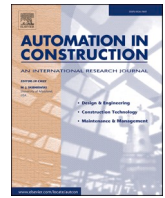




Contents lists available at ScienceDirect

## Automation in Construction

journal homepage: [www.elsevier.com/locate/autcon](http://www.elsevier.com/locate/autcon)

# Activity recognition of construction equipment using fractional random forest

Armin Kassemi Langroodi, Faridaddin Vahdatikhaki<sup>\*</sup>, Andre Doree

Department of Construction Management and Engineering, University of Twente, Horsttoren Z-204, Drienerlolaan 5, 7522 NB Enschede, the Netherlands

## ARTICLE INFO

### Keywords:

Activity recognition  
Machine learning  
Construction simulation  
Fractional feature augmentation  
Random forest

## ABSTRACT

The monitoring and tracking of construction equipment, e.g., excavators, is of great interest to improve the productivity, safety, and sustainability of construction projects. In recent years, digital technologies are leveraged to develop monitoring systems for construction equipment. These systems are commonly used to detect and/or track different pieces of equipment. However, the recent research work has indicated that the performance of the equipment monitoring system improves when they are able to also recognize/track the activities of the equipment (e.g., digging, compacting, etc.). Nevertheless, the current direction of research on equipment activity recognition is gravitating towards the use of deep learning methods. While very promising, the performance of deep learning methods is predicated on the comprehensiveness of the dataset used for training the model. Given the wide variations of construction equipment, in size and shape, the development of a comprehensive dataset can be challenging. This research hypothesizes that through the use of a robust feature augmentation method, shallow models, such as Random Forest, can yield a comparable performance without requiring a large and comprehensive dataset. Therefore, this research proposes a novel machine learning method based on the integration of Random Forest classifier with the fractional calculus-based feature augmentation technique to develop an accurate activity recognition model using a limited dataset. This method is implemented and applied to three case studies. In the first case study, the operations of two different models of excavators (one small-size and one medium-size) were tracked. By using the data from one excavator for the training and the data from the other one for testing, the impact of equipment size and operators' skill level on the performance of the proposed method is investigated. In the second case study, the data from an actual excavator was used to predict the activity of a scaled remotely controlled excavator. In the last case study, the proposed method was applied for rollers (as an example of non-articulating equipment). It is shown that the fractional feature augmentation method can have a positive impact on the performance of all machine learning methods studied in this research (i.e., Neural Network and Support Vector Machine). It is also shown that the proposed Fractional Random Forest method is able to provide comparable results to deep learning methods using considerably smaller training dataset.

## 1. Introduction

The construction industry has a high workplace incident rate compared to other industries [3,4]. The role of equipment in construction accidents and injuries is significant [1,2,5]. This is mainly due to the sheer size of construction equipment, the high congestion level of construction sites, and the frequent and unstructured interaction between workers and equipment [6,7]. Additionally, given the leading role and criticality of equipment operations, the productivity of construction operations heavily depends on the efficiency of the equipment work

[8–14]. It is also shown that construction equipment is responsible for a large volume of CO<sub>2</sub> emissions and therefore has detrimental environmental impacts [15]. Construction equipment, thus, plays an important role in the development of strategies to improve the sustainability of construction projects [16]. This accentuates the significance of enhancing the use of construction equipment to boost productivity, safety, and sustainability of construction projects [4,17].

The monitoring of operations of the construction heavy equipment is of great interest and importance because the data collected from these operations can, among others, be used to (a) provide feedback to

<sup>\*</sup> Corresponding author.

E-mail addresses: [a.kassemilangroodi@utwente.nl](mailto:a.kassemilangroodi@utwente.nl) (A.K. Langroodi), [f.vahdatikhaki@utwente.nl](mailto:f.vahdatikhaki@utwente.nl) (F. Vahdatikhaki), [a.g.doree@utwente.nl](mailto:a.g.doree@utwente.nl) (A. Doree).

<https://doi.org/10.1016/j.autcon.2020.103465>

Received 12 April 2020; Received in revised form 27 October 2020; Accepted 30 October 2020

Available online 3 December 2020

0926-5805/© 2020 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

operators and managers about safety and productivity [18], (b) gain better insight into the planning and simulation of construction operation [19], (c) develop materials for training [20], and (d) support innovation in equipment manufacturing.

Conventionally, the monitoring of construction equipment is done manually, which is a time-consuming and labor-intensive process [21,22]. In recent years, the construction industry started to exploit disruptive digital and sensor technologies to automate the monitoring of construction equipment. This includes efforts to develop and use a wide range of technologies, e.g., Global Position System (GPS), Ultra-Wideband (UWB), Radio Frequency Identification (RFID), and Vision-based solutions [7,18,19,23–39].

The monitoring of equipment can be done at different levels: At the most basic level, monitoring systems intend to detect different pieces of equipment [23,40], dangerous proximities [32,41,42] or entrance to/exit from predefined zones [37,43]. At this level, monitoring systems are indifferent to continuous behavior of specific equipment and only perform static analysis. At the next level, the monitoring systems are not only detecting equipment/events but also track equipment movement across the construction site [7,28,44–46]. Tracking-based systems require a Real-time Location System (RTLS), e.g., GPS or UWB. Other types of sensors such as proximity sensors can also be integrated into the system to combine the element of tracking and detection. While for the majority of the construction equipment (e.g., rollers, pavers, and graders) it is sufficient to track the equipment as a rigid body (i.e., a monolithic object), articulated equipment (e.g., excavators, cranes, and mobile cranes) can be tracked both as rigid bodies (i.e., using a single sensor) [18,46] and articulated bodies (i.e., using a sensor array to capture the motions along different Degrees of Freedom (DOFs) of the equipment) [7,34]. It is shown that pose estimation (i.e., tracking all DOFs of the equipment) is more effective for productivity, safety, and sustainability monitoring of articulated equipment [7,24,34,47–50]. While pose estimation is commonly referred to as the 3D tracking of articulated equipment, for the sake of brevity even 2D tracking of equipment as a rigid body will be referred to as pose estimation in the remainder of the paper.

It is demonstrated that automated equipment monitoring systems have great potentials for improving the productivity, safety, and sustainability of construction operations. However, the previous studies have demonstrated that equipment monitoring systems can be significantly improved if the tracking data are contextualized to detect and distinguish between different activities of the equipment, e.g., digging and swinging [18,19,51]. This is mainly because a great number of safety, productivity and sustainability-related measures depend on activities that are being performed by the equipment. For instance, the identification of utility strike dangers by an excavator depends not only on the location of the excavator and utilities but also whether or not the excavator is performing a digging task at that location. Additionally, the data about the activity of the equipment can be used to analyze construction operations at a higher level of abstraction and for managerial and strategic decision making. Such data can be used to estimate the progress of projects [36], update predictive (simulation) models about the productivity [19,31], perform predictive safety assessment [6], and develop data-driven simulation models [51,52].

Conventionally, the translation of the location/pose data into activity information is done either manually or using heuristics, which are very case-specific. This process is also labor- and cost-intensive [40,53]. In recent years, there has been an upsurge in the development and application of automated activity recognition methods [19,21,51–57]. These methods, mostly, use Machine Learning (ML) methods to analyze site images or sensory data collected from different equipment to classify different activities. In supervised ML methods, an indexed dataset is used to train the fittest classifier that can explain the variation in the dataset. The trained classifier is, then, applied to new testing datasets to evaluate the accuracy of the method. The existing ML-based equipment activity recognition support both homogenous single datasets, e.g., only

images [21], and heterogeneous datasets, e.g., the fusion of GPS and Inertial Measurement Units data [51].

Over the past few years, the ML-based equipment activity recognition methods have matured and developed considerably. But, to improve the performance of these methods (i.e., in terms of generalizability and also accuracy), the general research trend is noticeably gravitating towards the use of deep learning algorithms (e.g., Convolutional Neural Network and Long-Short-Term Memory) [53–56]. Generalizability refers to the extent to which the model can accurately predict the cases that are not part of the training dataset) As shown in these research work, these methods are successful in predicting the activities of the construction equipment with high accuracy. However, the main limitation of deep-learning methods is their dependency on large and comprehensive training datasets [58–61]. Given the diversity of construction equipment models, this would mean that a large set of data needs to be collected from different models of each equipment to ensure the accuracy of the activity recognition models. A recent study has clearly pointed out the limitations of deep learning methods for activity recognition and highlighted that the reliance on a large dataset is a major deterrent for the adoption of these methods in practice [62].

Given the above-mentioned challenge, this research is motivated by the ambition to explore an alternative approach in order to reduce the dependency of equipment activity recognition models on large and comprehensive training datasets without compromising the accuracy. Therefore, this research builds on recent findings about the effect of feature augmentation methods on improving the performance of machine learning methods [63–65] and hypothesizes that through the use of a robust feature augmentation method, a shallow learning model can yield a comparable performance to deep learning methods without requiring a large and comprehensive dataset. Feature augmentation refers to a set of techniques that try to expand the feature domain of a dataset by adding synthetic features to the original features. In this research, a fractional [66] Random Forest (RF) classifier is proposed. The choice of fractional feature augmentation as a feature augmentation method, and RF as the shallow learning method is justified by the following:

- (a) RF is chosen because of its inherent ability to better explain the complex pattern in the data by developing a large number of predictors and using a voting mechanism [67]. Being an ensemble classifier, RF is a strong classifier that is shown to be suitable for multi-dimensional and complex problems [68]. RF is also found to be a strong classifier for imbalanced datasets (i.e., the sample sizes of different classes are not the same) [69]. This pertains very well to the case of equipment activity recognition because the distribution of length of different activities for construction equipment is largely disproportional. Also, the literature provides ample evidence that RF classifiers offer high generalizability [70–72]. Previous research has indicated that RF outperforms other shallow classifiers for predicting construction equipment activities [73,74]. Most notably, Lee et al. [74] compared 17 different machine learning methods, including several deep learning methods, for construction activity recognition and observed that RF has the best performance.
- (b) Fractional calculus allows performing fractional order derivative or integration on data series. In the context of ML, fractional feature augmentation allows considering a wider range of features (e.g., input parameters) by performing the fractional derivative and/or integral on original features. For instance, if one of the features for training a model for estimating the activity of an excavator is the angular speed of the boom, the fractional feature augmentation can generate a range of features that are neither velocity nor acceleration, but can account for both. For instance, 0.5 derivatives of velocity offer a feature that contains elements of both velocity and acceleration. Similarly, fractional integration of the angular speed between 0 and 1 considers a

**Table 1**  
Overview of activity-recognition methods for construction equipment.

Category	Author	Target equipment	Input data source	Activity recognition approach	Number of activities	Accuracy	
Vision-based	Zou and Kim [78]	Excavator	Cameras installed on the site	Color-based tracing	2	96.70%	
	Gong et al. [79]	Excavator		Bayesian network	3	79%	
	Golparvar-Fard et al. [21]	Excavator and Truck		Support Vector Machine	4 (excavator)	76.0%	
					3 (excavator)	86.33%	
	Azar et al. [80]	Excavator and truck		Heuristics and Support Vector Machine	3 (truck)	98.33%	
	Kim et al. [90]	Excavator and truck		Heuristics	Identifies loading cycle	95%	
Audio-based	Kim & Chi [53]	Excavator	microphone array	Long short-term memory networks	4	91.27% precision	
	Cai et al. [56]	Varied		Long short-term memory networks	3	93.8% precision	
	Chen et al. [82]	Excavator	microphone array	3D residual neural network	3	95%	
	Cheng et al. [84]	Excavator, wheel loader, dozer & dumper		Support Vector Machine	2	93.8%	
	Cheng et al. [57]	Min excavator, backhoe, loader, dozer	microphone array	Support Vector Machine	2	17% ~ 98%	
	Lee et al. [74]	Grader, Truck, Excavator, Bulldozer	Videos and audio files	Support Vector Machine	2	>90%	
	Cao et al. [85]	Various excavation equipment	Microphone sensors	17 different deep and shallow learning methods	N/A	Up to 93%	
	Cao et al. [86]	Various excavation equipment	Microphone sensors	Support Vector Machine and extreme learning machine	N/A	Up to 99%	
	Motion-based	Vahdatikhaki and Hammad [19]	Excavator	UWB	Support Vector Machine and Extreme Learning Machine	N/A	Up to 88%
		Akhavian & Behzadan [51]	Loader	GPS, Gyroscope, Accelerometer	Heuristics	4	N/A
Axelsson and Daniel [88]		Truck	GPS, Gyroscope, Accelerometer	Neural Network	5	86.1%	
					4	81.3%	
Kim et al. [73]		Excavator	IMU	Neural Network, Random Forest, Naïve Bayes	5	98.59%	
Rashid and Louis [55]		Excavator and loader	IMU	Dynamic Time Warping	4	54% ~ 95%	
Slaton et al. [54]		Excavator and compactor	Accelerometer	Long short-term memory networks	9	87.3%	
Sherafat et al. [89]	Excavator	Microphone, IMU	Convolutional Neural Network	7	85% ~ 99%		
			Support Vector Machine	4	77.1%		
					92%		

range of features between pure velocity to pure displacement. Therefore, fractional feature augmentation allows increasing the number of features without requiring additional data collection. This can help better capture the complex pattern between equipment activities and its kinematics and thus improve the generalizability of the model. The integration of fractional feature augmentation with machine learning methods has been shown effective in other domains [75–77]. However, to the best of the authors' knowledge, the integration of ML, including RF, and fractional feature augmentation has never been studied for construction-related problems.

Based on the above premises, the main objective of this research is to develop a Fractional Random Forest (FRF) activity recognition method that can reduce the need for a comprehensive training dataset without compromising the accuracy and generalizability of the model. By addressing this issue, the proposed research contributes to making automated activity recognition of construction equipment easier to develop (by requiring smaller training dataset) and thus more accessible to practitioners and analysts.

The remainder of this paper is structured as follows: First, the relevant literature is reviewed and presented. Next, the proposed method is explained in detail. Then, the implementation of the proposed method and three case studies are presented. Finally, the contributions, conclusions, and limitations and future work are presented.

## 2. Related work

Given the benefits of activity-related data for safety, productivity and sustainability analysis of construction operations, many researchers have started to investigate this topic in the past few years. Table 1 summarizes these research works.

Three main categories of equipment activity recognition methods can be identified, namely vision-based, audio-based, and motion-based methods. The major difference between the three categories is the type of data used to estimate the activities of the equipment. While vision-based methods use images generated from the video recording of construction sites, audio-based methods, which have gained momentum in the past few years, use the sound captured from construction sites to infer the activity of different pieces of equipment based on their distinctive sound patterns. Motion-based (or Kinematic-based) methods utilize a wide range of sensors (e.g., IMUs, GPSs, accelerometers, etc.) to capture the motions of the equipment.

### 2.1. Vision-based methods

Zou and Kim [78] developed a color-based tracing method to identify the idle time of excavators. While the reported accuracy is high, this study only considered two activities, namely idle and working. For a wide range of safety and productivity-related analyses, this level of detail is too coarse to develop an accurate estimation of cycle times and proximity. Gong et al. [79] developed a Bayesian Network-based solution to distinguish between 3 different activities of an excavator and obtained an accuracy of 79%. Golparvar-Fard et al. [21] developed an activity recognition model using Support Vector Machine (SVM). Different accuracy levels were achieved for excavator and truck, namely 76% and 98.33%. The main limitation of this work is that it requires a priori knowledge about the duration and starting point of each activity. This would significantly overshadow the applicability of this method as part of an automated monitoring system given that this information needs to be gathered manually or through another method. Azar et al. [80] combined heuristics and SVM to identify the loading cycle of excavators. While very useful for productivity measurements, the method fails to provide detailed information about the time-stamped activity of

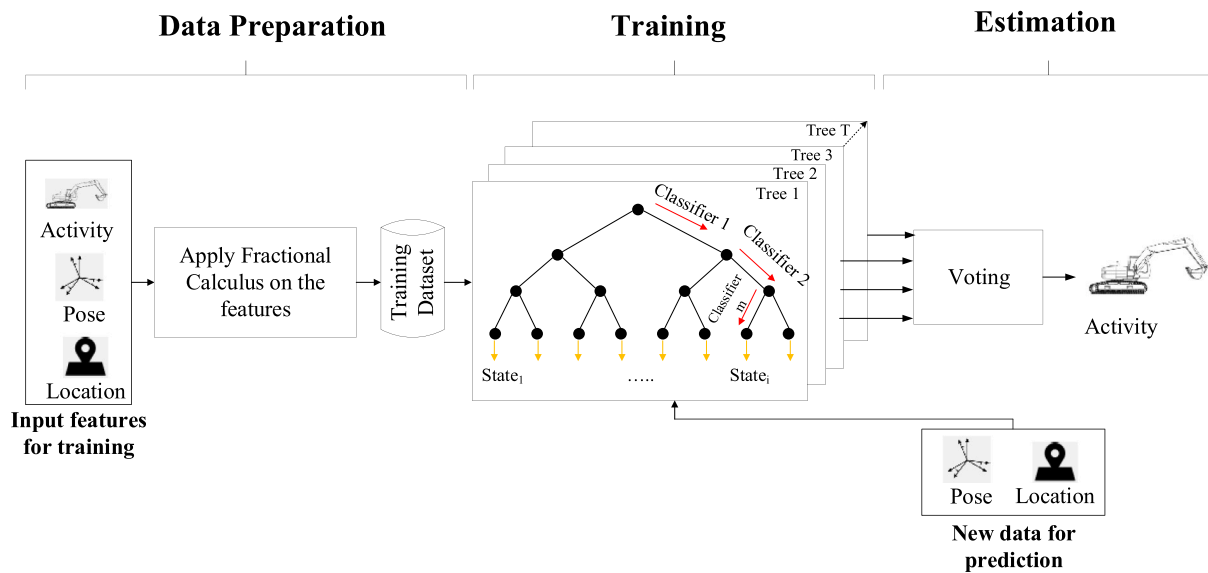


Fig. 1. Schematic representation of the proposed method.

the equipment. Kim et al. [81] also developed a vision-based heuristic method to distinguish between 3 activities of truck and excavators. However, heuristics-based methods generally fail to achieve high generalizability given the complexity of construction sites and operations and given the fact that similar operations can happen differently under different conditions. Kim and Chi [53], Cai et al. [56] and Chen et al. [82] have proposed deep learning methods based on the use of Long Short-term Memory Networks and Convolutional Neural Network and achieved an accuracy of above 90%.

While vision-based methods are gaining more tractions in recent years and proved to be very promising, several inherent limitations are still unaddressed. First, the accuracy of the developed model depends heavily on the comprehensiveness of the training dataset. Without a dataset that covers a wide range of (1) equipment type and shape, (2) angles of view and operational patterns, the accuracy of vision-based models can be lower when applied to new cases [56,82]. Second, vision-based methods are very sensitive to the weather condition, lighting, and cameras' field of view [83]. This would impact the practicality of the proposed methods in an uncontrolled environment.

## 2.2. Audio-based methods

Audio-based methods are the next category of construction equipment activity recognition solutions. This is a rather new approach towards identifying equipment activities. In general, these methods apply signal processing and machine learning techniques to detect specific audio pattern each activity or each type of equipment is producing. Cheng et al. [57,84] used the audio signal from equipment and applied SVM to recognize the activity of equipment based on the sound pattern. While high accuracy was obtained, the method only identifies activities at low levels of detail (i.e., splitting equipment activities to only major and minor). The contribution of this level of detail to improving safety, productivity and sustainability can be minimal. Others [74,85–87] have used audio-based methods to distinguish between different pieces of excavation equipment. Despite their good performance, these methods are not purported to recognize different activities of equipment, which in turn limits their applicability for a wide range of monitoring purposes. In general, four main limitations are observed with these methods: (1) audio-based solutions are yet not able to identify equipment activities at the same level of detail as vision and motion-based methods, (2) at the current level of development, these methods are only limited to detect one equipment/activity at the time. This means that at this moment

these methods are not able to work on congested construction sites where several pieces of equipment work simultaneously [83], (3) these methods are sensitive to the relative positioning of the data collection device to the source of the sound. Developing a robust multi-equipment classifier seems to require a large set of training data, (4) the sound pattern can be sensitive to geographical and geological features of the site, which makes the development of robust classifiers difficult.

## 2.3. Motion-based methods

The last category of activity recognition methods is motion-based solutions. In this category, a wide range of tracking sensors are used to collect the input data for activity recognition. Vahdatikhaki and Hammad [19] have proposed a heuristics-based method that leverages location data collected by UWB to determine the activities of an excavator. The proposed method is robust because of several layers of filtering mechanisms that are applied to ensure the output of activity recognition corresponds to the logical sequence of excavator operations. Nevertheless, the heuristic rules are sensitive to the context and generally perform with a high degree of variability in the accuracy. Akhavian and Behzadan [51] used GPS, gyroscope, and accelerometer embedded in smartphones to track the operations of loaders. They experimented with different types of classifiers for activity recognition and reported Neural Network (NN) as the best classifier. Kim et al. [73] used dynamic time-warping technique to improve the accuracy of activity recognition. The main limitation of the last two work is that the proposed classifiers are very sensitive to the input features and the initial configuration of the dataset. The proposed methods are not tested for cases outside the scope of the dataset. Axelsson and Daniel [88] compared different classifiers and observed that NN and RF are the best performing machine learning methods for activity recognition. This is one of the very few research works where the generalizability of models is tested for cases outside the scope of the training dataset. Nevertheless, because only the basic features coming from the sensory data are used, a major drop in the level of accuracy, i.e., from 95% to 54%, is observed when the model is applied to a new case. Sherafat et al. [89] proposed a hybrid kinematic-acoustic system the utilizes data from IMU and microphones concurrently. Slaton et al. [54] compared two deep learning methods for activity recognition of compactors and excavators. In this study, only accelerometer data are used. Compared to other studies, they reported lower accuracy, i.e., 77.6%, and pointed out the importance of having a more comprehensive dataset to improve the accuracy further. Rashid

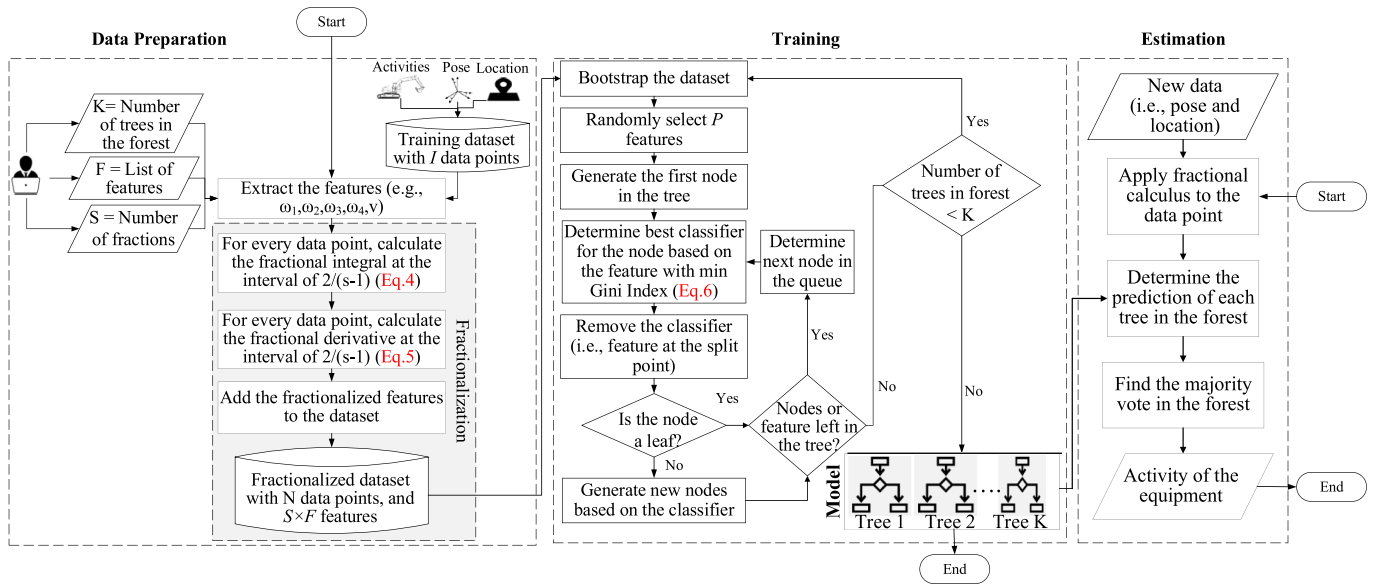


Fig. 2. Flowchart of the proposed Fractional Random Forest (FRF) method for the activity recognition of equipment.

and Louis [55] identified the limitations of the previous work in terms of sensitivity to the limited patterns present in the dataset and applied a series of data augmentation methods to improve the generalizability of the activity recognition methods. Significant improvement in the accuracy of the model is reported in this research. However, the data augmentation method is limited to expanding the scope of the existing features in the dataset, e.g., by stretching and compacting the time series representing individual activities such as swinging. In doing so, this research does not consider features other than those directly measured by sensors for the augmentation of the dataset.

Based on the above review, the research trends is ostensibly gravitating towards the application of deep learning methods using large training datasets. As stated in the introduction, this can pose applicability and generalizability issues given the wide variety of equipment models. Addressing the above-mentioned gap, this research argues that (1) the use of ensemble learning methods, e.g., Random Forest, can reduce the sensitivity of the model to the dataset because these methods tend to consider many different combinations of input features and use majority voting among classifiers. In this way, ensemble methods can better capture the intricate relationship between input features and develop more generalizable patterns, (2) the use of fractional feature augmentation allows (a) expanding the number of features without requiring additional data collection, (b) generating complex features that can potentially better explain the pattern in the data.

### 3. Proposed method

Fig. 1 presents the overview of the proposed activity recognition method. The proposed method consists of 3 main phases, namely data preparation, training, and estimation, which function based on the input data coming from IMUs and GPS attached to the equipment. In the data preparation stage, time-series equipment 2D location and/or 3D pose data labeled with activities of the equipment are used for the training of

the RF model. Next, fractional feature augmentation is applied to the location/pose data and new fractional features are generated and added to the training dataset. Once the training dataset is prepared, the bootstrapping technique is applied to generate a series of subset training datasets, as will be explained in Section 3.2. Then, an FRF classifier is trained based on the bootstrapped training dataset. As will be explained in Section 3.2, FRF generates several decision tree classifiers, each of which uses a randomly chosen set of features for the classification. At the end of the training, the model (i.e., the forest) contains  $K$  trees. In the estimation phase, a new data point (including location and/or pose data) is fed to the model and voting mechanism is triggered. During the voting, each tree presents its estimated activity of the equipment (at a given data point) based on how that tree classifies the data. Consequently, the majority voting of trees will be used to determine the final estimated activity.

The remainder of this section presents a more detailed explanation of the proposed method step-by-step, as shown in Fig. 2.

#### 3.1. Data preparation

The first step in the data preparation is to set the basic parameters required for FRF. These parameters include the number of trees ( $K$ ), a list of features ( $F$ ), and the number of fractions ( $S$ ). Number of trees would specify the number of individual decision trees, i.e., classifiers, that will be trained to form the forest. The list of features includes all the features that will be used as input for the model. It is important to mention that since fractional feature augmentation applies fractional integral and derivative on the input features, and to make sure that products of these operations have physical meaning, it is recommended to limit the nature of features to only velocity features. This is the case because applying integral and derivative on velocity yields displacement and acceleration, respectively. As shown in Fig. 3, any fractional integral and derivative between  $-1$  and  $1$  on velocity generate a complex feature that can

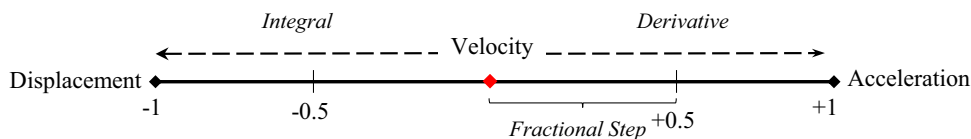


Fig. 3. Schematic representation of fractional integral and derivative on velocity.



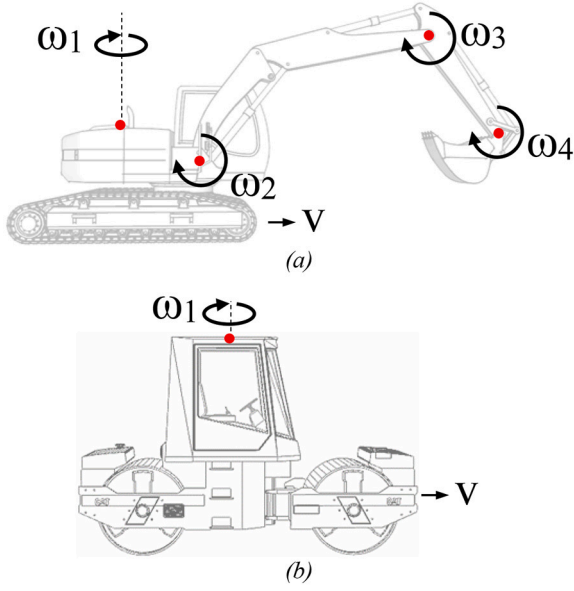


Fig. 4. Input features for (a) an excavator as an example of articulate equipment, and (b) a roller as an example rigid body equipment.

represent velocity-displacement and velocity-acceleration, respectively. The number of fractions ( $S$ ) represents the number of fractional steps between the range of  $[-1,1]$ . For instance,  $S = 5$  encompasses a fractional range of  $[-1, -0.5, 0, 0.5, 1]$ . The mathematical calculation of fractional feature augmentation will be presented in Section 3.1.1.

Once the basic parameters are set by the user, the required features need to be extracted from the dataset. Based on the explanation presented in the previous paragraph, the input features for examples of (1) an articulated equipment, i.e., an excavator, and (2) a rigid body equipment, i.e., roller, are shown in Fig. 4. In this example, each DOF of the excavator and roller is represented by a velocity vector. These include the magnitude of velocities of the superstructure ( $v$ ), the rotational velocity of the superstructure ( $\omega_1$ ), and the angular velocities of the arm system ( $\omega_2, \omega_3, \omega_4$ ) (only for the excavator). For a piece of equipment like a roller, rotational velocity helps better track the equipment heading and thus distinguish activities where the direction of movement is important (compacting forward/backward).

### 3.1.1. Fractional calculus

Once the required features are prepared, the next step in the proposed method is to apply fraction calculus on the features. In this research, the Rieman-Liouville method is used for fractional calculus

[90]. The general form of fractional integral and derivative according to Rieman-Liouville is presented in Eqs. (1) and (2), respectively [90].

$${}_a D_T^{-n} f(T) = \frac{1}{\Gamma(n)} \int_a^T (T-t)^{n-1} f(t) dt \quad (1)$$

$${}_a D_T^n f(T) = \frac{d^m}{dt^m} {}_a D_T^{-(m-n)} f(T) \quad (2)$$

where

${}_a D_T^n f(t)$ :  $n$ -fold integral(-)/derivative(+) of function  $f(t)$  in the range  $[a, a + t]$ .

$a$ : starting point of integral/derivative.

$T$ : the length of the range at which integral/derivative is being made.

$n$ : integral/derivative order in real number domain.

$\Gamma(n)$ : the Gamma Function

$(m) \lceil n \rceil$

In Eq. (1), the value of  $\Gamma(n)$  is determined as shown in Eq. (3) [90].

$$\Gamma(n) = \int_0^\infty x^{n-1} e^{-x} dx \quad (3)$$

The above-presented equations are presented for continuous space. However, in the context of the activity recognition, the integral and derivative space are discrete. In this space, the value of  $f(t)$  corresponds to the value of each feature shown in Fig. 4 at the time instance  $t$ . To better demonstrate the discrete form of Eqs. (2) and (3), the typical data array structure of features list for the example of the excavator can be used, as shown in Fig. 5. To be able to apply Eq. (2) on a discrete space, the first step is to determine the range at which the integral will be applied. In the example, let's assume the data is collected at the rate of  $\frac{1}{\tau}$  Hz. In this case, the integral will be applied over the range of  $T$ , where  $T > \tau$ . Given this assumption, Eq. (1) will translate to Eq. (4).

$${}_a D_T^{-n} f(T) = \frac{\tau}{\Gamma(n)} \sum_{t=a}^T (T-t)^{n-1} f(t) \quad (4)$$

In the discrete space and for derivative order  $0 < n \leq 1$  (i.e.,  $m = 1$ ), Eq. (2) can be reformulated as shown in Eq. (5).

$$\begin{aligned} D_{kT}^n f(kT) &= \left( \frac{{}_k D_{(k+1)T}^{-(1-n)} f((k+1)T) - {}_{k-1} D_{kT}^{-(1-n)} f(kT)}{T} \right) \\ &= \frac{\tau}{T \times \Gamma(n)} \left[ \sum_{t=kT}^{(k+1)T} ((k+1)T-t)^n f(t) - \sum_{t=(k-1)T}^k (T-t)^n f(t) \right] \end{aligned} \quad (5)$$

where  $k$  is an integer number. For  $>1$ , Eq. (5) should be repeated  $m$  times on successive time slots.

$T$												
$\tau$												
$V_1$	$V_2$	...	$V_T$	$V_{T+1}$	$V_{T+2}$	...	$V_{2T}$	.....	$V_{jT+1}$	$V_{jT+2}$	...	$V_{iT}$
$\omega_1^1$	$\omega_2^1$	...	$\omega_n^1$	$\omega_{T+1}^1$	$\omega_{T+2}^1$	...	$\omega_{2T}^1$	.....	$\omega_{jT+1}^1$	$\omega_{jT+2}^1$	...	$\omega_{iT}^1$
$\omega_1^2$	$\omega_2^2$	...	$\omega_n^2$	$\omega_{T+1}^2$	$\omega_{T+2}^2$	...	$\omega_{2T}^2$	.....	$\omega_{jT+1}^2$	$\omega_{jT+2}^2$	...	$\omega_{iT}^2$
$\omega_1^3$	$\omega_2^3$	...	$\omega_n^3$	$\omega_{T+1}^3$	$\omega_{T+2}^3$	...	$\omega_{2T}^3$	.....	$\omega_{jT+1}^3$	$\omega_{jT+2}^3$	...	$\omega_{iT}^3$
$\omega_1^4$	$\omega_2^4$	...	$\omega_n^4$	$\omega_{T+1}^4$	$\omega_{T+2}^4$	...	$\omega_{2T}^4$	.....	$\omega_{jT+1}^4$	$\omega_{jT+2}^4$	...	$\omega_{iT}^4$

Fig. 5. Data array structure of the original features list for the example shown in Figure.

**Table 2**  
Example of fractional calculus applied to the rotational velocity of an excavator.

Time (s)	$\omega^1$ (deg/s)		Time (s)	$D^{-1} \omega^1$	$D^{-0.5} \omega^1$	$\omega^1$	$D^{0.5} \omega^1$	$D^1 \omega^1$
0	3.26	→	1	6.73	5.80	6.34	0.43	3.08
0.5	7.88							
1	6.34							
1.5	7.11							
2	3.26							

Features	Time			
	$T$	$2T$	...	$iT$
$D^{-1}V_T$	$D^{-1}V_{2T}$	...	$D^{-1}V_{iT}$	
$D^{-0.5}V_T$	$D^{-0.5}V_{2T}$	...	$D^{-0.5}V_{iT}$	
$V_T$	$V_{2T}$	...	$V_{iT}$	
$D^{0.5}V_T$	$D^{0.5}V_{2T}$	...	$D^{0.5}V_{iT}$	
$D^1V_T$	$D^1V_{2T}$	...	$D^1V_{iT}$	
$D^{-1}\omega^1_T$	$D^{-1}\omega^1_{2T}$	...	$D^{-1}\omega^1_{iT}$	
$D^{-0.5}\omega^1_T$	$D^{-0.5}\omega^1_{2T}$	...	$D^{-0.5}\omega^1_{iT}$	
$\omega^1_T$	$\omega^1_{2T}$	...	$\omega^1_{iT}$	
$D^{0.5}\omega^1_T$	$D^{0.5}\omega^1_{2T}$	...	$D^{0.5}\omega^1_{iT}$	
$D^1\omega^1_T$	$D^1\omega^1_{2T}$	...	$D^1\omega^1_{iT}$	
⋮	⋮	⋮	⋮	
$D^{-1}\omega^4_T$	$D^{-1}\omega^4_{2T}$	...	$D^{-1}\omega^4_{iT}$	
$D^{-0.5}\omega^4_T$	$D^{-0.5}\omega^4_{2T}$	...	$D^{-0.5}\omega^4_{iT}$	
$\omega^4_T$	$\omega^4_{2T}$	...	$\omega^4_{iT}$	
$D^{0.5}\omega^4_T$	$D^{0.5}\omega^4_{2T}$	...	$D^{0.5}\omega^4_{iT}$	
$D^1\omega^4_T$	$D^1\omega^4_{2T}$	...	$D^1\omega^4_{iT}$	

**Fig. 6.** Data array structure of the fractional features list for the example shown in Fig. 5.

As shown in Eqs. (4) and (5), the fractional integral and derivative depend on  $S$  (i.e., the number of fractions). For instance, at  $S = 5$  and for the data point at time  $T$ , the following fractional features need to be calculated [ ${}_aD_T^{-1}f(T)$ ,  ${}_aD_T^{-0.5}f(T)$ ,  $f(T)$ ,  ${}_aD_T^{0.5}f(T)$ ,  $f(T)$ ,  ${}_aD_T^{0.5}f(T)$ ,  ${}_aD_T^1f(T)$ ]. To better put these equations in perspective, the rotational velocity of an excavator's superstructure ( $\omega_1$ ) over the period of 1 s with the update rate of 2 Hz can be used as an example, as shown in Table 2. Let's assume  $S = 5$  and integral is calculated over the range of  $T = 1$  s. The fractional integral and derivative of  $\omega_1$  are calculated based on Eqs. (4) and (5) as shown in Table 2.

There are three important points to highlight here: (1) after applying the fractional feature augmentation, the frequency of data is discounted from  $\frac{1}{T}$  Hz to  $\frac{1}{iT}$  Hz. This is because for the fractional feature to be calculated for the data at point  $kT$ , all the data in the range of  $((k-1)T, (k+1)T)$  are used. (2) after discounting the data frequency, the activity of the new data point (i.e., the label) is determined based on the dominant activity in the original period. (3) After applying the fractional feature augmentation for  $S$  fractions, the list of features will grow by  $S$  times. For instance, if the input feature list includes 5 features (as shown in Fig. 5), and if the number of fractions is set at 5, the final feature size will be 25. In this case, the ultimate list of features that will be used for the training of the RF model is in the form of the matrix presented in Fig. 6. The matrix for a rigid body equipment, e.g., roller, will be simpler in the sense that it would only have the fractional values of traversal velocity ( $V$ ) and rotational velocity ( $\omega_1$ ).

### 3.2. Training

The next phase of the proposed method is the training of the FRF model. The first step is to bootstrap the dataset. Bootstrapping [91] is a technique in which multiple classifiers are developed based on randomly sampled (with replacement) subsets of the original training dataset. The bootstrapped dataset has the same size as the original dataset (i.e.,  $i$  datapoints), however, because of the sampling with replacement, each bootstrap dataset may contain several repeated data points. Bootstrapping helps reduce the output error that might be caused by specific data points in the original training dataset [92].

In the next step, the Random Subspace Method (RSM) [93] is applied. In this method, a similar approach as bootstrapping is applied to the list of features, but without replacement. Therefore, this step is also known as feature bagging. More specifically, a random subset of the original features will be selected to train each tree in the forest. Therefore,  $P$  features will be randomly selected from  $S \times F$  features available in the fractionalized list of features, where  $P < S \times F$ . For instance, in the above example, a set of 5 features can be selected from the list of 25 fractionalized features to train each tree in the dataset.

To put the previous two steps into perspective, bootstrapping and feature bagging can be understood as a sampling of columns (with replacement) and rows (without replacement) of the matrix shown in Fig. 6, respectively.

Although from this point onward similar steps as of a basic decision tree algorithm are followed [94], a brief explanation is provided for completeness. To generate a tree, first all features are examined to find the feature that can best classify datasets into two nodes, as shown in Fig. 2. To examine the potential of each feature as a node, first, all the possible split points for that feature are considered. Then, the Gini index is used to determine the degree of impurity of the classification at that split point [94], as shown in Eq. (6). Once all the split points are explored, the point with the minimum Gini index is considered as the classifier based on that specific feature. Next, the same procedure is repeated until the potential classifier by each feature is determined. Finally, the feature with the minimum Gini index is selected as the node.

$$G_q = 1 - \sum_{i=1}^c P_i \tag{6}$$

where

$G_q$ : Gini index at node  $q$ .

$C$ : Number of classes (i.e., number of equipment activities).

$P_i$ : Probability of a data point being classified to class  $i$ .

To determine the end of each branch in the tree, the potential of each node (other than the root node) for being a leaf is first controlled before it is grown. A node is determined as a leaf when the Gini index of the node is smaller than any possible subsequent classifier that can be considered at that node. If the node is not a leaf, then the node is branched out and two new nodes are generated. The above procedure is followed until the tree cannot be grown any further, i.e., there are no more unbranched nodes or no feature is left. The procedure of growing a tree is repeated for  $K$  trees in the forest. In this way, at the end of the training phase, there are  $K$  independent decision trees that are trained



Fig. 7. (a) IMU for pose estimation and (b) GNSS for localization.

using unique bootstrapped datasets and bagged features, as shown in Fig. 3.

### 3.3. Estimation

The last phase of the method is to assess the activity of a new input data point, which comprises the pose of the equipment. When new input data is provided to the model, first the fractional calculus is applied to the data points to generate the same set of features as in the trained model. It should be highlighted that since fractionalization requires a range of data point, FRF is applied over the discounted update rate of  $\frac{1}{7}$  Hz. Once all the fractional features are generated, FRF asks each tree to determine the activity of the equipment corresponding to the specific datapoints. This procedure results in  $K$  estimates about the activity of the equipment corresponding to that time frame. Finally, the majority vote of trees is used to determine the verdict of FRF about the activity of the equipment.

## 4. Implementation and case studies

To demonstrate the feasibility and effectiveness of the proposed method, it is implemented and tested on three different case studies. The method is implemented in the Python environment on a workstation laptop with Core i7 2.80 GHz CPU and 16 GB RAM running on a Windows 10 operating system. As shown in Fig. 7, high-end industry-grade MTi 1-series Xsens IMUs were used in conjunction with the U-blox EVK-M8T GNSS for rotational and translation motion tracking, respectively. These sensors collect data at a rate of 100 Hz (the GNSS data is up-sampled at this frequency). This sensor network provides the same set of input data shown in Fig. 4, i.e., traversal velocity, angular velocities of the arm system and the rotational velocity of the superstructure.

During the installation of IMUs on the equipment, care should be taken to ensure that the internal coordinate system of the sensor is aligned with the local axes of the part it intends to track. This is important because IMUs are used to measure the relative velocity of each part along its local axes. If there is a misalignment between the IMU and the equipment part, the measurements will have an offset. This can result in significant errors when the excavator is working on topologically rough terrains. During the case studies, IMUs are placed in well-designed and aligned 3D printed casings that can keep the sensor parallel to the bottom plate of the casing. During the sensorization of the equipment, a spirit level is used to make sure the casing is aligned to the local axis of each component.

The collected data were preprocessed. The steps explained in the previous work of authors have been applied to remove outliers, synchronize the data and unify the time steps [19]. Also, Extended Kalman Filter was applied to smoothen GNSS and IMU data [95]. Moreover, the IMU sensors were equipped with Active Heading Stabilization software component to minimize the drifting error [96].

Given the main hypothesis of this research, i.e., FRF can serve as a generalizable and accurate activity recognition method without requiring a large training dataset, it was necessary to collect data for different types of equipment and operating conditions. To this end three case studies were considered. In the first case study, the method was applied on two different excavators (one small-size and one medium-size) on a construction field at a training school. This case study intends to demonstrate the robustness and generalizability of the proposed method. In the second case study, and to further substantiate the generalizability of the method for an extreme case, the model developed based on the actual excavator in the first case study was tested on a scaled remotely controlled excavator (scale 1/12). Finally, in the last case study, the method was tested on two rollers as examples of rigid body equipment. While the first two case studies were conducted in controlled environments, the last case study was conducted on an uncontrolled environment of an actual construction project.

In all case studies, the performance of the proposed FRF method is evaluated. However, to put this assessment in perspective, three baseline methods (i.e., Random Forest, Neural Network, and Support Vector Machine) are used to shed light on the following questions: (1) Which ML method performs the best?, and (2) To what extent can fractional feature augmentation improve the performance of each ML method?. Therefore, six different models are trained and tested: RF, FRF, NN, Fractional NN (FNN), SVM, and Fractional SVM (FSVM). In these case studies, Multi-layer Perceptron (MLP) NN is used. Through a set of iterative process, the best structure for the MLP mode is found when two hidden layers of 50 neurons are used. As for SVM, the Radial Basis Function (RBF) is used as the kernel. The shape of decision function is set as one-vs-one (OVO). During the fractionalization,  $T = 1$  s was used, as explained in Section 3.1.1. Also, the fractionalization step was set at 7. In the case of FRF, the number of trees was set at  $K = 1000$  and feature bagging fraction ( $P$ ) was set at 10.

In conformity with similar studies [55,82], three metrics are used for the assessment of the performance, namely accuracy, precision, and recall. Eqs. (7) to (9) show how these metrics are calculated.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (7)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (8)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (9)$$

### 4.1. First case study

#### 4.1.1. Data collection and preparation

In this case study, three different scenarios are developed to test the





(a)



(b)

Fig. 8. Equipment used in the case study.

Table 3  
The size of each dataset used in the first case study.

The size of each datasets (datapoints)			
Case CX80C with expert operator	Terex TC75 with expert operator	Case CX80C with novice operator	A mixed dataset
14,185	9379	6694	30,257

performance of the FRF method for cases where FRF is trained based on one dataset and then tested on (1) a different dataset representing the operation of a different piece of equipment with a different geometry, (2) a different dataset representing the operation of the same equipment but by an operator with a different skill level, and (3) the same dataset but on the portion that was excluded from the training.

Two different excavators (i.e., Terex TC75 and Case CX80C) are used for the data collection, as shown in Fig. 8. While Terex TC75 is considered as a small-size excavator, the customized Case CX80C used in this case study is categorized as the medium-size. Data is collected at an operator training school in the Netherlands, i.e., SOMA College [97]. This was considered an adequate testing environment because (1) it provided a controlled environment where specific scenarios can be simulated and (2) it provided access to expert and novice operators.

To cover the scenarios mentioned above, three sessions of data collection were organized. In the first session, an expert operator was

Table 4  
Mapping of datasets with different scenarios used in the case study.

Dataset scenario	Case CX80C with expert operator	Terex TC75 with expert operator	Case CX80C with novice operator	A mixed dataset
Varied Geometry	Training	Testing		
Varied Skill level	Training		Testing	
No variation				Training and testing



Fig. 9. VR interface used for data labeling.

asked to perform a simple digging operation using Case CX80C. In these operations, operators are asked to relocate to the digging point, excavate soil on an even ground and dump the material on a nearby stockpile. In the second session, the same expert operator was asked to operate Terex TC75 and perform the same operation but at a different nearby location. In the last session, a novice operator is requested to use Case CX80C to perform the same digging operation. For each scenario, about 5 min of data were collected. The sessions are intentionally kept short to test the hypothesis of the research in terms of good generalizability with small datasets. The sizes of the datasets, after filtering the data, are presented in Table 3. The recent studies that proposed deep learning methods for equipment activity recognition [54,55] used 125 K and 287 K datapoints for training the proposed models, respectively. Compared to these studies, the training dataset in this study is considerably smaller (i.e., only 11% of [54] and 5% of [55]).

These three sessions resulted in four different datasets, (1) operation of Case CX80C by an expert operator, (2) operation of Terex TC75 by an expert operator, (3) operation of Case CX80C by a novice operator, and (4) a dataset containing all the above operations. As stated above, these four datasets are used to investigate the performance of the FRF method on (1) varied geometry, (2) varied skill level, and (3) no variation (i.e., when a combined dataset is used for training and testing with 70:30 ratio). Table 4 shows the structure of these scenarios.

To annotate the datasets, a 3D data visualizer is developed to mimic the motions of the equipment in a Virtual Reality (VR) environment and then label the activity of each data point. The use of VR visualizer as a means for data labeling is preferred because it eliminates the synchronization error between the video recording and sensor data. Fig. 9 shows a snapshot of the VR visualizer used for data labeling. As shown in previous studies [51], the number of activities used in data labeling has an impact on the accuracy of the model. Therefore, two different sets of activities were considered. In the first set, 5 activities are considered, namely (1) Idle, (2) Relocating (i.e., the excavator moves on its tracks), (3) Swinging, (4) Digging, and (5) Filling (i.e., dumping the material on the truck). After the close observation of the data in the visualizer, it was noticed that the filling activity is very short and hardly distinguishable from the swinging activity (especially when the equipment is handled by an expert operator). Therefore, in the second set of activities, the filling was eliminated and all of its instances were replaced with swinging. Consequently, the second set comprises 4 activities (i.e., idle, relocating,

**Table 5**  
Accuracy, precision and recall of 5-activity models (best performances are highlighted in bold).

Model		Varied geometry			Varied skill			No variation		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall
5-activity models	RF	80.9%	68.2%	74.3%	84.1%	74.6%	74.2%	86.6%	80.0%	80.5%
	FRF	<b>86.2%</b>	<b>78.0%</b>	<b>77.0%</b>	<b>86.9%</b>	<b>82.9%</b>	70.6%	<b>90.4%</b>	<b>86.3%</b>	<b>84.1%</b>
	NN	77.9%	65.0%	71.4%	83.8%	72.1%	76.7%	82.3%	78.4%	78.0%
	FNN	82.1%	71.8%	75.8%	84.2%	72.9%	<b>78.9%</b>	84.5%	80.2%	79.3%
	SVM	73.4%	63.3%	69.0%	68.3%	66.2%	62.5%	83.1%	76.6%	75.2%
	FSVM	79.2%	71.3%	71.0%	69.9%	69.1%	67.7%	84.0%	80.4%	74.2%

		Varied Geometry					Varied Skill level					No Variation				
		Idle	Relocating	Swinging	Digging	Filling	Idle	Relocating	Swinging	Digging	Filling	Idle	Relocating	Swinging	Digging	Filling
RF	Idle	98.1%	0.0%	0.0%	1.4%	0.5%	80.4%	0.0%	0.5%	19.2%	0.0%	97.4%	0.1%	0.3%	2.2%	0.0%
	Relocating	0.1%	<b>99.5%</b>	0.3%	0.1%	0.0%	0.0%	<b>100.0%</b>	0.0%	0.0%	0.0%	0.5%	<b>99.5%</b>	0.0%	0.0%	0.0%
	Swinging	1.2%	0.0%	63.2%	26.9%	8.7%	0.1%	0.0%	50.7%	29.0%	20.2%	1.6%	0.2%	61.7%	17.5%	18.9%
	Digging	3.7%	0.0%	9.3%	<b>82.7%</b>	4.3%	1.9%	0.0%	7.4%	<b>89.1%</b>	1.6%	1.9%	0.1%	5.4%	<b>90.5%</b>	2.0%
	Filling	1.5%	0.0%	<b>50.0%</b>	20.6%	27.9%	0.3%	0.0%	21.0%	28.0%	<b>50.7%</b>	2.4%	0.0%	21.6%	22.4%	<b>53.5%</b>
FRF	Idle	98.3%	0.0%	0.0%	1.7%	0.0%	49.1%	0.0%	0.0%	50.9%	0.0%	98.8%	0.1%	0.6%	0.5%	0.0%
	Relocating	0.0%	<b>99.9%</b>	0.0%	0.0%	0.0%	0.0%	<b>100.0%</b>	0.0%	0.0%	0.0%	0.0%	<b>99.8%</b>	0.2%	0.0%	0.0%
	Swinging	0.1%	0.0%	64.9%	25.4%	9.6%	0.0%	0.0%	48.3%	28.9%	22.8%	0.2%	0.2%	70.4%	16.0%	13.2%
	Digging	1.1%	0.0%	6.3%	<b>91.8%</b>	0.8%	0.0%	0.0%	3.2%	<b>96.4%</b>	0.4%	0.9%	0.1%	3.4%	<b>95.3%</b>	0.3%
	Filling	0.4%	0.0%	<b>50.8%</b>	18.7%	30.2%	0.0%	0.0%	16.6%	24.3%	<b>59.1%</b>	0.1%	0.0%	20.3%	23.6%	<b>56.1%</b>
NN	Idle	99.1%	0.0%	0.0%	0.9%	0.0%	93.8%	0.0%	0.0%	6.2%	0.0%	98.2%	0.6%	0.1%	0.9%	0.2%
	Relocating	0.0%	<b>99.5%</b>	0.5%	0.1%	0.0%	0.0%	<b>99.9%</b>	0.0%	0.1%	0.0%	0.0%	<b>99.9%</b>	0.0%	0.1%	0.0%
	Swinging	1.6%	0.0%	50.9%	35.8%	11.8%	1.2%	0.0%	51.2%	27.5%	20.1%	1.6%	0.6%	48.9%	34.7%	14.3%
	Digging	4.0%	0.0%	11.6%	<b>79.3%</b>	5.1%	4.5%	0.0%	6.8%	<b>86.9%</b>	1.8%	2.9%	0.0%	4.9%	<b>86.8%</b>	5.3%
	Filling	1.9%	0.0%	<b>47.7%</b>	21.8%	28.6%	1.6%	0.0%	19.3%	27.5%	<b>51.6%</b>	1.3%	0.0%	23.7%	19.0%	<b>56.1%</b>
FNN	Idle	97.7%	0.0%	2.0%	0.3%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	96.9%	0.7%	0.1%	2.2%	0.1%
	Relocating	1.4%	<b>98.4%</b>	0.2%	0.0%	0.0%	0.0%	<b>99.7%</b>	0.3%	0.0%	0.0%	0.8%	<b>98.7%</b>	0.3%	0.2%	0.0%
	Swinging	0.3%	0.0%	56.4%	25.9%	17.5%	0.3%	0.1%	56.9%	24.4%	18.4%	0.3%	5.0%	49.3%	29.7%	15.7%
	Digging	1.5%	0.0%	11.1%	<b>84.8%</b>	2.6%	6.5%	0.7%	5.4%	<b>86.4%</b>	0.9%	0.9%	0.3%	4.3%	<b>91.8%</b>	2.7%
	Filling	2.3%	0.0%	<b>38.7%</b>	17.3%	41.6%	0.4%	0.0%	23.8%	24.3%	<b>51.5%</b>	0.0%	0.0%	25.5%	17.5%	<b>57.0%</b>
SVM	Idle	98.9%	0.0%	0.0%	0.8%	0.3%	88.9%	5.0%	0.0%	6.1%	0.0%	98.9%	0.4%	0.1%	0.6%	0.0%
	Relocating	0.9%	<b>81.4%</b>	16.3%	1.0%	0.5%	6.3%	<b>79.4%</b>	9.5%	4.3%	0.5%	4.9%	<b>87.4%</b>	6.3%	1.4%	0.1%
	Swinging	1.3%	0.7%	74.1%	19.3%	4.6%	1.9%	0.2%	34.8%	59.6%	3.5%	1.6%	4.2%	64.9%	19.2%	10.1%
	Digging	5.0%	0.2%	15.7%	<b>76.6%</b>	2.5%	1.2%	0.7%	4.1%	<b>93.0%</b>	1.1%	3.9%	0.1%	5.6%	<b>90.0%</b>	0.4%
	Filling	1.8%	0.3%	<b>66.9%</b>	17.7%	13.4%	1.7%	0.2%	22.4%	59.6%	16.1%	3.4%	0.6%	42.7%	21.2%	32.2%
FSVM	Idle	99.5%	0.0%	0.0%	0.5%	0.0%	90.6%	7.2%	0.0%	2.2%	0.0%	98.4%	1.6%	0.1%	0.0%	0.0%
	Relocating	2.5%	<b>81.4%</b>	14.3%	1.8%	0.1%	8.7%	<b>84.8%</b>	1.2%	5.3%	0.0%	3.6%	<b>90.0%</b>	4.1%	2.3%	0.0%
	Swinging	0.0%	0.2%	71.8%	22.0%	6.0%	0.0%	0.1%	74.0%	22.4%	3.5%	0.0%	5.8%	59.4%	29.6%	5.2%
	Digging	2.6%	0.1%	9.2%	<b>87.8%</b>	0.3%	0.0%	0.2%	20.0%	<b>79.0%</b>	0.7%	2.7%	0.2%	3.5%	<b>93.5%</b>	0.0%
	Filling	0.3%	0.0%	<b>67.4%</b>	17.6%	14.8%	0.0%	0.1%	83.3%	6.4%	10.2%	3.0%	1.4%	25.7%	43.9%	26.0%

**Fig. 10.** Confusion matrices of different models for different datasets considering 5 activities.

swinging, and digging).

#### 4.1.2. Results

As mentioned in Section 4, six different models were tested for 3 different scenarios once considering 5 activities and once considering 4 activities for the excavator. Tables 4 and 5 present the result of the performance analysis for all the scenarios. Also, Figs. 10 and 11 show the confusion matrices of the different models in terms of accuracy.

As a first observation, all models performed better (i.e., an average of 4.6% in accuracy) when only 4 activities are considered, as can be discerned in Tables 5 and 6. This is in line with the findings from the previous studies [21,51]. Also, all models performed worse when the testing dataset did not include the same scenario used in the training set (i.e., an average of 5.4% reduction in accuracy). This highlights the significance of testing activity recognition methods for the cases other than the one present in the training dataset to ensure generalizability.

Concerning the first question mentioned in Section 4.2, it can be observed that, in average, RF-based methods outperformed other methods. To elaborate on this matter, the performance margin of RF over NN and SVM is plotted in Fig. 12. As shown in this figure, RF outperformed NN in terms of accuracy by an average of 2.2%. The

performance margin is greater when the fractional variation of the two methods are compared, i.e., an average of 4% increase in accuracy. The accuracy margin of RF over SVM is greater with an average of 8.3% and 9.9% improvements in pure and fractional variations, respectively. The same applies to the precision of the models. RF-based methods always performed better than NN and SVM. Nonetheless, recall performance remained more or less the same between RF and NN-based methods. Again, RF methods performed much better than SVM methods in terms of recall. In general, it can be concluded that RF-based methods have better performance over other methods. It is also shown that RF-based methods have high generalizability with 86.2% and 86.9% accuracy for 5-activity classifiers for varied geometry and varied skill level, respectively. The generalizability was even higher for 4-activity classifiers, i.e., 90.4% and 92.7% for varied geometry and varied skill level, respectively.

To answer the second question in Section 4.2, the contribution of fractional feature augmentation on the improvement of generalizability was scrutinized. Fig. 13 shows the improvement achieved in terms of accuracy, precision, and recall when fractional feature augmentation was applied to different methods. As it is evident in the results, the fractional version of all methods outperformed the pure method. In the

		Varied Geometry				Varied Skill level				No Variation			
		Idle	Relocating	Swinging	Digging	Idle	Relocating	Swinging	Digging	Idle	Relocating	Swinging	Digging
RF	Idle	98.0%	0.0%	0.3%	1.7%	80.8%	0.0%	17.2%	2.0%	98.0%	0.1%	0.4%	1.5%
	Relocating	0.2%	99.4%	0.3%	0.1%	0.0%	100.0%	0.0%	0.0%	1.1%	98.6%	0.2%	0.0%
	Swinging	1.4%	0.0%	73.8%	24.8%	0.2%	0.0%	76.0%	23.8%	1.9%	0.0%	79.0%	19.1%
	Digging	3.4%	0.0%	15.8%	80.8%	1.9%	0.0%	11.2%	86.9%	2.0%	0.1%	8.6%	89.4%
FRF	Idle	97.2%	0.0%	2.3%	0.5%	96.3%	0.0%	0.2%	3.6%	99.1%	0.0%	0.4%	0.5%
	Relocating	0.0%	99.9%	0.1%	0.0%	0.0%	100.0%	0.0%	0.0%	0.1%	99.6%	0.4%	0.0%
	Swinging	0.0%	0.0%	75.0%	25.0%	0.0%	0.0%	72.9%	27.1%	0.0%	0.0%	83.2%	16.8%
	Digging	1.3%	0.0%	7.7%	91.0%	0.0%	0.0%	3.9%	96.1%	0.8%	0.1%	3.9%	95.2%
NN	Idle	98.6%	0.0%	0.2%	1.2%	98.3%	0.0%	0.0%	1.7%	98.0%	0.6%	0.2%	1.2%
	Relocating	0.0%	97.1%	2.8%	0.1%	0.0%	100.0%	0.0%	0.0%	0.0%	99.9%	0.0%	0.1%
	Swinging	1.4%	0.8%	66.7%	31.1%	1.6%	0.0%	70.7%	27.7%	1.4%	0.4%	65.1%	33.1%
	Digging	4.4%	0.1%	17.9%	77.6%	4.4%	0.0%	7.7%	87.9%	2.9%	0.0%	5.7%	91.4%
FNN	Idle	100.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	96.5%	0.5%	0.3%	2.7%
	Relocating	1.4%	91.8%	3.4%	3.4%	0.0%	99.8%	0.2%	0.0%	0.5%	98.6%	0.5%	0.5%
	Swinging	0.1%	0.6%	75.7%	23.5%	0.1%	0.0%	75.4%	24.5%	1.4%	2.2%	74.0%	22.3%
	Digging	1.4%	0.1%	10.5%	88.1%	5.1%	0.1%	6.3%	88.6%	0.9%	0.1%	7.5%	91.6%
SVM	Idle	99.1%	0.0%	0.2%	0.8%	88.9%	5.0%	0.0%	6.1%	98.9%	0.4%	0.1%	0.6%
	Relocating	0.9%	81.4%	16.9%	0.9%	6.3%	79.6%	10.3%	3.8%	4.9%	87.3%	6.5%	1.3%
	Swinging	1.5%	0.5%	82.0%	16.0%	1.8%	0.2%	82.5%	15.5%	2.4%	2.8%	79.5%	15.4%
	Digging	5.0%	0.2%	25.2%	69.7%	3.1%	1.9%	75.8%	19.2%	3.9%	0.1%	8.9%	87.1%
FSVM	Idle	99.5%	0.0%	0.0%	0.5%	90.6%	7.2%	0.0%	2.2%	98.4%	1.6%	0.1%	0.0%
	Relocating	2.5%	83.3%	12.4%	1.8%	8.7%	84.7%	2.1%	4.5%	3.6%	90.0%	4.1%	2.2%
	Swinging	0.1%	0.1%	85.5%	14.3%	0.0%	0.1%	90.6%	9.3%	1.2%	4.2%	84.9%	9.7%
	Digging	2.6%	0.1%	17.9%	79.3%	0.0%	0.2%	36.6%	63.1%	2.8%	0.2%	9.1%	88.0%

Fig. 11. Confusion matrices of different models for different datasets considering 4 activities.

Table 6

Accuracy, precision and recall of 4-activity models (best performances are highlighted in bold).

Model		Varied geometry			Varied skill			No variation		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall
4-activity models	RF	84.7%	80.0%	88.0%	87.7%	84.2%	85.9%	90.0%	89.7%	91.2%
	FRF	<b>90.4%</b>	<b>88.7%</b>	<b>90.8%</b>	<b>92.7%</b>	<b>94.4%</b>	<b>91.3%</b>	<b>94.1%</b>	<b>94.7%</b>	<b>94.3%</b>
	NN	81.1%	76.3%	85.1%	88.0%	82.4%	89.2%	88.0%	89.9%	88.7%
	FNN	87.1%	85.0%	89.0%	89.3%	83.7%	90.9%	89.6%	90.9%	90.2%
	SVM	76.1%	74.3%	83.1%	76.3%	62.4%	67.5%	87.2%	85.64%	88.2%
	FSVM	82.4%	79.7%	86.9%	76.6%	78.4%	82.2%	89.1%	87.4%	90.3%

case of RF, for instance, and considering 5-activity classifiers, fractional feature augmentation improved the accuracy by 5.3%, 2.8%, and 3.8% for varied geometry, varied skill level, and no variation scenarios, respectively. Even higher improvement was achieved for the 4-activity FRF method, with 5.7%, 5%, and 4.1% improvement in the accuracy for varied geometry, varied skill level, and no variation scenarios, respectively. However, the improvement to the accuracy caused by fractional feature augmentation is not limited to RF-based methods. Although to a lesser extent, similar patterns of improvement can be observed for NN and SVM-based methods. The contribution of fractional feature augmentation to precision is even higher than that of accuracy. For instance, an average improvement of 8.1% and 8% in the precision is observed when fractional feature augmentation is applied to 5-activity and 4-activity classifiers, respectively. Finally, it is also shown that fractional feature augmentation improved recall, although to a lesser extent compared to precision.

## 4.2. Second case study

### 4.2.1. Data collection and preparation

To further test the extent to which the proposed FRF method is generalizable, an extreme case test was designed in which the model trained based on the Case CX80C excavator was used to estimate the activities of a remotely controlled scaled equipment, which is scaled at 1/12. As shown in Fig. 14, IMUs and GNSS receiver are attached to this equipment to capture the motion data. A similar pattern of digging was simulated in the outdoor environment on the campus of the University of Twente. In the labeling of the dataset, the same 5 activities mentioned in the first case study were used. The data were preprocessed using the same methods as in case study 1 and the same activity recognition methods were applied.

### 4.2.2. Results

Table 7 summarizes the results of this case study. Fig. 15 presents the



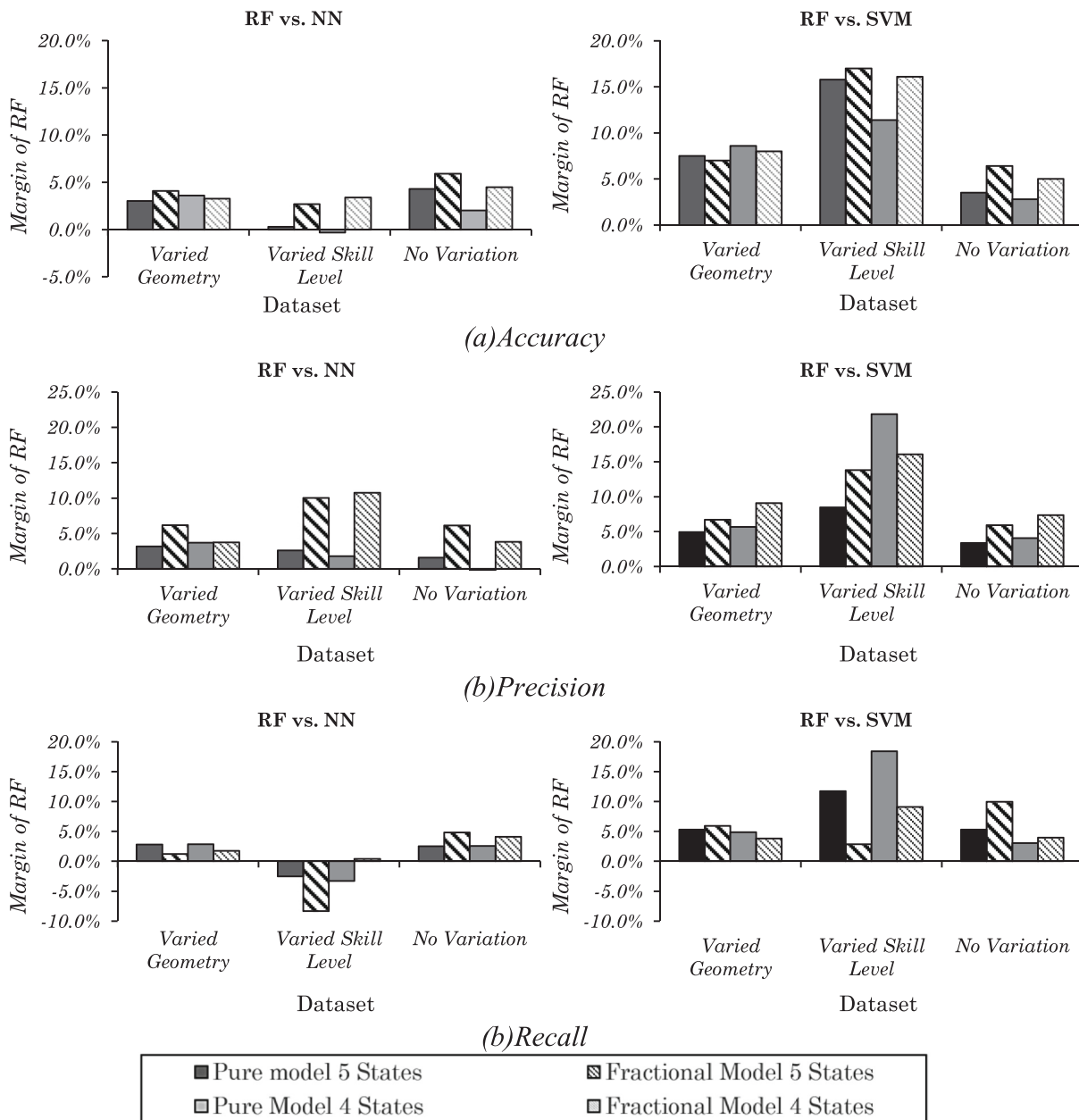


Fig. 12. Performance margin of RF over NN and SVM in terms of (a) accuracy, (b) precision, and (c) recall.

comparison of RF with other methods in terms of accuracy and also the analysis of the accuracy gain by applying fractional feature augmentation to different models. As shown in this figure and in conformity with the previous case study, FRF performs the best in terms of accuracy, precision and recall. The accuracy margins of RF over NN and SVM are about 7% and 25%, respectively. Similar to the previous case study, fractional feature augmentation has improved the accuracy performance of all models.

The significant observation in this case study is that while the velocity ranges of RC excavator is considerably different from that of an actual excavator, FRF was still able to estimate activities with high accuracy, i.e., 72.9%. Having said that, while it is still comparable to results shown in previous research [21,54,79], the accuracy of all models is noticeably less than the first case study. This is mainly because of disparities between the kinematic chain of the RC excavator and the actual excavator. Most importantly, the RC excavator used in this case study did not have a controllable DOF at the bucket. This introduced some anomalies in the activity pattern of the RC excavator.

### 4.3. Third case study

#### 4.3.1. Data collection and preparation

In the last case study, the proposed method was tested for two rollers, as examples of rigid body equipment, on an actual construction site in Amsterdam. The project entailed resurfacing of 3 different sections of a main street, as shown in Fig. 16(a). Therefore, rollers had to relocate several times to get ready for the next batch of compaction. The two rollers were Hamm HW90 model. As shown in Fig. 16(b), each roller was equipment with one GNSS receiver and one IMU, yielding the same data shown in Fig. 4(b). In this case study, a different GNSS receiver was used (i.e., Trimble SPS851 [98]). The data were collected over a window of  $\pm 30$  min with the update rate of 10 HZ, resulting in an average of 16,000 data points after preprocessing for each roller. The data were manually labeled for four activities, namely relocating, idle, moving forward (i.e., towards the paver) and moving backward (i.e., away from the paver). All models presented in the first case study were trained based on the data from one roller and then tested on the data from the other roller.



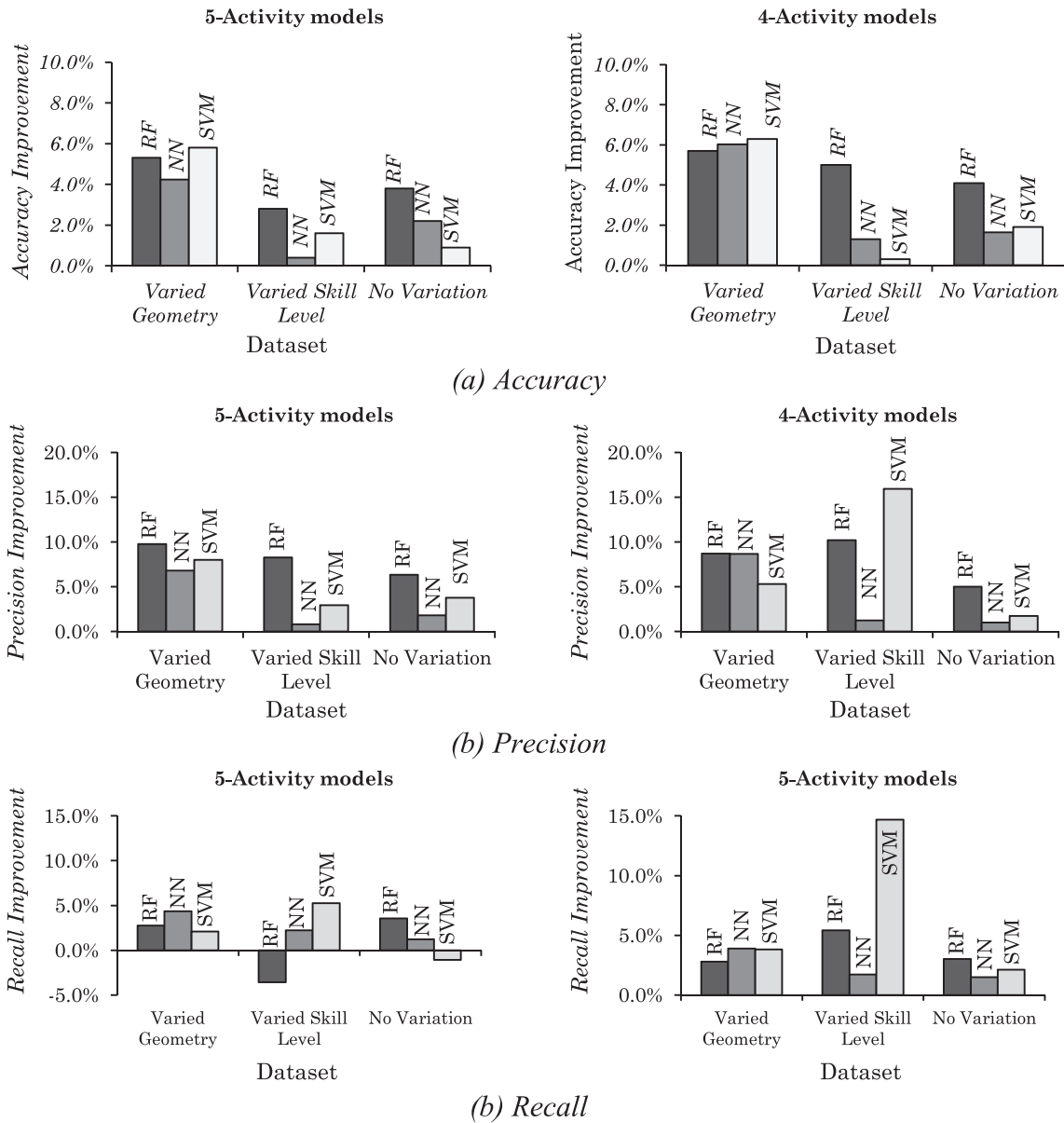


Fig. 13. Impact of applying fractional feature augmentation on (a) accuracy, (b) precision, and (c) recall of different models.

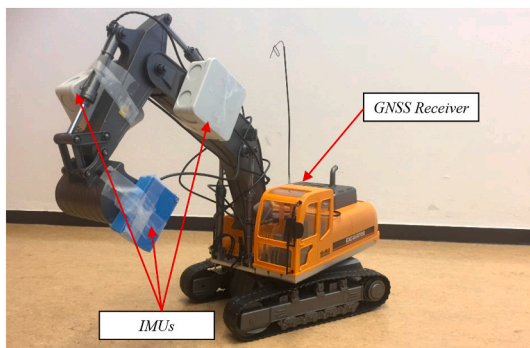


Fig. 14. Installation of sensors on the scaled excavator.

4.3.2. Results

Table 8 summarizes the results of this case study. Also, Fig. 17 presents a detailed analysis of the performance of the proposed method

Table 7

Accuracy, precision and recall of testing the models on scaled equipment (best performances are highlighted in bold).

Model	Accuracy	Precision	Recall
RF	67.3%	63.5%	69.1%
FRF	<b>72.9%</b>	<b>68.7%</b>	<b>71.7%</b>
NN	60.7%	58.6%	65.2%
FNN	65.7%	63.6%	69.5%
SVM	43.0%	44.7%	48.4%
FSVM	46.7%	50.0%	50.4%

against other baseline methods. As shown in the table, FRF outperformed other models in terms of accuracy (82.6%) and recall (94.6%). However, FNN performed the best in terms of precision. As shown in Fig. 17, fractional feature augmentation improved the performance of all models, although the improvement in the case of SVM is very marginal. The noticeable observation in this case study is that in the pure form, all models had similar accuracy. In the fractional mode,

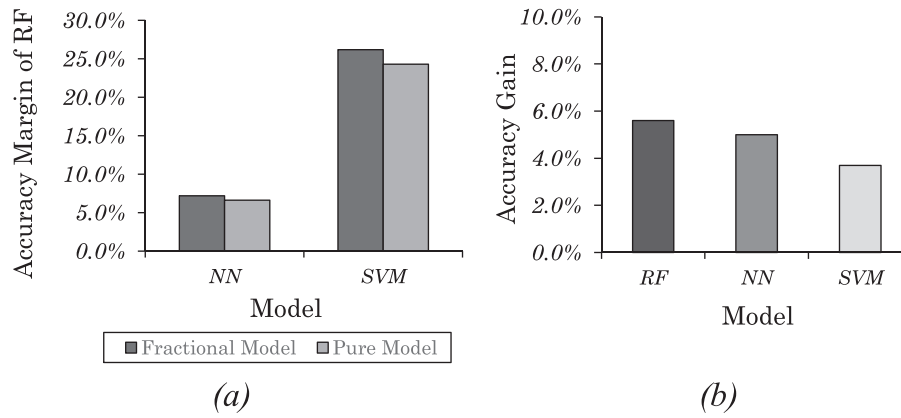


Fig. 15. (a) Accuracy margin of RF over NN and SVM, and (b) Impact of applying fractional feature augmentation on accuracy for RC excavator.

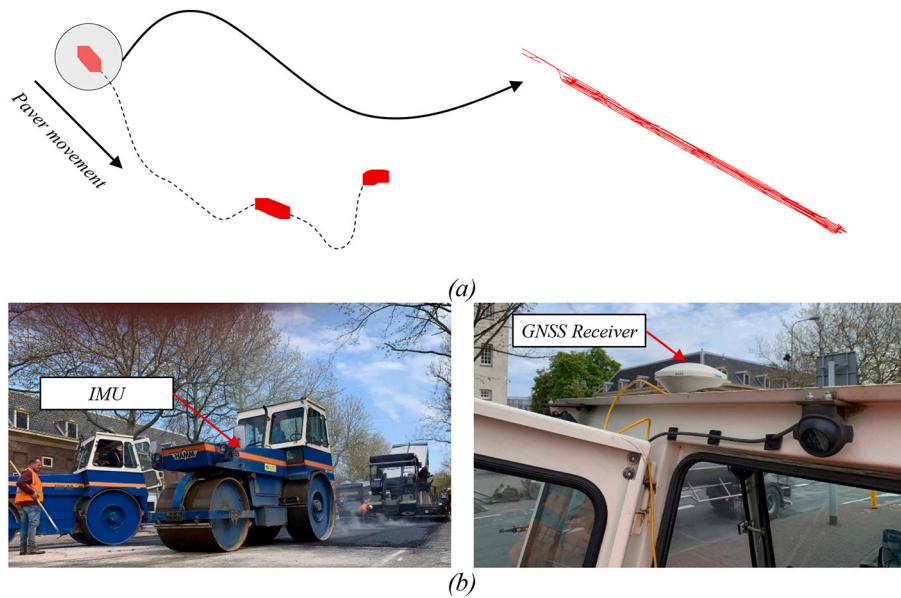


Fig. 16. (a) layout of the project, and (b) Installation of sensors on rollers.

Table 8

Accuracy, precision and recall of testing the models on rollers (best performances are highlighted in bold).

Model	4 activities			3 activities		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
RF	77.3%	87.1%	86.5%	99.9%	<b>100.00%</b>	<b>99.99%</b>
FRF	<b>82.6%</b>	81.6%	<b>94.6%</b>	<b>99.9%</b>	99.52%	99.89%
NN	77.3%	78.8%	80.0%	99.9%	99.90%	99.98%
FNN	80.7%	<b>89.4%</b>	88.9%	98.9%	99.26%	83.83%
SVM	77.9%	72.0%	90.0%	99.8%	99.41%	99.84%
FSVM	78.1%	76.3%	93.8%	99.2%	99.52%	99.44%

although FRF had a slightly better performance, the difference is not significant. These findings can be construed as an indication that for rigid body equipment, while fractional feature augmentation is as effective as in the case of articulated equipment, the extent to which RF contributes to improved accuracy is less. This is mainly because articulated equipment has a considerably more complex kinematic which makes the distinction between their activities more intricate. In the case of rigid body equipment, the underlying rules that define equipment activities are simpler and less dependent on the complex interaction between DOFs. Therefore, all classifiers are able to predict the activities

with similar accuracy.

Another important observation from this case study is that the formulation of orientation-based activities (e.g., moving forward and backward) would complicate the activity recognition solely based on velocity-related features. If these two activities are merged into a single activity, the performance would increase significantly for all models as shown in Table 8. Another possible approach is to consider the orientation-related features in the training of the models. It is expected that while the accuracy will increase for this case study, it would compromise the generalizability of the model because the orientation of movements in different projects is not the same.

### 5. Discussion

The results of the case studies corroborate the hypothesis of this research that the FRF can help develop an accurate activity recognition model without requiring large datasets. In the case studies, small datasets (i.e., approximately 10% of those used to develop recently proposed deep learning models) were used to train the FRF model, and then it was successfully tested on new cases. This attests to the high generalizability of the proposed method in spite of much smaller training datasets. With the accuracy of up to 94% for articulated equipment and 99% for rigid body equipment, the method is shown to have comparable accuracy to

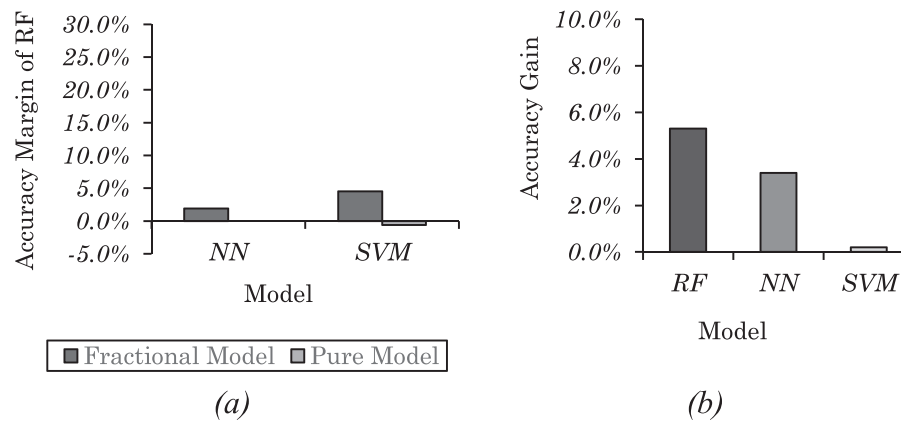


Fig. 17. (a) Accuracy margin of RF over NN and SVM, and (b) Impact of applying fractional feature augmentation on accuracy for Rollers with 4 activities.

recent research work [54,55,89].

Based on the above, the main contributions of this paper are as follows: (1) it is shown that accurate and generalizable activity recognition models can be developed without the need for large training datasets and the use of deep learning methods. This was achieved through the novel integration of the fractional feature augmentation method with RF. From the practical standpoint, by easing the need for comprehensive training datasets, this research helps develop accurate activity recognition models that can be used commercially in equipment monitoring systems; (2) it is shown that regardless of the machine learning method employed, the fractional feature augmentation method can improve the performance of activity recognition models. This feature augmentation method is shown to have the potential to become a common part of all the future developments in ML-based equipment activity recognition; (3) to the best of authors' knowledge, it is for the first time that the activity recognition classifiers have been rigorously tested for generalizability by investigating the impact of variation in the geometry of the equipment and the skill of operators. It is demonstrated that these variations indeed take a toll on the performance of the classifier. This signifies the importance of applying a similar testing regime in the future to ensure that activity recognition methods are robust and generalizable.

Based on the presented results from the case studies, the proposed method seems to be more effective and relevant for articulated equipment such as excavators and loaders. This is mainly because the kinematics of this type of equipment and how changes in their DOFs contribute to transition between activities are inherently more complex. As mentioned in the introduction, RF classifiers are known to be more suited for complex multi-dimensional problems. This is better manifested in the case of articulated equipment. The same applies, to a lesser extent, to the use of fractional feature augmentation. While the proposed method of augmenting the feature domain proved to be effective in all cases, the improvement in the case of articulated equipment seems to be more conspicuous. Again, this is because fractional integral and derivative on velocity are expected to better account for complex kinematics where the transition between activities cannot be attributed to a simple single-feature threshold, like the case of 3-activity roller.

This research complements the recent work on the augmentation of training dataset by Rashid and Louis [55] in that while this research focused on the expansion of feature domain through feature augmentation, the work of Rashid and Loius proposed a method (and showed the benefits thereof) for expanding the size of the training dataset through generation of synthetic data. It can be postulated that the combination of the proposed fractional feature augmentation method with the training data augmentation can significantly improve the generalizability of activity recognition classifiers in the future.

Another important point of discussion is about the requirements of the proposed method in terms of sensor data. Because the fractional feature augmentation discounts the update rate of the data, it is essential

that the data is collected at a high frequency. At the same time, because the pose estimation of articulated equipment mostly requires sensor data fusion, it is important to be able to synchronize the data. This especially pertains to location data which are normally collected at a lower frequency than IMU readings. While up-sampling techniques can be used, it is preferred to avoid high frequency discrepancy between different sensors to avoid erosion of accuracy. On the same note, given that any new data entry to the model (for the evaluation purpose) also need to be subject to fractional augmentation, the application of the proposed method for high-frequency real-time monitoring needs to be further studied.

## 6. Conclusions

The detection of the activities of construction equipment is important for the monitoring of a wide range of safety, productivity, and sustainability related practices. The existing activity recognition methods depend heavily on a comprehensive dataset, which is a limitation given the sheer number of variations in the size and shape of the equipment. This research proposed a novel Fractional Random Forest method to develop an accurate activity recognition model that can work with small datasets and remain generalizable. The proposed method was applied to three case studies where several scenarios were developed to test the generalizability of the proposed method.

Based on the analysis of the results, the following main conclusions can be reported: (1) With the accuracy of up to 94% for articulated equipment and 99% for rigid body equipment, the proposed FRF was able to deliver comparable performance to recent deep learning-based activity recognition methods with only a fraction of training dataset used in previous methods [54,55,89]. The low dependency on the comprehensive training data set provides the advantage of using a small set of training data to accurately predict the activities of a large set of equipment; (2) Compared to other baseline shallow learners (i.e., MLP and SVM), FRF has a better performance in terms of accuracy, precision, and recall; (3) the FRF model was able to predict activities of an actual piece of equipment in a different size/shape with an accuracy of 86.2%. In an extreme case of testing the model on a scaled RC equipment, FRF delivered an accuracy of 72.9% which is still comparable to results reported in the recent machine-learning based methods. This indicated a high generalizability of the proposed method; and finally (4) regardless of the machine learning method, fractional feature augmentation can help improve performance of activity recognition model.

In the future, the proposed method will be tested for a wider range of variations in equipment sizes and operators' skill level. Also, given that on different types of terrain and for different types of work the movement pattern of the equipment may change, it is beneficial to collect data for equipment working on different types of terrain and performing a wider range of tasks. Finally, the proposed FRF method can be further

enhanced by integrating it with the data augmentation method proposed by Rashid and Louis [55]. This will be addressed in the future study of the authors.

### Declaration of Competing Interest

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

### References

- [1] J. Hinze, X. Huang, L. Terry, The nature of struck-by accidents, *J. Constr. Eng. Manag.* 131 (2005) 262–268, [https://doi.org/10.1061/\(ASCE\)0733-9364\(2005\)131:2\(262\)](https://doi.org/10.1061/(ASCE)0733-9364(2005)131:2(262)).
- [2] J.W. Hinze, J. Teizer, M. Rinker, Visibility-related fatalities related to construction equipment, *Saf. Sci.* 49 (2011) 709–718, <https://doi.org/10.1016/j.ssci.2011.01.007>.
- [3] E. Kazan, M.A. Usman, Worker safety and injury severity analysis of earthmoving equipment accidents, *J. Saf. Res.* 65 (2018) 73–81, <https://doi.org/10.1016/J.JSR.2018.02.008>.
- [4] B.U. Ayhan, O.B. Tokdemir, Predicting the outcome of construction incidents, *Saf. Sci.* 113 (2019) 91–104, <https://doi.org/10.1016/J.SSCI.2018.11.001>.
- [5] J. Teizer, T. Cheng, Proximity hazard indicator for workers-on-foot near miss interactions with construction equipment and geo-referenced hazards, *Autom. Constr.* 60 (2015) 58–73, <https://doi.org/10.1016/j.autcon.2015.09.003>.
- [6] O. Golovina, J. Teizer, N. Pradhananga, Heat map generation for predictive safety planning: preventing struck-by and near miss interactions between workers-on-foot and construction equipment, *Autom. Constr.* 71 (2016) 99–115, <https://doi.org/10.1016/J.AUTCON.2016.03.008>.
- [7] F. Vahdatikhaki, A. Hammad, Risk-based look-ahead workspace generation for earthwork equipment using near real-time simulation, *Autom. Constr.* 58 (2015) 207–220, <https://doi.org/10.1016/J.AUTCON.2015.07.019>.
- [8] S.C. Ok, S.K. Sinha, Construction management and economics construction equipment productivity estimation using artificial neural network model construction equipment productivity estimation using artificial neural network model, *Constr. Manag. Econ.* 24 (10) (2007) 1029–1044, <https://doi.org/10.1080/01446190600851033>.
- [9] G. Kannan, Field studies in construction equipment economics and productivity, *J. Constr. Eng. Manag.* 137 (10) (2011) 823–828, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000335](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000335).
- [10] P.F. Kaming, P.O. Olomolaiye, G.D. Holt, F.C. Harris, Factors influencing construction time and cost overruns on high-rise projects in Indonesia, *Constr. Manag. Econ.* 15 (1) (2010) 83–94, <https://doi.org/10.1080/014461997373132>.
- [11] S.G. Naoum, Factors influencing labor productivity on construction sites, *Int. J. Product. Perform. Manag.* 65 (2016) 401–421, <https://doi.org/10.1108/IJPPM-03-2015-0045>.
- [12] A.T. Gurmu, A.A. Aibinu, Construction equipment management practices for improving labor productivity in multistory building construction projects, *J. Constr. Eng. Manag.* 143 (2017), [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001384](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001384).
- [13] M. Waris, M. Shahir Liew, M.F. Khamidi, A. Idrus, Criteria for the selection of sustainable onsite construction equipment, *Int. J. Sustain. Built Environ.* 3 (2014) 96–110, <https://doi.org/10.1016/j.ijsbe.2014.06.002>.
- [14] R.A. Rivas, J.D. Borchering, V. González, L.F. Alarcón, Analysis of factors influencing productivity using craftsman questionnaires: case study in a Chilean construction company, *J. Constr. Eng. Manag.* 137 (2011) 312–320, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000274](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000274).
- [15] A. Hajji, The use of construction equipment productivity rate model for estimating fuel use and carbon dioxide (CO<sub>2</sub>) emissions Case study: bulldozer, excavator and dump truck, *Int. J. Sustain. Eng.* 8 (2015) 111–121, <https://doi.org/10.1080/19397038.2014.962645>.
- [16] P. Lewis, M. Leming, W. Rasdorf, Impact of engine idling on fuel use and CO<sub>2</sub> emissions of nonroad diesel construction equipment, *J. Manag. Eng.* 28 (2012) 31–38, [https://doi.org/10.1061/\(ASCE\)ME.1943-5479.0000068](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000068).
- [17] World Economy Forum, Industry Agenda Shaping the Future of Construction a Breakthrough in Mindset and Technology Prepared in collaboration with the Boston Consulting Group. [http://www3.weforum.org/docs/WEF\\_Shaping\\_the\\_Future\\_of\\_Construction\\_full\\_report\\_pdf](http://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Construction_full_report_pdf), 2016 (accessed October 26, 2020).
- [18] F. Vahdatikhaki, A. Hammad, Dynamic equipment workspace generation for improving earthwork safety using real-time location system, *Adv. Eng. Inform.* 29 (2015) 459–471, <https://doi.org/10.1016/J.AEI.2015.03.002>.
- [19] F. Vahdatikhaki, A. Hammad, Framework for near real-time simulation of earthmoving projects using location tracking technologies, *Autom. Constr.* 42 (2014) 50–67, <https://doi.org/10.1016/J.AUTCON.2014.02.018>.
- [20] F. Vahdatikhaki, K. El Ammari, A.K. Langroodi, S. Miller, A. Hammad, A. Doree, Beyond data visualization: a context-realistic construction equipment training simulators, *Autom. Constr.* 106 (2019), <https://doi.org/10.1016/j.autcon.2019.102853>.
- [21] M. Golparvar-Fard, A. Heydarian, J.C. Niebles, Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers, *Adv. Eng. Inform.* 27 (2013) 652–663, <https://doi.org/10.1016/J.AEI.2013.09.001>.
- [22] R. Navon, Research in automated measurement of project performance indicators, *Autom. Constr.* 16 (2007) 176–188, <https://doi.org/10.1016/j.autcon.2006.03.003>.
- [23] M. Memarzadeh, M. Golparvar-Fard, J.C. Niebles, Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors, *Autom. Constr.* 32 (2013) 24–37, <https://doi.org/10.1016/J.AUTCON.2012.12.002>.
- [24] F. Vahdatikhaki, A. Hammad, H. Siddiqui, Optimization-based excavator pose estimation using real-time location systems, *Autom. Constr.* 56 (2015) 76–92, <https://doi.org/10.1016/J.AUTCON.2015.03.006>.
- [25] F. Vahdatikhaki, S.M. Langari, A. Taher, K. El Ammari, A. Hammad, Enhancing coordination and safety of earthwork equipment operations using multi-agent system, *Autom. Constr.* 81 (2017) 267–285, <https://doi.org/10.1016/J.AUTCON.2017.04.008>.
- [26] C. Zhang, A. Hammad, Multiagent approach for real-time collision avoidance and path replanning for cranes, *J. Comput. Civ. Eng.* 26 (2012) 782–794, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000181](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000181).
- [27] C. Zhang, A. Hammad, S. Rodriguez, Crane pose estimation using UWB real-time location system, *J. Comput. Civ. Eng.* 26 (2012) 625–637, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000172](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000172).
- [28] A. Alshibani, O. Moselhi, Productivity based method for forecasting cost & time of earthmoving operations using sampling GPS data, *J. Inform. Technol. Construct. (ITcon)*. 21 (2016) 39–56. <http://www.itcon.org/2016/3> (accessed October 26, 2020).
- [29] N. Pradhananga, J. Teizer, Automatic spatio-temporal analysis of construction site equipment operations using GPS data, *Autom. Constr.* 29 (2013) 107–122, <https://doi.org/10.1016/J.AUTCON.2012.09.004>.
- [30] R. Akhavan, S.M. Asce, A.H. Behzadan, A.M. Asce, Knowledge-based simulation modeling of construction Fleet operations using multimodal-process data mining, *J. Constr. Eng. Manag.* 139 (11) (2013), [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000775](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000775).
- [31] R. Akhavan, A.H. Behzadan, An integrated data collection and analysis framework for remote monitoring and planning of construction operations, *Adv. Eng. Inform.* 26 (2012) 749–761, <https://doi.org/10.1016/J.AEI.2012.04.004>.
- [32] E.D. Marks, J. Teizer, Method for testing proximity detection and alert technology for safe construction equipment operation, *Constr. Manag. Econ.* 31 (2013) 636–646, <https://doi.org/10.1080/01446193.2013.783705>.
- [33] Z. Zhu, X. Ren, Z. Chen, Integrated detection and tracking of workforce and equipment from construction jobsite videos, *Autom. Constr.* 81 (2017) 161–171, <https://doi.org/10.1016/J.AUTCON.2017.05.005>.
- [34] J. Wang, S. Razavi, Two 4D models effective in reducing false alarms for struck-by-equipment Hazard prevention, *J. Comput. Civ. Eng.* 30 (2016), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000589](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000589).
- [35] E. Rezaazadeh Azar, B. McCabe, Part based model and spatial-temporal reasoning to recognize hydraulic excavators in construction images and videos, *Autom. Constr.* 24 (2012) 194–202, <https://doi.org/10.1016/J.AUTCON.2012.03.003>.
- [36] E. Rezaazadeh Azar, B. McCabe, Automated visual recognition of dump trucks in construction videos, *J. Comput. Civ. Eng.* 26 (2012) 769–781, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000179](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000179).
- [37] A. Montaser, O. Moselhi, RFID+ for tracking earthmoving operations automated site data acquisition view project cost estimation view project, in: *The Proceedings of Construction Research Congress*, 2012, pp. 1011–1020, <https://doi.org/10.1061/9780784412329.102>.
- [38] T. Cheng, J. Teizer, Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications, *Autom. Constr.* 34 (2013) 3–15, <https://doi.org/10.1016/J.AUTCON.2012.10.017>.
- [39] T. Cheng, G.C. Migliaccio, J. Teizer, U.C. Gatti, Data fusion of real-time location sensing and physiological status monitoring for ergonomics analysis of construction workers, *J. Comput. Civ. Eng.* 27 (2013) 320–335, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000222](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000222).
- [40] S. Chi, C.H. Caldas, Automated object identification using optical video cameras on construction sites, *Comput. Aided Civil Infrastruct. Eng.* 26 (2011) 368–380, <https://doi.org/10.1111/j.1467-8667.2010.00690.x>.
- [41] S. Chae, Development of warning system for preventing collision accident on construction site, in: *26th International Symposium on Automation and Robotics in Construction*, 2009, p. 55. <http://www.irbnet.de/daten/iconda/CIB14818.pdf> (accessed October 26, 2020).
- [42] S. Chae, T. Yoshida, Application of RFID technology to prevention of collision accident with heavy equipment, *Autom. Constr.* 19 (2010) 368–374, <https://doi.org/10.1016/J.AUTCON.2009.12.008>.
- [43] S. Moon, B. Yang, Effective monitoring of the concrete pouring operation in an RFID-based environment, *J. Comput. Civ. Eng.* 24 (2010) 108–116, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000004](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000004).
- [44] J. Teizer, B.S. Allread, C.E. Fullerton, J. Hinze, Autonomous pro-active real-time construction worker and equipment operator proximity safety alert system, *Autom. Constr.* 19 (2010) 630–640, <https://doi.org/10.1016/j.autcon.2010.02.009>.
- [45] X. Luo, W.J. O'Brien, F. Leite, J.A. Goulet, Exploring approaches to improve the performance of autonomous monitoring with imperfect data in location-aware wireless sensor networks, *Adv. Eng. Inform.* 28 (4) (2014) 287–296, <https://doi.org/10.1016/j.aei.2014.08.004>.
- [46] J. Wang, S.N. Razavi, Low false alarm rate model for unsafe-proximity detection in construction, *J. Comput. Civ. Eng.* 30 (2016), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000470](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000470).



- [47] M.M. Soltani, Z. Zhu, A. Hammad, Skeleton estimation of excavator by detecting its parts, *Autom. Constr.* 82 (2017) 1–15, <https://doi.org/10.1016/J.AUTCON.2017.06.023>.
- [48] K.M. Lundeen, S. Dong, N. Fredricks, M. Akula, J. Seo, V.R. Kamat, Optical marker-based end effector pose estimation for articulated excavators, *Autom. Constr.* 65 (2016) 51–64, <https://doi.org/10.1016/j.autcon.2016.02.003>.
- [49] J. Xu, H.S. Yoon, Vision-based estimation of excavator manipulator pose for automated grading control, *Autom. Constr.* 98 (2019) 122–131, <https://doi.org/10.1016/j.autcon.2018.11.022>.
- [50] C.J. Liang, K.M. Lundeen, W. McGee, C.C. Menassa, S. Lee, V.R. Kamat, A vision-based marker-less pose estimation system for articulated construction robots, *Autom. Constr.* 104 (2019) 80–94, <https://doi.org/10.1016/J.AUTCON.2019.04.004>.
- [51] R. Akhavian, A.H. Behzadan, Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers, *Adv. Eng. Inform.* 29 (2015) 867–877, <https://doi.org/10.1016/j.aei.2015.03.001>.
- [52] R. Akhavian, A.H. Behzadan, Coupling Human Activity Recognition and Wearable Sensors for Data-driven Construction Simulation. <http://www.itcon.org/2018/1,2018> (accessed October 26, 2020).
- [53] J. Kim, S. Chi, Action recognition of earthmoving excavators based on sequential pattern analysis of visual features and operation cycles, *Autom. Constr.* 104 (2019) 255–264, <https://doi.org/10.1016/j.autcon.2019.03.025>.
- [54] T. Slaton, C. Hernandez, R. Akhavian, Construction activity recognition with convolutional recurrent networks, *Autom. Constr.* 113 (2020), <https://doi.org/10.1016/j.autcon.2020.103138>.
- [55] K.M. Rashid, J. Louis, Times-series data augmentation and deep learning for construction equipment activity recognition, *Adv. Eng. Inform.* 42 (2019), <https://doi.org/10.1016/J.AEI.2019.100944>.
- [56] J. Cai, Y. Zhang, H. Cai, Two-step long short-term memory method for identifying construction activities through positional and attentional cues, *Autom. Constr.* 106 (2019), <https://doi.org/10.1016/J.AUTCON.2019.102886>.
- [57] C.F. Cheng, A. Rashidi, M.A. Davenport, D. Anderson, Audio signal processing for activity recognition of construction heavy equipment, in: 33rd International Symposium on Automation and Robotics in Construction, 2016, in: [https://www.iaarc.org/publications/2016\\_proceedings\\_of\\_the\\_33rd\\_isarc\\_auburn\\_usa/audio\\_signal\\_processing\\_for\\_activity\\_recognition\\_of\\_construction\\_heavy\\_equipment.html](https://www.iaarc.org/publications/2016_proceedings_of_the_33rd_isarc_auburn_usa/audio_signal_processing_for_activity_recognition_of_construction_heavy_equipment.html) (accessed October 26, 2020).
- [58] C. Lam, C. Yu, L. Huang, D. Rubin, Retinal lesion detection with deep learning using image patches, *Investig. Ophthalmol. Vis. Sci.* 59 (2018) 590–596, <https://doi.org/10.1167/iov.17-22721>.
- [59] W. Zhao, Research on the deep learning of the small sample data based on transfer learning, in: AIP Conference Proceedings, 2017, <https://doi.org/10.1063/1.4992835>.
- [60] D. Perez, D. Banerjee, C. Kwan, M. Dao, Y. Shen, K. Koperski, G. Marchisio, J. Li, Deep learning for effective detection of excavated soil related to illegal tunnel activities, in: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference, UEMCON 2017, Institute of Electrical and Electronics Engineers Inc., 2017, pp. 626–632, <https://doi.org/10.1109/UEMCON.2017.8249062>.
- [61] D. Soekhoe, P. van der Putten, A. Plaat, On the impact of data set size in transfer learning using deep neural networks, in: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Springer Verlag, 2016, pp. 50–60, [https://doi.org/10.1007/978-3-319-46349-0\\_5](https://doi.org/10.1007/978-3-319-46349-0_5).
- [62] J. Wang, Y. Chen, S. Hao, X. Peng, L. Hu, Deep learning for sensor-based activity recognition: a survey, *Pattern Recogn. Lett.* 119 (2019) 3–11, <https://doi.org/10.1016/j.patrec.2018.02.010>.
- [63] T. DeVries, G.W. Taylor, Dataset augmentation in feature space, in: 5th International Conference on Learning Representations, ICLR 2017 - Workshop Track Proceedings, International Conference on Learning Representations, ICLR, 2019. <https://arxiv.org/abs/1702.05538v1> (accessed October 26, 2020).
- [64] Z. Chen, Y. Fu, Y. Zhang, Y.-G. Jiang, X. Xue, L. Sigal, Multi-level semantic feature augmentation for one-shot learning, *IEEE Trans. Image Process.* 28 (2018) 4594–4605, <https://doi.org/10.1109/TIP.2019.2910052>.
- [65] H. Wang, J. Gu, S. Wang, An effective intrusion detection framework based on SVM with feature augmentation, *Knowl.-Based Syst.* 136 (2017) 130–139, <https://doi.org/10.1016/j.knosys.2017.09.014>.
- [66] F. Ikeda, A numerical algorithm of discrete fractional calculus by using inhomogeneous sampling data, *Trans. Soc. Instrum. Control Eng.* 42 (2006) 941–948, <https://doi.org/10.9746/sicetr1965.42.941>.
- [67] M. Pal, Random forest classifier for remote sensing classification, *Int. J. Remote Sens.* 26 (2005) 217–222, <https://doi.org/10.1080/01431160412331269698>.
- [68] M.W. Ahmad, M. Mourshed, Y. Rezgui, Trees vs neurons: comparison between random forest and ANN for high-resolution prediction of building energy consumption, *Energy Build.* 147 (2017) 77–89, <https://doi.org/10.1016/j.enbuild.2017.04.038>.
- [69] I. Brown, C. Mues, An experimental comparison of classification algorithms for imbalanced credit scoring data sets, *Expert Syst. Appl.* 39 (2012) 3446–3453, <https://doi.org/10.1016/j.eswa.2011.09.033>.
- [70] S. Xia, G. Wang, Z. Chen, Y. Duan, Q. Liu, Complete random Forest based class noise filtering learning for improving the generalizability of classifiers, *IEEE Trans. Knowl. Data Eng.* 31 (2019) 2063–2078, <https://doi.org/10.1109/TKDE.2018.2873791>.
- [71] C.Q.X. Poh, C.U. Ubeynarayana, Y.M. Goh, Safety leading indicators for construction sites: a machine learning approach, *Autom. Constr.* 93 (2018) 375–386, <https://doi.org/10.1016/j.autcon.2018.03.022>.
- [72] Q. Han, C. Gui, J. Xu, G. Lacidogna, A generalized method to predict the compressive strength of high-performance concrete by improved random forest algorithm, *Constr. Build. Mater.* 226 (2019) 734–742, <https://doi.org/10.1016/j.conbuildmat.2019.07.315>.
- [73] H. Kim, C.R. Ahn, D. Engelhaupt, S. Lee, Application of dynamic time warping to the recognition of mixed equipment activities in cycle time measurement, *Autom. Constr.* 87 (2018) 225–234, <https://doi.org/10.1016/J.AUTCON.2017.12.014>.
- [74] Y.-C. Lee, M. Scarpiniti, A. Uncini, Advanced sound classifiers and performance analyses for accurate audio-based construction project monitoring, *J. Comput. Civ. Eng.* 34 (2020), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000911](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000911).
- [75] X. Hu, H. Yuan, C. Zou, Z. Li, L. Zhang, Co-estimation of state of charge and state of health for lithium-ion batteries based on fractional-order Calculus, *IEEE Trans. Veh. Technol.* 67 (2018) 10319–10329, <https://doi.org/10.1109/TVT.2018.2865664>.
- [76] R. Ibrahim, The fractional differential polynomial neural network for approximation of functions, *Entropy.* 15 (2013) 4188–4198, <https://doi.org/10.3390/e15104188>.
- [77] K. Chen, C. Li, R. Tang, Estimation of the nitrogen concentration of rubber tree using fractional calculus augmented NIR spectra, *Ind. Crop. Prod.* 108 (2017) 831–839, <https://doi.org/10.1016/j.indcrop.2017.06.069>.
- [78] J. Zou, H. Kim, Using hue, saturation, and value color space for hydraulic excavator idle time analysis, *J. Comput. Civ. Eng.* 21 (2007) 238–246, [https://doi.org/10.1061/\(ASCE\)0887-3801\(2007\)21:4\(238\)](https://doi.org/10.1061/(ASCE)0887-3801(2007)21:4(238)).
- [79] J. Gong, C.H. Caldas, C. Gordon, Learning and classifying actions of construction workers and equipment using bag-of-video-feature-words and Bayesian network models, *Adv. Eng. Inform.* 25 (2011) 771–782, <https://doi.org/10.1016/J.AEI.2011.06.002>.
- [80] E.R. Azar, S. Dickinson, B. McCabe, M. Asce, Server-customer interaction tracker: computer vision-based system to estimate dirt-loading cycles, *J. Constr. Eng. Manag.* 139 (7) (2013) 785–794, [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000652](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000652).
- [81] J. Kim, S. Chi, J. Seo, Interaction analysis for vision-based activity identification of earthmoving excavators and dump trucks, *Autom. Constr.* 87 (2018) 297–308, <https://doi.org/10.1016/J.AUTCON.2017.12.016>.
- [82] C. Chen, Z. Zhu, A. Hammad, Automated excavators activity recognition and productivity analysis from construction site surveillance videos, *Autom. Constr.* 110 (2020), <https://doi.org/10.1016/j.autcon.2019.103045>.
- [83] B. Sherafat, C.R. Ahn, R. Akhavian, A.H. Behzadan, M. Golparvar-Fard, H. Kim, Y. C. Lee, A. Rashidi, E.R. Azar, Automated methods for activity recognition of construction workers and equipment: state-of-the-art review, *J. Constr. Eng. Manag.* 146 (2020), [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0001843](https://doi.org/10.1061/(ASCE)CO.1943-7862.0001843).
- [84] C.F. Cheng, A. Rashidi, M.A. Davenport, D.V. Anderson, Activity analysis of construction equipment using audio signals and support vector machines, *Autom. Constr.* 81 (2017) 240–253, <https://doi.org/10.1016/J.AUTCON.2017.06.005>.
- [85] J. Cao, W. Wang, J. Wang, R. Wang, Excavation equipment recognition based on novel acoustic statistical features, *IEEE Trans. Cybernet.* 47 (2017), <https://doi.org/10.1109/TCYB.2016.2609999>.
- [86] J. Cao, W. Huang, T. Zhao, J. Wang, R. Wang, An enhance excavation equipments classification algorithm based on acoustic spectrum dynamic feature, *Multidim. Syst. Sign. Process.* 28 (2017) 921–943, <https://doi.org/10.1007/s11045-015-0374-z>.
- [87] J. Cao, T. Zhao, J. Wang, R. Wang, Y. Chen, Excavation equipment classification based on improved MFCC features and ELM, *Neurocomputing.* 261 (2017) 231–241, <https://doi.org/10.1016/j.neucom.2016.03.113>.
- [88] D. Wass, H. Axelsson, Machine Learning for Activity Recognition of Dumpers. <http://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1354995&dswid=-2037,2019> (accessed March 17, 2020).
- [89] B. Sherafat, A. Rashidi, Y.C. Lee, C.R. Ahn, A hybrid kinematic-acoustic system for automated activity detection of construction equipment, *Sensors* 19 (2019), <https://doi.org/10.3390/s19194286>.
- [90] S. Das, Functional Fractional Calculus for System Identification and Controls, Springer, Berlin Heidelberg, 2008, <https://doi.org/10.1007/978-3-540-72703-3>.
- [91] L. Breiman, Bagging predictors, *Mach. Learn.* 24 (1996) 123–140, <https://doi.org/10.1007/bf00058655>.
- [92] A.M. Prasad, L.R. Iverson, A. Liaw, Newer classification and regression tree techniques: bagging and random forests for ecological prediction, *Ecosystems.* 9 (2006) 181–199, <https://doi.org/10.1007/s10021-005-0054-1>.
- [93] T.K. Ho, The random subspace method for constructing decision forests, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (1998) 832–844, <https://doi.org/10.1109/34.709601>.
- [94] C.M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag, New York, 2006.
- [95] Z. Pentek, T. Hiller, T. Liewald, B. Kuhlmann, A. Czermak, IMU-based mounting parameter estimation on construction vehicles, in proceedings of 2017 DGON inertial sensors and systems, ISS 2017, Inst. Electric. Electron. Engin. Inc. (2017) 1–14, <https://doi.org/10.1109/InertialSensors.2017.8171504>.
- [96] Xsens, Active Heading Stabilization (AHS). <https://base.xsens.com/hc/en-us/articles/211809465-Active-Heading-Stabilization-AHS> (accessed October 26, 2020).
- [97] SOMA, College. <https://www.soma-college.nl/> (accessed October 26, 2020).
- [98] Trimble, Trimble SPS851. [https://www.trimble.com/support\\_trl.aspx?Nav=Collction-50145&pt=SPS751,2020](https://www.trimble.com/support_trl.aspx?Nav=Collction-50145&pt=SPS751,2020) and SPS851 (accessed July 23, 2020).