Automation in Construction 71 (2016) 198–209

Contents lists available at ScienceDirect

Automation in Construction

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

Smartphone-based construction workers' activity recognition and classification

Reza Akhavian [a,*], Amir H. Behzadan [b]

[a] School of Engineering, California State University, East Bay, 25800 Carlos Bee Blvd, Hayward, CA 94542, USA
[b] Department of Technology and Construction Management, Missouri State University, 901 S. National Avenue, Springfield, MO 65897, USA

ARTICLE INFO

Article history:
Received 3 October 2015
Received in revised form 24 July 2016
Accepted 13 August 2016
Available online 21 August 2016

Keywords:
Construction
Workers
Activity recognition
Productivity analysis
Smartphone sensors
Accelerometer
Gyroscope
Machine learning
Neural networks

ABSTRACT

Understanding the state, behavior, and surrounding context of construction workers is essential to effective project management and control. Exploiting the integrated sensors of ubiquitous mobile phones offers an unprecedented opportunity for an automated approach to workers' activity recognition. In addition, machine learning (ML) methodologies provide the complementary computational part of the process. In this paper, smartphones are used in an unobtrusive way to capture body movements by collecting data using embedded accelerometer and gyroscope sensors. Construction activities of various types have been simulated and collected data are used to train five different types of ML algorithms. Activity recognition accuracy analysis has been performed for all the different categories of activities and ML classifiers in user-dependent and -independent ways. Results indicate that neural networks outperform other classifiers by offering an accuracy ranging from 87% to 97% for user-dependent and 62% to 96% for user-independent categories.

Published by Elsevier B.V.

1. Introduction

Effective and timely analysis and tracking of workforce activities are essential to overall productivity measurement, progress evaluation, labor training programs, and safety and health management [1–3]. The construction industry, as a major contributor to the U.S. economy, has traditionally suffered from low productivity and high inefficiency stemmed from misallocating resources resulting in under-utilizing or over-utilizing them in the project. Arguably, the first step in alleviating this problem is to accurately monitor and evaluate the time spent on interconnected construction tasks involving labor force, and compare the results with project benchmarks in order to improve the amount of time and resources spent on work packages involved in typical construction activities [4]. In addition to its benefits to productivity monitoring, the outcome of this analysis can be used for stochastic process simulation input modeling, work sampling, and integrated detailed assessment and continuous workflow improvement. For instance, the authors have designed and implemented a data-driven construction simulation framework by tracking construction entities [5,6]. Joshua and Varghese [7] adopted a similar approach to facilitate the manual process of work sampling in construction projects.

Process monitoring and control provides a solid basis for tracking and measurements required for activity analysis. Recent advancements in automated data collection to track resources and measure work progress have shown promising prospects for streamlining crew activity analysis compared to the conventional (manual) approaches such as direct observations and survey-based methods. This is mostly because manual methods involving human observers are tedious, time consuming, and error-prone. Furthermore, large amounts of data should be collected in order to maintain the statistical significance of observations.

However, automated technologies for data acquisition are still being assessed in terms of their reliability and feasibility in construction domain applications. In one hand, vision-based techniques have been proposed and investigated by a number of researchers for automated activity analysis [8]. On the other hand, wireless sensor-based methodologies have been examined to collect spatio-temporal activity data [9]. While vision-based methods are often prone to extant occlusions and illumination variability in construction jobsites, sensor-based techniques do not require a clear line-of-sight (LOS) and extensive computations and can provide relatively low cost solutions (compared to laser-scanning for instance). Despite their advantages, a longstanding challenge and impediment to the widespread use of sensor-based data collection schemes is that traditional sensor installation and maintenance in construction jobsites is not a trivial task (if not at all impossible) due to prohibitive ambient factors such as dust, adverse weather conditions, and harsh working environments.

* Corresponding author.
E-mail addresses: reza.akhavian@csueastbay.edu (R. Akhavian), abehzadan@missouristate.edu (A.H. Behzadan).

http://dx.doi.org/10.1016/j.autcon.2016.08.015
0926-5805/Published by Elsevier B.V.

To remedy this situation, a relatively newer data collection technique has been trending which uses ubiquitous sensors that are readily available to and carried by most individuals on a daily basis. Such technologies are, for instance, provided through built-in sensors in most mobile phones. Mobile devices are advantageous over other activity recognition data collection platforms since they unobtrusively provide a self-sufficient data collection, computing, and storage scheme. Recently, several research projects within the construction domain have taken advantage of ubiquity of smartphones to design and prototype useful applications for construction workers on the jobsite [10,11]. Such applications in essence deliver information to the site personnel, while there is a great potential to infer information using the built-in sensors. A typical smartphone has an almost inclusive subset of these context-aware sensors including accelerometer, gyroscope, GPS, magnetometer, barometer, proximity sensors, light sensors, Bluetooth, Near Field Communication (NFC), and cameras [9]. In addition to academic endeavors, a number of construction equipment and tool manufacturers have started to produce rugged drop-proof, and dust- and water-resistant smartphones specifically designed for construction jobsites [12].

This paper presents a thorough evaluation of the performance of an activity analysis framework for recognition and classification of various construction worker activities using smartphone built-in sensors. In this research, data are collected from a variety of construction activities performed by construction workers and are annotated for feature extraction to train machine learning classifiers. Data-driven methodologies in activity recognition fall into one of the two major categories of generative or discriminative approaches. While in generative approach probabilistic models such as Baysian network are used to build a description of input, the discriminative approach models the mapping from inputs to outputs or data to activities [13]. Using generative models such as hidden Markov models (HMM) and dynamic Baysian network (DBN) is not within the scope of this research since they are not capable of capturing transitive dependences of the observations due to their very strict independence assumptions.

2. Literature review

2.1. Automated recognition of construction worker activities

Previous research for activity recognition and classification of construction workers mainly falls into the vision-based category. Microsoft Kinect, for example, was employed by some researchers for vision-based activity recognition in indoor and controlled environments [14,15]. In another set of studies, 2D videos are used to collect visual data for action recognition in construction sites. For example Favela, Tentori, Castro, Gonzalez, Moran and Martínez-García [3] used a wireless video camera to extract human poses from video to recognize construction workers' actions. In a different study, 3D range image camera was used for tracking and surveillance of construction workers for safety and health monitoring [16]. Gonsalves and Teizer [16] indicated that if their proposed system is used in conjunction with artificial neural network (ANN), the results would be more robust for prevention of fatal accidents and related health issues. In their study on construction workers' unsafe actions, Han and Lee [17] developed a framework for 3D human skeleton extraction from video to detect unsafe predefined motion templates. All of these frameworks, although presented successful results in their target domain, require installation of multiple cameras (up to 8 in some cases), have short operational range (maximum of 4 m for Kinect), and require a direct LOS for implementation. Such shortcomings have served as a major motivation to investigate alternative solutions that can potentially alleviate these problems.

Recently, researchers in construction engineering and management (CEM) have investigated the applications of senor-based worker activity analysis. For example, a data fusion approach using ultra-wide band (UWB) and Physiological Status Monitors (PSMs) for productivity [4] and ergonomics [18] analysis was proposed. In these studies, UWB and PSM data were fused and the result was categorized using a spatio-temporal reasoning approach. However, the level of detail in recognizing the activities was limited to identification of traveling, working, and idling states of workers and could not provide further insight into identified activities. Prior to this study, the integration of UWB, payload, and orientation (angle) data with spatio-temporal taxonomy-based reasoning was adopted by the authors for construction equipment activity analysis to support process visualization, remote monitoring and planning, and knowledge-based simulation input modeling [19–21]. More recently, the authors have used smartphone-based data collection and activity recognition for data-driven simulation of construction workers' activity by extracting process knowledge and activity durations [22]. Joshua and Varghese [7] were among the first researchers who explored the application of accelerometer in construction for work sampling. However, the scope of their work was limited to only a single bricklayer in a controlled environment. Moreover, their proposed framework used accelerometer as the sole source of motion data. Also, the necessity of installing wired sensors on the worker's body may introduce a constraint on the worker's freedom of movement.

2.2. Activity recognition using cellphone sensors

Detection and classification of human activities using wearable inertial measurement units (IMUs) consisting of accelerometer and gyroscope gained traction among computer science researchers in mid-2000′s with applications in different fields such as healthcare and sports [23–25]. In all such studies, data pertaining to human physical movements are captured using IMUs and different postures and dynamic transitions are detected by training classifiers. However, recent studies are mostly geared toward leveraging the ubiquity, ease of use, and self-sufficiency of mobile phones for human activity recognition [26–29]. In one study, Reddy, Mun, Burke, Estrin, Hansen and Srivastava [30] used decision tree and dynamic hidden Markov model (DHMM) to classify activities such as standing, walking upstairs, biking, driving a car, and jumping using accelerometer and GPS data. In another research, Sun, Zhang, Li, Guo and Li [28] used support vector machines (SVMs) to build a human daily physical activity recognition system using mobile phone accelerometers. More recently, mobile phone gyroscope has been also employed in addition to accelerometer for activity recognition. For example, using accelerometer and gyroscope data and hierarchical SVM, Kim, Cho and Kim [31] classified daily activities to sitting, walking up- and downstairs, biking, and having no motion. Moreover, Martín, Bernardos, Iglesias and Casar [32] used decision table, decision tree, and naïve Bayes to classify data from various smartphone sensors such as accelerometer and gyroscope to classify daily activities into standing, sitting, jogging, and walking upstairs.

Despite its great potential for construction automation, and considering the existing interest in construction workers' activity recognition, the application of such emerging data collection platforms has not been fully investigated within the CEM domain. In the research presented in this paper, signature patterns observed in the signals received from wearable IMUs of ubiquitous smartphones are analyzed to recognize activities performed by different construction workers.

3. Research objectives and contributions to the body of knowledge

As stated in the previous Section, existing work on activity recognition within the CEM domain has primarily focused on vision-based systems while a very limited number of studies aimed at developing multimodal sensor-based data collection schemes. Hence, the presented study in this paper contributes to the body of knowledge by investigating construction worker activity recognition through (1) using the sensors embedded in mobile phones to (2) identify complex activities that consist of more than one task by (3) deploying combined features of accelerometer and gyroscope (i.e. IMU) data. In particular, this research provides new insight into the accuracy of recognizing construction workers' complex and continues activities through different learning algorithms where more than one task is performed by a worker, using mobile built-in IMUs.

4. Methodology

In this study, data are collected using mobile phone accelerometer and gyroscope sensors. Collected raw sensory data are segmented into windows containing certain number of data points. Next, key statistical features are calculated within each window. Furthermore, each segment is labeled based on the corresponding activity class performed at the time identified by the timestamp of the collected data. In order to train a predictive model, five classifiers of different types are used to recognize activities performed in the data collection experiments. Fig. 1 depicts the steps from data collection to activity recognition. All data processing including the statistical computation of features and training, testing, and validation of the classifiers were performed in Matlab using in-house codes.

4.1. Data acquisition using mobile phones

Wearable sensors are small size mobile sensors designed to be worn on the body. Most such wearable mobile sensors can be found in existing smartphones. Accelerometer, gyroscope, ambient temperature sensor, light sensor, barometer, proximity sensor, and GPS are some of the sensing technologies that are built-in on most of the commercially available smartphones. Accelerometer sensors measure the acceleration of the device. The reading can be in one, two, or all three axes of X, Y, and Z. The raw data is represented as a set of vectors and returned together with a timestamp of the reading. Gyroscope is a sensor that measures the rotation rate of the device by detecting the roll, pitch, and yaw motions of the smartphone about the X, Y, and Z axes. Similar to accelerometer, readings are presented as time-stamped vectors. When the mobile device is attached to a human body involved in different activities, these two sensors generate different (and unique) patterns in their transmitted signals.

4.2. Data preparation

When collecting data for a long period of time, it can be observed that sometimes the sensors temporarily lag or fail to properly collect and store data for fractions of a second to a few seconds and in return, compensate for the missing data points by collecting data at a rate higher than the assigned frequency. In such cases, a preprocessing technique to fill in for missing data points and remove redundant ones can help insuring a continues and orderly dataset. Also, since the raw data are often collected with a high sampling rate, segmentation of the data helps in data compression and prepares data for feature extraction [33]. If segmentation is performed considering an overlap between adjacent windows, it reduces the error caused by the transition state noise [34]. The length of the window size depends on the sampling frequency and the nature of activities targeted for classification from which data is collected [34].

4.3. Feature extraction

Feature is an attribute of the raw data that should be calculated [33]. In data analytics applications, statistical time- and frequency-domain features generated in each window are used as the input of the training process [23]. The ability to extract appropriate features depends on the application domain and can steer the process of retaining the relevant information. Most previous studies on activity recognition have used almost the same features for training the models and classification of activities [35].

4.4. Data annotation

Following data segmentation and feature extraction, the corresponding activity class labels should be assigned to each window. This serves as the ground truth for the learning algorithm and can be retrieved from a video, recorded at the time of the experiment.

4.5. Supervised learning

In supervised learning classification, class labels are provided to the learning algorithms to generate a model or function that matches the

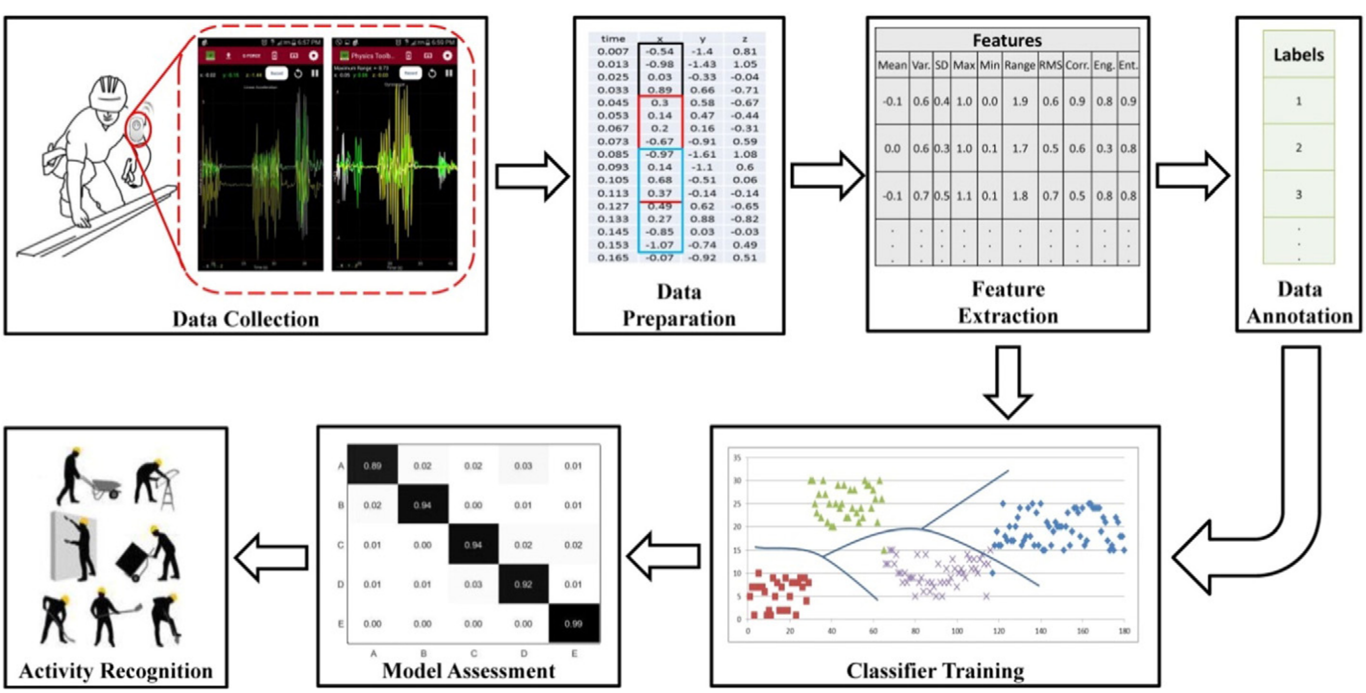

Fig. 1. Framework for construction worker activity recognition using mobile sensors.

input (i.e. features) to the output (i.e. activity classes) [23]. The goal is to infer a function using examples for which the class labels are known (i.e. training data). The performance of this function is evaluated by measuring the accuracy in predicting the class labels of unseen examples. Researchers have used different types of supervised classification methods for activity recognition [28,30,31].

4.6. Model assessment

In order to determine the reliability of the trained model in detecting new examples of activity classes, part of the training dataset is used for testing the model. It is recommended that the test set is independent of the training set, meaning that the data that are used for testing have not been among the training data. For example, randomly chosen 10% of the training data can be left out so that the training is performed on the remaining 90% of data. Assessment of the model provides an opportunity for its fine-tuning so that certain variables (e.g. regularization factor to prevent over-fitting in neural networks) in the algorithm can be revised to yield the best possible model.

4.7. Activity recognition

Once the model is trained and its parameters are finalized, it can be used for recognizing activities for which it has been trained. While data is being collected to determine the activities according to a trained classifier, such data can be stored in a dataset repository and be added to the existing training data, so that the model is further trained with a richer training dataset.

5. Experiment setup and data analysis

In this research, experiments were conducted in an outdoor workspace where different activities performed by multiple construction workers were imitated. These activities included sawing, hammering, turning a wrench, loading sections into wheelbarrows, pushing loaded wheelbarrows, dumping sections from wheelbarrows, and returning with empty wheelbarrows. Activities were performed in 3 different categories in order to assess certain circumstances (described in the following Subsections) in the outcome of classification. A commercially available armband was used to secure a smartphone on the upper arm of the dominant hand of each worker. Recent research on the selection of accelerometer location on bricklayer's body for activity recognition has shown that according to the body movement of the worker while performing different bricklaying activities, among 15 potential locations for wearing an accelerometer, the lower left arm and the upper right arm are the two best locations that yield the highest information gain [36]. In this study, the lower arm was not selected for recognition of the activities of interest since the workers stated that it would preclude convenient execution of some activities. Consequently, the selection of the upper arm was expected to provide accurate and consistent results compared to other locations on the body. Fig. 2 shows some snapshots of the construction workers wearing mobile phones on their upper arms while performing assigned activities in the experiments conducted in this research. As it appears in Fig. 2, the second worker's armband is located slightly above the elbow. The movements produced are the same as the upper arm and as long as the position of the device does not change significantly during the experiment, the training and later testing are still valid. It should be noted that all four human subjects were Construction Engineering and Management students who had basic prior experience working in various construction jobsites.

5.1. Data collection

Smartphone built-in sensors and sensor logging applications in both Android and iOS operating systems were used for data collection. Most of the current iOS and Android smartphones are equipped with 3 degree-of-freedom (DoF) sensors including three-axis accelerometer and gyroscope. For example, Apple iPhones are equipped with STMicroelectronics LIS331DLH accelerometer and the L3G4200D gyroscope. There are plenty of free applications available on both Apple Store and Google Play that enable data collection and spreadsheet logging. The spreadsheet files are automatically saved following the data collection and transferred to a remote computer for processing. The sampling frequency was set at 100 Hz. This frequency is neither too low to miss any movement corresponding to the target activities, nor too high to result in a large size for the collected data file. This sampling frequency has been also used in previous studies for accelerometer based activity recognition [37,38]. Data was collected in all 3 axes (X, Y, Z) from accelerometer and gyroscope. Construction workers were asked to do their assigned activities for a certain period of time while waiting for a few seconds in between each instance of their assigned activities. Each activity was performed by two subjects for later user-independent evaluations. Two subjects performed only sawing. In this case, the goal of activity recognition was to differentiate between the time they were *sawing* and the time they were *not sawing*. Two other subjects performed hammering and turning a wrench. In this case, the activity recognition was intended to detect the time they were *hammering*, the time they were *turning the wrench*, and the time there were not doing any of the two activities. Finally, the last two subjects were responsible for pushing the wheelbarrow and loading/unloading the sections. Therefore, the activities to be recognized in this case were *loading* sections into a wheelbarrow, *pushing* a loaded wheelbarrow, *dumping* sections from a wheelbarrow, and *returning* with an empty wheelbarrow. Workers in the experiments were not instructed to perform any activity in any specific way; rather they were only tasked with completing their own job in their natural body pose and movements.

Time-stamped data were logged into comma separated values (CSV) spreadsheets. The entire experiment was videotaped for data annotation.

Fig. 2. Data collection experiments (mobile devices are marked with dashed circles).

Time-stamped data from accelerometer and gyroscope were also synchronized and subsequently matched with the timer of the video during the data annotation process. This step is to ensure that the timing of the individual sensory data points within the dataset is mapped precisely to that of the recorded video so that a data point labeled at t_1 represents an activity performed at t_2 if and only if $t_1 = t_2$ in the experiment.

5.2. Data analysis

Table 1 shows the number of data points collected per sensor per axis. Since classifications are conducted in 3 activity categories, the numbers of collected data points are tabulated and reported for each category. Category 1 includes only one distinguishable activity, *sawing*, to assess the performance of the classifiers in detecting value-adding versus non-value-adding instances in a pool of accelerometer and gyroscope data. The result of classification in this category contributes to the overall performance of the developed activity recognition system when used for productivity measurement. In this category, *sawing* is categorized against *idling*. Category 2 includes instances of consecutive *hammering* and *turning a wrench* as two completely separate but adjacent activities with almost similar corresponding movements of the worker's arm. These two activities are also classified against *idling* to assess the accuracy of the developed activity recognition system in differentiating between activities that produce similar physical body motions. Finally, in category 3, four activities that produce different distinguishable body movements are categorized. These activities include *loading sections into a wheelbarrow*, *pushing a loaded wheelbarrow*, *dumping sections from a wheelbarrow*, and *returning an empty wheelbarrow*, that were also categorized against *idling*. Multiplication of the number of data points by 6 will result in all data points collected from two sensors in three axes.

The smartphone clock is used as the reference clock for all the applications that collect data using smartphone's embedded sensors. Therefore, it is fair to assume that all sensor data are collected with a uniform sampling rate and each data point from any given sensor can be paired with a data point from another sensor, as long as their timestamps match. However, due to the inevitable drift of the sensors in high frequency data collection mode, a pre-processing step is necessary. This is to make sure that data points are aligned with each other according to the timestamp and eventually belong to the same timestamp in the recorded video of the same experiment.

In order to make up for the missing data and remove redundant data collected in a higher rate than the assigned sampling frequency, the timestamps of the adjacent collected data points were examined. Considering the 100 Hz sampling frequency, the normal difference between the two adjacent timestamps must be around 0.01 s. Therefore, in the data preparation phase, if this difference is greater than 0.015 s, the X, Y, and Z values of the missing data point were interpolated as the average of the two adjacent data points. As for the redundant collected data, any data point collected within less than 0.005 s of the last collected data point was removed. This assures the compatibility of the collected data with the recorded videotape for data annotation. As far as data segmentation was concerned, every 128 data points were segmented in one window and considering the 100 Hz sampling frequency, each window amounts to 1.28 s of data collection. The choice of 128 data points was due to conversion of the time domain to the frequency domain using fast Fourier transform (FFT) in which the window size should be a power of 2 [18,19]. If the window size is not a power of 2, zeros will be added to the end of the window or it would be truncated to become a power of 2. With regard to the overlapping of the adjacent windows, previous studies for accelerometer-based activity recognition have suggested a 50% overlap between the adjacent windows [23,39] and hence, a 50% overlap was also considered for data analysis in this research. 50% overlap means that the second half of the first section will be overlapped with the first half of the second section. The overlapping assures that no value-adding activity that may be split into two consecutive sections will go unnoticed.

Table 1
Collected data points per sensor per axis in each activity category.

Category	Activity	Number of data points per sensor per axis
1	Sawing	120,755
2	Hammering + turning a wrench	149,682
3	Loading + hauling + unloading + returning	337,800

Moreover, common features used for activity recognition found in literature [35] were selected in this study and extracted from the raw data. In particular, mean, maximum, minimum, variance, root mean square (RMS), interquartile range (IQR) and correlation between each two pairs of axes comprised the seven time-domain features and spectral energy and entropy were the two frequency domain features. Mean, maximum, minimum, and variance are simply calculating the same for the data points in a window. RMS is the square root of the arithmetic mean of the squares of the values in a section. IQR is the difference between the first and third quartile on the data point values. Also, the correlation of the mean of the section data points from each two pair of axes is calculated. Finally, spectral energy describes the distribution of the signal's energy by the frequency and spectral entropy measures the irregularity of the signal by calculating the normalized information entropy of the discrete FFT component magnitudes [24]. Considering data collection in three axes of the two sensors and nine independent features extracted per sensor per axis, a total of 54 features were extracted from all collected data. Labeling windows was performed manually according to the recorded video of the data collection experiment. The extracted features include but are not limited to statistical time-domain features such as mean, maximum, and RMS, as well as statistical frequency-domain such as signal energy and entropy.

5.3. Classifier training

The performance of five different classification techniques in accurately detecting worker activities was systematically evalauted. In particular, neural network, decision tree, K-nearest neighbor (KNN), logistic regression, and support vector machine (SVM) were employed for classification. Decision tree, KNN, and SVM have been previously used for activity recognition [23,37,38] so they were also selected in this study. However, neural network and logistic regression were examined to a much lesser extent [40].

5.3.1. Neural network

The architecture of the neural network used for recognizing the activities is depicted in Fig. 3.

As shown in this figure, the network consists of one input, one hidden, and one output layer. Considering the 54 features that serve as the input of the neural network, the input layer has $m = 54$ units. The hidden layer consists of $p = 25$ units; this is selected considering the

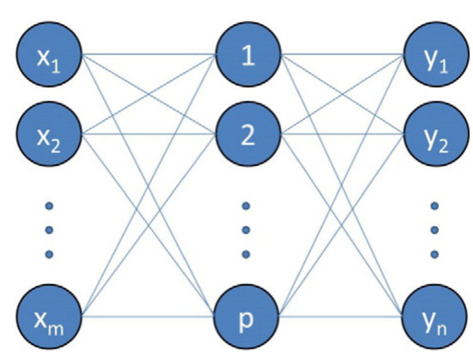

Fig. 3. The architecture of the neural network used in this research.

average of the sum of the units in the input and output layers [41]. The number of units for the output layer is equal to the number of activity classes, n in each case. Given the large feature space and in order to prevent overfitting, regularization was used. Using a regularization parameter, the magnitude of the model weights decreases, so that the model will not suffer from high variance to fail to generalize to the new unseen examples [42]. The activation function (i.e. hypothesis) used for minimizing the cost function in the training process is a Sigmoid function shown in Eq. (1),

$$h_{\Phi}(x) = \frac{1}{1+e^{-\Phi x}} \tag{1}$$

in which $h_{\Phi}(x)$ is the activation function (i.e. hypothesis), Φ is a matrix of model weights (i.e. parameters), and x is the features matrix. In this study, in order to minimize the cost function, the most commonly used neural network training method, namely feed-forward backpropagation is used. Considering a set of randomly chosen initial weights, the backpropagation algorithm calculates the error of the activation function in detecting the true classes and tries to minimize this error by taking subsequent partial derivatives of the cost function with respect to the model weights [43].

5.3.2. Decision tree

Decision tree is one of the most powerful yet simplest algorithms for classification [44]. The decision tree method that is used in this research is classification and regression tree (CART). CART partitions the training examples in the feature space into rectangle regions (a.k.a. nodes) and assigns each class to a region. The process begins with all classes spread over the feature space and examines all possible binary splits on every feature [44]. A split is selected if it has the best optimization criterion which is the Gini diversity index in this research, as shown in Eq. (2),

$$I_G(f) = 1-\sum_{i=1}^{k} f_i^2 \tag{2}$$

in which I_G is the Gini index, f_i is the fraction of items labeled with value i and k is the number of classes. The process of splitting is repeated iteratively for all nodes until they are *pure*. A node is considered *pure* if it contains only observations of one class, implying a Gini index of zero, or that there are fewer than 10 observations to split.

5.3.3. K-nearest neighbor (KNN)

Similar to the decision tree and unlike the neural network, KNN is a simple algorithm. Training examples identified by their labels are spread over the feature space. A new example is assigned to a class that is most common amongst its K nearest examples considering the Euclidean distance that is used as the metric in this research and as appears in Eq.(3),

$$D = \sqrt{\left(x_i^{(1)}-x_{new}^{(1)}\right)^2+\left(x_i^{(2)}-x_{new}^{(2)}\right)^2+\ldots+\left(x_i^{(d)}-x_{new}^{(d)}\right)^2} \tag{3}$$

in which D is the Euclidean distance, x_i is an existing example data point which has the least distance with the new example, x_{new} is the new example to be classified, and d is the dimension of the feature space.

5.3.4. Logistic regression

Logistic regression is a type of regression problems in which the output is discretized for classification [45]. Logistic regression seeks to form a hypothesis function that maps the input (i.e. training data) to the output (i.e. class labels) by estimating the conditional probability of an example belonging to class k given that the example actually belongs to the class k. This is accomplished by minimizing a cost function using a hypothesis function and correct classes to find the parameters of the mapping model [45]. The hypothesis function used in this research is the same as the activation function introduced in Eq. (1) (the Sigmoid function) and thus the cost function to minimize is as shown in Eq. (4),

$$J(\theta) = -\frac{1}{m}\,[\sum_{i=1}^{m} y^{(i)} \log h_\theta\left(x^{(i)}\right)+\left(1-y^{(i)}\right)\log\left(1-h_\theta\left(x^{(i)}\right)\right) \tag{4}$$

in which $J(\theta)$ is the cost function, m is the number of training examples, θ is the feature, $h(\theta)$ is the hypothesis function, $x^{(i)}$ is the ith training example, and $y^{(i)}$ is the corresponding correct label. Once the cost function is minimized using any mathematical method such as the Gradient Decent [45] and parameters are found, the hypothesis will be formed. In multiclass classification, the one-versus-all method is used to determine if a new example belongs to the class k [45]. Therefore, considering k classes, k hypothesis functions will be evaluated for each new example and the one that results in the maximum hypothesis is selected.

5.3.5. Support vector machine (SVM)

Compared to decision tree and KNN, SVM is considered as a more powerful classification algorithm. Although it has been widely used in vision-based pattern recognition and classification problems, some researchers [44] used it for classifying daily activities and thus its performance is also assessed in this research. In a nutshell, SVM tries to maximize the margin around hyperplanes that separate different classes from each other. SVM can benefit from a maximum margin hyperplane in a transformed feature space using kernel function to create non-linear classifiers. The kernel function used for non-linear classification in this research is Gaussian radial basis function (rbf) which has been successfully applied in the past to activity recognition problems [23]. Further description of SVM models are out of the scope of this study but can be found in [44].

6. Results and discussion

The performance of the classifiers is assessed in two ways. First, the training accuracy of each classifier was calculated. This means that all collected data points were used for both training and testing which provided an overall insight into the performance of a host of classification algorithms in recognizing construction worker activities using accelerometer and gyroscope data. Next, a more robust approach in evaluation of classifiers was adopted. In particular, 10-fold stratified cross validation was used and the results of the 10 replications of the training and testing were averaged out to report the overall accuracy. In k-fold cross validation, data are divided into k parts with (almost) equal number of data points. Next, in k recursive steps, one part is left out for testing and the remaining k-1 parts are used for training. In “stratified” version of k-fold cross validation, the k fold segmentation is done in a way that the proportion of the data from every class in each of the k parts remains the same as that of the entire training data [45]. It is

Table 2
Classification accuracy (%) for category 1 activities.

Category1		Neural network	Decision tree	KNN	Logistic regression	SVM
Training	Subject I	100.00	99.36	98.08	98.72	98.19
	Subject II	99.25	99.15	97.34	98.08	97.33
10-Fold CV	Subject I	96.77	96.06	95.95	96.05	96.91
	Subject II	97.02	95.42	96.27	96.70	96.59

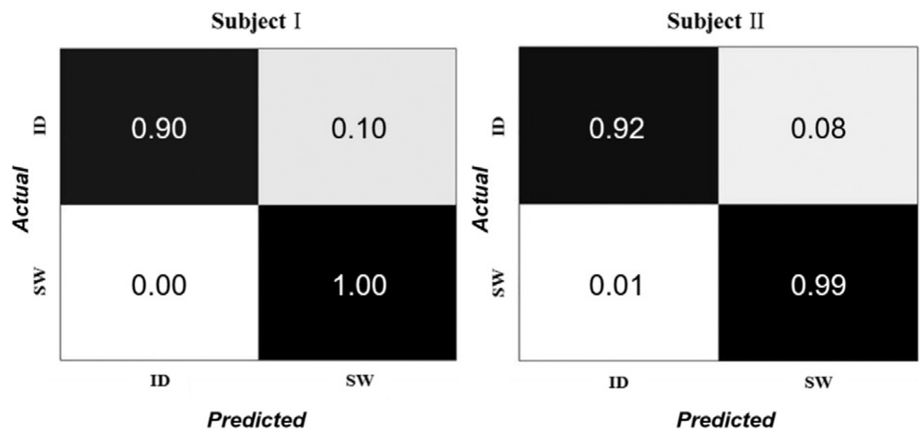

Fig. 4. Confusion matrices of 10-fold cross validation of neural network classification for category 1 activities.

worth mentioning that in the following Subsections, accuracy is measured as the ratio of the sum of true positive and true negative over the total instances.

6.1. Category 1

The classification accuracies are reported for 3 activity categories listed in Table 1. The following activity codes are used in reporting the results: in the first category, activity *sawing (SW)* and *being idle (ID)* are classified. In the second category, activities *hammering (HM)*, *turning a wrench (TW)*, and *being idle (ID)* are classified. Finally, in the third category classification is performed on the activities *loading sections into wheelbarrow (LW)*, *pushing a loaded wheelbarrow (PW)*, *dumping sections from wheelbarrow (DS)*, *returning an empty wheelbarrow (RW)*, and *being idle (ID)*. Table 2 shows the results of training and 10-fold cross validation classification accuracy of both subjects performing activities of category 1.

According to Table 2, over 99% training accuracy was achieved for both subjects in category 1 using neural network classifier. This confirms the hypothesis that IMU data pertaining to a single activity performed by different workers contain highly distinguishable patterns. However, training accuracy is not an appropriate measure to assess the ability of using such data for new instances of the same activity. Nevertheless, the stratified 10-fold cross validation results confirm that regardless of the nature of classification algorithm, a single activity can be recognized with over 95% accuracy using all five classifiers. A thorough exploration of classification results within each category can help understanding the accuracies of each one of the activities versus the non-value-adding (i.e. *idling*) state. To achieve this, the confusion matrices of 10-fold stratified activity classifications for both subjects resulted from the best classifier (i.e. neural network) are shown in Fig. 4. In this confusion matrix, the rows show the percentage of actual instances and columns indicate the percentage of predicted instance of activities labeled. For example, in Fig. 4 for Subject 1, row ID, in 90% of the instances where the subject was *Idle* the predicted instance was *Idle* too. However in 10%, it was predicted as *Sawing*. Fig. 4 indicates more than 90% accuracy in correct detection of the instances of the two activities.

6.2. Category 2

Since it is very likely that a construction worker performs more than one highly distinguishable activity at a time, activities performed in category 2 are designed such that they produce almost the same physical arm movement. Table 3 shows the training and 10-fold cross validation classification accuracy results of both subjects performing activities of category 2.

Similar to category 1, the training accuracies are high particularly for the neural network classifier and the decision tree. CART decision trees are not very stable and a small change in the training data can change the result drastically. Moreover, a decision tree is actually expected to have a training accuracy of around 100% anyway due to its training nature. However, as appears in the outcome of the 10-fold cross validation, neural network presents an average of around 90% accuracy for both subjects. This is while all other classification methods performed almost the same with a slight superiority of KNN relative to the other algorithms. This result is particularly important considering the fact that the two activities in category 2 (i.e. *hammering* and *turning a wrench*) produce almost similar physical movements in a worker's arm. Fig. 5 shows how these two activities are classified using 10-fold cross validation of the result obtained from neural network.

As appeared in Fig. 5, both activities have been in fact classified with a high accuracy and the major contributor to lowering the overall accuracy was the idling state. This can be justified by the fact that the non-value-adding state may include different forms of physical movements in case different activities are performed. In other words, the ID class includes various movements of different types so that relative to other two activities, more instances have been misclassified.

6.3. Category 3

In the third category, a mixture of different distinguishable activities performed by construction workers is included to evaluate the performance of the developed activity recognition system in recognizing them. Table 4 shows the training and 10-fold cross validation classification accuracy results of both subjects performing activities of category 3.

Table 3
Classification accuracy (%) for category 2 activities.

Category2		Neural network	Decision tree	KNN	Logistic regression	SVM
Training	Subject I	98.62	97.07	93.81	88.14	87.28
	Subject II	93.30	94.67	91.67	84.03	83.43
10-Fold CV	Subject I	93.19	85.83	87.80	86.42	85.34
	Subject II	86.64	78.20	83.35	81.02	81.72

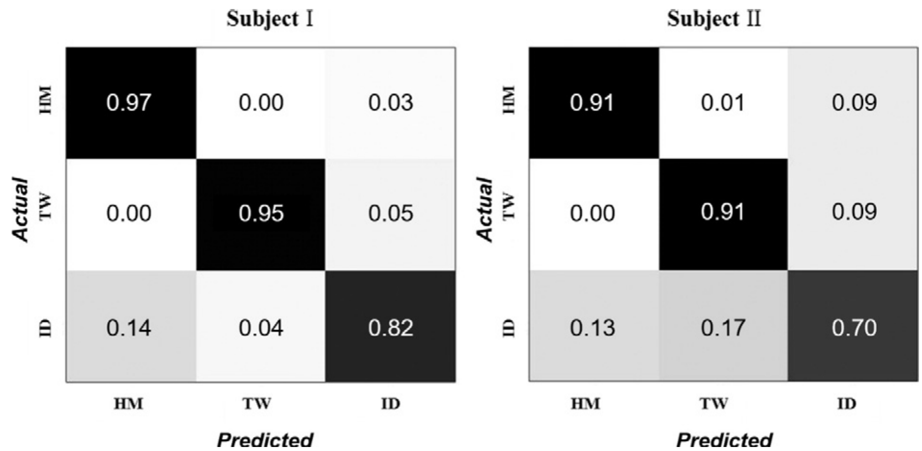

Fig. 5. Confusion matrices of 10-fold cross validation of neural network classification for category 2 activities.

According to Table 4, again decision tree yielded a high accuracy in training while as expected, its performance is not the same in 10-fold cross validation evaluation. However, except for the decision tree and SVM, all other classifiers, namely neural network, KNN, and logistic regression resulted in around 90% average accuracy for both subjects. Similar to the other two categories, the feedforward back-propagation implementation of the neural network resulted in the highest accuracy among all. Fig. 6 shows how different activities in this category are classified using 10-fold cross validation of the result obtained from neural network.

Based on the confusion matrices of Fig. 6, the non-value-adding or *idling* state was classified properly in both cases. The most confused activities are LW and RW, particularly for the first subject, and LW, PW, and RW for the second subject. This might be due to the fact that LW and RW result in similar body movement patterns, while as confirmed in the two presented cases, different humans perform various activities with slightly different body movements (function of body height, body shape, …). This may result in some confusion between two or more activities in each case.

6.4. Combined data

After classifying the activities within each category based on the individual data received from each subject, the data collected from both subjects were combined to perform another round of classification. This evaluation allows further investigation of whether appending new data collected in future instances to existing data warehouse would result in acceptable classification and recognition of activities. Table 5 shows the result of the classification of combined data in all three categories.

According to Table 5, all categories have training accuracies of more than 90% in at least one classification algorithm. This promising result indicates that there exist classifiers that can categorize activities of different natures using combined data collected from wearable IMUs in different instances. In case of new examples, considering the robust 10-fold cross validation technique, while logistic regression's and to a larger extent, KNN's performance is very close to that of neural network, again neural network outperforms all the other classifiers. Fig. 7 shows the confusion matrices of neural network for combined data of all three categories. According to Fig. 7, some of the classes such as ID in Category 2 and LW in Category 3 are not classified with as much accuracy as the other activities. Obviously, part of this error is attributed to the overall accuracy of classification that requires further improvement. In case of Category 2 and the ID class, it seems that Subject 2 (according to Subsection 6.2) introduces the majority of error as a result of what has been discussed in Subsection 6.2. This, however, needs to be taken into consideration for further improvement of the accuracy of detecting ID class in future. In case of Category 3, the LW is mostly confused with the PW class which again is more associated with Subject 2 movements. This error can be the result of similarity of movement in Subject 2 as well the adjacency of the two activities that may have resulted in false recognitions. It should be noted that all such errors are subject to further refinements of the detection and classification framework.

6.5. Subject-independent evaluation

The last evaluation of classifiers' performances is conducted for the case of using data from one subject as the training set to classify activities of the second subject. This assessment is particularly important when trained model with the existing data is sought to be used for newly collected data. Table 6 shows the results of training each classifier using the data collected from subject I/II and tested on the data collected from subject II/I. In each category, the "I on II" row indicates that the classifiers were trained using the subject I data and tested on subject II data, and the "II on I" row indicates that the classifiers were trained using the subject II data and tested on subject I data. Comparing different classifiers, it is apparent from Table 6 that KNN has the best classification accuracy which is even slightly better than neural network in this case. This is true for all the categories and thus indicates the power of KNN (despite its simplicity) in generalizing a trained model to new examples. Comparing different activity categories in this scenario, while classification of category 1 activities with only one distinguishable activity results in an accuracy of more than 96%, classification of activities in the other two categories have resulted in less accuracies. In particular, category 2 with two similar activities shows a less accurate performance. Nevertheless, while category 3 classification was performed

Table 4
Classification accuracy (%) for category 3 activities.

Category3		Neural network	Decision tree	KNN	Logistic regression	SVM
Training	Subject I	94.80	97.11	95.75	90.37	85.82
	Subject II	90.37	96.58	94.96	87.83	79.12
10-Fold CV	Subject I	92.01	87.95	90.75	90.75	84.42
	Subject II	88.90	87.12	86.74	86.51	78.55

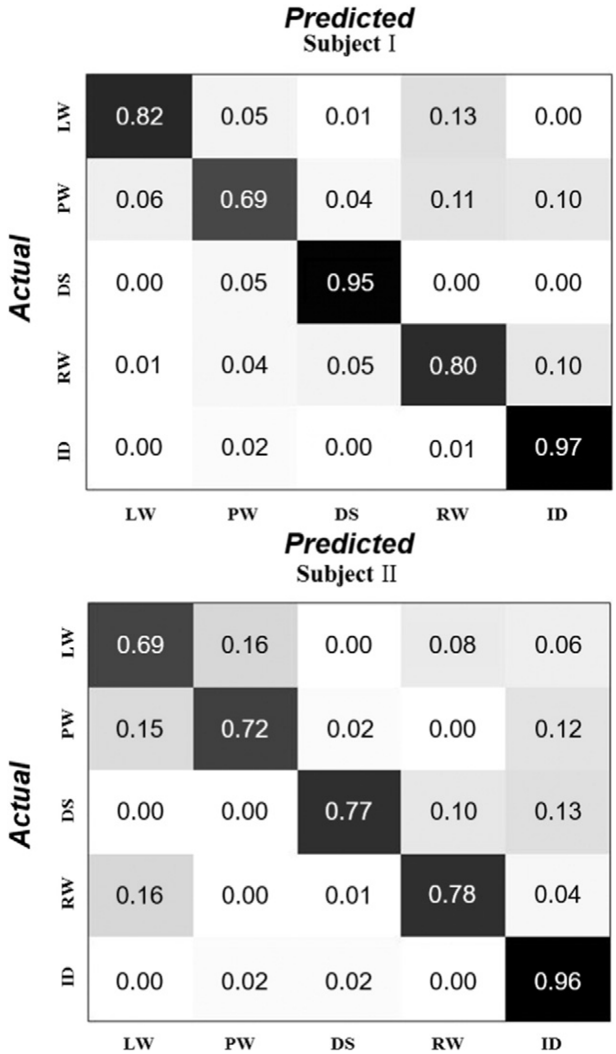

Fig. 6. Confusion matrices of 10-fold cross validation of neural network classification for category 3 activities.

using 5 different classes, an accuracy of around 80% in the best case (i.e. KNN) shows promising results when a rich data warehouse is available.

7. Summary and conclusions

In spite of its importance, automated recognition of worker activities on construction jobsites has not been given due attention in CEM literature. While some efforts have been made in the past to develop vision-based techniques for automated tracking and recognition of construction entities, the state-of-the-art in employing IMU sensors with wide variety of applications in other domains has not been yet explored within the CEM context. This paper introduces a novel methodology for designing and testing a low-cost pervasive construction worker activity recognition system capable of detecting activities of various natures that are typical to construction jobsites. Towards this goal, built-in sensors of ubiquitous smartphones were employed to assess the potential of wearable systems for activity recognition. Smartphones were affixed to workers' arms using sport armbands, and accelerometer and gyroscope data were collected from multiple construction workers involved in different types of activities.

The high levels of training accuracies achieved by testing several classification algorithms including neural network, decision tree, K-nearest neighbor (KNN), logistic regression, and support vector machine (SVM) confirmed the hypothesis that different classification algorithms can detect patterns that exist within signals produced by IMUs while different construction tasks are performed. Through 10-fold stratified cross validation, algorithms were trained with 90% of the available data and the trained models were tested on the remaining 10%. In different categories of activities, around and over 90% accuracy was achieved. This promising result indicates that built-in smartphone sensors have high potential to be used as integrated data collection and activity recognition platforms in construction environments. However, it should be noted that the results might be affected by the fact that workers in the experiments were not actual construction workers and the experiments were conducted in a controlled environment. Thus, detection of various (subtly different) ways of carrying a field task may not have been reflected in the results achieved. While the results show a promising prospective to employ ubiquitous smartphones for construction activity recognition, there are some implementation details that can potentially affect the results in real-world settings. Therefore, a direction for future work will be to explore potential scenarios that may introduce anomalies in the data and to investigate the system under such conditions. For example, sudden movements, or potential cases where a smartphone is not worn and/or does not function properly should be meticulously considered in a holistic framework of activity recognition. Another example of real world implementation issues is limited battery life and storage capacity of smartphones. Although such issues will be eventually addressed over time and as the hardware design technology catches up with the rapid pace of application development, one potential solution to this problem could be using a dynamic data collection frequency that varies over time depending on the resolution of the worker's movements and/or significance of certain body motions to activity recognition.

Further investigations were conducted by combining the data from multiple subjects. In the first two categories with less activities to be classified, accuracies of more than 90% were achieved which indicate that combination of data collected from different workers can result in promising outcome for activity recognition. When the number of activities increased and more similar activities were sought to be classified (i.e. category 3) the recognition accuracy fell to 70%–80%. In the last assessment, data from each subject were used to train two different classifiers. The trained models were then tested using the data collected from another subject. While this scenario introduced the most challenging situation, KNN was able to present around 95%, 75%, and 80% accuracies. It is worth mentioning that in terms of computational time, KNN is highly superior to neural network as it is much less complex because there is no need for an optimization process with high iteration numbers. KNN simply compares the test data to the training data and that is why it is also referred to as a "lazy learner" [44].

Overall, results indicated that the CEM domain similar to other sectors such as health and fitness, medicine, and elderly care can benefit

Table 5
Classification accuracy (%) for combined data of subjects I and II in all three activity categories.

Combined data for subjects I & II		Neural network	Decision tree	KNN	Logistic regression	SVM
Training	Category 1	99.75	99.04	97.71	97.39	97.23
	Category 2	91.67	95.49	92.27	82.86	82.86
	Category 3	89.49	96.48	92.46	86.41	78.62
10-Fold CV	Category 1	96.27	95.58	96.22	96.54	96.64
	Category 2	87.78	78.57	87.73	82.23	82.18
	Category 3	88.17	85.62	87.68	85.84	78.34

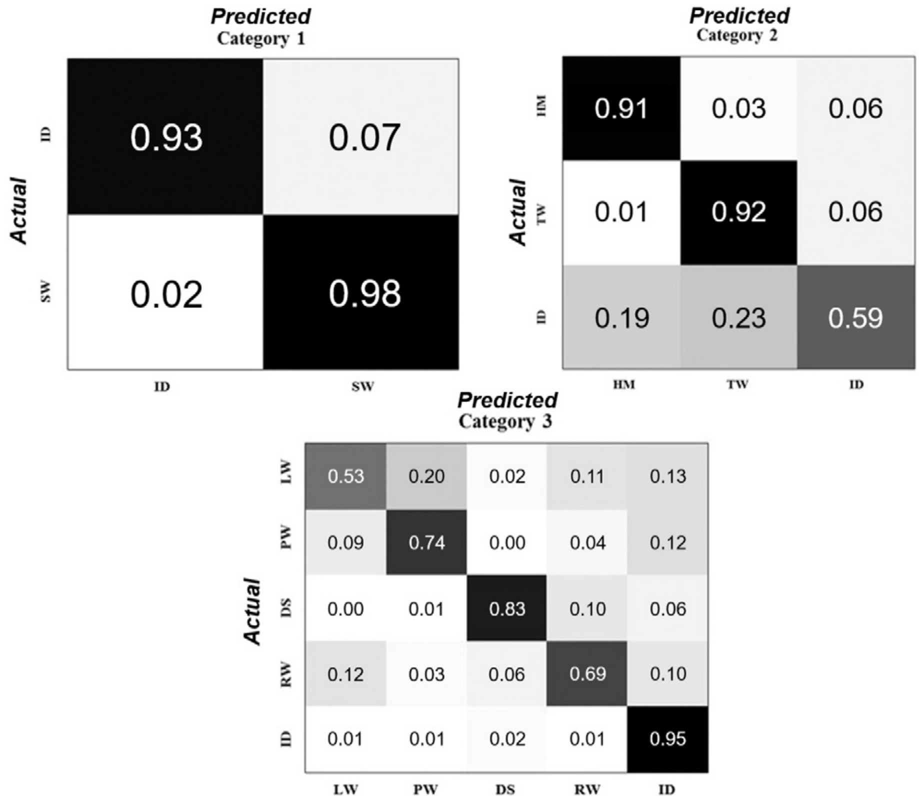

Fig. 7. Confusion matrices of 10-fold cross validation of neural network classification for combined data of subjects I and II in all three activity categories.

from the applications of activity recognition on construction jobsites. Some application areas include productivity measurement, progress evaluation, labor training programs, and safety and health management.

8. Future work

While the results show a promising prospective to employ ubiquitous smartphones for construction activity recognition, there are some implementation details that can potentially affect the results in real-world settings. Therefore, a direction for future work will be to explore potential scenarios that may introduce anomalies in the data and to investigate the system under such conditions. For example, sudden movements, or potential cases where a smartphone is not worn and/or does not function properly should be meticulously considered in a holistic framework of activity recognition. Another example of real world implementation issues is limited battery life and storage capacity of smartphones. Although such issues will be eventually addressed over time and as the hardware design technology catches up with the rapid pace of application development, one potential solution to this problem could be using a dynamic data collection frequency that varies over time depending on the resolution of the worker's movements and/or significance of certain body motions to activity recognition.

Another potential direction for future work in this area will be to explore whether the results achieved so far can be used for automatically extracting process knowledge such as activity durations and precedence logic for the purpose of ubiquitously updating and maintaining simulation models corresponding to field operations. In addition, another branch of future work rooted in the current research is automated identification of unsafe workers' body postures in physically demanding construction activities. Work-related Musculoskeletal Disorder (WMSD), back, knee, and shoulders injuries are among the most common injuries that can be prevented or reduced by complying with Occupational Safety and Health Administration (OSHA) or the National Institute for Occupational Safety and Health (NIOSH) standards and rules [46].

Productivity measurement and improvement is another direction for future work of this study. There has been a great deal of research on different techniques for productivity evaluation, tracking, and improvement in construction industry such as the construction industry institute (CII) productivity measurement methods [47], the construction productivity metric system (CPMS) [48], activity/work sampling [49,50], and recent studies targeting the relationship between task-level productivity and physical movements such as the study conducted by Gatti, Migliaccio, Bogus and Schneider [51]. In particular, using the collected data it is possible to calculate the proportion of time dedicated by each worker to each activity. For example, Fig. 8 shows pie charts

Table 6
Accuracy (%) of classifiers trained with data from one subject and tested on data from another subject.

		Neural network	Decision tree	KNN	Logistic regression	SVM
Category 1	I on II	94.24	94.78	96.05	93.71	94.04
	II on I	95.73	92.00	96.42	96.26	93.07
Category 2	I on II	62.10	63.05	68.20	64.93	63.30
	II on I	73.65	55.10	78.30	80.53	80.93
Category 3	I on II	78.85	73.66	79.23	76.62	72.45
	II on I	77.86	57.31	79.79	78.92	71.49

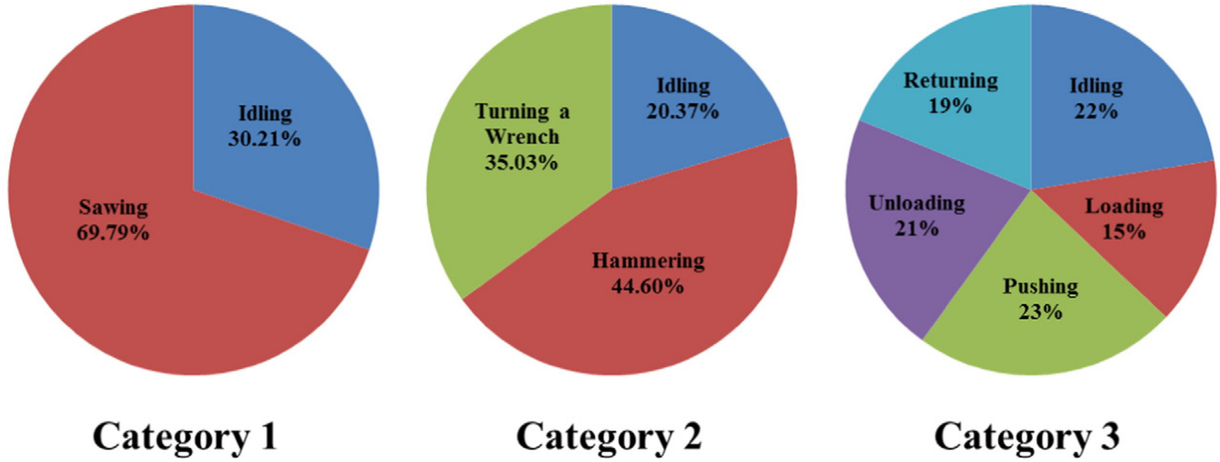

Fig. 8. Discovered time allocation proportions in the conducted experiments, for productivity measurement.

indicating the proportions of time dedicated to each activity in the experiments conducted in this research, as discovered by the designed activity recognition system. It is worth mentioning that although imbalanced classes of activities affect classification accuracy, the developed system is capable of differentiating dominant activities (those that take more time) and other activities. This is evidenced by the first pie chart in Fig. 8.

The discovered knowledge presented in this Figure is of great importance to the process of productivity measurement and improvement. Particular to the activity/work sampling, this information can help automate the process, thus significantly reducing the manpower required for manual analysis and potential errors and inconsistencies associated with manual observations.

References

[1] T. Stiefmeier, D. Roggen, G. Ogris, P. Lukowicz, G. Tröster, Wearable activity tracking in car manufacturing, IEEE Pervasive Comput. 2 (2008) 42–50, http://dx.doi.org/10.1109/mprv.2008.40.

[2] T. Stiefmeier, C. Lombriser, D. Roggen, H. Junker, G. Ogris, G. Tröster, Event-based activity tracking in work environments, in: Proceeding of Applied Wearable Computing (IFAWC), 2006 3rd International Forum on, VDE 2006, pp. 1–10.

[3] J. Favela, M. Tentori, L.A. Castro, V.M. Gonzalez, E.B. Moran, A.I. Martínez-García, Activity recognition for context-aware hospital applications: issues and opportunities for the deployment of pervasive networks, Mob. Netw. Appl. 12 (2–3) (2007) 155–171, http://dx.doi.org/10.1007/s11036-007-0013-5.

[4] T. Cheng, J. Teizer, G.C. Migliaccio, U.C. Gatti, Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data, Autom. Constr. 29 (2013) 24–39, http://dx.doi.org/10.1016/j.autcon.2012.08.003.

[5] R. Akhavian, A.H. Behzadan, Evaluation of queuing systems for knowledge-based simulation of construction processes, Autom. Constr. 47 (2014) 37–49, http://dx.doi.org/10.1016/j.autcon.2014.07.007.

[6] R. Akhavian, A.H. Behzadan, Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers, Adv. Eng. Inform. (2015) http://dx.doi.org/10.1016/j.aei.2015.03.001.

[7] L. Joshua, K. Varghese, Accelerometer-based activity recognition in construction, J. Comput. Civ. Eng. 25 (5) (2010) 370–379, http://dx.doi.org/10.1061/(asce)cp.1943-5487.0000097.

[8] J. Gong, C.H. Caldas, An intelligent video computing method for automated productivity analysis of cyclic construction operations, in: Proceeding of Proc. The 2009 ASCE International Workshop on Computing in Civil Engineering 2009, pp. 64–73.

[9] C. Kim, T. Park, H. Lim, H. Kim, On-site construction management using mobile computing technology, Autom. Constr. 35 (2013) 415–423, http://dx.doi.org/10.1016/j.autcon.2013.05.027.

[10] C. De Dominicis, A. Depari, A. Flammini, S. Rinaldi, E. Sisinni, Smartphone based localization solution for construction site management, Proceeding of Sensors Applications Symposium (SAS), 2013 IEEE, IEEE 2013, pp. 43–48.

[11] R. Chen, Ubiquitous Positioning and Mobile Location-based Services in Smart Phones, IGI Global, 2012.

[12] ENR, http://www.enr.com/articles/9486-equipment-tracks-trends-news-briefs?v=preview (accessed August 15).

[13] L. Chen, J. Hoey, C.D. Nugent, D.J. Cook, Z. Yu, Sensor-based activity recognition, IEEE Trans. Syst. Man Cybern. Part C 42 (6) (2012) 790–808, http://dx.doi.org/10.1109/tsmcc.2012.2198883.

[14] T. Dutta, Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace, Appl. Ergon. 43 (4) (2012) 645–649, http://dx.doi.org/10.1016/j.apergo.2011.09.011.

[15] 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 (6) (2012) 635–644, http://dx.doi.org/10.1061/(asce)cp.1943-5487.0000279.

[16] R. Gonsalves, J. Teizer, Human motion analysis using 3D range imaging technology, Proceeding of Proceedings of the 26th International Symposium on Automation and Robotics in Construction, Austin, Texas 2009, pp. 76–85.

[17] S. Han, S. Lee, A vision-based motion capture and recognition framework for behavior-based safety management, Autom. Constr. 35 (2013) 131–141, http://dx.doi.org/10.1016/j.autcon.2013.05.001.

[18] T. Cheng, G. Migliaccio, J. Teizer, U. Gatti, Data fusion of real-time location sensing (RTLS) and physiological status monitoring (PSM) for ergonomics analysis of construction workers, ASCE J. Comput. Civ. Eng. (2012) http://dx.doi.org/10.1061/(asce)cp.1943-5487.0000222.

[19] R. Akhavian, A.H. Behzadan, Dynamic simulation of construction activities using real time field data collection, Proceeding of 18th Workshop of Intelligent Computing in Engineering and Architecture (EG-ICE), Enschede, Netherlands 2011, pp. 1–10.

[20] R. Akhavian, A.H. Behzadan, An integrated data collection and analysis framework for remote monitoring and planning of construction operations, Adv. Eng. Inform. 26 (4) (2012) 749–761, http://dx.doi.org/10.1016/j.aei.2012.04.004.

[21] R. Akhavian, A.H. Behzadan, Knowledge-based simulation modeling of construction fleet operations using multimodal-process data mining, J. Constr. Eng. Manag. 139 (11) (2013) 04013021, http://dx.doi.org/10.1061/(asce)co.1943-7862.0000775.

[22] R. Akhavian, A. Behzadan, Wearable sensor-based activity recognition for data-driven simulation of construction workers' activities, Proceeding of Proceedings of the 2015 Winter Simulation Conference, IEEE Press 2015, pp. 3333–3344.

[23] N. Ravi, N. Dandekar, P. Mysore, M.L. Littman, Activity recognition from accelerometer data, Proceeding of AAAI 2005, pp. 1541–1546.

[24] L. Bao, S.S. Intille, Activity recognition from user-annotated acceleration data, Pervasive Computing, Springer 2004, pp. 1–17.

[25] U. Maurer, A. Smailagic, D.P. Siewiorek, M. Deisher, Activity recognition and monitoring using multiple sensors on different body positions, Proceeding of International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06), IEEE 2006, pp. 4–116.

[26] T. Brezmes, J.-L. Gorricho, J. Cotrina, Activity recognition from accelerometer data on a mobile phone, Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living, Springer 2009, pp. 796–799.

[27] J.R. Kwapisz, G.M. Weiss, S.A. Moore, Activity recognition using cell phone accelerometers, ACM SigKDD Explor. Newsl. 12 (2) (2011) 74–82, http://dx.doi.org/10.1145/1964897.1964918.

[28] L. Sun, D. Zhang, B. Li, B. Guo, S. Li, Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations, Proceeding of International Conference on Ubiquitous Intelligence and Computing, Springer 2010, pp. 548–562.

[29] A. Bayat, M. Pomplun, D.A. Tran, A study on human activity recognition using accelerometer data from smartphones, Procedia Comput. Sci. 34 (2014) 450–457, http://dx.doi.org/10.1016/j.procs.2014.07.009.

[30] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, M. Srivastava, Using mobile phones to determine transportation modes, ACM Trans. Sens. Netw. (TOSN) 6 (2) (2010) 13, http://dx.doi.org/10.1145/1689239.1689243.

[31] T.-S. Kim, J.-H. Cho, J.T. Kim, Mobile motion sensor-based human activity recognition and energy expenditure estimation in building environments, Sustainability in Energy and Buildings, Springer 2013, pp. 987–993.

[32] H. Martín, A.M. Bernardos, J. Iglesias, J.R. Casar, Activity logging using lightweight classification techniques in mobile devices, Pers. Ubiquit. Comput. 17 (4) (2013) 675–695, http://dx.doi.org/10.1007/s00779-012-0515-4.

[33] M. Khan, S.I. Ahamed, M. Rahman, R.O. Smith, A feature extraction method for realtime human activity recognition on cell phones, Proceeding of Proceedings of

3rd International Symposium on Quality of Life Technology (isQoLT 2011). Toronto, Canada, 2011.
[34] X. Su, H. Tong, P. Ji, Activity recognition with smartphone sensors, Tsinghua Sci. Technol. 19 (3) (2014) 235–249, http://dx.doi.org/10.1109/ccnc.2013.6488584.
[35] M. Shoaib, S. Bosch, O.D. Incel, H. Scholten, P.J. Havinga, A survey of online activity recognition using mobile phones, Sensors 15 (1) (2015) 2059–2085, http://dx.doi.org/10.3390/s150102059.
[36] L. Joshua, K. Varghese, Selection of accelerometer location on bricklayers using decision trees, Comput.-Aided Civ. Infrastruct. Eng. 28 (5) (2013) 372–388, http://dx.doi.org/10.1111/mice.12002.
[37] Z. Yan, V. Subbaraju, D. Chakraborty, A. Misra, K. Aberer, Energy-efficient continuous activity recognition on mobile phones: an activity-adaptive approach, Proceeding of 2012 16th international symposium on wearable computers, Ieee 2012, pp. 17–24.
[38] M. Kose, O.D. Incel, C. Ersoy, Online human activity recognition on smart phones, Proceeding of Workshop on Mobile Sensing: From Smartphones and Wearables to Big Data 2012, pp. 11–15.
[39] S. Chernbumroong, A.S. Atkins, H. Yu, Activity classification using a single wrist-worn accelerometer, Proceeding of Software, Knowledge Information, Industrial Management and Applications (SKIMA), 2011 5th International Conference on, IEEE 2011, pp. 1–6.
[40] J. Staudenmayer, D. Pober, S. Crouter, D. Bassett, P. Freedson, An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer, J. Appl. Physiol. 107 (4) (2009) 1300–1307, http://dx.doi.org/10.1152/japplphysiol.00465.2009.
[41] J. Heaton, Introduction to Neural Networks With Java, Heaton Research, Inc., 2008
[42] S.S. Haykin, Neural Networks and Learning Machines, Pearson Education, Upper Saddle River, 2009.
[43] M.H. Hassoun, Fundamentals of Artificial Neural Networks, MIT press, 1995.
[44] C.M. Bishop, Pattern Recognition and Machine Learning, springer, New York, 2006.
[45] J. Friedman, T. Hastie, R. Tibshirani, The Elements of Statistical Learning, Springer series in statistics Springer, Berlin, 2001.
[46] E.R. Vieira, S. Kumar, Working postures: a literature review, J. Occup. Rehabil. 14 (2) (2004) 143–159, http://dx.doi.org/10.1023/b:joor.0000018330.46029.05.
[47] C.I. Institute, Productivity Measurement: An Introduction, Austin, TX, 1990.
[48] H.-S. Park, S.R. Thomas, R.L. Tucker, Benchmarking of construction productivity, J. Constr. Eng. Manag. 131 (7) (2005) 772–778, http://dx.doi.org/10.1108/14635779610149151.
[49] H.R. Thomas, Labor productivity and work sampling: The bottom line, J. Constr. Eng. Manag. 117 (3) (1991) 423–444, http://dx.doi.org/10.1061/(asce)0733-9364(1991)117:3(423).
[50] H.R. Thomas, J. Daily, Crew performance measurement via activity sampling, J. Constr. Eng. Manag. 109 (3) (1983) 309–320, http://dx.doi.org/10.1061/(asce)0733-9364(1983)109:3(309).
[51] U.C. Gatti, G.C. Migliaccio, S.M. Bogus, S. Schneider, An exploratory study of the relationship between construction workforce physical strain and task level productivity, Constr. Manag. Econ. 32 (6) (2014) 548–564, http://dx.doi.org/10.1080/01446193.2013.831463.