Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo and M. Loper, eds.

AUTOMATED ACTIVE AND IDLE TIME MEASUREMENT IN MODULAR CONSTRUCTION FACTORY USING INERTIAL MEASUREMENT UNIT AND DEEP LEARNING FOR DYNAMIC SIMULATION INPUT

Khandakar M. Rashid
Joseph Louis

School of Civil and Construction Engineering
Oregon State University
Corvallis, Oregon 97330, USA

ABSTRACT

Modular construction is gaining popularity in the USA for several advantages over stick-built methods in terms of reduced waste and time. However, productivity monitoring is an essential part to utilize the full potential of modular construction methods. This paper proposes a framework to automatically measure active and idle time at various workstations in modular construction factories, which essentially dictates the efficiency of production. This cycle time information can be used as inputs for dynamic prediction using simulation modeling. Vibration data were collected from workstations using inertial measurement units (IMUs), and a deep learning network was used to extract active and idle time from the vibration data. The result of this study showed that the proposed methodology can automatically calculate the active and idle time at various workstations with a 2.7% average error. This presents the potential of utilizing sensors and AI with simulation modeling for production monitoring and control.

1 INTRODUCTION AND BACKGROUND

Modular construction techniques are increasingly gaining popularity over the traditional stick-built methods for higher quality in production, and reduced construction time (Lawson et al. 2011). One of such setups is the volumetric construction, ensuring maximum modularized construction inside the factory and shipping the assembled unit to the project location to be installed on-site (Kawecki 2010). The construction process of these modular units in a factory closely resembles a manufacturing production line, where different workstations are dedicated to a specific type of activity (e.g., building floors and walls, installing walls, installing insulation, etc.), and through which the modular unit would travel. Thus each component of the modular unit (e.g., wall, floor, ceiling, etc.), as well as the modular unit spends a different amount of time (i.e., cycle time) at each station based on their particular design specifications. Typically, the workstations of a modular construction factory can be divided into two categories, off-line stations, and online stations. The off-line stations are typically dedicated to the actual creation of panelized components such as walls and floors; while on-line stations are part of the assembly line for the volumetric unit, where various pre-made components are added and assembled to the modular unit. Therefore, the overall productivity of the factory depends on the productivity of both types of workstations. A delay at any of the workstations has the potential to cause bottlenecks in production that can adversely affect the factor's ability to meet demand. Thus, to plan workstation layout and allocate resources to them, it is essential to know how much time is required to complete work at each station and understand how that time is divided between various activities performed and idle times at the stations. Traditionally, the activity cycle times or active/idle times are measured manually using stopwatches and visual observation, or by studying pre-recorded video footage of the operation. These manual approaches are time-consuming, laborious, and error-prone. Moreover, the introduction of new construction and assembly methods renders previous data outdated and thus requires

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

more manual effort to determine the impact of the new process. Thus, automatically measuring the cycle time, and active/idle time at various workstations (i.e., online and offline stations) can play a significant role in reducing the manual effort, and thereby assist in effective production planning and monitoring systems. Eventually this cycle time data can be used as simulation inputs for dynamic prediction of completion time. To that end, this paper presents an automatic active and idle time measurement technique using IMU and a deep learning algorithm. Typically, the major activities (e.g., sanding the floor, placing a wall on the floor, etc.) in the modular construction factory generates various vibration signal patterns. The IMU unit is used to collect those raw vibrations, and then feed them into a deep learning network, long short term memory (LSTM) to identify active and idle states of the workstations. LSTM is a variant of recurrent neural network (RNN) specifically designed to handle time-series sequence data. Eventually, the trained LSTM network is used to predict the active and idle time at different workstations.

In on-site construction, there is an extensive body of work related to identifying different activities of equipment and workers using real-time location system (RTLS) (El-Omari and Moselhi 2011; Ergen et al. 2007), inertial measurement units (IMUs) (Akhavian and Behzadan 2012; Mathur et al. 2015; Rashid and Louis 2019), and computer vision techniques (Golparvar-Fard et al. 2013; Heydarian et al. 2012; Nath et al. 2018). In contrast, offsite construction has not seen such extensive implementation of automated technologies. However, radio frequency identification (RFID) technology was explored in an offsite construction factory to track various components of the building to calculate the cycle time. Azimi et al. (2011) illustrated an automated project monitoring and control framework using high-level architecture and RFID. Altaf et al. (2018) developed a production planning and control system using RFID, data mining, and simulation-based optimization in a panelized home production factory. Moreover, 3D/4D visualization tool was also explored to better perform different activities inside a modular construction factory (Jureidini et al. 2016). Lean tools and techniques have been implemented in modular construction facilities to explore their feasibility in reducing production time and waste (Moghadam 2014; Nahmens and Ikuma 2012; Yu et al. 2013).

Even though tracking components of the building with RFID technology demonstrated potential in calculating cycle time correctly, the location tracking system is not capable of identifying if there is any work done or the building unit is sitting idle at a particular workstation. For example, a floor can spend a couple of hours in a wall installation station, but maybe only half of that time was active and value-adding. Thus, to develop an efficient production planning and monitoring system, it is essential to know the cycle time as well as the active and idle time in the factory. To achieve that goal, this paper utilizes IMU and deep learning networks to automatically measure the active and idle time at workstations in modular construction factories. The following section discusses the methodological steps undertaken in this paper to achieve the abovementioned goal.

2 METHODOLOGY

This section presents the methodological steps of this study shown in Figure 1. Raw data collected from the IMUs are first segmented using the appropriate window size. The raw dataset consists of a 3-axis accelerometer. Gyroscope and magnetometer data are not used in this study as they do not contain any value-adding information in classifying active or idle states. Video data are used to label the dataset with ground truth. After segmentation, the dataset is split into training, validation, and testing set using 70%, 10%, and 20% ratios respectively. Raw acceleration data are used as inputs in the LSTM network. No feature extraction is performed in this study to harness the capability of the LSTM network to automatically extract low-level features from the dataset. An LSTM network architecture with five layers is used to train the classification model. Validation data are used to fine-tune the hyperparameter of the model, such as learning rate, mini-batch size, number of epochs, etc. Training data are used to train the model with adjusted hyperparameters. Testing data are used to evaluate the trained model and to predict the testing data. Accuracy, precision, recall, and F-1 score are used as performance metrics to evaluate the model. Then, these predictions are passed through a window filtering for active and idle time calculation. The following section discusses details regarding segmentation, LSTM architecture, and window filtering.

Figure 1. Flowchart of the methodological steps.

Choosing the appropriate window size to segment data is a particularly challenging task in this study. Different activities of the modular house building produce discrete vibration patterns containing idle signal patterns in between. For example, installing a wall on the floor contains aligning the wall and nailing the bottom plate to the floor. In this case, the first impact of the wall in the floor generates vibration, while the fine-tuning of the alignment before the nailing starts may not generate any signal pattern. Even though these in-between activities may not generate any vibration, they should be considered as an active state. Thus, finding an appropriate window size that correctly corresponds to the state of the workstation is paramount for robust model training. Figure 2 shows a random x-axis vibration of a two-minute segment. This segment contains the active, idle, and moving state of the modular unit. We can see that shorter window size, such as *Window A* is too small to capture enough vibration spikes to correctly represent the active state, while *Window B* is too big as it contains a large portion of two different states. Thus, after carefully analyzing the activities from the video, two different window sizes are selected, 15 seconds for online stations and 25 seconds for offline stations.

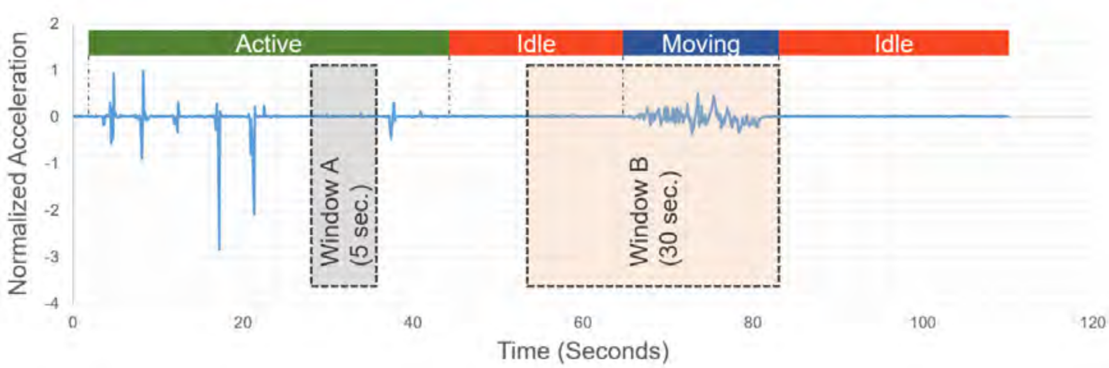


Figure 2. Challenge in selecting the appropriate window size.

A bidirectional LSTM network is used to classify different states of the workstation. Three states (e.g., active, idle, and moving) for online workstations and two states (e.g., active and idle) for offline workstations are considered. The RNN contains five layers as shown in Figure 3.

Figure 3. The network architecture of the LSTM model.

Raw 3-axis accelerometer data are used in the *sequence input layer*. The *LSTM layer* contains 100 hidden units to learn long-term temporal dependencies between time steps of sequence data in terms of weight matrix and bias vector. The *fully connected layer* multiplies the inputs by the weight matrix and adds the bias vector. The *softmax layer* applies a neural transfer function to the input. And finally, the classification output layer computes the cross-entropy loss for multi-class classification problems with mutually exclusive classes (3 for online stations and 2 for offline stations). Mini batch size of 100 is used with a maximum of 40 epochs and a learning rate of 0.001.

3 CASE STUDY AND RESULTS

The dataset contains acceleration data for both online and offline workstations. For online workstations, six modular units were mounted with IMUs under the floor. For each unit, one full working day (i.e., 8 hours) data were collected. For offline workstations, three IMU units were attached under three tables where partial walls are built. Four hours of vibration data were collected for each of the offline stations. Figure 4 shows the IMU attachment positions in online and offline stations. The vibration data were collected with a sampling frequency of 10 Hz.

(a) Online workstation. (b) Offline workstation.

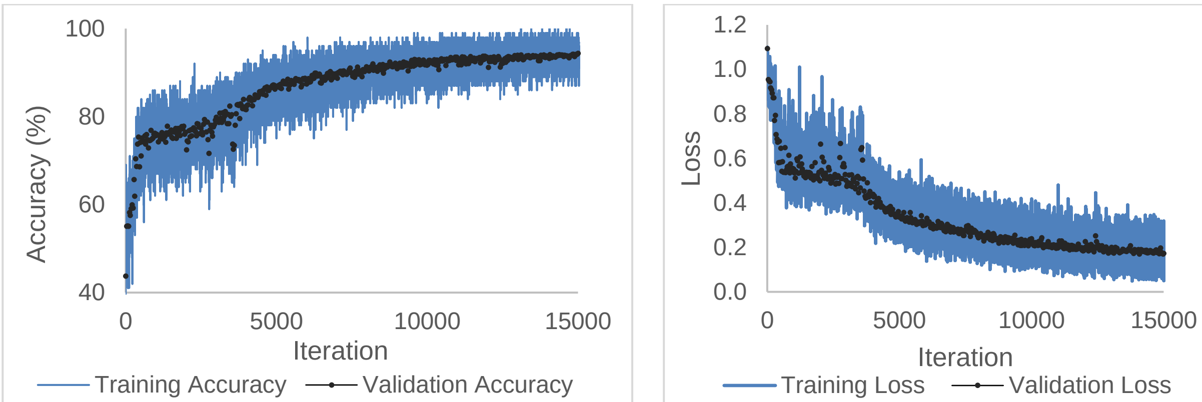
Figure 4. Location of the IMU attachment.

The dataset was then labeled using the reference videos and segmented into sequences using two different window sizes (15 seconds for online and 25 seconds for offline workstations), both with 50% overlap. However, as moving happens only when the modular units travel from one station to another, IMU data regarding *moving* activity are few compared to *active* or *idle*. This can create a bias in the training process. Thus, time-series data augmentation techniques presented in (Rashid and Louis 2019) were used to augment *moving* data. Table 1 shows the data distribution in training, validation, and testing sets.

Table 1. Training, validation, and testing data distribution for online and offline stations.

Class Labels	**Number of Sequences**					
	Online Stations *(15 sec. window size)*			Offline Stations *(25 sec. window size)*		
	Training	Validation	Testing	Training	Validation	Testing
Active	12040	1720	3440	1376	197	393
Idle	20503	2929	5858	459	66	131
Moving	4715	674	1347	N/A		

Online workstations were labeled with three states; active, idle, and moving. As modular units pass through various online stations in the assembly line, it is important to know when a unit moves from one station to the next. This moving state will help to calculate the idle and active time at the different online stations. As offline workstations are stationary, only active and idle states are labeled. The labeled data were split into training, validation, and testing datasets with 70%, 10%, and 20% ratios. Raw acceleration data were used as inputs to the LSTM layer with 100 hidden units. This layer mapped the input sequence into 100 features. The training and validation data set were used to train the model with fine-tuned hyperparameters. Figure 5 shows the training progress of the LSTM network of the online workstations.



(a) Training and validation accuracy. (b) Training and validation loss.

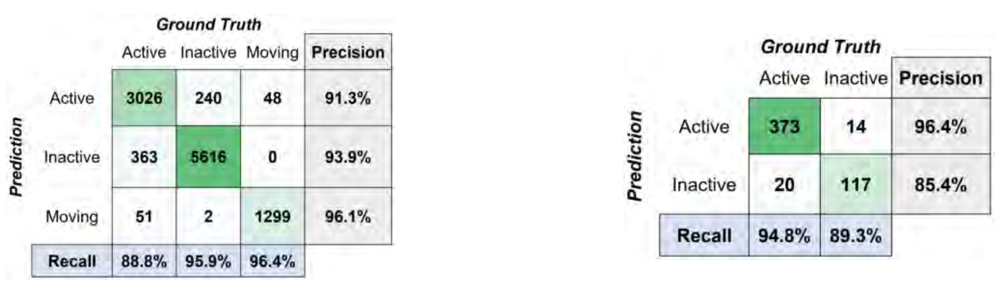
Figure 5. Training progress of the LSTM network for online workstations.

The validation accuracy for online stations was 94.1% and for offline stations was 94.8%. After training the LSTM, the testing data was used to evaluate the performance of the model. Table 2 shows the evaluation results of the model tested with the testing data.

Table 2. Accuracy, precision, recall, and F-1 score of the LSTM model for online and offline stations.

	Accuracy	**Precision**	**Recall**	**F-1 Score**
Online Stations	93.4%	93.4%	93.8%	93.6%
Offline Stations	93.5%	92.1%	90.9%	91.5%

The trained LSTM models demonstrated a 93.6% F-1 score for online stations and a 91.5% F-1 score for offline stations. Even though the F-1 score represents the overall performance of the network, they do not provide information regarding the misclassification of different classes. Thus, a confusion matrix was used to identify the classes that are misclassified as shown in Figure 6.



(a) Online workstations. (b) Offline workstations.

Figure 6. Confusion matrix of LSTM network.

The confusion matrix of online stations shows that *Moving* demonstrates the highest precision and recall. Moreover, *Moving* is mostly misclassified with *Active*, which is understandable as some *Activities* may have a very similar signal pattern as *Moving*.

Next, one random section was selected from the dataset for both online and offline workstations to calculate active and idle time from the trained model. An 8.5-hour dataset for the offline station was used as input to the trained model. Figure 7 shows the ground truth and the prediction for the online stations. We can see there were two *Moving* instances in the ground truth, where five were predicted by the model. A closer look reveals that each of the three misclassified *Moving* instances occurred during the *Active* class. The original dataset contained very few *Moving* classes, as, after a couple of hours of activities, the modular units are moved to the next station using an electric pusher, which takes about 15 to 30 seconds. To balance the dataset, augmentation techniques were used to generate synthetic *Moving* data. A similar plot was drawn for the offline stations in Figure 8. Figure 8 contains only the *Idle* and *Active* state of the offline workstations.

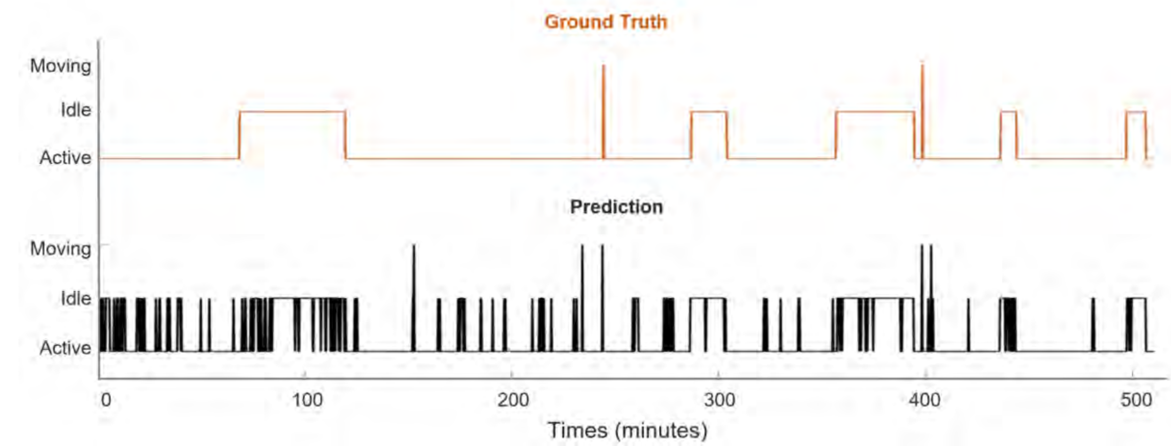


Figure 7. Ground truth and prediction of the LSTM model for online stations.

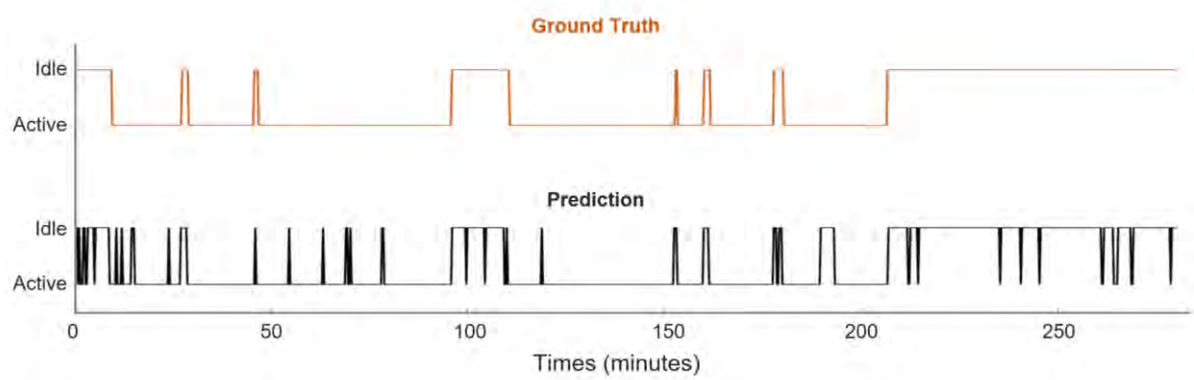


Figure 8. Ground truth and prediction of the LSTM model for offline stations.

Finally, idle and active time was calculated using the predictions shown in Table 3. We see that, for offline stations, out of the 8.5 hours 387.2 minutes were active in reality, and the prediction was 380 minutes with a 1.8% error. Similarly, the online workstation showed a 1.4% error in calculating active time from the prediction.

Table 3. Active and idle time calculation from the trained LSTM model.

	Time (minutes)			
	Offline Stations *(8.5 hours)*		**Online Stations** *(4.6 hours)*	
	Active	Idle	Active	Idle
Ground Truth	387.20	123.30	175.00	105.88
Prediction	380.00	129.70	177.50	103.33
Error	1.80%	5.19%	1.40%	2.40%

4 DISCUSSION AND CONCLUSION

Tracking the active and idle time at various workstations in a modular construction factory can be a key step towards productivity assessment. To address this, this paper investigated the potential of vibration generated from the activities performed to identify the active and idle time using a deep learning approach. One of the major challenges was to choose the appropriate window size as an *active* sequence can have multiple *idle* durations in between, and considering those as idle time is not purely logical. Thus, a larger window size (i.e., 15 seconds for online stations and 25 seconds for offline stations) was considered in this study. The trained LSTM network demonstrated a 93.6% F-1 score for online stations and a 91.5% F-1 score for offline stations. The prediction of the LSTM network showed the capability of automatically measuring active and idle time with an average error of 2.7%. Measuring the cycle time of separate workstations proved to be challenging as a single misclassified *moving* state can yield to potentially higher error rate in active and idle time calculation. This work can be further extended by using this cycle time information into a simulation model for the dynamic prediction of completion time. Moreover, combining location tracking systems, such as RFID with the proposed system, can be used to detect the movement of the unit from one station to another and a deep learning model can calculate the active and idle time within each workstation. The primary limitation of this work is, some activities such as painting, sanding has the higher potential of not generating enough vibration to distinguish between active and idle state. Future research will be extended for other workstations where little or no vibration is generated by utilizing computer vision and machine hearing techniques. Overall, the proposed framework demonstrated the potential of using inexpensive sensors and artificial intelligence to track active and idle time in modular construction factories, which can be used for effective production planning, monitoring, and control system.

REFERENCES

Akhavian, R., and Behzadan, A. H. 2012. "An integrated data collection and analysis framework for remote monitoring and planning of construction operations." *Advanced Engineering Informatics*, 26(4), 749–761.

Altaf, M. S., Bouferguene, A., Liu, H., Al-Hussein, M., and Yu, H. 2018. "Integrated production planning and control system for a panelized home prefabrication facility using simulation and RFID." *Automation in Construction*, 85(September 2017), 369–383.

Azimi, R., Lee, S., Abourizk, S. M., and Alvanchi, A. 2011. "A framework for an automated and integrated project monitoring and control system for steel fabrication projects." *Automation in Construction*, 20(1), 88–97.

El-Omari, S., and Moselhi, O. 2011. "Integrating automated data acquisition technologies for progress reporting of construction projects." *Automation in Construction*, 20(6), 699–705.

Ergen, E., Akinci, B., East, B., and Kirby, J. 2007. "Tracking Components and Maintenance History within a Facility Utilizing Radio Frequency Identification Technology." *Journal of Computing in Civil Engineering*, 21(1), 11–20.

Golparvar-Fard, M., Heydarian, A., and Niebles, J. C. 2013. "Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers." *Advanced Engineering Informatics*, 27(4), 652–663.

Heydarian, A., Golparvar-Fard, M., and Niebles, J. C. 2012. "Automated Visual Recognition of Construction Equipment Actions Using Spatio-Temporal Features and Multiple Binary Support Vector Machines." *Construction Research Congress*, ASCE, May 7th, West Lafayette, Indiana, 889–898.

Jureidini, F., Afifi, M., Zhang, R., and Al-Hussein, M. 2016. "3D / 4D Visualization Framework for Modelling off-site Productivity of Modular Construction." *Modular and Offsite Construction (MOC) Summit*, September 29th, Alberta, Canada, 83–89.

Kawecki, L. R. 2010. "Environmental performance of modular fabrication: Calculating the carbon footprint of energy used in the construction of modular home." Doctoral Dissertation, School of Construction, Arizona State University. https://repository.asu.edu/items/8739

Lawson, R. M., Ogden, R. G., and Bergin, R. 2011. "Application of Modular Construction in High-Rise Buildings." *Journal of Architectural Engineering*, 18(2), 148–154.

Mathur, N., Aria, S. S., Adams, T., Ahn, C. R., and Lee, S. 2015. "Automated Cycle Time Measurement and Analysis of Excavator's Loading Operation Using Smart Phone-Embedded IMU Sensors." *International Workshop on Computing in Civil Engineering*, ASCE, June 21st, Austin, TX, 215–222.

Moghadam, M. 2014. "Lean-Mod: An Approach to Modular Construction Manufacturing Production Efficiency Improvement." Doctoral Dissertation, Department of Civil and Environmental Engineering, University of Alberta. https://doi.org/10.7939/R3NH4F

Nahmens, I., and Ikuma, L. H. 2012. "Effects of lean construction on sustainability of modular homebuilding." *Journal of Architectural Engineering*, 18(2), 155–163.

Nath, N. D., Chaspari, T., and Behzadan, A. H. 2018. "Automated ergonomic risk monitoring using body-mounted sensors and machine learning." *Advanced Engineering Informatics*, 38(2018), 514–526.

Rashid, K. M., and Louis, J. 2019. "Times-series data augmentation and deep learning for construction equipment activity recognition." *Advanced Engineering Informatics*, 42(April), 100944.

Yu, H., Al-Hussein, M., Al-Jibouri, S., and Telyas, A. 2013. "Lean transformation in a modular building company: A case for implementation." *Journal of Management in Engineering*, 29(1), 103–111.

AUTHOR BIOGRAPHIES

KHANDAKAR M. RASHID is a Ph.D. student in the School of Civil and Construction Engineering at Oregon State University. He completed his M.Sc. in Project Management from Missouri State University and a B.Sc. in Civil Engineering from Bangladesh University of Engineering and Technology. His research interests include sensing, data analytics, simulation modeling and optimization, and artificial intelligence applications in construction. His email address is rashidk@oregonstate.edu and mamunlinkon2011@gmail.com. His website is https://mamunlinkon2011.wixsite.com/khandakarrashid

JOSEPH LOUIS is an assistant professor of Civil and Construction Engineering at Oregon State University. He earned his Ph.D. and M.Sc. from the School of Civil Engineering at Purdue University. His research interests include simulation, visualization, and automation within the context of construction operations to improve the safety, productivity, and sustainability of construction. His email address is joseph.louis@oregonstate.edu. His websites are http://research.engr.oregonstate.edu/louis/research and https://josephlouis.org/