

Article

From Activity Recognition to Simulation: The Impact of Granularity on Production Models in Heavy Civil Engineering

Anne Fischer ^{1,*}, Alexandre Beiderwellen Bedrikow ¹, Iris D. Tommelein ², Konrad Nübel ³
and Johannes Fottner ¹

¹ Chair of Materials Handling Material Flow Logistics, TUM School of Engineering and Design, Technical University of Munich, 85748 Garching, Germany

² Civil and Environment Department, Project Production Systems Laboratory (P2SL), University of California, Berkeley, CA 94720-1712, USA

³ Chair of Construction Process Management, TUM School of Engineering and Design, Technical University of Munich, 80333 Munich, Germany

* Correspondence: anne.fischer@tum.de

Abstract: As in manufacturing with its Industry 4.0 transformation, the enormous potential of artificial intelligence (AI) is also being recognized in the construction industry. Specifically, the equipment-intensive construction industry can benefit from using AI. AI applications can leverage the data recorded by the numerous sensors on machines and mirror them in a digital twin. Analyzing the digital twin can help optimize processes on the construction site and increase productivity. We present a case from special foundation engineering: the machine production of bored piles. We introduce a hierarchical classification for activity recognition and apply a hybrid deep learning model based on convolutional and recurrent neural networks. Then, based on the results from the activity detection, we use discrete-event simulation to predict construction progress. We highlight the difficulty of defining the appropriate modeling granularity. While activity detection requires equipment movement, simulation requires knowledge of the production flow. Therefore, we present a flow-based production model that can be captured in a modularized process catalog. Overall, this paper aims to illustrate modeling using digital-twin technologies to increase construction process improvement in practice.

Keywords: digital twin in construction; heavy civil engineering equipment; process reference model; discrete-event simulation; deep learning; activity recognition



Citation: Fischer, A.; Beiderwellen Bedrikow, A.; Tommelein, I.D.; Nübel, K.; Fottner, J. From Activity Recognition to Simulation: The Impact of Granularity on Production Models in Heavy Civil Engineering. *Algorithms* **2023**, *16*, 212. <https://doi.org/10.3390/a16040212>

Academic Editor: Frank Werner

Received: 20 February 2023

Revised: 26 March 2023

Accepted: 11 April 2023

Published: 18 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In Germany, the introduction of digitization in the manufacturing industry called Industry 4.0 initiated the fourth industrial revolution [1]. Industry 4.0 includes the implementation of sensors and embedded systems on equipment to make equipment “smart”. This digitization, also known as the Internet of Things (IoT), helps to interconnect decentralized systems to form a “system of systems” (aka. a Cyber Physical System (CPS)) and to link the real world to its digital representation (aka. a Digital Twin (DT)) [2].

Like the manufacturing industry, the construction industry can benefit from the application of Industry 4.0 and its technologies, which is known as Construction 4.0 [3–5]. DTs in Construction (DTC) are often equated with Building Information Modeling (BIM) but missing feedback loops to the construction site [6]. In heavy civil engineering, research is being conducted on DTCs using construction equipment for updating and optimizing construction operations. Besides supporting data transformation, these DTCs use two key technologies [7]: (1) data-driven discrete-event simulation (DES) to digitally represent construction equipment operations and (2) artificial intelligence (AI) to analyze the DES input data coming from the equipment.

Different modeling approaches have shown how supervised machine learning (ML) can help to analyze this input equipment data. An objective has been to automatically recognize activities for equipment such as front-end loaders [8,9], hydraulic excavators [7,9–12], compactors [12], and drill rigs [13,14]. Researchers have used different methods to collect data, such as vision- or motion-based methods, and to analyze them, e.g., using logistic regression, k-nearest neighbor, decision trees, support vector machines, or artificial neural networks (ANN). As part of deep learning (DL), ANN fulfills the requirements for minimal pre-processing as well as robustness to outliers and sensor noise and recognizes many activities [9,12,14].

In all cases, classifying activities has been challenging, because no level of detail or granularity is a priori defined. Literature is lacking in studies that focus on activity modeling. The existing studies on activity recognition have aimed to improve DL techniques instead of improving data acquisition. This paper aims to address activity modeling using a real-world use case as an example to highlight the current challenges and then provide a solution in the form of a productivity model. Our research questions (RQ) are as follows:

RQ1: What impact does production model granularity have on activity recognition?

RQ2: How does production model granularity affect the application of DTC?

RQ3: What is needed to adopt production models for DTCs in heavy civil engineering?

To answer these questions, we start by giving an overview of the current DTC approaches including the modeling part of data-driven DES and activity recognition. Our overview's focus is on equipment-intensive construction operations; however, we recognized that there is much research into recognizing human activities [15]. Then, we present a case study in special foundation engineering, namely, the machine production of bored piles. This study underlines the challenges arising with activity recognition as well as the possibilities and limitations of data-driven DES in heavy civil engineering. It discusses the impact of process descriptions based on real-time data. As a basis for the adoption of DTC applications, a flow-based production model is proposed.

2. Related Work

2.1. DTC and Data-Driven DES Modeling

DTs are platforms that offer promising opportunities for monitoring and prediction in the construction industry [2]. While no universally agreed-upon definition exists of what a DT is [6], in general, it is about the data flow between the real and virtual worlds [16].

To digitally represent construction equipment operations, data-driven DES is widely reported in the literature, e.g., for tunneling [17], earthmoving [7], or pile drilling [18]. DES models simplify construction systems by discretizing activities so that they are bounded by start and end events. This enables investigations of the system's behavior under different parameters. To consider all the dynamics and uncertainties, a DES needs to be updated with as-built data to support reliable decisions based on the results. Akhavian and Bezhadan [19] identified three pieces of knowledge that can be extracted from the equipment: (1) state, (2) operational logic, and (3) layout arrangement. The dynamic updating of DES requires the pre-processing of data from the construction site, for which different statistical probability distribution functions may be used (e.g., Exponential, Gamma, Lognormal, Normal, and Weibull) and evaluated by goodness-of-fit tests (chi-Square test; Kolmogorov–Smirnov; Anderson–Darling) [20].

Finding a suitable (fit for purpose) level of modeling detail is challenging. For example, Liu et al. [21] presented a data-driven Monte Carlo simulation framework to predict equipment life-cycle costs. They used K-means clustering and expectation–maximization to distinguish between the equipment within a fleet. Louis and Dunston [22] replicated and modeled the discovered process as a Petri net in the DES model software jStrobe (a STROBOSCOPE [23] development). Kim et al. [24] also presented an earthmoving analysis but with the help of image recognition and using a process modeled in the DES software WebCYLONE (a CYCLONE [25,26] development) to optimize resource allocation. Kargul et al. [20] used the intralogistics DES software plant simulation [27] to model a

Petri net. They considered three stages of construction progress (25%, 50%, and 75% complete) to show the impact of data-driven modeling. Besides Petri nets, Fischer et al. conducted data-driven DES on agent-based modeling (macro-simulation) to schedule project interdependencies [28] and process modeling (micro-simulation) to investigate the influence of material flow [29]. They introduced a DTC using both DES models for adaptive planning to serve as a decision-making tool (Figure 1).

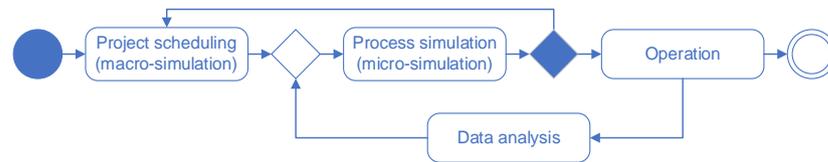


Figure 1. Production feedback control system of DTC (adapted from Fischer et al. [29].)

Rashid and Louis [7] conducted a study to show data integration into a DES model. Based on the activities recognized in their previous work [9], they calculated the activities’ duration, pre-processed it into distribution functions, and fit it using the chi-square goodness-of-fit test.

Several studies show the importance of data-driven DES. A variety of different modeling levels have been considered depending on the modeling purpose. In combination with data processing, frameworks exist to connect the DES model with the real equipment as needed to create a true DTC.

2.2. Activity Recognition Modeling

Activity recognition helps to automatically detect the state data of equipment that can be used to update DES models. It is a classification problem that can be addressed with supervised ML. Supervised ML requires labels to train and test the algorithms. However, defining the appropriate labels is a challenge. An early study facing the labeling problem was conducted by Akhavian and Behzadan [8]. They classified activities at different levels of detail (LoDs) to investigate the influence of the proposed labels. The number of LoDs defines how often the labeling dataset is divided into classes. Similarly, it defines the granularity at which the labels are described (Figure 2).

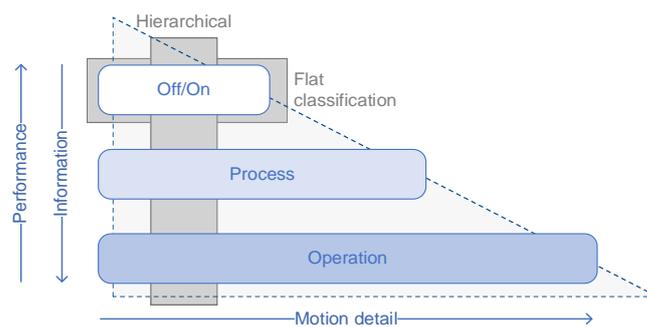


Figure 2. Classification problem for activity recognition modeling (white: LoD1; light blue: LoD2; and blue: LoD3) (adapted from [8,30]).

For example, LoD1 describes “engine on” or “engine off” whereas LoD2 divides “engine on” into “idle” and “busy” [8]. Challenges arise with activities that have similar signal patterns from vibration and angular velocity measurements, e.g., “scooping”, “dumping”, and “moving”. Therefore, these authors combine specific activities, i.e., “moving and scooping” and “moving and dumping”, to increase the model’s performance. They concluded that the less granular the LoD, the higher the model’s accuracy (ANN: 98.6, 81.3, and 86.1% for LoD3, LoD4, and LoD5). Rashid and Louis [11] confirmed this conclusion. They investigated the use of flat classification up to nine LoDs, with nine being the number of

activities to be recognized. Instead of relying on equipment vibrations, which can be highly equipment-dependent and may also be influenced by external forces, they investigated the recognition of the activity-specific equipment motions by placing Inertial Measurement Units (IMUs) on equipment joints. Their results are slightly better (ANN: 100, 98, 97, 95, and 92.1% for LoD2, LoD3, LoD4, LoD7, and LoD9) [11].

Harichandran et al. [30] extended this flat classification with a hierarchical classification to increase performance (Figure 2). This is of interest in case the classes under investigation have a specific hierarchy such that they can be grouped into metaclasses. The use of a hierarchical approach makes it possible to exploit the tree structure of the classes and thereby reduce the number of classes considered by a classifier. Instead of directly classifying a high number of activities (flat classification on the operation level in Figure 2), which may require highly complex models that need to detect subtle differences in the data, local classifiers need to distinguish a much smaller number of classes (hierarchical classification). A popular approach for hierarchical classification is the local classifier per parent node approach. A separate multi-class classifier, which is specialized to distinguish only between child nodes within a group, is trained for each subgroup. Use of this approach may result in an accuracy improvement of up to 15%. However, this result is only comparable with the previous studies on the construction equipment to a limited extent, as it refers to a case study analyzing the accelerometers of an automated construction system. Furthermore, questions arise with the use of classifiers that are connected in a series (hierarchical classification) regarding the decreased number of data points after each classification, or the influence of errors carried with each classification.

Overall, Akhavian and Behzadan [8] and Harichandran et al. [30] conclude that the choice of a classification depends on the type of equipment, its operation, and the purpose of the data analysis. Their generalized framework focuses on the activity modeling part including equipment states, operations, and hierarchical relationships. Furthermore, the purpose of activity recognition must be clarified for operation recognition, e.g., fuel consumption (LoD1), emission rate (LoD2), overall productivity (LoD3), and cycle time (LoD4). They state that the data must include enough processed information depending on the purpose.

To sum up, the claim that the research in the field of activity recognition in heavy civil engineering does not focus on activity modeling is only partially true. Studies show that model accuracy highly depends on the modeled granularity of the activities. Data-driven classification approaches have been introduced to overcome this challenge. However, the activity modeling frameworks that we were able to identify in the literature are too shallow. The studies mentioned that activity modeling requires process knowledge, but a conceptualization of the production system is missing.

2.3. Production System Models

According to Koskela [31], production system models in the construction industry reflect the transformation view (rather than the flow or value view) by modeling the process of product transformation from inputs to outputs. A product is characterized by, e.g., functionality, configuration, and geometry. According to the work breakdown taxonomy, the processes are divided hierarchically into smaller subprocesses, which consist of specific activities, e.g., resource assignment and sequencing of activities.

Activities are executed by the workers or equipment [31]. They may be value-adding and non-value-adding, such as waiting, putting away, or moving material. Activities can be broken down further, e.g., building a pile requires drilling holes, fabricating rebar cages, placing concrete, and lifting equipment cylinders [32]. Besides transformation and value, Koskela [33] emphasizes the flow between and within processes. A lean objective is to increase workflow reliability.

Tommelein [34] emphasizes that production system models must address product and process variability in order to lend themselves to improvement. This variability must be understood in detail to reveal and eliminate waste. The seven wastes in the sense of

lean are (1) overproduction, (2) defects, (3) transportation, (4) overprocessing, (5) inventory, (6) waiting, and (7) motion [35]. So, while equipment-driven operations typically have been designed around the optimization of equipment use, and the cost of equipment outweighs other costs in their operation, factors other than equipment cost minimization play a role in the overall process of optimization [36].

Different metrics exist to measure production flow. Kalsaas [37], for example, adapts the Overall Equipment Efficiency (OEE) in equipment-intensive construction projects but emphasizes that only using this metric is not sufficient. There is a need to consider the entire production system. With respect to production system models, the products and processes need to be investigated separately. Their variabilities and flow-based interdependencies must be understood before trying to improve the system.

2.4. Research Gap and Objective

Frameworks for application DTCs in heavy civil engineering using data-driven DES exist. However, only a few exist in combination with DL-based activity recognition. The maximum number of activities so far is nine and the excavator is the most commonly investigated equipment. The generalizability of the procedures used in existing studies to more complex construction equipment can thus be questioned. Furthermore, most studies in activity recognition found that their models' accuracy decreases with the increase in information about the production. The countermeasure has been to improve their proposed DL models, e.g., by applying more complex ML algorithms, instead of improving their classification models. A supposedly new method for modeling classification problems was introduced (hierarchical classification), but the associated study failed to discuss the actual granularity of modeling. It is commonly assumed that the more detailed the movements, the more information one receives about the production system. However, from the production system's point of view, it is worthwhile to analyze idle times to optimize the system. Thus, this paper first aims to fill the gap of the trade-off between the requirements of the DES and the DL models by demonstrating the influence the granularity of activity classification has on algorithm performance based on a real case study and, second, it introduces a flow-based production model.

3. Methodology

We conducted two parameter studies on AI-based activity recognition and DES to validate the proposed DTC (Figure 1). The input data are from a completed construction project on special foundation engineering in Rosenheim, Germany. This project involved the construction of a bypass road including two bridges near the German–Austrian border. The project consists of 32 bridge piers, each including 5 to 17 large diameter bored piles of the same type, and ranging from 26 m to 50 m in length. Due to challenging soil conditions [38], the project owner mandated comprehensive process documentation from the pile producer, BAUER Group. This documentation provided the data for our research. We used three data sources:

1. **Activity data:** While producing the pile, workers manually recorded activities on site with a tool provided by *fielddata.io* (a German start-up, now acquired by the BAUER Group). They had a choice of 27 predefined activities. The tool was connected via the equipment's Wi-Fi to have the same time stamps as the sensor data.
2. **Equipment sensor data:** We used data from sensors already installed on the equipment and sent via telematics to the proprietary platform every 1 Hz. Measurements included pump pressure, rotary torque, winch forces, and mast inclination (Table 1 and Figure 3 (black circles)).

Table 1. Sensor data available from the Kelly drill rig.

Sensor	Unit	Sensor	Unit
Depth	m	Main winch rope speed	cm/min
Torque of rotary drive	kNm	Pressure pump 1–4	bar
Speed of rotary drive	Rpm	Torque of Kelly bar	%
Main winch rope force	t	Auxiliary winch rope force	t
Crowd force	t	Crowd depth	m
Casing length	m	Boring threshold	m
Status rig	-	Torque steps	-
Main winch gear mode	-	Inclination X, Y	deg



Figure 3. Rotary drilling rig at the test site of Bauer Group in Schrobenhausen, Germany (pictures by Fischer, A.); components (number in white circle) and sensors (black circles): (1) under carriage; (2) upper carriage (pressure pumps); (3) main winch (force, speed, and tool depth); (4) auxiliary winch (force); (5) crowd winch (force); (6) kinematic system; (7) mast (inclination); (8) mast head; (9) Kelly bar; (10) rotary drive (torque, speed); (11) bucket drill tool; (12) auger drill tool; (13) casing string; and (14) casing shoes.

3. **Production log:** Every pile was documented in a handwritten report. This report gave insight into the bored pile sequence and start and end times. Thus, the duration of the following seven subprocesses is derived: (1) drill, (2) idle between drill and reinforce, (3) reinforce, (4) idle between reinforce and install contractor pipe to fill in concrete, (5) install contractor pipe, (6) idle between install contractor pipe and concrete, and (7) concrete. Data from 232 bored piles were analyzed.

We used the activity and equipment sensor data to test and train the DL models for automatic activity recognition. We used the production log input for the DES model.

4. Kelly Pile Production System

The Kelly drilling method used for pile production uses a rotary drill rig (“rig” in short) with different attachment tools and additional equipment to drill piles up to 3 m in diameter and more than 100 m deep [39] (Figure 3).

The application of activity recognition requires the subdivision of the production process into recurring standard activities. Therefore, we next describe the most common steps in the pile production process using the Kelly drilling method. We are aware that these steps vary depending on the construction project. The results are based on interviews conducted in previous work by the authors [36].

Kelly pile production consists of three main steps: (1) drill, (2) reinforce, and (3) concrete (Figure 4).

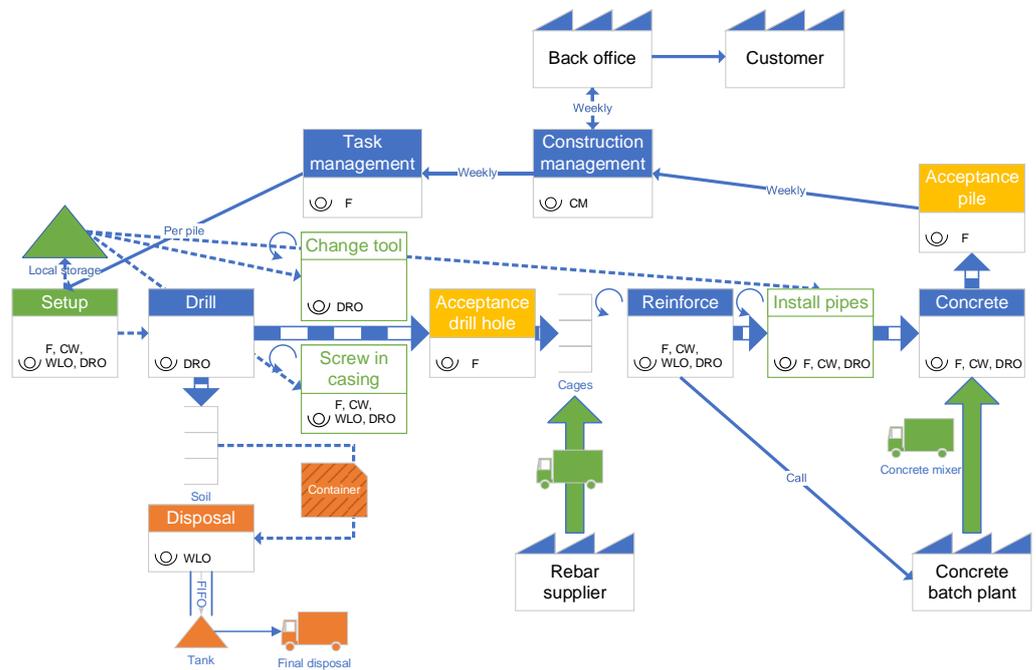


Figure 4. Value stream map for pile production using the Kelly drilling method [36] (blue: production; yellow: quality checks; green: procurement logistics; orange: disposal logistics) (CM: construction manager; F: foreman; CW: construction worker; DRO: drill rig operator; WLO: wheel loader operator).

Once the rig and any needed auxiliary equipment such as drilling tools, casings, and concrete delivery pipes are set up, the alternating steps of the drilling process start. The equipment picks up the appropriate tool, e.g., an auger (Figure 3 middle), then slews and positions itself toward the drilling attachment point. Lowered with the help of the telescopic Kelly bar, the drilling tool drills as the rotary drive applies torque to the locked Kelly. Once the tool is filled with soil cuttings, it is pulled out and emptied, usually in a container, to ease soil removal from the site. In turn, a wheel loader takes the drill cuttings to a disposal site for further processing. Casing sections prevent the hole from collapsing. The rotary drive turns in casings separately and casings are manually fixed by workers (Figure 3 right). The first casing has teeth for better progress in the soil. The deeper the drilling tool, the lower the performance due to longer run-in/out times and higher surface friction. Thus, casing oscillators are used when additional torque force is needed. For reinforcement, the rebar cage is attached to the auxiliary cable of the rig and raised. The rig swivels to the drilling attachment point to lower the reinforcement cage in the drilled opening. Before concrete placement starts, the delivery pipes are assembled, lowered into the drill hole, and joined together. Concrete is placed directly through the concrete mixer discharge or a concrete pump/bucket (requiring additional steps). Casing and delivery pipes are removed in alternation, often requiring extra power from a casing oscillator.

Nübel et al. [40] define three KPIs in pile production: (1) the pile length produced per day (or piles per day); (2) variations in the planned vs. actual output (process quality); and (3) the inclination of the piles (product quality). Pile production resembles a single-line production system. It is characterized by the use of highly specialized equipment and skilled operators [36]. Deep domain knowledge of geotechnical engineering and process technology is required to handle the complexity of the production system.

5. Activity Recognition

5.1. Deep Learning Models

Fischer et al. [29] showed three DL models for activity recognition: (1) activity recognition via telematics data is a solid alternative to the existing motion-based method; (2) recurrent neural network (RNN) models consisting of long-short term memory (LSTM) consider temporal dependencies in the modeling, which is most suitable for this use case; and (3) adding convolutional neural network (CNN) enables good feature extraction, outlier filtering, and smoothing method. Their proposed hybrid models on RNN and CNN, the DeepConvLSTM, and the bidirectional variant DeepConvBiLSTM were also developed and tested on the use case of Rosenheim. An average accuracy of up to 96.1% for 27 activities was achieved.

However, the question arises on the generalization capability of the results. The dataset was split randomly using the Scikit-Learn function in Python: 56% training, 14% validation, and 30% test data. Regarding the relatively low frequency and slow state changes in the equipment motions, two observable samples can be similar but are located once in the training set and once in the test set. Studies on addressing this issue were conducted in the previous work of Beiderwellen Bedrikow [41]. The dataset was, therefore, coherently split by production days. The results based on this data splitting methodology led to a decrease in the average accuracy of up to 51.9%. One reason for this result can be the high complexity of the data modeling recognizing 27 activities. As mentioned, comparable studies in the literature only examine activities of less than 10 (LoD9).

In this paper, we, therefore, investigate the influence on data modeling by hierarchical classification according to Harichandran et al. [30]. Table 2 shows the DL models implemented and tested by Beiderwellen Bedrikow [41]. The DL models were implemented using the Tensorflow package in Python.

Table 2. Architecture of the DL models used for the parameter study.

	MLP	DeepConvLSTM	DeepConvBiLSTM
Architecture	5xDense-Softmax	3xCNN–2xLSTM-Softmax	3xCNN–2xBiLSTM-Softmax
Temporal window	no	Overlapping sliding window	Overlapping sliding window
Window size	-	16 s	16 s

The baseline model extends an ANN with only one layer (based on previous work from the authors [14,41]) to a multilayer perceptron (MLP) consisting of 5 dense layers with 128 neurons each. While this model does not account for temporal dependencies, it is used as a comparison to investigate the influence of temporal relationships in activity recognition.

RNN is used to consider the temporal dependencies of the data, providing the outputs of the previous layer and the outputs of its layer at the last time. One of the significant types of RNNs is the LSTM network, which was also used by Rashid and Louis [11]. In addition to the inputs that are present in simple RNNs, LSTMs have an additional long-term state, which can store long-term dependencies [42].

The proposed hybrid framework for human and construction machinery activity recognition combines the short-time feature extraction capabilities of convolutional layers with the long-time temporal dependencies modeling capabilities of LSTM layers. The original architecture developed by Ordóñez and Roggen [43] consists of one input layer, four convolutional layers, two LSTM layers, and one Softmax activation layer. A slightly modified version of this architecture is used by Slaton et al. [12]. They add batch normalization layers between the convolutional layers and a dropout layer between the convolutional and recurrent layers.

The unidirectional model used in this paper is based on this modified architecture with minor changes to reduce overfitting. In addition, we investigate the influence of a bidirectional architecture, which processes the inputs in both directions. Bidirectional

RNNs have numerous applications in the field of Natural Language Processing (NLP) [44]. Xu et al. [38] explored the use of bidirectional RNNs for human activity recognition, in addition to its applications in NLP. The utilization of bidirectional layers results in a slight improvement in accuracy. We replace the two LSTM layers from the unidirectional architecture to two bidirectional LSTM layers. The window size of both time dependent DL models is 16 s, as it achieves satisfactory results [14].

The neural network for all three models was trained for 100 epochs using early stopping, a learning rate of 0.001, reducing the learning rate on plateaus, and a batch size of 256, with the Adam optimizer.

We rely on the accuracy and F1 score of the model’s predictions to evaluate model quality. When multiple classes are involved, the metrics for the complete dataset are obtained by taking the average of the metrics for each class. The specific formulas for calculating each metric are described in [45]. For ease of reading, we printed the results of the DeepConvBiLSTM. The results of the other models are in Appendix A.

5.2. Hierarchical Classification Study

The following study is based on the hierarchical classification, according to Harichandran et al. [30]. For each parent node (e.g., work) a classifier is trained, which can classify among the child nodes (concrete, reinforce, and drill). Instead of twenty-seven activities to be classified (flat classification), only up to seven activities need to be considered by the classifier.

The division into individual groups is based on the division by the construction company using the manual data collection app from the start-up company fielddata.io. The classification is fundamentally oriented to the differentiation of the process steps to reveal value-adding and non-value adding steps, e.g., to identify the delay of concrete trucks. We identified three LoDs (Figure 5). In LoD1, a distinction is made only between work and idle. Work class contains all activities in which the drill is actively involved. Its child nodes are the main process steps of pile production: drill, reinforce, and concrete (compare Figure 4). The parent node Idle distinguishes between downtime and secondary process time. These subgroups form the second LoD (LoD2). Finally, the third LoD (LoD3) divides the rough process steps into more detailed steps.

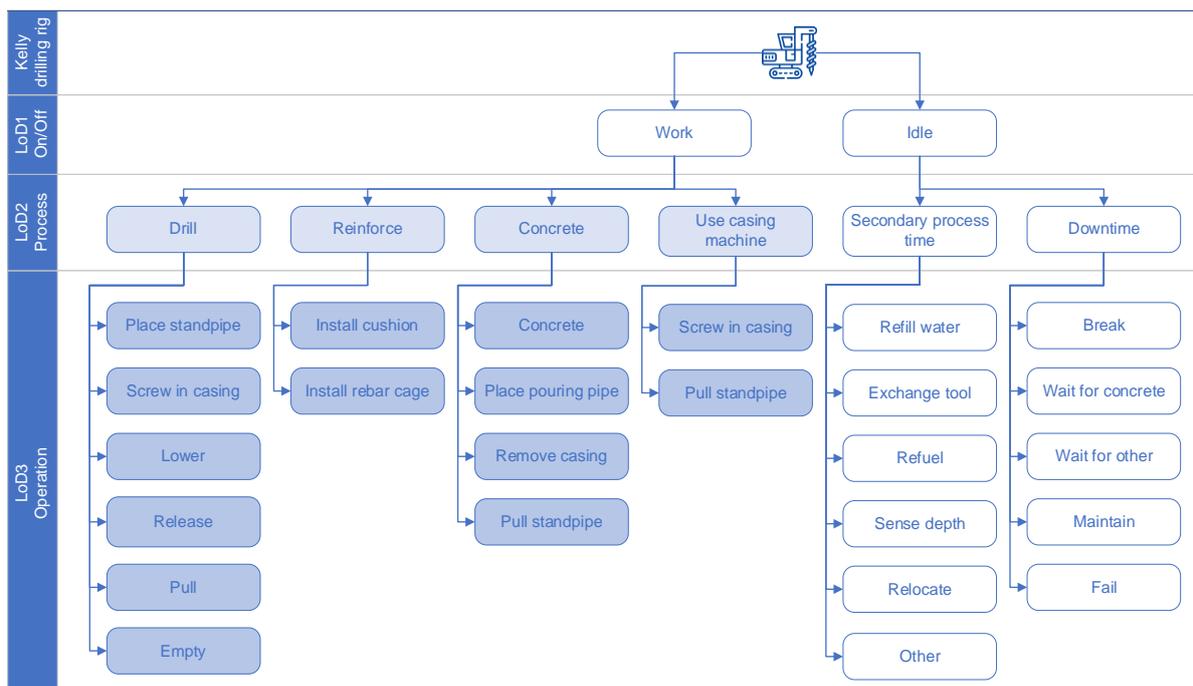


Figure 5. Overview of the LoDs (white: LoD1; light blue: LoD2; and blue: LoD3).

5.2.1. LoD1—Work vs. Idle

Table 3 displays each class’s average F1 and F1 scores when the models are applied to the test set. The baseline MLP and the two hybrid model variants have a similar average F1 score of about 0.83. Among the individual labels, the work label consistently has the highest F1 score of around 0.86 in all three cases. The F1 score for the idle label is 0.80 for the MLP and DeepConvBiLSTM and 0.82 for the DeepConvLSTM.

Table 3. F1 scores for the selected activities.

Level of Detail (LoD)	Parent Node	MLP	DeepConvLSTM	DeepConvBiLSTM
LoD1	-	0.83	0.84	0.83
LoD2	Work	0.42	0.58	0.62
	Idle	0.66	0.74	0.73
LoD3	Drill	0.59	0.85	0.85
	Reinforce	0.68	0.68	0.68
	Concrete	0.20	0.24	0.24
	Idle	0.39	0.39	0.41

Figure 6 shows the confusion plot for the DeepConvBiLSTM model (the confusion plots for the MLP and the DeepConvLSTM are shown in Appendix A). It reveals differences in performance between the models. The MLP has an accuracy of 73% for the idle class and 70% for the work class, while the two hybrid models show an increase in accuracy. The work class accuracy improves to 90% in the hybrid models, with only a slight increase for the idle class.

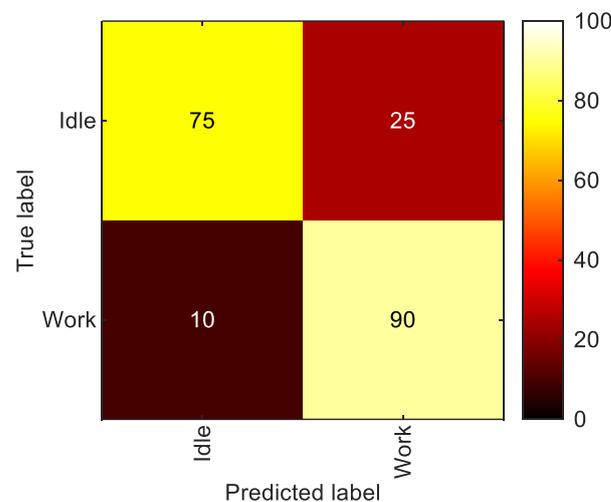


Figure 6. LoD1 confusion matrix for the DeepConvBiLSTM.

The hybrid models perform well in identifying the work class, with only 10% of samples that belong to the work class classified as idle. However, the main challenge lies in identifying the idle class as 25% of the samples idle class classified as work. One reason is the limited number of samples in the dataset, i.e., less than a quarter of the total dataset. Data augmentation methods can solve this imbalance of the datasets. Another reason is the similarity of the classes, which makes classifying difficult.

5.2.2. LoD2—Process Steps

The second level of detail focuses on the activities within the work and idle groups. On the one hand, the model for the work group is designed to handle the casing machine, concrete, drill, and reinforce classes. On the other hand, the model for the idle group is responsible for identifying the secondary process time and downtime activities.

The F1 scores for each label and the average F1 score are shown in Table 2, while Figure 7 displays the confusion matrices for the individual models' predictions for the DeepConvBiLSTM. The confusion plots for the MLP and the DeepConvLSTM can be found in Appendix A.

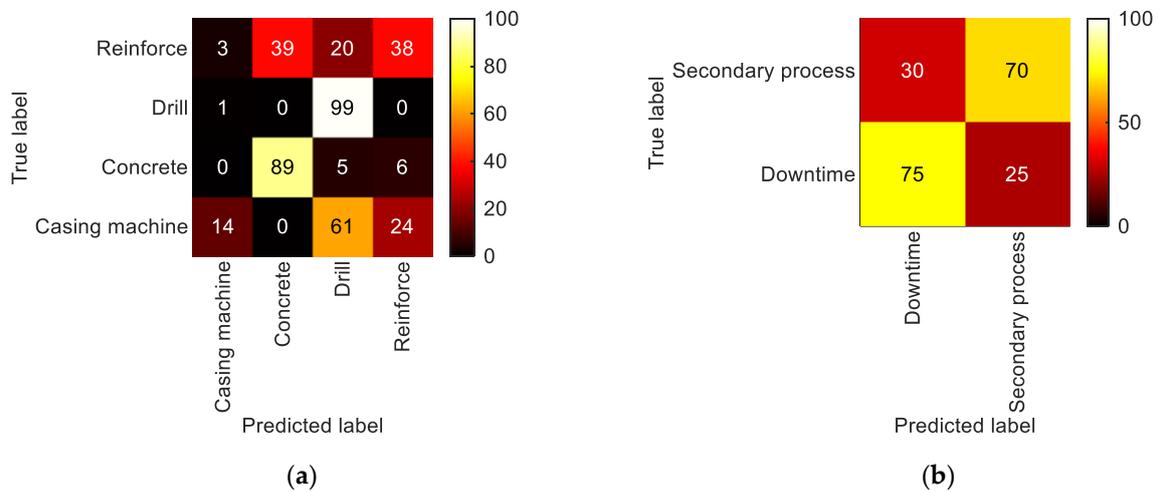


Figure 7. LoD2 confusion matrix for the DeepConvBiLSTM: (a) work; (b) idle.

The models differ noticeably in performance. For example, the baseline MLP model has an average F1 score of 0.42, with significant variations among the individual classes. The class drill has a high F1 score of 0.89, with 99% of the samples labeled as drill being correctly classified as such. However, the model tends to classify samples as drill, leading to low F1 values for the concrete and reinforce classes.

Both hybrid models show better overall performance, with the average F1 score increasing to 0.58 for the DeepConvLSTM and 0.62 for the DeepConvBiLSTM. The confusion matrices for these models show more entries concentrated on the secondary diagonal compared to the MLP model. The concrete class has an improved accuracy of 87% and 89% for the DeepConvLSTM and the DeepConvBiLSTM. The use of bidirectional LSTM layers leads to an overall improvement in performance, as seen in the increase in F1 scores for all classes, including casing machine and reinforce.

The second group at this level of detail is the idle group, which encompasses downtime and secondary process related to drill and concrete. The differentiation between downtime and secondary processes activities is addressed more effectively at a different level of detail. The baseline MLP model correctly classifies 70% of the downtime class samples and 63% of the secondary processes class samples, resulting in a mean F1 score of 0.66. Accuracy improves to 75% and 73% for the two hybrid models, and the F1 score rises to approximately 0.73.

5.2.3. LoD3—Detailed Process Steps

Classifying within the concrete superclass (LoD2) is challenging, as all three models tend to predict the two most common activities, *concrete* and *place pouring pipe* (LoD3).

The performance of the three models is similar when classifying the reinforce process step (LoD2), which includes *install rebar cage* and *install cushion* activities (LoD3). Only the confusion matrices for MLP and DeepConvLSTM are presented since the models barely differ. The average F1 value is 0.68 for MLP and DeepConvLSTM and 0.65 for DeepConvBiLSTM. These activities can be performed by machine or crane, which makes it difficult for the models to capture different behaviors for the same label. For example, if crane installs a rebar cage, no equipment activity is detected, leading to misclassifications.

Investigation of the process step *release* shows that the baseline MLP model has an average F1 score of 0.59. Among the six activities considered, three are recognized well,

two moderately well, and one poorly. The *place standpipe* activity is classified the best, with an F1 score of 0.89. The *release* activity is recognized with a score of 0.81 but is often misclassified as *pull*. The *screw in casing* activity also has a high F1 score of 0.81.

In contrast, the *pull* activity has a low F1 score of 0.6 due to many false positives. Finally, the *empty* activity has the second-worst F1 score of 0.43 and is often classified as *pull*. The *lower* activity is not considered in the modeling, and no samples are correctly classified.

Adopting the hybrid models results in a significant increase in the average F1-score to 0.85 for both DeepConvLSTM and DeepConvBiLSTM. All individual labels show an improvement in the F1 score. The F1 score for the *place standpipe* class increases to 0.93 with improved accuracy and fewer false positives. Although the accuracy of the *release* class decreases, the F1 score remains unchanged due to a decrease in false positives. The hybrid models address the difficulty in classifying transitions between two successive activities and improve the classification of the *screw in casing* and *empty* activities. The *lower* class, which the MLP did not recognize, sees a drastic increase in recognition to 67% with the hybrid models but still remains the worst recognized activity.

The F1 score for the baseline MLP and the two hybrid models is around 0.4, indicating no advantage of using the hybrid models over the MLP. Only the activity *refill water* can be accurately detected among the secondary processes, with an F1 score of 0.82 for all models. Again, the choice of labels plays a significant role here. Figure 8 shows that activities such as *relocate*, *depth sensing*, *refuel*, and *other* are not effectively differentiated as they describe only the process and lack any indication of machine movement.

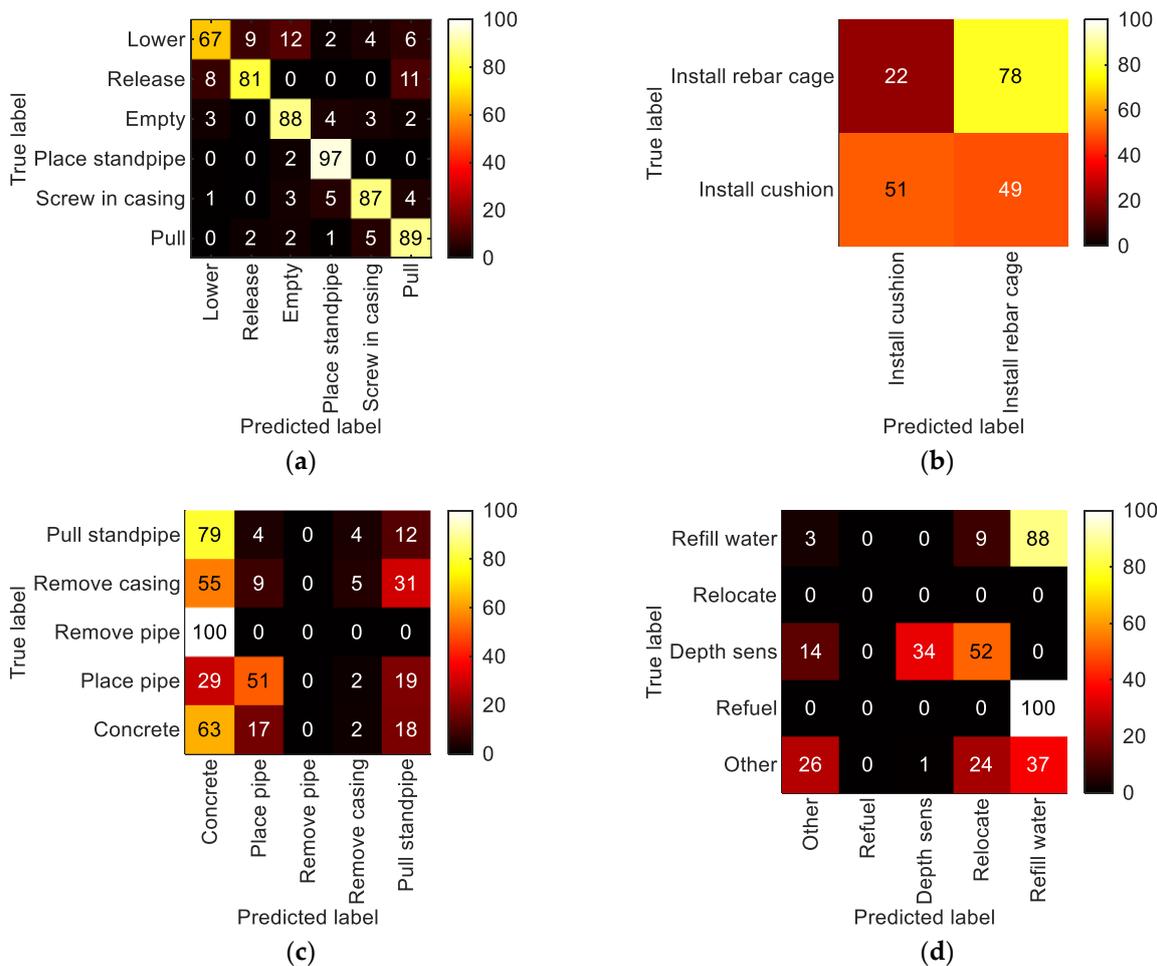


Figure 8. LoD3 confusion matrix for the DeepConvBiLSTM: (a) drill, (b) reinforce; (c) concrete; and (d) secondary processes.

5.3. Conclusion Regarding Activity Recognition

Compared to the flat classification study from Fischer et al. [14], the hierarchical classification underperforms. The average accuracy of 27 activities is 52.0% instead of 96.1%. However, even though the results fit quite well for the drilling operation (84.8%), the reinforce (67.5%), concrete (26.2%), and secondary processes (29.6%) super classes are classified with a low accuracy. These results do not show that the flat classification is recommended. As mentioned above, difficulties emerge due to generalization capabilities.

The results in this paper show the limitations of the labeling strategies: Selection of labels poses a significant problem for activity recognition. Certain activities proved to be very difficult to classify. Most misclassifications took place between specific activities. Figure 9 shows the raw data. As a simple example, the activities *concrete* and *wait for concrete* can be considered. All sensor signals are constant except for some existing noise (Figure 9a,b). There is a difference compared to the behavior during drilling (Figure 9c). In both activities (*concrete* and *wait for concrete*), the drilling rig just passively waits while the process is performed by other equipment, such as concrete trucks, and workers. A differentiation of the activities based on the equipment data is thus only barely possible. However, if activities in which the equipment is actively involved (LoD3), such as *release*, *screw in casing*, or *empty*, were considered, then satisfactory results were achieved in all cases.

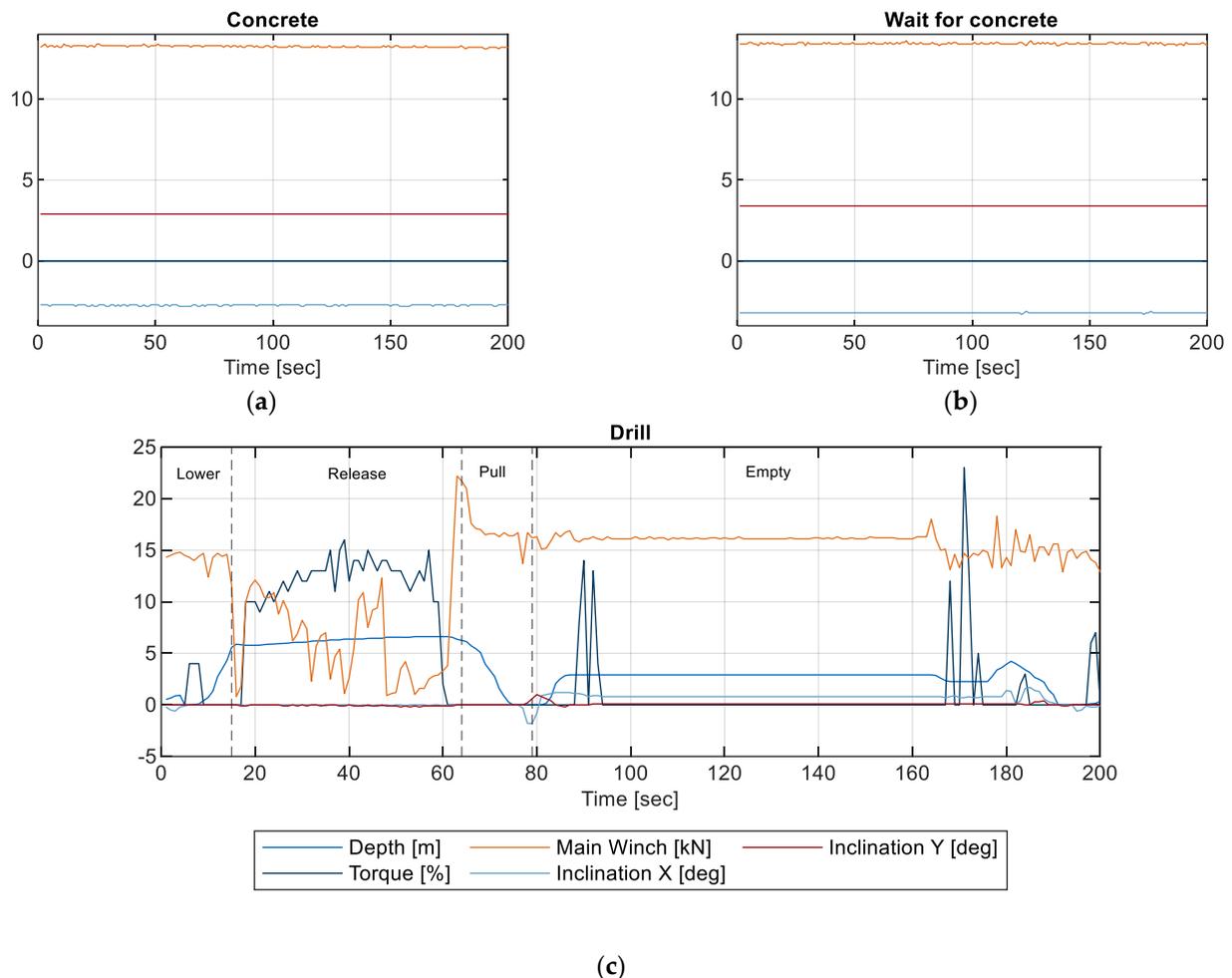


Figure 9. Raw data from the Kelly drill rig for different activities: (a) concrete; (b) pause; and (c) drill.

As shown in Figure 9, drill (LoD2) is described in every detail in *lower*, *release*, *pull*, and *empty* (LoD3). Thus, LoD3 labels, such as *install rebar cage* or *install cushion* can be improved when described in more detail.

The root cause of these problems is that equipment behavior is not taken into account in selecting the labeling strategy, which in this case, is based purely on the process steps and on a process view, without considering the activity recognition technology to be used. As a result, different labels are assigned in cases where the equipment does not exhibit different behaviors. It can be pointed out that the basic equipment behavior must be considered during the selection of the labels for automatic activity recognition.

6. Data-Driven DES

6.1. DES Model

The recognized activity data serve as input for the DES. This data-driven DES is the complement of the DTC approach, using data to help inform decision-making for the production system. In this paper, the DES model aims to forecast the construction project’s end time. To investigate the impact of adapting the input data, we conducted three studies with three different percentages of as-built data, i.e., 25%, 50%, and 75% of construction progress. This work is based on the previous work of Fischer et al. [36] but here considers the length of the piles by performing linear regression.

The DES model is implemented in Python. It is modeled as a simple Petri net, where each of the seven subprocesses from the production log is depicted as a single station with a defined duration or processing time characterized by a probability distribution (Figure 10).



Figure 10. DES model (circles: Source/Sink; white: LoD1; light blue: LoD2; and blue: LoD3).

The processing time duration of each single station is calculated from the production logs. The results are compared to the related bored pile length, as shown in Figure 11 for the subprocess drill. The linear regressions and their residuals are calculated for each process (Table 4). We assume a normal distribution of the residuals. However, it is important to mention that the simulation only takes into account the working time, and the total processing time should be interpreted in working time and not as calendar days.

In predicting the production time, the simulation randomly selects a duration that matches the input parameters of the best-fitted distribution function. Performing several runs, i.e., 10,000, and increasing the percentage of as-built data with construction progress, increases the probability of a good prediction of the production time. The result is the cumulative duration of the pile production. This forecast is compared to the real construction progress and varies due to the duration time randomly picked from the specified distributions.

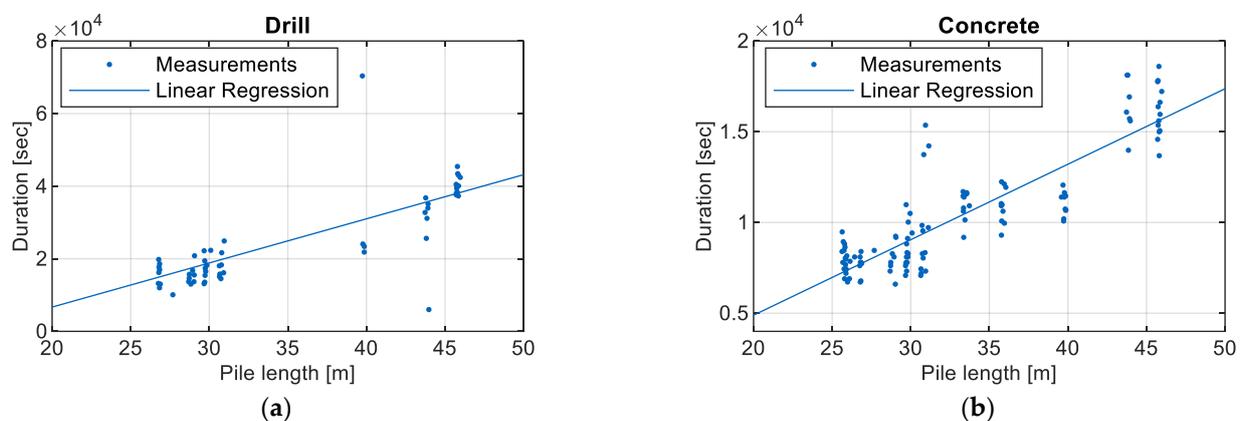


Figure 11. Linear regression for the processes: (a) drill; (b) concrete.

Table 4. Parameters for the linear regression and residuals depending on the process.

Process	Parameter	25%	50%	75%
Drill	Mu	-5.33×10^{-12}	7.21×10^{-12}	-2.17×10^{-12}
	Sigma	7995.83	7787.97	7090.17
	Slope	1211.12	1216.27	1173.93
	Intercept	-16,781.52	-17,638.90	-17,837.04
Idle 1	Mu	2.82×10^{-13}	-7.37×10^{-13}	-5.02×10^{-13}
	Sigma	3619.61	4555.91	5119.83
	Slope	-163.18	-26.71	-32.85
	Intercept	12,496.47	7209.08	7381.52
Reinforce	Mu	7.84×10^{-15}	-3.18×10^{-13}	-6.27×10^{-13}
	Sigma	500.08	581.13	545.05
	Slope	105.71	83.95	75.64
	Intercept	-2299.57	-1480.55	-1184.78
Idle 2	Mu	1.39×10^{-13}	3.4×10^{-13}	1.36×10^{-13}
	Sigma	762.92	1869.31	1636.55
	Slope	-45.05	-41.68	-20.96
	Intercept	2596.29	2519.07	1639.18
Install contractor pipe	Mu	-2.59×10^{-13}	-3.72×10^{-13}	-3.55×10^{-13}
	Sigma	785×10^{-1}	691.22	650.05
	Slope	72.25	58.94	56.68
	Intercept	-1039.17	-494.96	-301.61
Idle 3	Mu	6.27×10^{-14}	-1.09×10^{-12}	-3.35×10^{-13}
	Sigma	2238.13	2218.63	27,756.29
	Slope	-30.32	89.56	129.42
	Intercept	3359.87	-266.50	-1134.46
Concrete	Mu	0	-1.11×10^{-12}	-2.7×10^{-12}
	Sigma	1407.71	1534.68	1448.72
	Slope	267.39	415.97	436.95
	Intercept	1505.92	-3429.70	-4101.32

6.2. Forecast Study

Figure 12 shows the distributions for the expected construction time for the three volumes of data considered, along with the actual duration of the pile production. Figure 13 shows their cumulative durations. The average predicted total durations are presented in Table 5 as a function of the construction progress.

Table 5. Results of simulation study compared to actual duration of pile production.

Project Completion: Ratio of as-Built Data	Prediction of Total Pile Production:	
	Mean	Standard Deviation
25%	128.2 days (+7.76%)	1.4 days
50%	127.5 days (+7.14%)	1.2 days
75%	123.6 days (+3.86%)	0.8 days
100%	119.0 days	-

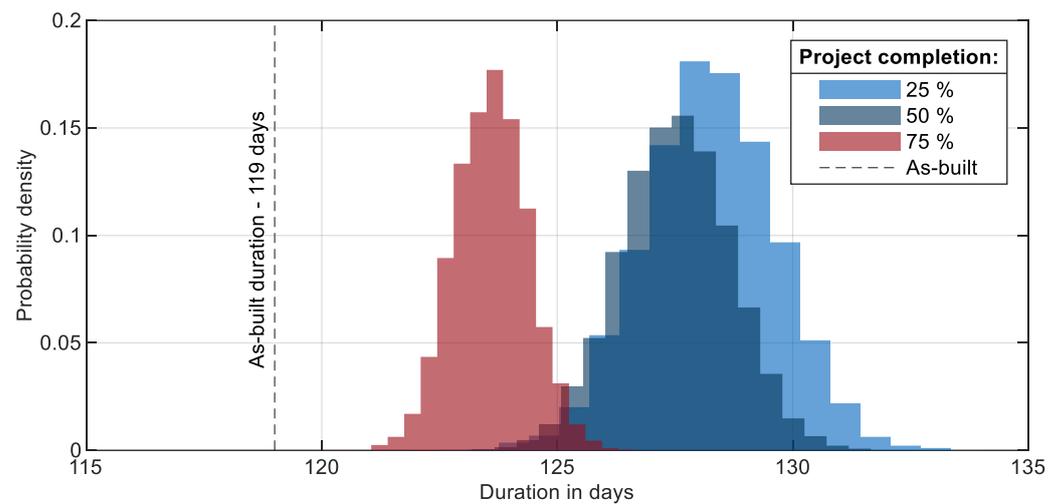


Figure 12. Results from the simulation runs: the forecasts of the construction end time get closer to the as-built duration as more piles are produced.

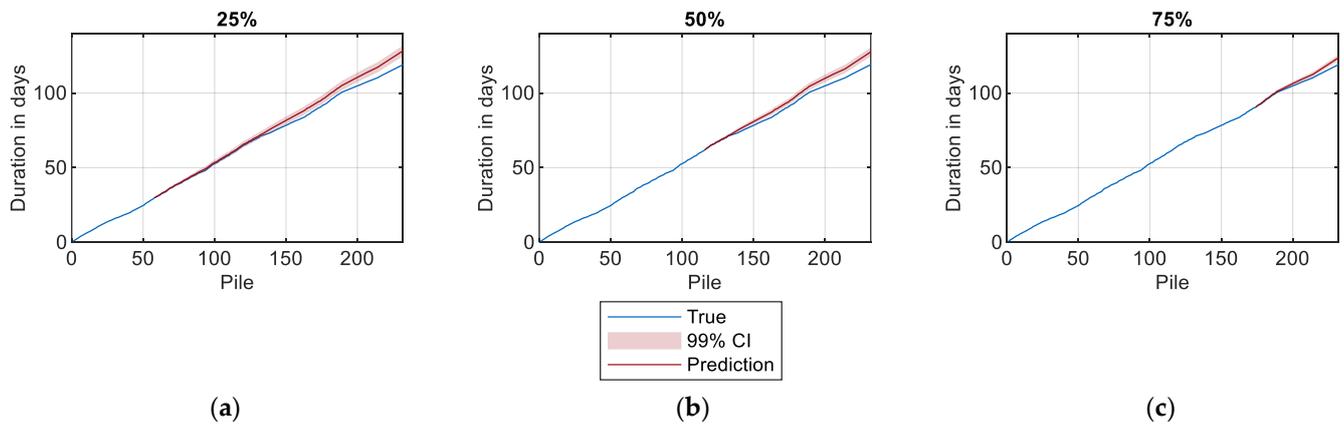


Figure 13. Cumulative duration of pile production during the construction project: as-built (blue) vs. as-predicted by the simulation runs (red) based on (a) 25%, (b) 50%, and (c) 75% as-built data (CI: confidence interval).

As the construction process progresses, the amount of as-built data used as simulation input could be increased (25%, 50%, to 75%). As a result, the distributions of the predictions get closer to the actual duration. The results in Table 5 show that, at 25% construction progress, the simulated data deviated only about 9 days from the as-built duration. At 50%, this deviation became smaller. At 75%, it decreased to about 5 days. Furthermore, uncertainty decreases with continuous updating of the data. The standard deviation around the mean of the predicted value decreased from 1.4 days at 25% of construction progress to 0.8 days at 75%.

6.3. Conclusion Regarding Data-Driven DES

Overall, the data-driven DES shows how one can use a DTC based on construction equipment data. The simulation results support the observation that, as the amount of information increases, the results improve and approach the real construction time. Since the data comes from a real construction project, each pile’s actual processing time duration and production steps vary. Statistics help to find dependencies, e.g., pile production time depends on pile length. Due to uncertainty and variability in the production data it is important to update the DES with as-built data in order to make reliable decisions.

However, the prediction overestimates the as-built duration. A reason may be the increased throughput due to the learning curve. Fischer et al. [36] found out that it takes

at least one week to create a team, and the more well-rehearsed the team is, the greater their performance will be, e.g., in reducing the setup of the rig. Another reason for this may be the distribution of the individual pile lengths over the construction project. In contrast to the use case (Figure 14a), if piles were sorted by the ascending pile length (Figure 14b), then the time required to construct the first 25% of them would be significantly less than that required to construct the last 25%. Therefore, the simulation would underestimate the as-built duration as the input data would constantly represent shorter piles.

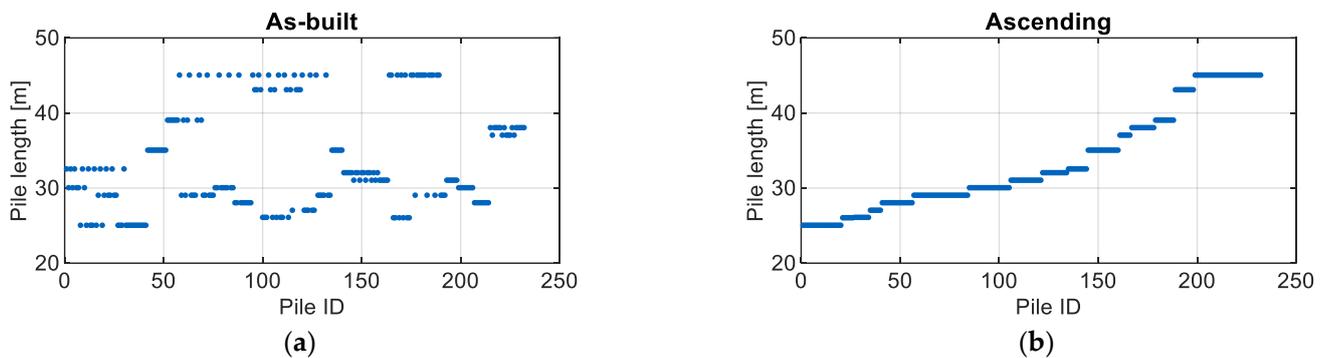


Figure 14. Construction sequence of piles with respect to their length: (a) as-built; (b) ascending.

The DES model presented here is limited in regard to capturing specifics of the construction process itself; it is a simple Petri net. A more detailed model is required to question the current production flow, including material, equipment, or work shifts, and investigate different impacts to optimize the process.

7. Discussion

The methodology used in this paper shows the impact that production model granularity has on DTC applications. We next discuss the findings in relation to the research questions.

RQ1: What impact does production model granularity have on activity recognition?

We investigated a specific use case, namely, the Kelly pile production. The hierarchical classification study of activity recognition shows the conflict between the granularity of the training and the test data. Related literature has shown that the model's accuracy decreases with an increase in the labels' granularity. They conclude that they have an insufficient knowledge about the production system itself. In this paper, although the given labels are based on deep expert knowledge, the results are not satisfactory either. They show that the granularity of the activity description must refer to equipment motions, which defines the requirement for a deep understanding of the production system. For example, activities where equipment is not involved cannot be distinguished sufficiently, e.g., *Concrete vs. Waiting for concrete*.

RQ2: How does production model granularity affect the application of DTC?

In our DTC approach, the activity recognition aims to feed the DES with as-built data. A data-driven simulation study shows how this DTC enables construction forecasts. Challenges arise on the assumed granularity of the production system model. We used a simple Petri net without interdependencies throughout the production flow. On the contrary, we integrated the recognized idle times. Thus, we cannot make any statement on the influence that material flow, such as soil disposal or concrete delivery, or operation variabilities, such as tool change or casing screw, have on the production system. However, these influences are essential for its optimization.

RQ3: What is needed to adopt production models for DTCs in heavy civil engineering?

Based on the previous RQs, we propose that DTCs require a production model that reflects the system's specific components (Figure 15).

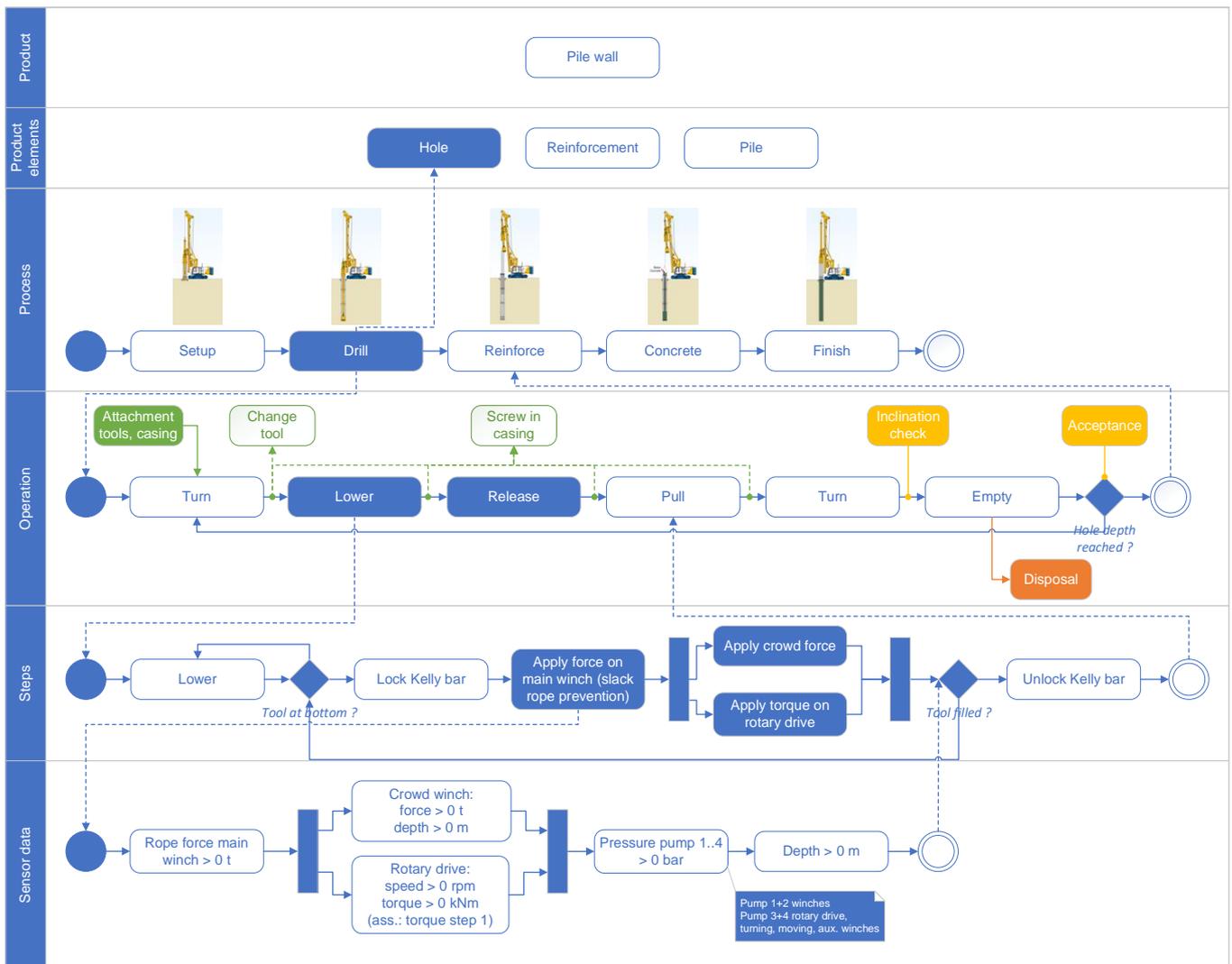


Figure 15. Activity diagram of the Kelly pile production system and its flows (material supply (green)/removal (orange), information (yellow)) focusing on the process step ‘Drill’ (blue).

The two components of the DTC, activity recognition and the data-driven DES, serve different purposes: Whereas activity recognition analyzes sensor data, requiring equipment motion steps, the DES analyzes performance requiring value-adding and non-value-adding subprocesses. Although these components have specific requirements, they pursue the same purpose: making the production flow reliable. This reliability is focused on making the production flow, or the sequence of its processes, visible and understandable to models. The next step is then to scale these models toward a catalog of production models (Figure 16). Nübel describes this scaling process generically in an evolution, such as power lines. Considering a top-down planning approach, project planning starts by adapting a reference project, deriving the flow-based production model, and modifying it with time. This evolution leads to an enlarged database capturing different variants of the models.

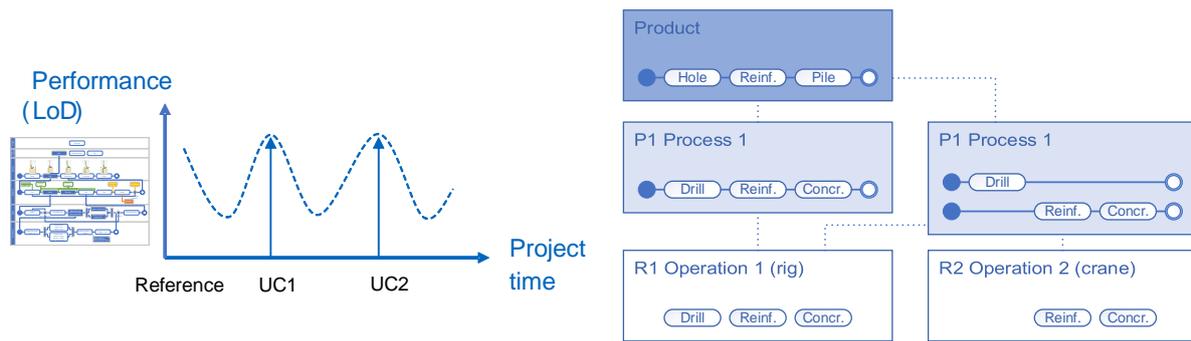


Figure 16. Nübel's catalog of construction processes (UC: use case; LoD: level of detail).

8. Conclusions

This paper focused on digital-twin integration research in the equipment-intensive construction industry. While other studies investigated excavators, we gave insights into the Kelly drill rig for pile production, which is more complex equipment regarding production requirements and data analytics (more than ten labels and features). Based on two studies, one on activity recognition and one on data-driven DES, the different requirements for the model's granularity are presented. On the other hand, activity recognition modeling needs to be more related to the equipment's motions than the process level. On the other hand, DES models can help to rethink the given production system if they represent its flow. These results reveal the gap among the different roles of a DTC and contribute to a flow-based technology production model that combines both a data-driven and a production-driven perspective. The objective is to maintain a modularized catalog of these models. Future research will evaluate the proposed methodology, e.g., to consider deviation detection and bottleneck identification in construction operations (as proposed by Rashid and Louis [7]). The proposed DTC needs to be implemented in an application including dynamic data transmission and integration from the drilling rig to the DES and back. Overall, the results need to be validated on an on-going construction project to show the practicability of DTCs. In addition, the fusion of sensor data and image data, or the rules of state transitions, must be considered to overcome the limitations that arise during activities where the equipment is not constantly involved in the production process, e.g., with concrete.

Author Contributions: Conceptualization, A.F., A.B.B., I.D.T. and K.N.; methodology, A.F., A.B.B., I.D.T. and K.N.; software, A.B.B.; validation, A.F.; formal analysis, A.F. and A.B.B.; investigation, A.F. and A.B.B.; resources, A.F. and A.B.B.; data curation, A.F.; writing—original draft preparation, A.F. and A.B.B.; writing—review and editing, I.D.T. and K.N.; visualization, A.F. and A.B.B.; supervision, I.D.T., K.N. and J.F.; project administration, J.F.; funding acquisition, I.D.T. and J.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung) BMBF grant number 02P17D249, the Bavarian Collaborative Research Program of the Bavarian State Government grant number DIK0415/08, and the Project Production Systems Laboratory (P2SL, p2sl.berkeley.edu) at UC Berkeley. The publishing fee for this open access journal paper was funded by the TUM Publishing Fund.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available as it was generated by the industry partner BAUER Group.

Acknowledgments: All support is gratefully acknowledged.

Conflicts of Interest: Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of BMBF, Bavarian State Government, or members of P2SL.

Appendix A

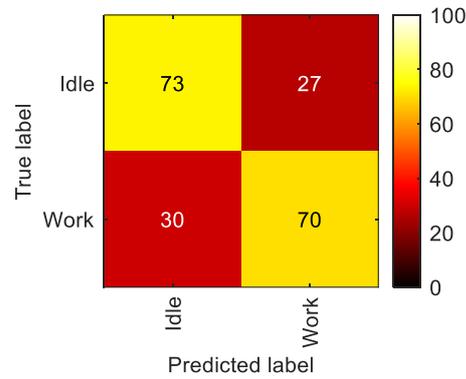


Figure A1. LoD1 confusion matrix for the MLP.

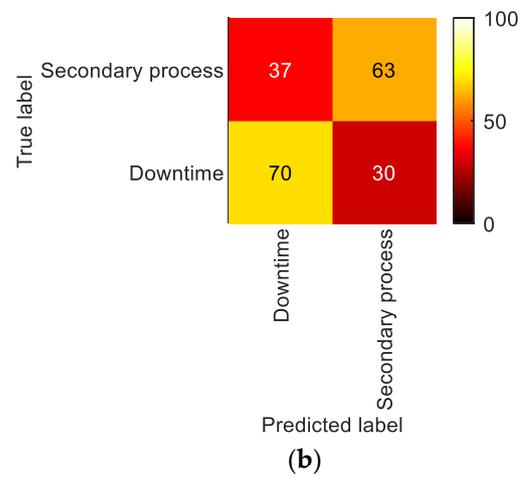
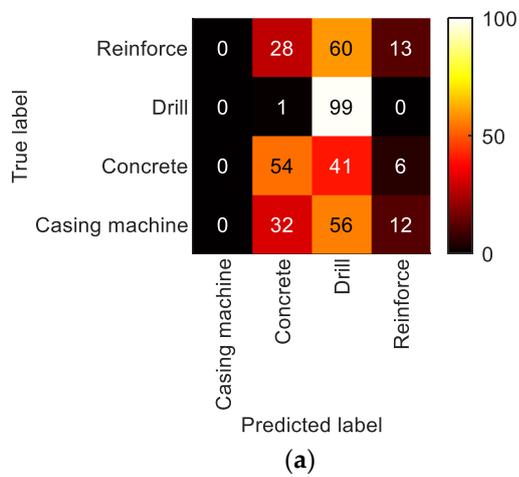


Figure A2. LoD2 confusion matrix for the MLP: (a) work; (b) idle.

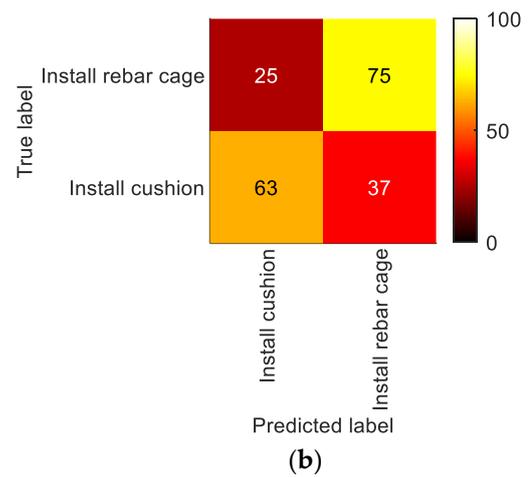
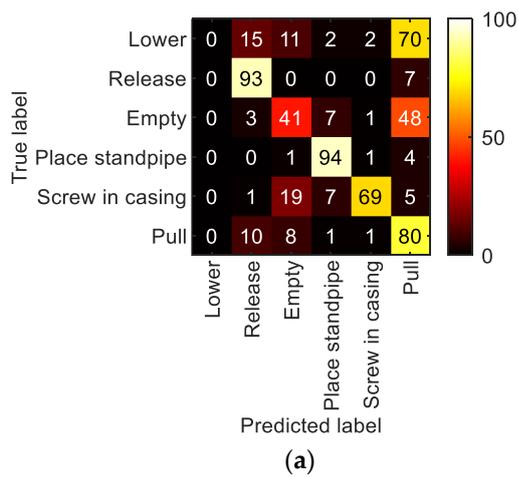


Figure A3. Cont.

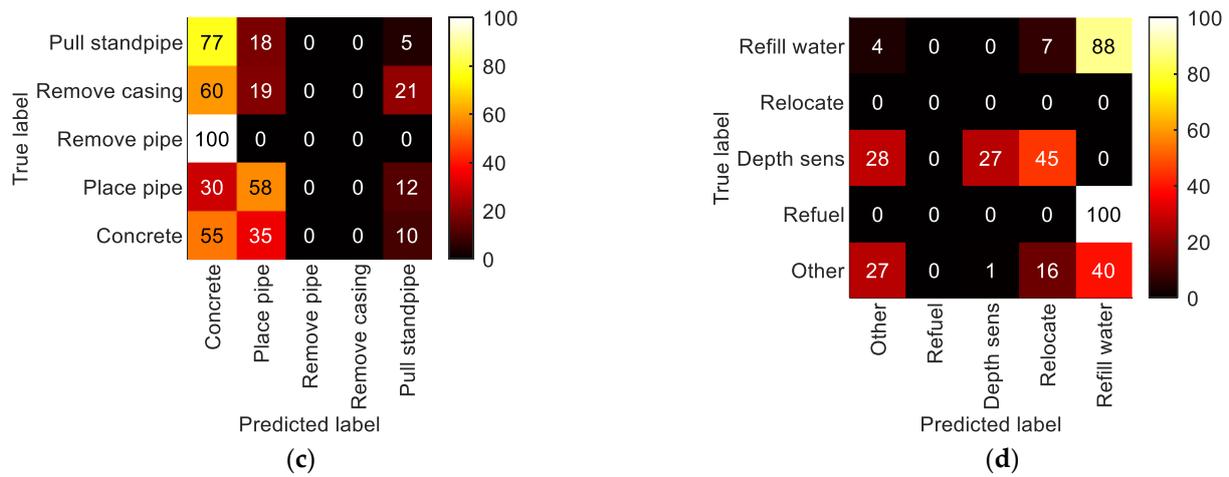


Figure A3. LoD3 confusion matrix for the MLP: (a) drill, (b) reinforce; (c) concrete; and (d) secondary processes.

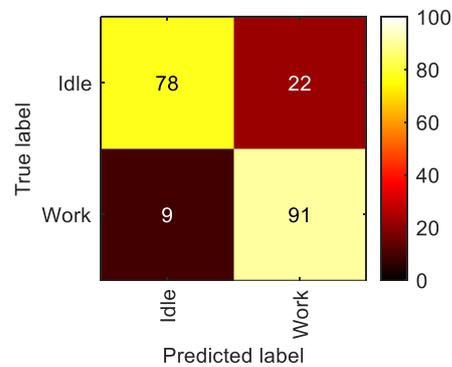


Figure A4. LoD1 confusion matrix for the DeepConvLSTM.

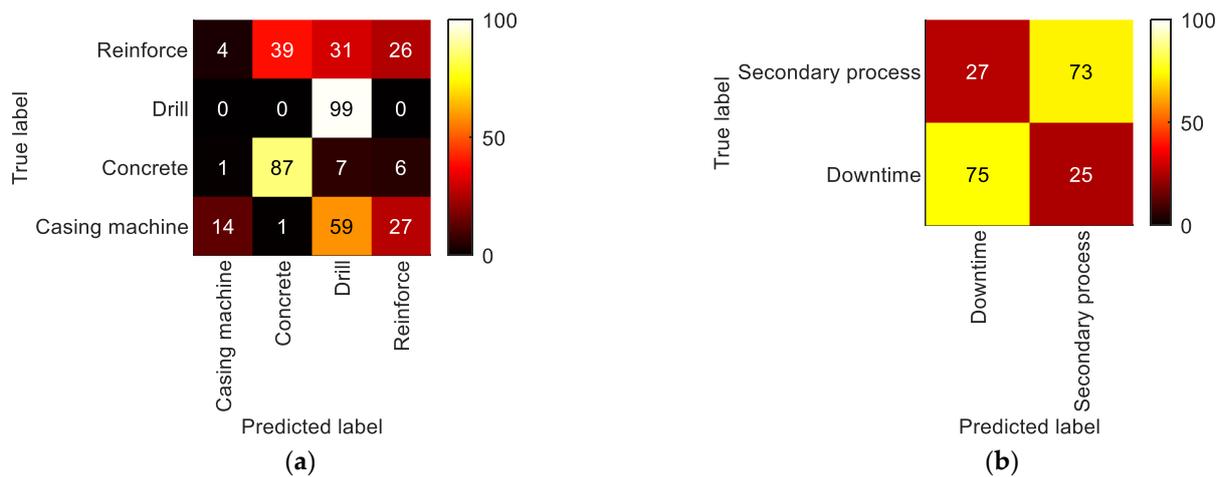


Figure A5. LoD2 confusion matrix for the DeepConvLSTM: (a) work; (b) idle.

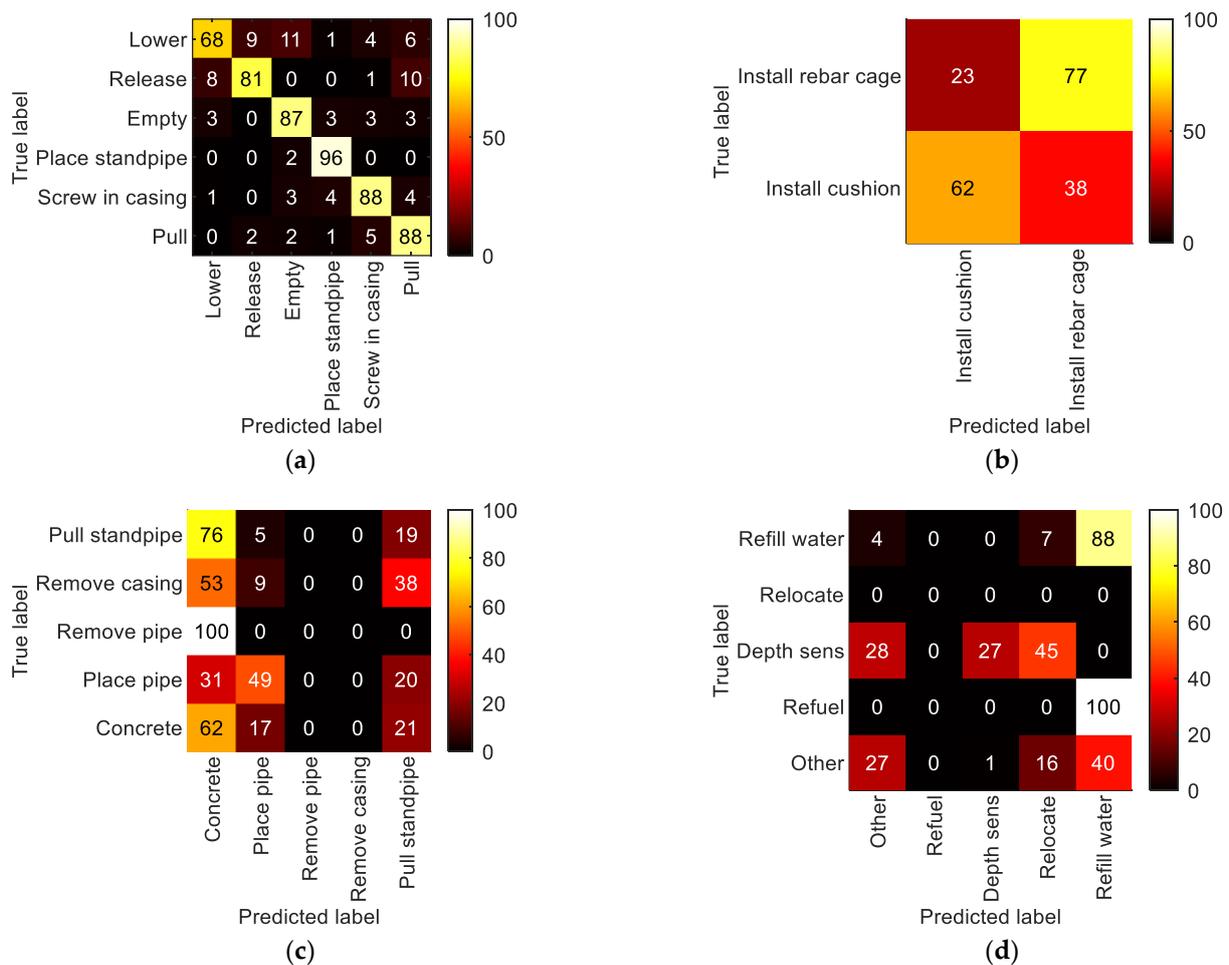


Figure A6. LoD3 confusion matrix for the DeepConvLSTM: (a) drill, (b) reinforce; (c) concrete; and (d) secondary processes.

References

- Lasi, H.; Fettke, P.; Kemper, H.-G.; Feld, T.; Hoffmann, M. Industry 4.0. *Bus. Inf. Syst. Eng.* **2014**, *6*, 239–242. [CrossRef]
- Fuller, A.; Fan, Z.; Day, C.; Barlow, C. Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* **2020**, *8*, 108952–108971. [CrossRef]
- Oesterreich, T.D.; Teuteberg, F. Understanding the implications of digitisation and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. *Comput. Ind.* **2016**, *83*, 121–139. [CrossRef]
- Turner, C.J.; Oyekan, J.; Stergioulas, L.; Griffin, D. Utilizing Industry 4.0 on the Construction Site: Challenges and Opportunities. *IEEE Trans. Ind. Inf.* **2021**, *17*, 746–756. [CrossRef]
- Hu, W.; Lim, K.Y.H.; Cai, Y. Digital Twin and Industry 4.0 Enablers in Building and Construction: A Survey. *Buildings* **2022**, *12*, 2004. [CrossRef]
- Sacks, R.; Brilakis, I.; Pikas, E.; Xie, H.S.; Girolami, M. Construction with digital twin information systems. *Data-Cent. Eng. (DCE)* **2020**, *1*, e14. [CrossRef]
- Rashid, K.M.; Louis, J. Integrating Process Mining with Discrete-Event Simulation for Dynamic Productivity Estimation in Heavy Civil Construction Operations. *Algorithms* **2022**, *15*, 173. [CrossRef]
- Akhavian, R.; Behzadan, A.H. Construction equipment activity recognition for simulation input modeling using mobile sensors and machine learning classifiers. *Adv. Eng. Inform.* **2015**, *29*, 867–877. [CrossRef]
- Rashid, K.M.; Louis, J. Times-series data augmentation and deep learning for construction equipment activity recognition. *Adv. Eng. Inform.* **2019**, *42*, 100944. [CrossRef]
- Golparvar-Fard, M.; Heydarian, A.; Niebles, J.C. Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. *Adv. Eng. Inform.* **2013**, *27*, 652–663. [CrossRef]
- Rashid, K.M.; Louis, J. Automated Activity Identification for Construction Equipment Using Motion Data from Articulated Members. *Front. Built Environ.* **2020**, *5*, 144. [CrossRef]

12. Slaton, T.; Hernandez, C.; Akhavian, R. Construction activity recognition with convolutional recurrent networks. *Autom. Constr.* **2020**, *113*, 103138. [[CrossRef](#)]
13. Fischer, A.; Liang, M.; Orschlet, V.; Bi, H.; Kessler, S.; Fottner, J. Detecting Equipment Activities by Using Machine Learning Algorithms. In Proceedings of the 17th IFAC Symposium on Information Control Problems in Manufacturing (INCOM 2021), Budapest, Hungary, 7–9 June 2021.
14. Fischer, A.; Bedrikow, A.B.; Kessler, S.; Fottner, J. Equipment data-based activity recognition of construction machinery. In Proceedings of the 2021 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, UK, 21–23 June 2021; pp. 1–6.
15. Sherafat, B.; Ahn, C.R.; Akhavian, R.; Behzadan, A.H.; Golparvar-Fard, M.; Kim, H.; Lee, Y.-C.; Rashidi, A.; Azar, E.R. Automated Methods for Activity Recognition of Construction Workers and Equipment: State-of-the-Art Review. *J. Constr. Eng. Manag.* **2020**, *146*, 3120002. [[CrossRef](#)]
16. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihm, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine* **2018**, *51*, 1016–1022. [[CrossRef](#)]
17. AbouRizk, S. Role of Simulation in Construction Engineering and Management. *J. Constr. Eng. Manag.* **2010**, *136*, 1140–1153. [[CrossRef](#)]
18. Fischer, A.; Li, Z.; Kessler, S.; Fottner, J. Importance of secondary processes in heavy equipment resource scheduling using hybrid simulation. In Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC), Dubai, United Arab Emirates, 4 August 2021; Feng, C., Linner, T., Brilakis, I., Castro, D., Chen, P.-H., Cho, Y., Du, J., Ergon, S., Garcia de Soto, B., Gasparik, J., et al., Eds.; International Association for Automation and Robotics in Construction (IAARC): Pittsburgh, PA, USA, 2021.
19. Akhavian, R.; Behzadan, A.H. Automated knowledge discovery and data-driven simulation model generation of construction operations. In Proceedings of the 2013 Winter Simulations Conference (WSC), Washington, DC, USA, 8–11 December 2013; pp. 3030–3041.
20. Kargul, A.; Günthner, W.; Bügler, M.; Borrmann, A. Web based field data analysis and data-driven simulation application for construction performance prediction. *Electron. J. Inf. Technol. Constr.* **2015**, *20*, 479–494.
21. Liu, C.; AbouRizk, S.; Morley, D.; Lei, Z. Data-Driven Simulation-Based Analytics for Heavy Equipment Life-Cycle Cost. *J. Constr. Eng. Manag.* **2020**, *146*, 04020038. [[CrossRef](#)]
22. Louis, J.; Dunston, P.S. Methodology for Real-Time Monitoring of Construction Operations Using Finite State Machines and Discrete-Event Operation Models. *J. Constr. Eng. Manag.* **2017**, *143*, 04016106. [[CrossRef](#)]
23. Martinez, J.C. Stroboscope: State and Resource Based Simulation of Construction Processes. Ph.D. Dissertation, University of Michigan, Ann Arbor, MI, USA, 1996.
24. Kim, H.; Bang, S.; Jeong, H.; Ham, Y.; Kim, H. Analyzing context and productivity of tunnel earthmoving processes using imaging and simulation. *Autom. Constr.* **2018**, *92*, 188–198. [[CrossRef](#)]
25. Halpin, D.W.; Riggs, L.S. *Planning and Analysis of Construction Operations*; Wiley: New York, NY, USA, 1992.
26. Halpin, D.W. An Investigation of the Use of Simulation Networks for Modeling Construction Operations. Ph.D. Dissertation, University of Illinois, Champaign, IL, USA, 1973.
27. Siemens. Tecnomatix Digital Manufacturing Software. Available online: <https://plm.sw.siemens.com/de-DE/tecnomatix/> (accessed on 29 December 2022).
28. Fischer, A.; Li, Z.; Wenzler, F.; Kessler, S.; Fottner, J. Cyclic Update of Project Scheduling by Using Telematics Data. *IFAC-PapersOnLine* **2021**, *54*, 217–222. [[CrossRef](#)]
29. Fischer, A.; Balakrishnan, G.; Kessler, S.; Fottner, J. Begleitende Prozesssimulation für das Kellybohrverfahren [Accompanying Process Simulation for the Kelly Drilling Process]. In *8. Fachtagung Baumaschinentechnik 2020*; TU Dresden: Dresden, Germany, 2020.
30. Harichandran, A.; Raphael, B.; Mukherjee, A. A hierarchical machine learning framework for the identification of automated construction. *ITcon* **2021**, *26*, 591–623. [[CrossRef](#)]
31. Koskela, L. An Exploration towards a Production Theory and Its Application to Construction. Ph.D. Dissertation, Technical Research Centre of Finland, Espoo, Finland, 2000.
32. Ballard, H.G. The Last Planner System of Production Control. Ph.D. Dissertation, University of Birmingham, Birmingham, UK, 2000.
33. Koskela, L. Application of the New Production Philosophy to Construction. *Stanf. Univ. Tech. Rep.* **1992**, *72*, 39.
34. Tommelein, I.D. Impact of Variability and Uncertainty on Product and Process Development. In *Construction Congress VI*; Walsh, K.D., Ed.; American Society of Civil Engineers: Reston, VA, USA, 2000; pp. 969–976.
35. Ohno, T. *Toyota Production System: Beyond Large-Scale Production*; Productivity Press: Portland, IN, USA, 1988.
36. Fischer, A.; Grimm, N.; Tommelein, I.D.; Kessler, S.; Fottner, J. Variety in Variability in Heavy Civil Engineering. In Proceedings of the 29th Annual Conference of the International Group for Lean Construction (IGLC), Lima, Peru, 12–18 July 2021.
37. Kalsaas, B.T. Further Work on Measuring Workflow in Construction Site Production. In Proceedings of the 20th Annual Conference of the International Group for Lean Construction, San Diego, CA, USA, 18–20 July 2012.
38. Xu, C.; Chai, D.; He, J.; Zhang, X.; Duan, S. InnoHAR: A Deep Neural Network for Complex Human Activity Recognition. *IEEE Access* **2019**, *7*, 9893–9902. [[CrossRef](#)]

39. 1536:2010+A1:2015, 93.020; Execution of Special Geotechnical Work—Bored Piles: German Version EN. DIN Deutsches Institut für Normung: Berlin, Germany, 2015.
40. Nübel, K.; Geiss, A.; Sommer, F.; Pielmeier, M.; Heinrich, M.; Rehfeld, B. Produktionsplanung und Produktionssteuerung im Spezialtiefbau [Production planning and production control in special foundation engineering]: Prozessorientierter Ablauf von Bauprojekten im Spezialtiefbau. *Bauing. VDI Bautech.* **2015**, 101–107.
41. Beiderwellen Bedrikow, A. Equipment Data-Based Activity Recognition of Construction Machinery. Bachelor Thesis, Chair of materials handling, material flow, logistics. Technical University of Munich (TUM), Garching, Germany, 2021.
42. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
43. Ordóñez, F.J.; Roggen, D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)]
44. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media: Newton, MA, USA, 2019.
45. Goodfellow, I.; Courville, A.; Bengio, Y. *Deep Learning: Das umfassende Handbuch: Grundlagen, aktuelle Verfahren und Algorithmen, neue Forschungsansätze [The Comprehensive Handbook: Fundamentals, Current Methods and Algorithms, New Research Approaches.]*, 1st ed.; Verlags GmbH & Co. KG: Frechen, Germany, 2018.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.