

RESEARCH ARTICLE

Calculation of the productivity of construction gypsum plaster worker using support vector machine algorithm

İbrahim Karataş^{id}, Abdulkadir Budak^{id}

Osmaniye Korkut Ata University, Department of Civil Engineering, Osmaniye, Türkiye

Article History

Received 07 July 2024

Accepted 07 October 2024

Keywords

Construction technology

Worker activities

Support Vector Machine

Worker productivity estimation

Artificial Intelligence

Abstract

Calculating the productivity of workers in traditional construction projects can be a daunting and time-consuming task. However, thanks to the advancements in technology and scientific research, measuring worker productivity can be automated. Therefore, this study aims to determine the productivity of gypsum plaster workers by collecting motion data with a sensor on their arms and using the support vector machine algorithm for analysis. Based on the estimation results, motion productivity is determined by calculating the ratio of the worker's working time to the total time. On the other hand, worker productivity is calculated by determining the ratio of the amount of work completed to the product of the number of workers and the total working hours. Finally, theoretical productivity is calculated by dividing motion productivity by the calculated worker productivity. According to the results of the analysis, whether the worker is working or not has been estimated correctly by 95.8%. On the other hand, the mean daily theoretical productivity has been determined to be 10.20 m²/man-hour. In this way, workers' activities can be automatically detected with a certain accuracy and their productivity can be calculated. This helps in the effective management of the construction site.

1. Introduction

The construction industry is a vital sector in the global economy, but it often faces challenges such as low productivity resulting in low quality, high costs, and time overruns. Measuring and estimating worker productivity in this industry can be challenging due to the variability of worker productivity depending on the labor force used. This creates difficulties in accurately estimating the worker productivity of the construction sector. In construction planning and process, measuring and estimating worker productivity is crucial. Determining worker productivity status helps

manage the process effectively. Incorporating new technologies and advanced methods to measure and forecast worker productivity is crucial for addressing concerns and improving the construction industry's efficiency and effectiveness. The integration of modern technology is key to enhancing worker productivity in construction, as per various studies [1-3]. According to a study in the USA, the construction sector has one of the lowest scores on the digitalisation index [4]. Today, with the convenience of digitalisation, scientific studies have been conducted to integrate technology into the construction industry [5-9, 21].

Correspondence İbrahim Karataş

 ibrahimkaratas@osmaniye.edu.tr

eISSN 2630-5771 © 2024 Authors. Publishing services by Golden Light Publishing®.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

In 2018, McKinsey & Company published a report stating that the construction industry will increasingly rely on advanced technologies. These technologies encompass digital twins, artificial intelligence, the Internet of Things (IoT), 3D printers, robotics, real-time monitoring and control systems, document management, and 3D modelling and BIM technologies [10]. The integration of cutting-edge technologies such as artificial intelligence, internet of things, real-time monitoring, and control can serve as a catalyst for automating the recognition of labor activities and calculating worker productivity in the construction sector. This advancement has the potential to significantly streamline processes and increase efficiency on construction sites.

In 2016, a study by Akhavan and Behzadan explored the use of an artificial intelligence model that could automatically detect worker movements using data from smartphones. The study analysed the effectiveness of various AI models, including artificial neural network, decision trees, nearest neighbour, logistic regression, and support vector machines, in identifying the time spent on each activity, including idle time. The findings suggest that utilizing machine learning techniques to analyse worker activity and time can be a valuable tool in evaluating productivity. The models created by the study can be directly applied to measure productivity [11]. Sanhudo et al. (2021) conducted a thorough study to assess the accuracy and usability of accelerometer data and machine learning algorithms in categorizing construction workers' activities. To achieve this, a realistic circuit was created to capture ten distinct activities that occur during wall building in a controlled laboratory environment. The workers were outfitted with accelerometer sensors to collect data for each activity. The data collected from the sensors were used to train and compare 13 different machine learning algorithms including decision tree, k nearest neighbour, logistic regression, multilayer perceptions, and support vector machines (SVM). Additionally, ensemble machine learning algorithms such as random forest, extreme random trees, adaptive boosting, and gradient boosting

machines were also employed for comparison. [12]. In Son's (2017) study, the author examined the possibility of measuring worker productivity by calculating the ratio of active working hours to total working hours, based on data obtained from workers. To enhance productivity in a construction site setting, the author suggests that gathering physical and location data every second can be beneficial since productivity data can be difficult to collect and measure in real-time. The study defines worker productivity as the ratio of time spent on direct work by a construction worker to total work time. Furthermore, the study focuses on measuring the individual productivity of each worker rather than the project's overall productivity [13]. In their study, Karatas and Budak (2021) [14] collected 3-axis accelerometer, gyroscope, and magnetometer data from workers using sensors to estimate the activities involved in a construction job. The data were then used to train and model basic machine learning algorithms such as logistic regression, Support Vector Classifier Decision tree, and K Nearest Neighbours. The results showed that the Support Vector Classifier algorithm provided the best prediction with an accuracy rate of 90%. The accuracy rates for the other algorithms were 87% for K Nearest Neighbours, 80% for logistic regression, and 80% for Decision tree.

The recognition of construction activities involves the utilization of specialized equipment. Specifically, the process entails the measurement of acceleration, angular velocity, and gravitational forces using a combination of three sensors - accelerometer, gyroscope, and magnetometer. The data obtained from these sensors are then analyzed to identify the specific construction activities taking place. It is worth noting that the accuracy of the results is dependent on the quality of the equipment used and the expertise of the personnel involved [15]. Some studies have used only accelerometer data for activity recognition in construction work [12, 16, 17]. The inclusion of both accelerometer sensors in the X, Y, and Z axes and gyroscope sensors that measure angular velocity allows for the recognition of construction activities. These sensors can detect rolling, pitching, and yawing movements

in the three axes, thereby expanding the range of worker activity movement characteristics. As a result, it becomes simpler to differentiate between each activity [18]. Therefore, some studies have used a combination of accelerometer and gyroscope data for construction worker activity identification analyses [11, 18, 19]. In addition [20] stated in their study that gravitational force measurements data collected with the magnetometer sensor would be useful to be used with other sensors in the recognition of labour activities. The aim of this study is to examine the information collected from employees through accelerometer, gyroscope, and magnetometer sensors. Through the use of artificial intelligence models, their movements and behaviours will be identified, enabling the computation of worker productivity to be automated.

In previous studies, data was mostly collected in laboratory settings using accelerometers or a combination of accelerometers and gyroscopes. These studies focused solely on identifying workers' movements. However, our study collected data from workers in a real construction site environment, using accelerometers, gyroscopes,

and magnetometers to capture more detailed movements. We not only calculated the workers' movement efficiency but also their theoretical efficiency. This approach sets our study apart from previous literature.

2. Research Methodology

As technology advances, data collection systems are becoming increasingly efficient at measuring worker productivity. Compared to traditional methods like direct observation or surveys, which can be time-consuming, tedious, and prone to errors, these automated systems offer a more effective solution. In light of this, our study seeks to utilize machine learning to analyze data collected via sensors attached to workers' arms, providing predictions of their activities and accurately calculating productivity. This study employs the support vector machine (SVM) machine learning method to accurately predict the work status of a Gypsum Plaster worker. To achieve this, we followed a step-by-step approach illustrated in the Fig. 1 flowchart.

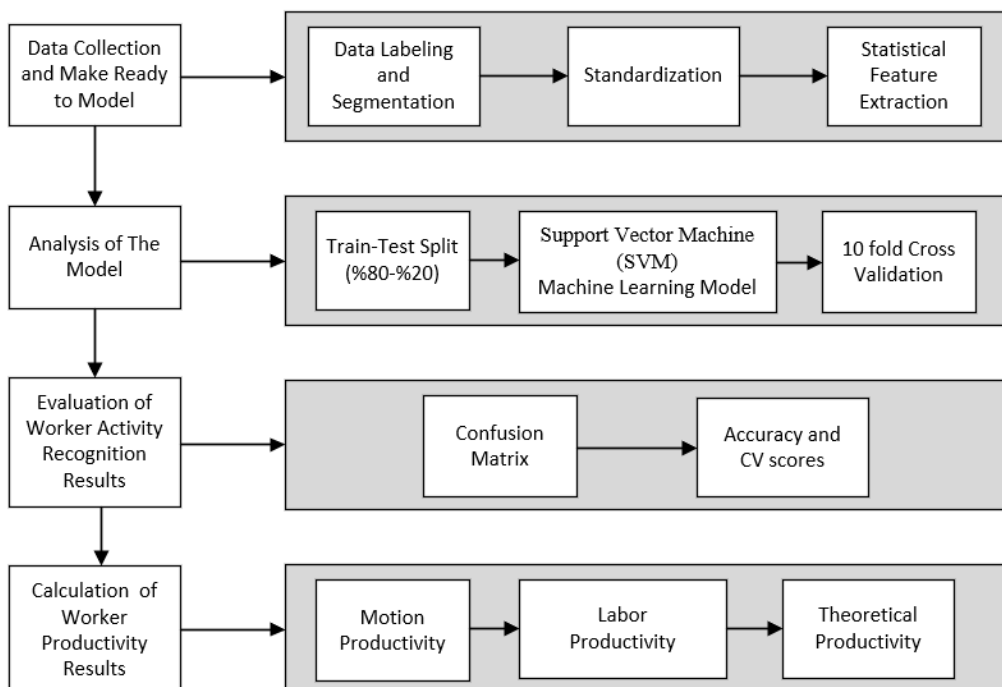


Fig. 1. Flowchart for the study

We began by collecting and categorizing raw data from the worker at the construction site. Subsequently, we meticulously processed, sorted, and extracted its essential attributes to ensure the data was model-ready. The initial step involves dividing the data into two distinct sets: training and test data. Subsequently, the training data will be meticulously analysed to determine the success of predictions in comparison with the test data. Once this is accomplished, the prediction model will be integrated with the data collected during the day to forecast worker activities. Finally, the theoretical efficiency will be calculated based on movement efficiencies and worker efficiencies.

In this study, we gathered data on gypsum plastering activities during the construction of a reinforced concrete school. Fig. 2 shows a sensor attached to the worker's arm with an armband. The skilled worker plasterer in Fig. 3 wore this sensor, and it recorded 3-axis acceleration, gyroscope, and magnetometer data throughout the process. Data

was collected throughout the day at 2-hour intervals between 8:00-10:00, 10:00-12:00, 13:00-15:00 and 15:00-17:00. To label the data according to the activity it belonged to, the workers' work at the construction site was observed and noted. The data was gathered with the assistance of a data acquisition device and subsequently analyzed utilizing the Python programming language. To elevate the predictive potential of the machine learning models, the data underwent normalization through the z-score standardization method subsequent to its collection. Based on the data collected, the accelerometer, gyroscope, and magnetometer sensors gather information at a rate of 25 Hz, which means that 25 data points are obtained per second. The accelerometer provides data in gravitational acceleration (g) units, the gyroscope sensor records values in degrees per second (deg/s) related to angle, and the magnetometer measures magnetic current density in Tesla (T).

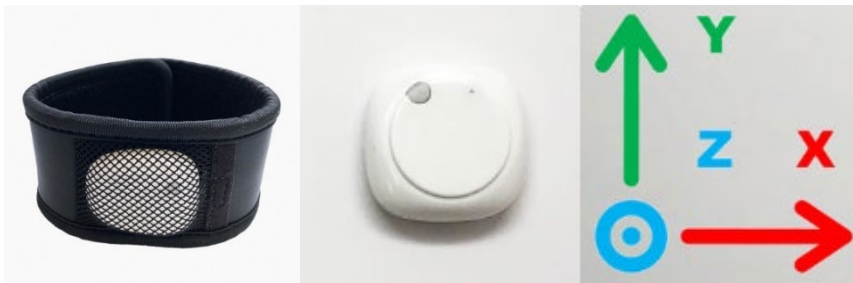


Fig. 2. Sensor attached to the worker's arm with an armband and x, y, and z directions



Fig. 3. Data collection with a sensor attached to the gypsum plaster workers' arm

Initially, the data will undergo separation and feature extraction processes to obtain the necessary features for prediction. The data obtained were partitioned into windows of a specific duration, and the statistical characteristics of the data within each window were computed. For this study, a window size of 4 seconds and an overlap rate of 75% were chosen. The 4-second windows were then shifted, as depicted in Fig. 4, and all windows were computed with a 75% overlap. The new features were calculated by analyzing the statistical properties of the data within the specified windows. These statistical features include the sum, median, mean, length, standard deviation, variance, square root (quadratic) mean, maximum and minimum of the values in each window of the data in the x, y, and z directions taken from the accelerometer, gyroscope, and magnetometer sensors as shown in Table 1. A total of 81 variables were created using this method. After this process, the new data obtained will be ready to be used for the model.

The data that was prepared for the model was initially divided into 80% training data and 20% test data. The training data was then analyzed, and prediction accuracies were calculated by comparing them with the test data. For this study, the machine learning model used was the Support Vector Machine (SVM) algorithm. After developing the SVM model, we utilized the 10-fold cross-

validation technique to ensure the model was unbiased and minimize data bias. This involved dividing the training data into 10 layers and using one layer as validation data while the remaining 9 layers were used as validation-training data. The prediction accuracy values were averaged for each layer to calculate the cross-validation prediction values.

The aim of gypsum plastering is to gauge worker productivity using state-of-the-art Artificial Intelligence techniques. In order to achieve this, data was gathered from accelerometers, gyroscopes, and magnetometers along the x, y, and z-axes over an 8-hour period. However, only the 2-hour measurements were scrutinized. For the remaining 6 hours, a Support Vector Machine (SVM) learning algorithm was employed to predict the measurements and create a model for analysis. This model aims to classify whether a worker is working based on their movement. Using this trained classification model, we predicted the working and non-working conditions of the worker with the remaining 6 hours of working data. Then, we calculated the times belonging to these classes.

The resulting model was then used to assess overall productivity. In order to utilize the SVM machine learning model to forecast construction activities, it is essential to prepare the data collected throughout the day in two-hour intervals.

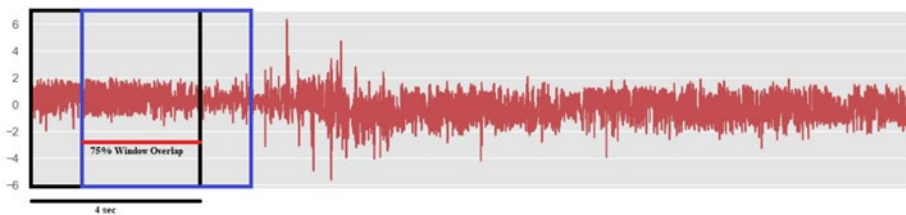


Fig. 4. Data segmentation process

Table 1. Features and description created with statistical feature extraction

Statistical Features Names	Description
Sum values	Calculates the sum over the seperated windows values
Median	Calculates the median the seperated windows values
Mean	Calculates the mean the seperated windows values
Length	Calculates the length the seperated windows values
Standard deviation	Calculates the standard deviation the seperated windows values
Variance	Calculates the variance the seperated windows values
Root mean score	Calculates the quadratic mean the seperated windows values
Maximum	Calculates the maximum value the seperated windows values
Minimum	Calculates the minimum value the seperated windows values

Once the data is prepared, it is analysed each two-hour measurement using the prediction model to obtain anticipated activity times. These predicted times will then be compared to the observed activity times. To assess the effectiveness of each activity during the 2-hour working periods, we will calculate their motion productivity based on the predicted and observed values. Motion productivity will be calculated by dividing the working time of each activity by the total time, which includes both working time and waiting time. The formula for calculating motion productivity is shown in Equation 1.

$$\text{Motion Productivity (\%)} = \frac{\text{Working Time (sec)}}{\text{Total Time (sec)}} \times 100 \quad (1)$$

The workers' motions and the amount of work done in 2-hour periods were measured at construction sites. This is done to calculate worker productivity values during these periods. Equation 2 is used to calculate worker productivity values.

$$\text{Worker Productivity} = \frac{\text{Total Output}}{\text{Total Input}} \quad (2)$$

After calculating motion and worker productivities, the theoretical productivity values

that a worker would achieve if it were worked at full capacity were calculated according to Equation 3.

$$\frac{\text{Theoretical Productivity}}{\text{Labor Productivity}} = \frac{\text{Motion Productivity}}{\text{Motion Productivity}} \quad (3)$$

3. Results and Discussions

A SVM machine learning model was designed to accurately detect the working status of a gypsum plaster at a construction site. The model was meticulously scrutinized, and the resulting confusion matrices have been conveniently presented in Fig. 5. The analysis reveals that the model was able to predict with 96% accuracy (704 out of 740) the total waiting activities. Furthermore, the model accurately predicted 97% (622 out of 644) of the total gypsum activities.

Confusion matrices are a useful tool to determine the predictive success of any activity in SVM model. In this study, the predictive success of the SVM model is presented in the Table 2, considering the overall prediction success.

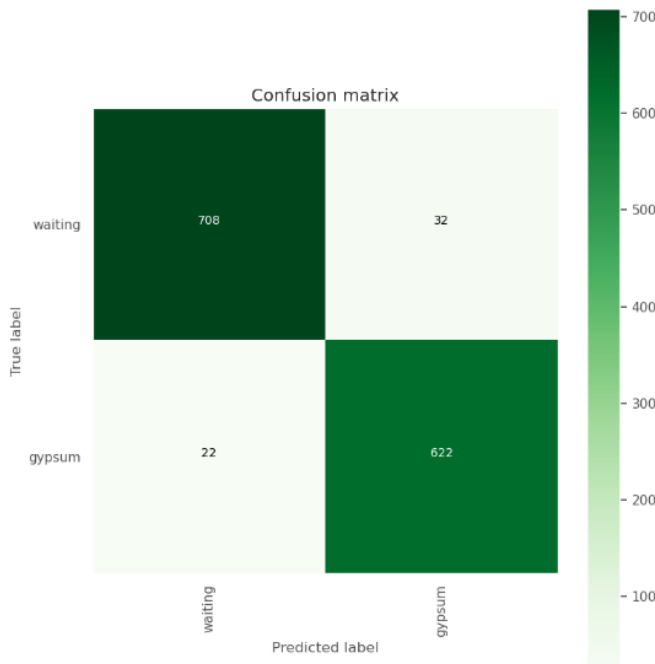


Fig. 5. Confusion matrix of SVM model

Table 2. Results of activity recognition

	Training Accuracy	10-fold CV Accuracy	Number of Data
SVM (C:100)	0.961	0.958	1384

Accuracy scores of the models were evaluated in two ways. Firstly, the model was assessed using training accuracy scores, and then accuracy values were generated using the 10-fold cross-validation (CV) technique, which is a more accurate and robust method for evaluating models. Upon thorough analysis of the results, it has been determined that the CV accuracy value of the model is considerably lower by a factor of 10. Nonetheless, it has been discovered that an evaluation based on these values would yield greater accuracy and robustness. Consequently, it has been observed that the SVM model boasts an impressive accuracy rate of 95.8% in predicting the working status of a plaster worker.

Once the prediction results of the SVM model were deemed acceptable, the data collected throughout the day was prepared and fed into the SVM model to obtain the prediction results. During the data collection process from sensors, the system also kept track of the time period in which the master worked on a particular job and the amount of work he accomplished. The aim is to calculate workers' estimated productivity throughout the day and compare it with their worker productivity. The

results are presented in Table 3. It shows that the observed and predicted working hours and motion productivity are very similar in all working hours. The estimated motion productivity during 8-10, 10-12, 13-15, and 15-17 was approximately 47%, 79%, 88%, and 63%, respectively.

Theoretical productivities were calculated by dividing measured worker productivity by calculated motion productivity. Table 4 shows the theoretical productivity of gypsum activity for each time frame. Theoretical productivity indicate how productive activities would be if they worked continuously during the designated working hours, without any waiting periods. Based on the results, the theoretical productivities of the activities were calculated to be 8.50, 11.33, 7.42, and 13.56, respectively.

When the productivity calculations of the gypsum plastering activity are analysed separately, the theoretical productivity and partial productivity of the construction activities in each time zone are visually evaluated and how much the workers work in proportion to their daily working capacity is shown in Fig. 6.

Table 3. Results of motion productivity

Activity	Working Hours	Work	Time (sec)		Motion Productivity (%)	
			Predict	Observed	Predict	Observed
Gypsum Plaster	8-10	Gypsum	3257	3231.1	47.05	47.41
		Waiting	3665	3583.69		
	10-12	Gypsum	5494	5430	79.44	79.74
		Waiting	1422	1379.5		
	13-15	Gypsum	6036	5954.89	87.59	87.73
		Waiting	855	833.048		
	15-17	Gypsum	4336	4282.88	62.68	62.88
		Waiting	2582	2528.43		

Table 4. Results of theoretical productivity

Activity	Working Hours	Worker Productivity (m ² /1 man-hour)	Motion Productivity	Theoretical Productivity	Unit
Gypsum Plaster	8-10	4.00	0.4705	8.50	m ² /man*hour
	10-12	9.00	0.7944	11.33	
	13-15	6.50	0.8759	7.42	
	15-17	8.50	0.6268	13.56	

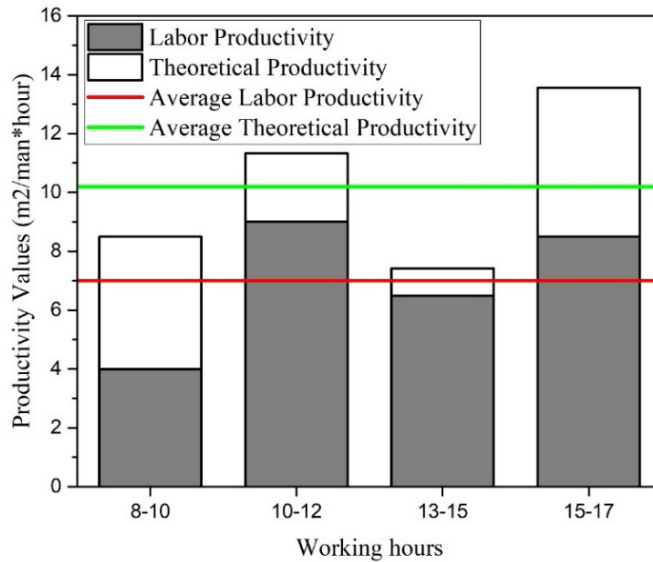


Fig. 6. Daily productivity values of the worker

Based on the findings, the gypsum activity demonstrated higher productivity during the 10-12 and 15-17 hour periods compared to other hours of the day. When comparing productivity values during these two periods, the worker productivity were found to be almost equal. However, the theoretical productivity between 15-17 hours is quite high. In other words, it is seen that the laborer works less in terms of working hours in this time period. On the other hand, when the productivity values between 8-10 and 13-15 hours are compared, the theoretical productivity is almost the same in these hours. It was found that the worker productivity value was higher between 13-15 hours of work. This means that although the worker's working capacity was similar during these time periods, they worked more in terms of working hours between 13-15 hours. In consideration of the daily productivity of a gypsum plaster worker, it is observed that should the worker continue to work without any breaks, the worker productivity can be increased from 7 m²/man*hour per day to a theoretical productivity of 10.20 m²/man*hour. In light of the fact that laborers are required to take a half-hour break prior to and following lunch over the course of the day, the maximal theoretical efficiency in gypsum plastering activity computes to 8.93 m²/man*hour.

4. Conclusions

Automatically recognizing labor activities and measuring productivity in real-time can have a positive impact on increasing productivity and aiding project managers in site control and management. By using sensors to capture motion data, the SVM algorithm analyse the data and determine with 95.8% accuracy whether a worker is present and working on the construction site. By monitoring the workers in real-time, it becomes possible to have better control over the construction site. Automatic recognition of their activities also allows for the calculation of their motion productivity. The theoretical efficiency is determined by comparing the worker productivity obtained from the construction site with the motion productivity. The concept of theoretical productivity pertains to a worker's ability to perform tasks based on their movements. In essence, enhancing worker productivity can be achieved by precisely gauging the work performed by workers and their theoretical productivity. This approach will enable us to identify the opportune time and method to augment worker productivity. Through this, we can optimize worker productivity at construction sites and guarantee more streamlined and effective project management. In

this case, the theoretical and partial productivity calculations of the activities should be calculated on the basis of the workers working on the basis of the activity, which will lead to more realistic and applicable results. Since the productivity of each activity will vary depending on the worker, it is clear that the productivity of which activity can be increased or not can vary. As a result, making an automatic productivity calculation based on the worker will have a significant positive effect on productivity. This study only collected data from a

single worker, focusing solely on plastering activity over the course of an 8-hour workday with measurements taken every 2 hours. In future studies, it would be beneficial to evaluate the entire team working on an activity rather than just one worker. This methodology can also be applied to other activities on construction sites. Furthermore, data collection and analysis could be conducted throughout the activity rather than just for one day, allowing for more effective and automated productivity analysis in project management.

Declaration

Funding

This research was partially funded by Osmaniye Korkut Ata University Scientific Research Projects Unit under grant number OKÜBAP-2022-PT2-008.

Author Contributions

İ Karataş: Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Writing-Original Draft, Writing-Review & Editing. A. Budak: Conceptualization, Methodology, Resources, Writing-Review & Editing, Visualization.

Acknowledgments

Not applicable.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Ethics Committee Permission

Not applicable.

Conflict of Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- [1] Budak A, Karatas I (2022) Impact of the BIM system in construction management services in developing countries: Case of Turkey. *Pamukkale University Journal of Engineering Sciences* 28:828-839.
- [2] Karatas I, Budak A (2023) Investigating the impact of lean-BIM synergy on labor productivity in the construction execution phase. *Journal of Engineering Research* 11:322-333.
- [3] Karatas I, Budak A (2023) Influence on construction waste of the lean-building information modelling interaction. *Proceedings of the Institution of Civil Engineers - Engineering Sustainability* 40:1-14.
- [4] Manyika J, Ramaswamy S, Khanna S, Sarrazin H, Pinkus G, Sethupathy G, Yaffe A (2015) *Digital America: A Tale of the Haves and Have-Mores*.
- [5] Xu Y, Zhou Y, Sekula P, Ding L (2021) Machine learning in construction: From shallow to deep learning. *Developments in the Built Environment* 6:100045.
- [6] Karatas I, Budak A (2024) Development and comparative of a new meta-ensemble machine learning model in predicting construction labor productivity. *Engineering, Construction and Architectural Management* 31:1123-1144.
- [7] Akinosho TD, Oyedele LO, Bilal M, Ajayi AO, Delgado MD, Akinade OO, Ahmed AA (2020) Deep learning in the construction industry: A review of present status and future innovations. *Journal of Building Engineering* 32:101827.

- [8] Begić H, Galić M (2021) A Systematic Review of Construction 4.0 in the Context of the BIM 4.0 Premise. *Buildings* 11:337.
- [9] Calvetti D, Mêda P, Chichorro Gonçalves M, Sousa H (2020) Worker 4.0: The Future of Sensored Construction Sites. *Buildings* 10:169.
- [10] Blanco JL, Mullin A, Pandya K, Parsons M, Ribeirinho MJ (2018) Seizing opportunity in today's construction technology ecosystem. McKinsey & Company.
- [11] Akhavian R, Behzadan AH (2016) Smartphone-based construction workers' activity recognition and classification. *Automation in Construction* 71:198-209.
- [12] Sanhudo L, Calvetti D, Martins JP, Ramos NM, Mêda P, Gonçalves MC, Sousa H (2021) Activity classification using accelerometers and machine learning for complex construction worker activities. *Journal of Building Engineering* 35:102001.
- [13] Son W (2017) Exploring the Feasibility of Measuring Individual Labor Productivity Using a Wearable Activity Tracker. Master Dissertation, University of Texas.
- [14] Karatas I, Budak A (2021) Prediction of labor activity recognition in construction with machine learning algorithms. *Icontech International Journal of Surveys, Engineering, Technology* 5:38-47.
- [15] Fang Y, Cho YK, Chen J (2016) A framework for real-time pro-active safety assistance for mobile crane lifting operations. *Automation in Construction* 72:367-379.
- [16] Ryu J, Seo J, Jebelli H, Lee S (2019) Automated action recognition using an accelerometer-embedded wristband-type activity tracker. *Journal of Construction Engineering and Management* 145(1):4018114.
- [17] Joshua L, Varghese K (2014) Automated recognition of construction labour activity using accelerometers in field situations. *International Journal of Productivity and Performance Management* 63:841-862.
- [18] Zhang M, Chen S, Zhao X, Yang Z (2018) Research on construction workers' activity recognition based on smartphone. *Sensors* 18(8):2667.
- [19] Akhavian R, Behzadan AH (2018) Coupling human activity recognition and wearable sensors for data-driven construction simulation. *Journal of Information Technology in Construction* 23:1-15.
- [20] Garcia-Gonzalez D, Rivero D, Fernandez-Blanco E, Luaces MR (2020) A public domain dataset for real-life human activity recognition using smartphone sensors. *Sensors* 20(8):2200.
- [21] Guvel ST, Karatas I (2023) A novel special length rebar order approach based on AI optimization techniques for reduction of rebar cutting waste. *Journal of Construction Engineering, Management & Innovation* 6(4):285-296.