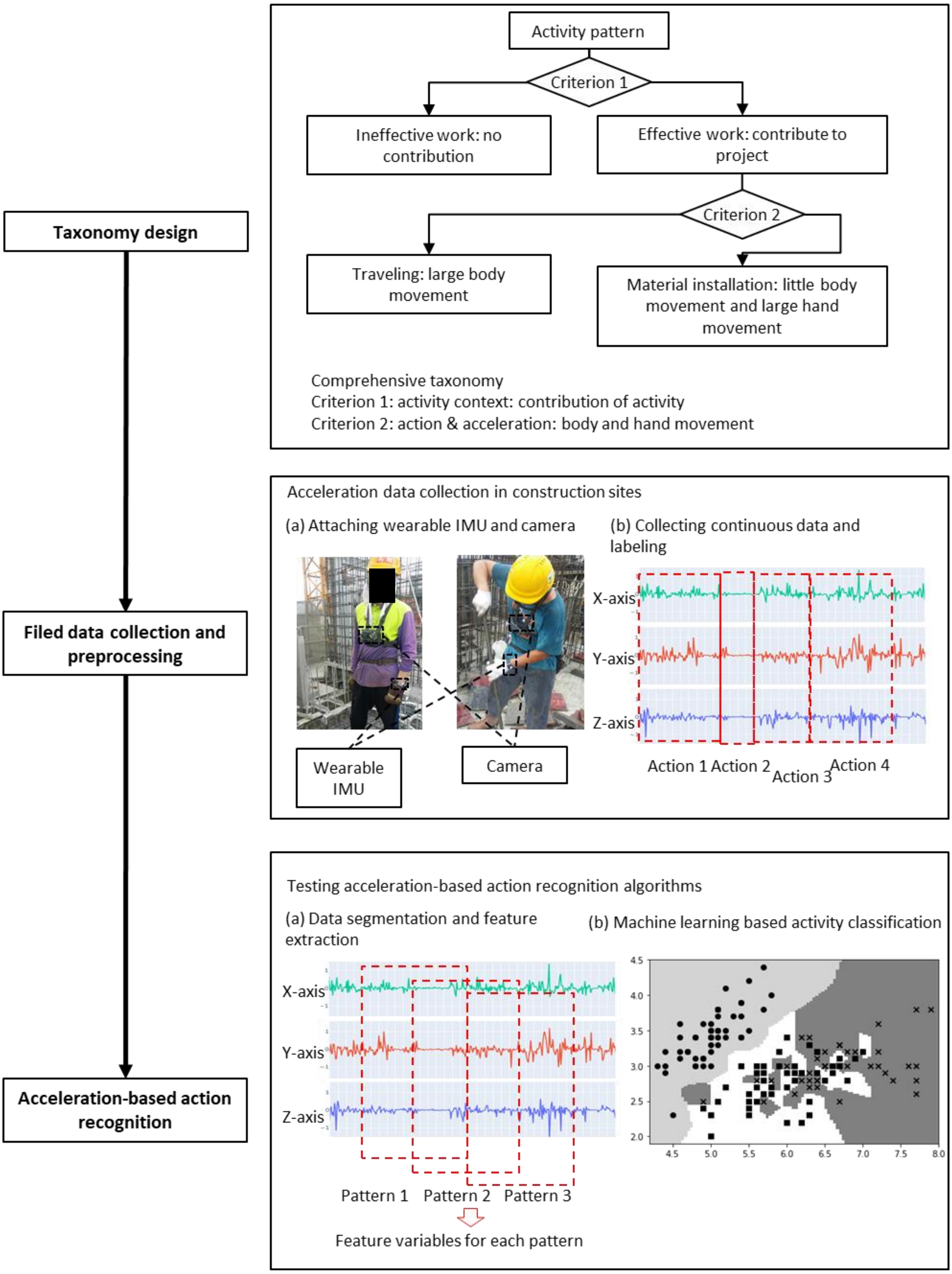


Figure 1. Research framework

3.1 Activity taxonomy design

The proposed comprehensive activity taxonomy consists of three hierarchical levels of activities to effectively extract activity-related information and to better understand the work context performed by a construction worker (Table 2). The first criterion for categorizing activities is whether the activity is relevant to the production process, and the activities are classified into “idling” (e.g., standing and sitting) or “work” at Level 1. As an offspring activity category of “work,” activities at Level 2 are defined in accordance with activity-related movements, depending on whether they involve hand-dominant or whole body-dominant movements. As the acceleration data are collected from a smartwatch, the signals will be more dominantly affected by hand movements and less affected by whole-body movements. By classifying Level 2 activities into “traveling” that involves horizontal whole-body movements and “material installation” that is associated with hand-dominant activities, the acceleration signals from the two activities can be more distinguishable. Three activity categories at Level 2 that include “stationary,” “traveling,” and “material installation” can provide information to evaluate work efficiency of the operations to be monitored. For example, the longer time spent on “material installation” may indicate that the operation will be more efficient for producing outputs. The activities at Level 3 focus more on understanding the work context that will help identify productivity inhibitors. For this purpose, “traveling” is further classified into “transportation” and “transferring materials and tools” at Level 3, and “material installation” is divided into four subactivities, including “material preparation,” “material connecting,” “material placing,” and “supplement work.” Detecting the problematic activities that can lead to inefficiency in activities at Level 2 is possible by further classifying activities at Level 3. However, as the activity categories at Level 3 are based on general work contexts, they can be applicable to any other construction operations that involve delivering and installing materials for certain building components. However, some activities, including intermittent or supportive activities for other activities at Level 3, are unclearly classified on the basis of work contexts. These activities are included in “supplement work”. Table 2 shows examples of basic tasks that can be included in activities at Level 3 for rebar work and formwork that are operations to be tested in this study. For example, “material preparation” that refers to producing components for further operation can include several basic tasks, such as cutting, bending, and drilling. “Material connecting” is the assembling tasks, including fixing, tiling, screwing, and knocking, and material placing represents the lifting and adjusting of associated components. “Supplement work” includes all supportive movements that occur during the installation process. The basic tasks shown in **Error! Reference source not found.** can be used for a better categorization of the activity taxonomy in this study and for precisely recognizing in the labeling procedure. Such segmentation of tasks for activities in Level 3 will help understand the context of activities but will also increase the uncertainty of an automated activity classification using a wearable sensor. Specifically, the classification of Level 3 activities is questionable due to the similarity and dissimilarity of acceleration signals from different activities. For instance, knocking and cutting movements are the offspring activities of “material installation” that will generate cyclic acceleration data with repetitive hand movements. Consequently, distinguishing the Level 3 activities for “material installation” solely by hand movements is difficult because each category of the operation comprises dynamic and complex hand movements. Transportation and transferring of materials/tools will have different hand movements. The hand will swing periodically in “transportation” activities (e.g., walking) or sway (e.g., adjusting tool while walking) mildly (e.g., holding material steady while walking) in material or tool “transferring’ activities.” These facts lead this research to investigate the activity classification performance with a proposed activity taxonomy (Table 2).

Level 1 Activity	**Level 2 Activity**	**Level 3 Activity**	**Basic task**
Idling	Stationary (Ineffective)	Standing/sitting	Standing, sitting
Work	Traveling (Supportive work)	Transportation	Horizontal, vertical and inclined movement, jumping, striding, going upstairs/downstairs, climbing up/down a ladder
		Transferring materials and tools	Carrying materials in horizontal, vertical and inclined movement, carrying materials while going upstairs/downstairs and climbing ladders, dynamical wrist movement while traveling
	Material Installation (Effective work)	Material preparation	Rebar work: cutting, bending Formwork: cutting, measuring, and drawing
		Material connecting	Rebar work: fixing, tying, installing stirrup Formwork: screwing, drilling, knocking, removing nails
		Material placing	Rebar work: placing, adjusting, lifting Formwork: Attaching, adjusting, lifting formwork
		Supplement work	Lifting materials and tools, squatting, standing up, rotating trunk, transition movement

Table 2. Activity taxonomy

3.2 Collecting and preprocessing of data

Figure 2. Site photos for data collection

Data collection was performed during formwork and rebar work (Figure 2) to study the validity of the proposed activity taxonomy and the performance of the acceleration-based activity recognition approach. Nineteen individual periods were involved in the data collection, and each period lasted a whole workday. A large-scale dataset that included 498 h of videos and 2.8 billion samples of acceleration data was constructed from 18 construction workers. Each participant was equipped with an Apple Watch, which was embedded with a sensor, in the dominant hand to record cumulatively 3D acceleration data through a self-developed WatchOS app. The frequency of data collection was set to 100 Hz, indicating that the wearable sensor recorded 100 acceleration data sets for each second. A chest-mounted GoPro camera was used to record their hand movements simultaneously for data labeling. The videos were recorded in 30 FPS, allowing the ground truth of activity information to be captured and stored in a stable and durable manner. The data collections were conducted for two sessions per day (i.e., morning session and afternoon session), and each session lasted 2 h or so. The equipment was taken off during the lunch break because the device needed to be calibrated again before starting the afternoon session. The collected acceleration signals were labeled for each data point based on the researchers' observation on video recordings. Each video frame was labeled by using one of the activities defined at each level of the proposed activity taxonomy based on the observer's judgement. Corresponding acceleration signals were labeled by comparing time information for each data point. In some video scenes, workers' hand activities were unclearly captured. In this case, the activities were determined on the basis of the observations of overall sequences of activities. However, one of the challenges for data labeling is to judge the boundary of consecutive activities. The boundary was determined on the basis of the starting time of the following activity for consistent labeling. If there are significant transitions between two consecutive activities, then these transitional activities were labeled as "supplement work" considering their work contexts. Unqualified data, such as collected under poor light conditions and data recorded during the break in the restroom, were excluded from further processing to avoid possible confusion caused by bad judgment on ongoing activity.

3.3 Machine learning-based activity recognition

Traditional machine learning and deep learning algorithms were applied to test the applicability of the proposed activity taxonomy. A sliding window technique was applied when segmenting labeled acceleration signals into patterns of equal size because any human activities should last for a particular duration ~~(Banos et al., 2014)~~(Banos et al., 2014). The length of the window was determined by considering the nature of construction activity. On the basis of the experience of previous research (Ryu et al., 2019), this study tested multiple window lengths (i.e., 0.5, 1.0, 1.5, 2.0, 2.5, 3, 3.5, and 4.0 s) and determined the optimal window length in accordance with the classification accuracy. The activity labels of each segmented data were determined on the basis of the majority voting rule when data points with multiple activity labels were found within the window (Ballabio et al., 2019).

For classifiers of different activities, this study investigated traditional feature-based machine learning and deep learning approaches for performance comparison. As traditional machine learning classifieds, we selected three classifiers that had been widely applied for activity recognition, namely, 1) ensemble bagged trees (Dietterich, 2000), 2) support vector machine (Hsu & Lin, 2002), and 3) k-nearest neighbor (Sutton, 2012). The Classification Learner app in MATLAB (2019a, MathWorks) was utilized to train and test the models for identifing the best-performing classifier and corresponding hyperparameters, aiming to validate the feasibility of the proposed taxonomy. Typical features applied in the activity recognition were time-domain features and frequency-domain features (Preece et al., 2009). Time-domain features interpret the statistical characteristics of motion signals, including but not limited to the mean, maximum, median, and variance of the signals (Figo et al., 2010). Specifically, this study used eight time-domain features that consist of mean value, minimum value, maximum value, range, standard deviation, kurtosis, correlation, and skewness of acceleration signals in X, Y, and Z axis. Two frequency-domain features, energy and entropy, were used to capture the acceleration streams in terms of frequency, which evaluate action complexity in acceleration-based activity analysis (Ryu et al., 2019). Fast Fourier transform was applied to extract frequency-domain features from raw signals (Preece et al., 2009).

This study tested deep learning algorithms that had the comparative benefits of eliminating the need for hand-crafted features and can save time and effort in the selection and optimization of features and the reduction of human bias (Krizhevsky et al., 2012). This study implemented a bidirectional long short-term memory (BiLSTM), one of the deep learning algorithms known to provide reliable classification performance for acceleration-based action recognition (Yang et al., 2020). The designed architecture is shown in Figure 3.

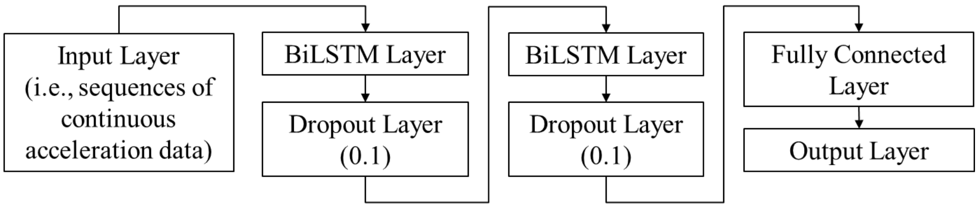


Figure 3. Architecture of the deep learning algorithm

Two types of cross-validation techniques were applied to evaluate the performance of classifiers for each level of activities: 1) leave-one-out cross-validation (LOOCV) and 2) leave-one-subject-out cross-validation (LOSOCV). For leave-one-out cross-validation, the whole data set was randomly separated into five exclusive subsets of equal sizes. Each subset was utilized as testing data for each trial of validation, and the remaining datasets were used for training the machine learning models. The average prediction accuracy of the five validation tests was regarded as the classification performance of the designed algorithm, indicating the overall accuracy of the trained model (Refaeilzadeh et al., 2009). To investigate subject-to-subject variation, we conducted the LOSOCV, which selects one worker's data as testing data once a time and the data from other workers for training the models (Berrar, 2019). The classification models were trained and tested by using different levels of activity data (Levels 1, 2, and 3) to examine whether the classification results at each level will be accurate and reliable for understanding productivity issues during construction operations. The action classification results at each level of the work taxonomy are presented by using the confusion matrices, where each row represents actual classes, and each column corresponds to predicted classes (Mantyjarvi et al., 2001). In the confusion matrix, recall quantifies the fraction of positive observations that are correctly predicted, and precision calculates the ratios of correct predictions that are actually positive (Davis & Goadrich, 2006).

In addition to randomly selecting the training and testing data, this study tested the algorithms with continuous data. In particular, the continuous pattern of acceleration data was used for training models, and the trained model was evaluated with strictly continuous acceleration signals. As continuous acceleration signals reflect real construction tasks better than randomly selected data, the prediction results are supposed to show more realistic action recognition performance in practice. Postprocessing techniques were applied to benefit from this additional information of continuous data (Gil-Martín et al., 2020). On the basis of our preliminary examination of the results, some errors were frequently observed in the middle of ongoing work for a specific activity, and the misclassified data were relatively short, lasting only for 1 or 2 s. Considering the context of the construction activities, this intermittent class found in the classification results will be likely an error. Thus, if the predicted class of the activity 1) lasts less than the unit length of sliding window and 2) the class is observed in the middle of other continuously lasting activities, then this intermittent class was regarded as a misclassified class, and the class was modified as adjacent classes. After the postprocessing procedure, the study then calculated how much time was spent on each activity, which can potentially help evaluate the productivity of each worker.

4. RESULT

4.1 Accuracy of the trained models

Window size is a crucial parameter for accelerometer-based activity recognition. This study investigated the window size through pretesting, and the optimal window size was decided as 1.5 s after multiple tests. Among the various machine learning algorithms we used, a bagged-tree ensemble model showed the best classification performance at the pretesting. Table 3 shows the overall accuracy of classification results for three levels of activities according to 1) classifiers (i.e., traditional machine learning and deep learning algorithms), 2) validation methods (i.e., LOOCV and LOSOCV) and 3) data sampling (i.e., discrete data and continuous data). According to the results from LOOCV, Level 1 classification shows have excellent performance over 90% of accuracy while the deep learning model (i.e., BiLSTM) shows slightly better accuracy than the machine learning model (i.e., Ensemble Bagged Trees). At Level 2, the classification results from LOOCV range from 80%~90%, and again the deep learning model showed better accuracy especially for formwork. At Level 3, the deep learning model showed significantly higher classification performance than the traditional machine learning, indicating the use of the deep learning algorithms would be recommended to classify complex construction activities. However, the overall accuracy at Level 3 was about 77.0% and 74.9% for formwork and rebar work respectively even when using the deep learning model. When testing the classifiers using continuous data (i.e., LOOCV with continuous data) or LOSOCV, the overall accuracy tends to significantly drop, compared with the results from LOOCV. This may indicate that significant variations may exist in the collected data according to the time when the data collected and the subjects.

Work division			Formwork			Rebar work		
Testing data selection			LOOCV with discrete data	LOOCV with continuous data	LOSOCV	LOOCV with discrete data	LOOCV with continuous data	LOSOCV
Prediction Accuracy	Machine Learning (Ensemble Bagged Trees)	Level 1 activity	96.2%	95.3%	93.7%	95.7%	96.1%	93.5%
		Level 2 activity	83.8%	81.2%	78.5%	79.5%	74.6%	76.6%
		Level 3 activity	61.3%	50.3%	42.9%	57.1%	45.3%	44.7%
	Deep Learning (BiLSTM)	Level 1 activity	98.7%	98.9%	94.7%	98.6%	98.3%	97.2%
		Level 2 activity	90.6%	81.6%	77.8%	86.6%	79.3%	77.2%
		Level 3 activity	77.1%	55.7%	49.0%	74.9%	57.7%	55.6%

*LOOCV: leave-one-out cross-validation, LOSOCV: leave-one-subject-out cross-validation

Table 3. Overview of prediction accuracy

Based on the results from LOOCV with discrete data, the confusion matrices of all formwork and rebar work activities, the predicted category, actual category, precision, recall, and F1 score of each activity are presented in **Error! Reference source not found.** and Table 5. As shown in the Level 2 confusion matrix, the majority of incorrect predictions of traveling are reported as coming from rebar installation or form installation. For instance, Level 2 classification in **Error! Reference source not found.** shows that 82.6% of the predictions are form installation but such results actually belong to traveling. The prediction errors (98.0%) of form installation are misclassification between form installation and traveling. Given such consequences, the most significant errors are caused by confusion between traveling and rebar or form installation at Level 2 activities. In the Level 3 confusion matrices, the fractions of activity that are misclassified as supplement work are 73.4%, 80.5%, 82.6%, 76.5%, and 83.7% in the negative predictions of form placing, form connecting, form preparation, transferring materials and tools, and transportation, respectively. As shown in Table 5, the same issue is also observed in the activity recognition for rebar work. In this regard, supplement work at Level 3 activities is the most dynamic activity that caused considerable confusion with traveling-related activities and other material installation activities. Such facts might imply that the confusion between form or rebar installation and traveling at Level 2 is mainly due to the confusion between supplement work and traveling-related activities at Level 3.

Level 1 Activity	Predicted category	W	I	Recall (%)
True category	W	25894	380	98.6
	I	1183	12759	91.5
	Precision (%)	95.6	97.1	
	F1 Score	1.0	0.9	

* W: Work, I: Idling

Level 2 Activity	Predicted category	W_FI	W_TR	I_SS	Recall (%)
True category	W_FI	19889	545	459	95.2
	W_TR	4431	845	89	15.8
	I_SS	995	16	12947	92.8
	Precision (%)	78.6	60.1	95.9	
	F1 Score	0.9	0.3	0.9	

* W_FI: Form installation, W_TR: Traveling, I_SS: Stand/sit

Level 3 Activity	Predicted category	W_FI_SP	W_FI_PL	W_FI_CT	W_FI_PA	W_TR_MT	W_TR_SP	I_SS_ST	Recall (%)
True category	W_FI_SP	7863	2	529	738	36	468	372	78.6
	W_FI_PL	643	13	73	100	2	16	42	1.5
	W_FI_CT	2818	1	1557	374	8	121	180	30.8
	W_FI_PA	3039	2	277	1005	9	118	234	21.5
	W_TR_MT	1256	0	73	110	82	131	72	4.8
	W_TR_SP	2338	1	149	196	1 4	834	109	22.9
	I_SS_ST	477	0	58	126	5	25	13283	95.1
	Precision (%)	42.3	68.4	57.0	37.6	52.6	48.0	92.9	
	F1 Score	0.6	0.0	0.4	0.3	0.1	0.3	0.9	

* W_FI_SP: Supplement work, W_FI_PL: Form placing, W_FI_CT: Form connecting, W_FI_PA: Form preparation, W_TR_MT: Transferring materials and tools, W_TR_SP: Transportation, I_SS_ST: Standing/Sitting

Table 4. Confusion matrix of formwork activity classification

Level 1 Activity	Predicted category	W	I	Recall (%)
True category	W	25001	301	98.8
	I	812	9622	92.2
	Precision (%)	96.9	97.0	
	F1 Score	1.0	1.0	

* W: Work, I: Idling

Level 2 Activity	Predicted category	W_RI	W_TR	I_SS	Recall (%)
True category	W_RI	16568	1200	303	91.7
	W_TR	4911	2212	104	30.6
	I_SS	682	42	9714	93.1
	Precision (%)	74.8	64.0	96.0	
	F1 Score	0.8	0.4	1.0	

* W_RI: Rebar installation, W_TR: Traveling, I_SS: Stand/sit

Level 3 Activity	Predicted category	W_RI_SP	W_RI_PL	W_RI_CT	W_RI_PA	W_TR_MT	W_TR_SP	I_SS_ST	Recall (%)
True category	W_RI_SP	1498	73	362	26	3	434	103	59.9
	W_RI_PL	414	240	211	14	1	169	46	21.9
	W_RI_CT	665	86	688	22	0	236	69	39.0.
	W_RI_PA	307	39	148	92	0	172	52	11.4
	W_TR_MT	73	5	13	3	21	98	3	9.7
	W_TR_SP	696	43	205	17	2	1188	56	53.8
	I_SS_ST	135	13	39	9	0	53	3059	92.5
	Precision (%)	39.5	48.1	41.3	50.0	77.8	50.6	90.3	
	F1 Score	0.5	0.3	0.4	0.2	0.2	0.5	0.9	

* W_RI_SP: Supplement work, W_RI_PL: Rebar placing, W_RI_CT: Rebar connecting, W_RI_PA: Rebar preparation, W_TR_MT: Transferring materials and tools, W_TR_SP: Transportation, I_SS_ST: Standing/Sitting

Table 5. Confusion matrix of rebar work activity classification

4.2 Activity time estimation

The activity time estimation was performed to further examine the applicability of the action recognition approach for more detailed activity analysis in the construction field. For the performance measurement, the duration of each activity was first calculated on the basis of the recorded video data. The average duration of formwork was 2.8 h, and the average length of a rebar work was 1.9 h. This study cumulated the prediction results to measure the time spent on each activity category. With the estimated duration of each activity and the ground truth, the performance of activity time estimation was calculated. As shown in Table 6, the average estimation accuracies of Level 1 activities are 99.5% and 99.4% for formwork and rebar work, respectively. The trained models can determine the working time of formwork and rebar work with accuracies of 96.6% and 92.0%, respectively. The estimation accuracies of Level 3 activities are 65.2% and 74.4%. Such results imply the feasibility of monitoring the progress of each activity by utilizing wearable data from the construction environment. In particular, the proposed time estimation method contributes to the precise distinguishing between effective and ineffective work, and such facts offer an opportunity of implementing countermeasures to the activity in question.

Work division	Sample #		Time (hour)		Accuracy (%)
			Work	Idling	
Formwork	1	Ground truth	1.9	0.9	99.2
		Estimation	1.9	0.9	
	2	Ground truth	1.9	0.9	99.8
		Estimation	1.9	0.9	
	3	Ground truth	1.9	0.9	99.6
		Estimation	1.9	0.9	
	4	Ground truth	1.9	0.9	99.9
		Estimation	1.9	0.9	
	5	Ground truth	1.7	1.1	98.7
		Estimation	1.7	1.1	
				Average	99.5
Rebar work	1	Ground truth	1.1	0.9	99.3
		Estimation	1.1	0.9	
	2	Ground truth	1.1	0.9	99.7
		Estimation	1.1	0.9	
	3	Ground truth	1.1	0.8	99.3
		Estimation	1.1	0.9	
	4	Ground truth	1.1	0.8	99.0
		Estimation	1.1	0.8	
	5	Ground truth	1.1	0.8	99.8
		Estimation	1.1	0.8	
				Average	99.4

Table 6. Spending time estimation of Level 1 activity

Work division	Sample #		Time (hour) W_MI*	Time (hour) W_TR	Time (hour) I_SS	Accuracy (%)
Formwork	1	Ground truth	1.5	0.4	0.9	97.2
		Estimation	1.5	0.4	0.9	
	2	Ground truth	1.5	0.4	0.9	97.4
		Estimation	1.5	0.4	0.9	
	3	Ground truth	1.5	0.4	0.9	93.1
		Estimation	1.5	0.3	0.9	
	4	Ground truth	1.4	0.4	0.9	99.0
		Estimation	1.4	0.4	0.9	
	5	Ground truth	1.3	0.4	1.1	96.5
		Estimation	1.4	0.3	1.1	
					Average	96.6
Rebar work	1	Ground truth	0.7	0.4	0.9	96.8
		Estimation	0.7	0.3	0.9	
	2	Ground truth	0.7	0.4	0.9	95.2
		Estimation	0.7	0.3	0.9	
	3	Ground truth	0.7	0.4	0.8	96.2
		Estimation	0.8	0.3	0.8	
	4	Ground truth	0.7	0.4	0.8	89.2
		Estimation	0.8	0.3	0.8	
	5	Ground truth	0.7	0.4	0.8	82.6
		Estimation	0.9	0.3	0.8	
					Average	92.0

* W_MI: Material (formwork and rebar) installation, W_TR: Traveling, I_SS: Stand/sit

Table 7. Spending time estimation of Level 2 activity

Work division	Sample #		Time (hour) W_FI_SP*	Time (hour) W_FI_PL	Time (hour) W_FI_CT	Time (hour) W_FI_PA	Time (hour) W_TR_MT	Time (hour) W_TR_SP	Time (hour) I_SS_ST	Accuracy (%)
Form work	1	Ground truth	0.6	0.0	0.4	0.5	0.2	0.2	0.9	59.3
		Estimation	1.0	0.0	0.3	0.2	0.1	0.3	0.9	

	2	Ground truth	0.6	0.0	0.4	0.5	0.2	0.2	0.9	67.0
		Estimation	0.9	0.0	0.4	0.2	0.1	0.3	0.9	
	3	Ground truth	0.6	0.0	0.4	0.5	0.2	0.2	0.9	61.0
		Estimation	0.9	0.0	0.3	0.2	0.1	0.3	0.9	
	4	Ground truth	0.6	0.1	0.4	0.4	0.2	0.2	0.9	67.8
		Estimation	0.7	0.0	0.3	0.3	0.1	0.4	1.0	
	5	Ground truth	0.0	0.5	0.0	0.2	0.5	0.1	0.3	70.4
		Estimation	0.0	0.7	0.1	0.4	0.2	0.1	0.3	
									Average	65.2
Rebar work	Sample #		W_RI_SP**	W_RI_PL	W_RI_CT	W_RI_PA	W_TR_MT	W_TR_SP	I_SS_ST	Accuracy (%)
	1	Ground truth	0.2	0.1	0.2	0.2	0.0	0.4	0.9	84.6
		Estimation	0.2	0.0	0.2	0.2	0.0	0.4	0.9	
	2	Ground truth	0.2	0.1	0.2	0.2	0.0	0.4	0.9	91.4
		Estimation	0.2	0.1	0.2	0.2	0.0	0.3	0.9	
	3	Ground truth	0.2	0.2	0.2	0.2	0.0	0.4	0.8	66.2
		Estimation	0.0	0.4	0.1	0.2	0.0	0.0	0.4	
	4	Ground truth	0.2	0.2	0.2	0.2	0.0	0.4	0.8	68.2
		Estimation	0.0	0.3	0.1	0.3	0.0	0.0	0.4	
	5	Ground truth	0.2	0.2	0.2	0.2	0.0	0.4	0.8	61.8
		Estimation	0.0	0.4	0.2	0.3	0.0	0.0	0.3	
									Average	74.4

* W_FI_SP: Supplement work, W_FI_PL: Form placing, W_FI_CT: Form connecting, W_FI_PA: Form preparation, W_TR_MT: Transferring materials and tools, W_TR_SP: Transportation, I_SS_ST: Standing/Sitting

** W_RI_SP: Supplement work, W_RI_PL: Rebar placing, W_RI_CT: Rebar connecting, W_RI_PA: Rebar preparation, W_TR_MT: Transferring materials and tools, W_TR_SP: Transportation, I_SS_ST: Standing/Sitting

Table 8**.** Spending time estimation of Level 3 activity

5. DISCUSSION

5.1 Feasibility of acceleration-based activity recognition in the construction field

Previous research showed the potential of acceleration-based activity recognition to recognize diverse construction activities. However, the applicability of field activity detection has not been validated in terms of 1) the reliability of activity recognition in field conditions and 2) the defining of construction activities. The activity recognition algorithms in previous studies have been tested with discrete or independent data that ignore the noise and sequence characteristics of continuous acceleration signals collected from construction job sites. Construction activities in previous research are categorized on the basis of single standards, such as the nature of movement or contribution of tasks. Therefore, the derived classification results have limitations on providing information for measuring the efficiency of construction workers or for finding low productivity areas in the construction field concerned. We propose a new taxonomy to address these issues with consideration of movement and work context and subsequently validate it by using extensive field data.

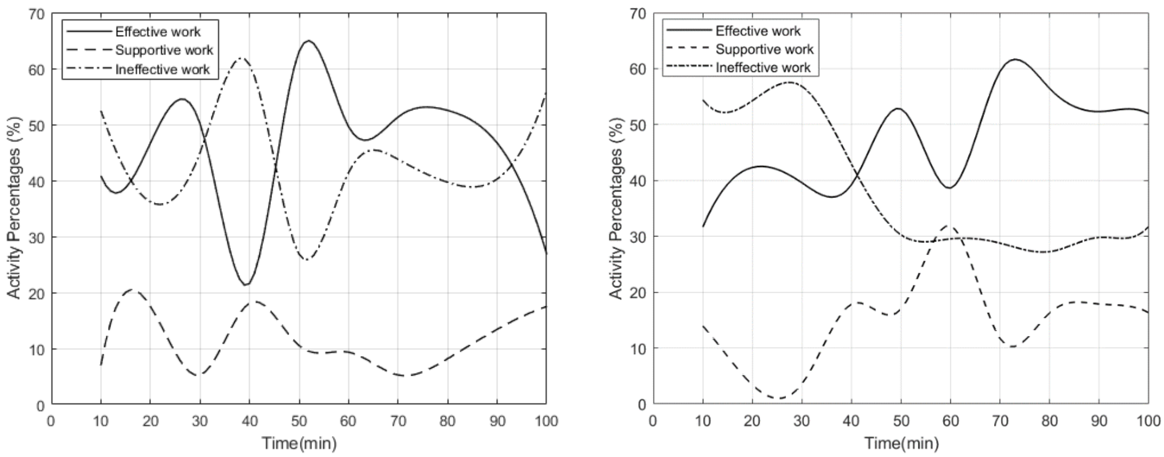
The understanding of the exclusive characteristics of different human activities is challenging due to the complex nature of human activities, which can induce classification confusion. Therefore, defining activities with a clear and comprehensive understanding of their nature is necessary for developing useful activity taxonomy (Bulling et al., 2014). Previous attempts in activity definition have primarily oriented toward a single principle (e.g., nature of the movement or contribution of work), and classifications of construction activities based on such principle have been validated in many previous studies. (Akhavian & Behzadan, 2016; Joshua & Varghese, 2014; Ryu et al., 2019; Weiss et al., 2016). Although movement-based activity taxonomy has a high classification accuracy, it still has several limitations when dealing with practical problems. First, depending on the context, similar movements can be delivered from different activities. In this case, the classification algorithms will perform poorly, especially when the activities being classified have largely similar characteristics of movements. Second, a movement-based activity taxonomy (e.g., lifting, sitting, and walking) cannot deliver sufficient information to solve practical problems, such as the identifying of low productivity operations in the field.

To overcome these issues, several studies have introduced a context-based activity taxonomy that categorizes construction activities based on their contributions to the project (Forde & Buchholz, 2004; Hallowell & Gambatese, 2009; Joshua & Varghese, 2014) for evaluating productivity in a rough manner. However, most construction activities consist of diverse tasks (e.g., effective work of an ironworker includes fetching, adjusting, and tying rebar). Previous context-based activity taxonomies are insufficient to reveal the root causes of low productivity due to the lack of detailed information about ongoing activities. In an attempt to solve such problems, this study considered movement- and context-based taxonomy when defining an activity. Theoretically, acceleration signals collected from the dominant hand are regarded as an integrated response of whole-body movements and hand movements (Ryu et al., 2019). Therefore, a different combination of body and hand movement is an intuitive standard for identifying activities that share a distinct acceleration response. However, activities that have similar movements (e.g., lifting material from the ground, squatting, and standing up) are difficult to be accurately identified in accordance with the movement-based system. The context standard was introduced to enrich the textural information of activity and to extend the classification categories. In this regard, the capability of activity recognition for identifying low productivity issues is enhanced.

The construction activities are formatted as a three-level taxonomy with a hierarchical structure (**Error! Reference source not found.**), which allows classifying specific activities by zooming in or out the action level and identifying the optimal classification level by trading off between performance (i.e., accuracy) and outcomes (i.e., information extracted from the results) (Blanke & Schiele, 2010; Krishnan et al., 2013). On the basis of the result shown in *LOOCV: leave-one-out cross-validation, LOSOCV: leave-one-subject-out cross-validation

Table 3, the neural network algorithms can train more powerful classifiers. The classification accuracy at Level 1 (i.e., "idling" and "work") shows over 90% accuracy because "idling" involves mostly no movement on hands, which can be easily distinguished from "work," which involves significant arm and body movements, and has substantial changes in acceleration signals. At Level 2, we further divide "work" into two subcategories, 1) traveling, and 2) installing tasks considering that they have different work contexts (e.g., traveling is a supportive activity, and installing material is a value-added task) and body movements (e.g., "traveling" involves abundant body movements and few cyclic movements from hands, and "installing" involves abundant hand movements and few body movements). The classification accuracy at Level 2 is over 80%, and the algorithm can differentiate between horizontal whole-body movements (e.g., "traveling") and hand-dominant activities (e.g., "material installation"). In accordance with the confusion matrix at this level (**Error! Reference source not found.** and Table 5), the most significant errors result from the confusion between "traveling" and "material installation" because "material installation" frequently involves a temporal allocation (e.g., moving 1–2 m to pick up materials), which has a large similarity with "traveling" (e.g., moving to another work zone). The accuracy of Level 3 activity classification is lower than that of Level 1 and Level 2, showing 50%–60% accuracy because more detailed work contexts were contained. The classification results show that the significant confusion within the offspring categories of Level 2 activity, "material installation" occurs. This finding may indicate that the proposed algorithm cannot recognize the considerable interclass variability in Level 3 activities due to the similar nature of body and hand movements for these activities. As the types of activities at Level 3 were more frequently changed during the operation, the acceleration signals may include the noise data from transition patterns between activities. However, in terms of measuring spending time for Level 3 activities, the accuracy increased up to approximately 75% (Table 8), showing the potential for being used to understand the productivity issues during construction operations.

The classification results at Level 2 are accurate, allowing to identify productivity issues by providing meaningful information, such as the time expenditure of workers. For instance, two continuous patterns of acceleration data were sampled from two form workers who were at the same site and worked simultaneously. The activity percentage values were calculated on the basis of the spending time estimation method in Section 4.2, and the percentages were plotted in a time series domain, as shown in Figure 4. In particular, the activity percentages of the two form workers were calculated on the basis of 10 min. The productivity of form worker No. 2 was higher in the selected 100 min because his effective work rate remained at a relatively high level without any huge drop by comparing Figure 4 (a) and Figure 4 (b). The cause of the low productivity issues can be exposed. Taking form worker No.1 in Figure 4 (a) as an instance, the effective work rate dropped during the time from 30 min to 40 min, and the ineffective rate increased extremely at the same period. This finding indicates that the increasing proportion of ineffective work is the cause of the low productivity issue in the selected period. The root cause of low productivity issue of form worker No. 2 from 50 min to 60 min can be recognized as the increasing percentage of supportive work by using the same method. Considering the ineffective work is not dominant and the effective work rate remains at 40%, the worker was on short travel between two installation trades.



a. Form worker No.1 b. Form worker No.2

Figure 4. Time series line plot illustrating activity percentage in every 10 min

5.2 Remaining challenges to enhance the classification performance

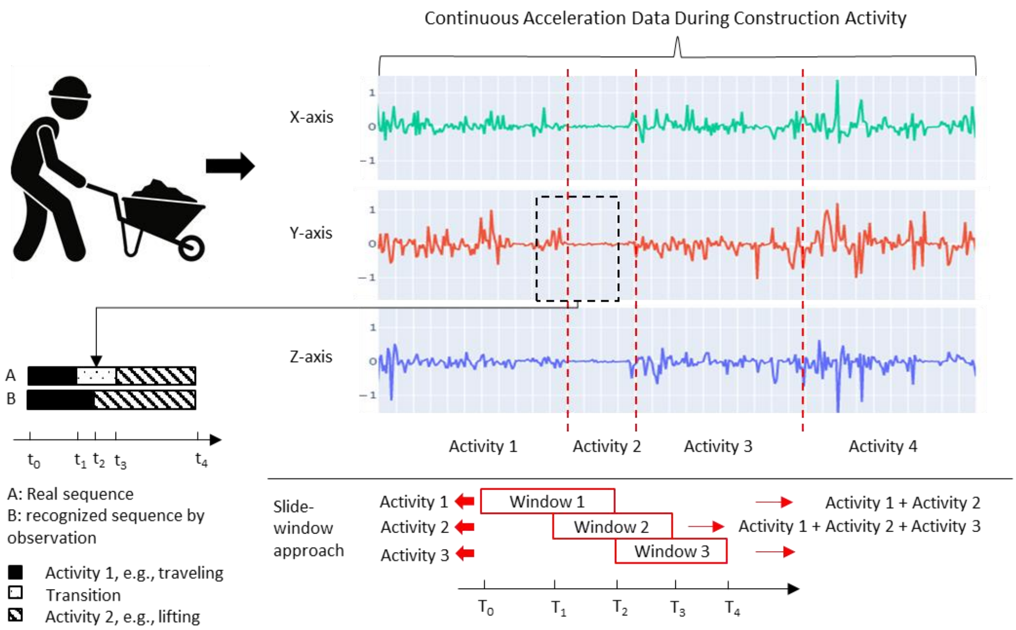
Although the classification result at Level 2 activity can distinguish low productivity issues, it is insufficient to expose the root cause. In this regard, the Level 3 activity is necessary for finding the cause of the delay. However, the current performance of Level 3 activity classification does not satisfy the demand in the construction field because recognizing a sequence of activities from an uncontrolled environment (i.e., construction field) is challenging. In addition to the human variability, several remaining challenges exist, and they are 1) difficulty in handling the transition effect between activities, 2) inaccurate segmentation of time-series movement data, and 3) information loss during the machine learning process. The first challenge deals with the transition moment in continuous human activities (Minnen et al., 2006). In Figure 5 (a), sequence A refers to a real activity stream, which indicates that a transition pattern (i.e., pattern from t_1 to t_3) shall exist between two explicit activities (e.g., traveling and lifting) considering that human activity changes gradually. However, such transition has been disregarded in this study because 1) the duration of the transition activities is relatively short compared with other activities that are explicitly defined in the taxonomy in **Error! Reference source not found.** (Lara & Labrador, 2012); 2) the temporal boundaries of transitions are difficult to determine by human observation because the transition activity and its neighboring activities share similar movements as recorded in videos. A sample of a labeled sequence (i.e., sequence B) can be found in Figure 5 (a), which shows that activity 1 lasts from t_1 to t_2, and the following activity (i.e., activity 2) lasts from t_2 to t_4. A comparison between the real sequence (i.e., sequence A) and the recognized sequence (i.e., sequence B) shows that the two transition patterns (i.e., activity from t_1 to t_2 and activity from t_2 to t_3) are mistakenly recognized as activity 1 and activity 2, respectively. Considering the transition effect is widespread in the continuous activity patterns, the massive mislabeling of the activity category induces significant errors when training the dataset and the ground truth. Thus, the misclassification rate is considerably high, and the classification system is unacceptable for field productivity evaluation.

One of the alternatives is to regard "transition" as an extra activity to address this issue (Zhang et al., 2010). In previous research, Rednic et al. (2013) used a transition filter to improve the classification accuracy and stability. On the basis of the assumption that more recent posture has a higher correlation with the actual posture, the weighted-voting methods can filter out unreasonable postural vibrates located in the high-frequency domain. The filtering process is validated as useful for increasing the certainty of the transition boundaries. However, the improvement in accuracy is limited. Rather than setting clear-cut boundaries, some researchers (Abonyi et al., 2005) introduced the idea of fuzzy clustering (i.e., data points can belong to more than one cluster) that helps to determine the fuzzy boundaries of time-series data (e.g., the continuous acceleration data). Fuzzy segmentation (i.e., setting fuzzy boundaries for the activity pattern) is then adopted in the activity recognition to overcome the transition effect (Zhang et al., 2014). The researchers defined the fuzzy boundaries with Gaussian membership and a time variable, and translated the segmenting issue into an optimizing problem. The bias caused by the transition effect can be restricted by solving the optimization problem. In future research, we will apply the proposed approaches and test the feasibility of reducing transition effect in continuous field data.

In the classification of human activities, continuous sensor data are segmented into sequences for the feature extraction process. However, the setting of data windows of activities without introducing any classification errors is still a challenging task (Bao & Intille, 2004). A sliding window technique for data segmentation was primarily applied, investigated, and validated in previous research (Bulling et al., 2014). Similar to previous studies, we used a sliding window technique with fixed window size. As shown in Figure 5 (b), the acceleration data collected during construction activities (i.e., activities from $T0$ to $T5$) are segmented into three windows (i.e., independent activity pattern). Specifically, window 1 lasts from $T0$ to $T2$, window 2 lasts from $T1$ to $T3$, and window 3 lasts from $T2$ to $T4$. The durations of the windows (i.e., $T0$ to $T2$, $T1$ to $T3$, $T2$ to $T4$) are constant, and the overlapping between two consequent windows is set to 50%. However, the use of the fixed-size sliding window can induce considerable misclassification due to two causes of errors (Gu et al., 2009). The duration of the different activity categories is diverse due to the different natures of human movement. The spending time of the same type of activity can vibrate during the work. In these regards, a fixed-size window cannot purely and fully include a single type of activity, leading to extreme errors when preparing training data and testing data. Therefore, enhancing classification performance by window size optimization is difficult (Huynh & Schiele, 2005). Previous research demonstrated that the algorithms can perform better if the features and length of windows are considered as separate activity categories.

The multiclass problem is another observed issue related to the sliding window approach (Yao et al., 2018). As shown in Figure 5 (b), multiple categories of activity can be found in the same window (e.g., window 1 consists of activity 1 and activity 2; and window 2 includes activity 1, activity 2, and activity 3). However, following the majority voting principle, a single activity label should be assigned to each data window, which can bring about a significant loss of activity information and result in considerable misclassification. The ground truth of the activity may be disturbed because the true label is different from the label selected for the window. For instance, the data of activity 2 were labeled as activity 1 in the segmenting process in window 1 in Figure 5 (b). Therefore, the data of activity 1 were accidentally polluted by the activity 2 data, resulting in the misleading of the algorithms. Laguna et al. (2011) proposed a dynamic segmenting approach to address these limitations. In this approach, the starting and end times of the window and the window length are concluded as core parameters to determine the windows dynamically. Therefore, changes in activities are integrated into formulas as a significant variable for indicating the beginning and ending points of window. The results show that the dynamic window approach effectively reduces classification confusion. Yao et al. (2018) proposed a dense labeling scheme that labels each individual data point rather than labeling the data segment . Each data point can be regarded as a “window” that includes only one datum. The data point is assigned a unique label that will not be adjusted by any vote-based filtering. Therefore, the problems of information loss and label confusion caused by the sliding window method can be overcome.

The last issue of the current model is that the sequential characteristic of continuous construction activity is still ignored. In a sequential activity for construction (i.e., activities that occur in a certain order), an activity can affect the action that occurs after it. For instance, if the prior activity is “sitting,” then the subsequent behavior cannot be “walking” or “running” because the activity “standing up” cannot be avoided between “sitting” and “walking.” A transition from “walking” to “standing up” is also impossible based on the context. In this study, such unreasonable sequences are frequently observed from the classification model, resulting in significant errors. To overcome this issue, Panahandeh et al. (2013) introduced the continuous hidden Markov model (HMM) to analyze gait phase and joint activity via IMU measurements. Five individual activities, namely, going upstairs, going downstairs, running, standing, and walking, are discussed in the study. The HMM model integrates the activity influence through two objects: 1) discrete chain of activities, which reflects the order and relationship between activities, 2) probability density functions of the future variables, which add the influence on the classification algorithms. The final classification accuracy of this probabilistic activity ranges from 90% to 99%, indicating a great potential for solving the classifying continuous human activity classification problem. Future research can test the continuous HMM with the field-collected data to reduce any unreasonable sequences existing in the classification results.



a. transition effect b. sliding window approach

Figure 5. Illustration of errors induced by transition effect and segment method

6. CONCLUSION

This study investigated the validity of action recognition algorithms with a newly proposed comprehensive and universally applicable work taxonomy that was designed considering movement and construction contexts. In particular, the performance of the proposed approach was studied by using acceleration data collected in a construction site during unstructured ongoing concrete work. Acceleration signals during formwork and rebar work were labeled with activities defined at three hierarchical levels based on the proposed activity taxonomy and used for testing traditional machine learning- and deep learning-based action recognition algorithms. The testing results show that the classification performance for Level 1 activities for formwork and rebar work is relatively reliable with higher than 95% accuracy, and the prediction accuracies range from 74.6% to 83.8% for Level 2 activity classification. The classification accuracies for Level 3 activities vary from 45.3% to 61.3%.

The classification results for activities at Level 1 and Level 2 demonstrate that 1) the proposed taxonomy can convey comprehensive activity information (i.e., activity context information and movement information) and reduce confusion among the categories in the same level, and 2) the performance of acceleration-based activity recognition algorithm is acceptable when dealing with noisy data (i.e., long-term and continuous data collected directly from the construction site). However, the rather low accuracy for activities at Level 3 may indicate the limitation of the use of acceleration signals for micro-level activity analysis. This study evaluated the spending time estimation of long-term continuous signals collected from the field, which reported high accuracies in measuring the activity duration of Level 1 and Level 2 activity. On the basis of the duration data, the time spent ratio of each activity can be evaluated through the timeline. Therefore, evaluating the work efficiency is possible by comparing it with the benchmark. The root cause of the low-efficiency problem can be exposed by analyzing the time spent ratio, which will help optimize the construction trade for improving productivity.

The measuring of workers' activities can provide quantitative evidence for identifying productivity issues from the perspective of individual workers. Acceleration-based action recognition is regarded as a useful means for automated activity analysis, but it suffers from a nonstandardized definition of activities and a lack of validity in a practical setting. This study may provide a solid foundation for automated activity analysis by proposing a practical approach on how to define and analyze construction activities using acceleration data. The comprehensive validation of action recognition algorithms using unstructured field data in this study can convince practitioners about the reliability of acceleration-based action recognition for Level 1 and Level 2 activities in practice.

Acknowledgments

This research study was supported by the General Research Fund (No. 15223018) from the Research Grants Council, Hong Kong, and a grant (21CTAP-C151784-03) from Technology Advancement Research Program funded by the Ministry of Land, Infrastructure, and Transport of Korean government.

7. **REFERENCES**

Abonyi, J., Feil, B., Nemeth, S., & Arva, P. (2005). Modified Gath–Geva clustering for fuzzy segmentation of multivariate time-series. *Fuzzy Sets and Systems*, *149*(1), 39-56.

Akhavian, R., & Behzadan, A. H. (2016). Smartphone-based construction workers' activity recognition and classification. *Automation in Construction*, *71*, 198-209.

Ballabio, D., Todeschini, R., & Consonni, V. (2019). Recent advances in high-level fusion methods to classify multiple analytical chemical data. *Data Handling in Science and Technology*, *31*, 129-155.

Bangaru, S. S., Wang, C., Busam, S. A., & Aghazadeh, F. (2021a). ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. *Automation in Construction*, *126*, 103653.

Bangaru, S. S., Wang, C., Busam, S. A., & Aghazadeh, F. J. A. i. C. (2021b). ANN-based automated scaffold builder activity recognition through wearable EMG and IMU sensors. *126*, 103653.

Banos, O., Galvez, J.-M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors*, *14*(4), 6474-6499.

Bao, L., & Intille, S. S. (2004). Activity recognition from user-annotated acceleration data. International conference on pervasive computing,

Berrar, D. (2019). Cross-Validation. In.

Blanke, U., & Schiele, B. (2010). Remember and transfer what you have learned-recognizing composite activities based on activity spotting. International Symposium on Wearable Computers (ISWC) 2010,

Buchholz, B., Paquet, V., Wellman, H., & Forde, M. (2003). Quantification of ergonomic hazards for ironworkers performing concrete reinforcement tasks during heavy highway construction. *AIHA journal*, *64*(2), 243-250.

Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, *46*(3), 1-33.

Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., Yu, Z. J. I. T. o. S., Man,, & Cybernetics, P. C. (2012). Sensor-based activity recognition. *42*(6), 790-808.

Cheng, T., Teizer, J., Migliaccio, G. C., & Gatti, U. C. (2013). Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data. *Automation in Construction*, *29*, 24-39.

Construction Industry Institute (CII) (2010). *Guide to activity analysis*. Univ. of Texas, Austin, TX.

Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. Proceedings of the 23rd international conference on Machine learning,

Dietterich, T. G. (2000). Ensemble methods in machine learning. International workshop on multiple classifier systems, Everett, J. G., & Slocum, A. H. (1994). Automation and robotics opportunities: construction versus manufacturing. *Journal of Construction Engineering and Management*, *120*(2), 443-452.

Figo, D., Diniz, P. C., Ferreira, D. R., & Cardoso, J. M. (2010). Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing*, *14*(7), 645-662.

Forde, M. S., & Buchholz, B. (2004). Task content and physical ergonomic risk factors in construction ironwork. *International Journal of Industrial Ergonomics*, *34*(4), 319-333.

Gil-Martín, M., San-Segundo, R., Fernández-Martínez, F., & Ferreiros-López, J. (2020). Improving physical activity recognition using a new deep learning architecture and post-processing techniques. *Engineering Applications of Artificial Intelligence*, *92*, 103679.

Golestani, N., & Moghaddam, M. (2020). Human activity recognition using magnetic induction-based motion signals and deep recurrent neural networks. In: Google Patents.

Gouett, M. C., Haas, C. T., Goodrum, P. M., & Caldas, C. H. (2011). Activity analysis for direct-work rate improvement in construction. *Journal of Construction Engineering and Management*, *137*(12), 1117-1124.

Gu, T., Wu, Z., Tao, X., Pung, H. K., & Lu, J. (2009). epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. 2009 IEEE International Conference on Pervasive Computing and Communications,

Gumaei, A., Hassan, M. M., Alelaiwi, A., & Alsalman, H. J. I. A. (2019). A hybrid deep learning model for human activity recognition using multimodal body sensing data. *7*, 99152-99160.

Hallowell, M. R., & Gambatese, J. A. (2009). Activity-based safety risk quantification for concrete formwork construction. *Journal of Construction Engineering and Management*, *135*(10), 990-998.

Hassan, M. M., Uddin, M. Z., Mohamed, A., & Almogren, A. J. F. G. C. S. (2018). A robust human activity recognition system using smartphone sensors and deep learning. *81*, 307-313.

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks*, *13*(2), 415-425.

Huynh, T., & Schiele, B. (2005). Analyzing features for activity recognition. Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies,

Hwang, S., & Lee, S. (2017). Wristband-type wearable health devices to measure construction workers' physical demands. *Automation in Construction*, *83*, 330-340.

Ignatov, A. J. A. S. C. (2018). Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. *62*, 915-922.

Jarkas, A. M. (2010). Critical investigation into the applicability of the learning curve theory to rebar fixing labor productivity. *Journal of Construction Engineering and Management*, *136*(12), 1279-1288.

Joshua, L., & Varghese, K. (2011). Accelerometer-based activity recognition in construction. *Journal of computing in civil engineering*, *25*(5), 370-379.

Joshua, L., & Varghese, K. (2014). Automated recognition of construction labour activity using accelerometers in field situations. *International Journal of Productivity and Performance Management*, *63*(7), 841-862.

Khosrowpour, A., Niebles, J. C., & Golparvar-Fard, M. J. A. i. C. (2014). Vision-based workface assessment using depth images for activity analysis of interior construction operations. *48*, 74-87.

Kim, K., & Cho, Y. K. (2020). Effective inertial sensor quantity and locations on a body for deep learning-based worker's motion recognition. *Automation in Construction*, *113*, 103126.

Krishnan, N., Cook, D. J., & Wemlinger, Z. (2013). Learning a taxonomy of predefined and discovered activity patterns. *Journal of ambient intelligence and smart environments*, *5*(6), 621-637.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Proceedings of the 25th International Conference on Neural Information Processing Systems-Volume 1 (NIPS'12).

Kwapisz, J. R., Weiss, G. M., & Moore, S. A. (2011). Activity recognition using cell phone accelerometers. *ACM SigKDD Explorations Newsletter*, *12*(2), 74-82.

Laguna, J. O., Olaya, A. G., & Borrajo, D. (2011). A dynamic sliding window approach for activity recognition. International Conference on User Modeling, Adaptation, and Personalization,

Lara, O. D., & Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, *15*(3), 1192-1209.

Lee, S.-M., Yoon, S. M., & Cho, H. (2017). Human activity recognition from accelerometer data using Convolutional Neural Network. 2017 ieee international conference on big data and smart computing (bigcomp),

Mantyjarvi, J., Himberg, J., & Seppanen, T. (2001). Recognizing human motion with multiple acceleration sensors. 2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236),

Minnen, D., Westeyn, T., Starner, T., Ward, J., & Lukowicz, P. (2006). Performance metrics and evaluation issues for continuous activity recognition. *Performance metrics for intelligent systems. NIST, Gaithersburg*, 141-148.

Ng, S. T., & Tang, Z. (2010). Labour-intensive construction sub-contractors: Their critical success factors. *International Journal of Project Management*, *28*(7), 732-740.

Panahandeh, G., Mohammadiha, N., Leijon, A., & Händel, P. (2013). Continuous hidden Markov model for pedestrian activity classification and gait analysis. *IEEE Transactions on Instrumentation and Measurement*, *62*(5), 1073-1083.

Preece, S. J., Goulermas, J. Y., Kenney, L. P., Howard, D., Meijer, K., & Crompton, R. (2009). Activity identification using body-mounted sensors—a review of classification techniques. *Physiological measurement*, *30*(4), R1.

Rednic, R., Gaura, E., Kemp, J., & Brusey, J. (2013). Fielded autonomous posture classification systems: design and realistic evaluation. 2013 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing,

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-Validation. *Encyclopedia of database systems, 5*, 532-538.

Ryu, J., Seo, J., Jebelli, H., & Lee, S. (2019). Automated action recognition using an accelerometer-embedded wristband-type activity tracker. *Journal of construction engineering and management*, *145*(1), 04018114.

Sanhudo, L., Calvetti, D., Martins, J. P., Ramos, N. M., Mêda, P., Gonçalves, M. C., & Sousa, H. J. J. o. B. E. (2021). Activity classification using accelerometers and machine learning for complex construction worker activities. *35*, 102001.

Sutton, O. (2012). Introduction to k nearest neighbour classification and condensed nearest neighbour data reduction. *University lectures, University of Leicester*, 1-10.

Thomas, H. R., & Daily, J. (1983). Crew performance measurement via activity sampling. *Journal of Construction Engineering and Management*, *109*(3), 309-320.

Wang, J., Chen, Y., Hao, S., Peng, X., & Hu, L. J. P. R. L. (2019). Deep learning for sensor-based activity recognition: A survey. *119*, 3-11.

Weiss, G. M., Timko, J. L., Gallagher, C. M., Yoneda, K., & Schreiber, A. J. (2016). Smartwatch-based activity recognition: A machine learning approach. 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI),

Xu, C., Chai, D., He, J., Zhang, X., & Duan, S. J. I. A. (2019). InnoHAR: A deep neural network for complex human activity recognition. *7*, 9893-9902.

Yang, J.-Y., Wang, J.-S., & Chen, Y.-P. (2008). Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern recognition letters*, *29*(16), 2213-2220.

Yang, K., Ahn, C. R., & Kim, H. J. A. E. I. (2020). Deep learning-based classification of work-related physical load levels in construction. *45*, 101104.

Yao, R., Lin, G., Shi, Q., & Ranasinghe, D. C. (2018). Efficient dense labelling of human activity sequences from wearables using fully convolutional networks. *Pattern Recognition*, *78*, 252-266.

Zhang, H., Zhou, W., & Parker, L. E. (2014). Fuzzy segmentation and recognition of continuous human activities. 2014 IEEE International Conference on Robotics and Automation (ICRA),

Zhang, S., McCullagh, P., Nugent, C., & Zheng, H. (2010). Activity monitoring using a smart phone's accelerometer with hierarchical classification. 2010 sixth international conference on intelligent environments,