Digital Twin for Production Estimation, Scheduling and Real-Time Monitoring in Offsite Construction

by

Fatima Alsakka

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Construction Engineering and Management

Department of Civil and Environmental Engineering
University of Alberta

Abstract

The offsite construction industry continues to rely on experience-based average production rates (i.e., average quantity per unit of time) to estimate and schedule production operations. This approach is hindered by various sources of production variability, such as machine breakdowns and material shortages, often resulting in high production estimation and scheduling errors; in fact, as described herein, using average production rates may result in overly optimistic production schedules, leading to missing schedule deadlines, cost overruns, and, most critically, an overburdened workforce. In this context, this thesis proposes a digital twin to enable dynamic production estimation, scheduling, and real-time monitoring of production operations in offsite construction with more accuracy compared to the current practice. The proposed digital twin comprises three major subsystems: (1) an estimation and scheduling subsystem, which estimates variable cycle times as a function of various factors that influence them and virtually mimics operations to estimate production time and generate production schedules; (2) a computer-vision-based data acquisition subsystem that enables the continuous collection of data necessary for regular tuning of the estimation models, accommodating new sources of variability; and (3) a real-time monitoring subsystem to monitor production operations in real time, tracking progress on production schedules and enabling the generation of updated schedules promptly in response to any deviations from the actual operations.

To support the development of these subsystems and their requisite functionalities, four main research objectives are pursued: (1) develop and examine a system that deploys computer-vision technology for the automated and accurate acquisition of cycle time data in a timely and cost-effective manner; (2) devise a methodical approach for the identification and understanding of the factors driving cycle time variability, and evaluate how this identification process improves the accuracy of cycle time estimation; (3) design and develop a data- and knowledge-driven system that estimates cycle times in consideration of various influencing factors and using automatically collected data to increase the estimation accuracy compared to traditional estimation methods; and (4) devise a feasible design of a digital twin that enables dynamic and more accurate production estimation, scheduling, and real-time monitoring in offsite construction factories. A diverse array of methods and technologies, including computer vision, 3D simulation, machine-learning-based prediction, statistical modelling, ultrasonic sensors, semi-structured interviews, direct observation, and literature reviews, are deployed and integrated to achieve these objectives.

A prototype of the digital twin is developed for a wall framing workstation within a panelized construction factory. The results show that average errors of less than 1 minute in data acquisition, a 36% reduction in cycle time estimation errors, and an 81% reduction in deviations between the production schedule and actual production are achieved compared to the current practice of relying on experience-based average production rates.

Preface

This thesis is an original work by Fatima Alsakka and follows a paper-based format. Excluding the introductory chapter (i.e., Chapter 1) and the concluding chapter (i.e., Chapter 7), the publication status of each chapter is as follows:

- A version of Chapter 2 has been published as "F. Alsakka, S. Assaf, I. El-Chami, M. Al-Hussein, Computer vision applications in offsite construction, Automation in Construction. 154 (2023) Article no. 104980. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104980."
- A version of Chapter 3 has been published as "F. Alsakka, I. El-Chami, H. Yu, M. Al-Hussein, Computer vision-based process time data acquisition for offsite construction, Automation in Construction. 149 (2023) Article no. 104803. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104803."
- A version of Chapter 4 has been published as "F. Alsakka, H. Yu, F. Hamzeh, M. Al-Hussein, Factors influencing cycle times in offsite construction, in: Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC), Lille, France, 2023: pp. 723–734. https://doi.org/https://doi.org/10.24928/2023/0123."
- A version of Chapter 5 has been submitted for publication as "F. Alsakka, H. Yu, F. Hamzeh, M. Al-Hussein, Data- and knowledge-driven cycle time estimation in offsite construction, Journal of Computers and Industrial Engineering." A preprint of this chapter will be made available online prior to the completion of the peer review process.
- A version of Chapter 6 has been submitted for publication as "F. Alsakka, I. El-Chami, H. Yu, F. Hamzeh, M. Al-Hussein, Digital twin for production estimation, scheduling and real-time monitoring in offsite construction, Journal of Operations Management."

Fatima Alsakka holds primary responsibility for the conceptualization of the above studies, development, analyses, interpretation, and manuscript writing. Dr. Mohamed Al-Hussein had a supervisory role, providing necessary resources for conducting the studies and contributing to the conceptualization and manuscript writing. Dr. Farook Hamzeh had a supervisory role and contributed to the conceptualization and writing of the corresponding manuscripts. Dr. Haitao Yu, as an industry partner, had a supervisory role, facilitated data collection, oversaw the development of the work, and validated the study findings. Dr. Ibrahim El-Chami served in a supervisory capacity, providing technical assistance for the use of computer vision and ultrasonic sensors, and contributing to the conceptualization and writing of the corresponding manuscripts. Sena Assaf served as a second reviewer in the scoping review manuscript and contributed to the manuscript writing.

The research presented in this thesis obtained research ethics approval from the University of Alberta Research Ethics Boards under the project name "A Digital Twin-Centric System for Dynamic Production Estimation and Scheduling in Wood Construction Manufacturing," with the approval number Pro00121305, dated November 19, 2022.

Acknowledgements

I would like to express my heartfelt appreciation to all those who have been instrumental in the completion of this thesis, including:

- Dr. Mohamed Al-Hussein, my supervisor, for his unwavering support and guidance, for inspiring the very idea behind this research, and for providing all the resources without which this research would not have been possible.
- Dr. Farook Hamzeh, my co-supervisor, for playing a key role in my academic growth since the very start of my academic journey, for his constant motivation for me to give my best, and for his continuous guidance and support.
- Dr. Haitao Yu, our industry partner, whose guidance and support were indispensable in bringing this research to fruition.
- Dr. Yasser Mohamed, for teaching me essential skills necessary for conducting this research and for generously sharing his wisdom and support.
- Dr. Ibrahim El-Chami, for his invaluable guidance and contributions to this research.
- Mr. Jonathan Tomalty, our technical editor, for his generous support throughout my studies, invaluable contributions, and meticulous technical review of our publications.
- Trina Cattral, our graduate program advisor, for her kind help and support throughout my entire program of studies.
- Dr. Luis Alarcón and Dr. Vicente Gonzalez, my examining committee members, for their review of my work and their valuable insights.
- My friends in Edmonton, who became my second family and made my PhD journey enjoyable and enriching.
- My family, to whom I owe everything, and whose support and love have brought me to where I am today.
- Last but not least, my beloved partner, for his constant encouragement in helping me grow into a better version of myself.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The construction sector is undergoing a notable shift towards the widespread adoption of offsite construction methods, a paradigm sometimes referred to as construction manufacturing or prefabricated construction. This trend is expected to continue, with the global market projected to expand from around \$130 billion in 2020 to a potential \$230 billion by 2030 (Khandelwal, 2021). This shift towards construction manufacturing involves the application of theories and practices that form the foundation of manufacturing systems. In simple terms, a manufacturing system is a "combination of humans, machinery, and equipment that are bound by a common material and information flow" (Caggiano, 2014). Accordingly, workers, machinery, and equipment in offsite construction factories are typically positioned in fixed locations at workstations, with each workstation assigned a well-defined production process (e.g., wall framing, window/door installation) or set of productions tasks that are part of a process. Despite this shift towards manufacturing, the industry still lacks well-established management tools for estimating production time (i.e., the total time needed to produce building components) and developing production schedules accordingly. In current practice, the tasks of estimating production time and developing production schedules are typically performed using average production rates—e.g., linear metres of wall panels per minute (Alsakka et al., 2023d)—derived from experience gained from previous projects. However, there are two main issues with this approach:

- Production time is contingent upon cycle times (CTs) of processes at workstations, where CT is defined as the time spanning from the start of a process cycle until its end. Meanwhile, CTs of production processes are highly variable due to the nonstandard design of building components processed at workstations in construction manufacturing. To illustrate this, a recent study of a panelized construction factory where wall panels, floor panels, roof components, and staircases are fabricated for on-site installation revealed a wide range of CTs for wood wall framing operations, with values ranging from as little as 2 minutes to as much as 58 minutes (Alsakka et al., 2023d). This high variability stems from numerous influencing factors associated with the component design, workforce, machinery, materials, workstation configuration, production line arrangement, factory workflows, and external conditions (Alsakka et al., 2023c). Given the wide range of CTs and the multitude of factors influencing them, an approach that uses average rates (e.g., square footage per minute) is likely to fall short of capturing the true production scenario. Notably, in the aforementioned case study, the use of average rates led to overly optimistic CT estimates (Alsakka et al., n.d.), (which, in turn, would result in optimistic production time estimates and impractical schedules).
- Manufacturing systems are designed and operated to satisfy specific business objectives (Chryssolouris, 2006). As such, they incorporate specific applications of science and technology (through their processes and machinery) with direct relevance to particular products (Parnaby, 1979). Put differently, these systems are tailored to meet the precise requirements of the products they are designed to produce. Consequently, substantial alterations in product requirements have the potential to disrupt manufacturing operations, making prior production experience largely irrelevant as a point of reference in estimating and scheduling the production of the product at hand.

  An example of this issue was observed in the lightweight wood panel production facility at which the research presented in this thesis was conducted. On the wall panel production line at this facility, a multi-function bridge is used for nailing sheathing boards (e.g., plywood, oriented strand board) to exterior walls in an automated fashion. The multi-function bridge is configured to nail materials with a specific degree of stiffness (dictating the shooting pressure of the nailing gun) that is a common setting in

> automated multi-function bridges currently used in offsite construction. Meanwhile, a new practice has emerged in the industry for “zero-lot-line” houses, whereby only drywall (i.e., no sheathing) is required for some of the exterior walls. The multi-function bridge was not suitable for executing nailing tasks on drywall, so nailing of these exterior walls had to be done manually. This constituted a major disruption to the standard workflow. This resulted in four to five workers operating at the sheathing workstation rather than two, schedule delays, as well as worker demoralization and burnout. This was largely attributable to a lack of data concerning the CTs of the manual nailing process and how they compared to those of the multi-function bridge. In particular, the experience that management personnel were drawing upon in their decision making in this instance was based on the use of the multi-function bridge for nailing operations in past production, so they were not able to accurately estimate the CTs of the new processes based on their experience. Although the fabrication sequence of panels could be optimized to alleviate this problem (e.g., alternate between interior walls and walls with OSB sheathing, on one hand, and exterior walls with drywall only, on the other hand), they were not able to effectively update their schedules without reliable production time measurements.

As Parnaby (1979) argued decades ago in this regard, to maintain stability of operations against variability, a supply of information and continuous data acquisition are required. Indeed, the variability of CTs over time must be continuously monitored in order to capture the effect of new sources of variability. Hence, automated data acquisition is an important dimension of the solution to the challenge of achieving reliable production time estimation and scheduling. However, another important dimension of the solution is to identify the underlying factors driving this variability. With data automatically acquired on CTs and knowledge of the factors influencing them, the average production rates traditionally used for production estimation can be replaced with a more robust machine-learning-based estimation method that is driven by both data and knowledge and that can be regularly fine-tuned to account for new sources of variability in operations. However, machine-learning methods, while powerful, are not guaranteed to fully capture or explain all sources of variability, and this deficiency is likely to result in estimation errors. Such errors may accumulate over time and, if not addressed, result in significant deviations between production schedules and actual operations. Digital twins have the potential to address this limitation, gain real-time insights into operational activities, and achieve effective real-time monitoring of production schedules, thereby enabling a more dynamic method of production scheduling in which schedules are updated in response to changes. This thesis thus proposes a solution for dynamic production estimation, scheduling, and real-time monitoring encompassing the following four dimensions: automated data acquisition (Dimension I); factors influencing CTs (Dimension II); data- and knowledge-driven CT estimation (Dimension III); and dynamic scheduling and digital twin design (Dimension IV).

1.2 Background and research needs

The following subsections provide background information and overviews of the research needs related to each of these four dimensions. More detailed discussions are provided in the remaining chapters of this thesis.

1.2.1 Dimension I: Automated data acquisition

A diverse range of tools and methods for data acquisition are available today. Among these, a notable technology that has been gaining momentum within the offsite construction industry for automated data acquisition is computer vision. Computer vision is a field of artificial intelligence in which computing systems are used to extract meaningful information from visual components (IBM, 2022), such as digital images and videos (e.g., CCTV), in an attempt to mimic the way in which the human brain perceives and understands visual information (Huang, 1996). Computer-vision technology enables the accurate identification and classification of objects, leading to data-driven recommendations or actionable insights within various systems. However, depending on the design of the given computer-vision-based methodology used to achieve these goals, the typical process of building computer-vision models often requires extensive up-front set-up efforts. Typically, a dataset comprising images or videos of the targeted object must first be created. The dataset must then be annotated, and this could entail classifying images or drawing bounding boxes around the objects of interest and classifying them. The annotated dataset is then used to train a computer-vision model, and this task also requires significant computational efforts. The annotated dataset must also be of sufficient size and quality to ensure adequate performance. In a word, the task of developing the computer-vision models is an iterative process that requires significant time, effort, and resources. In what follows, computer-vision models built following the steps described above are referred to as custom vision models. As per Microsoft's definitions, computer-vision models are pre-trained and process images based on users' input (Microsoft, 2022a), while custom vision models are trained by the user (Microsoft, 2022b).

The significant set-up effort required on the part of construction researchers in order for custom vision to be successfully deployed is evident in productivity- and efficiency-related applications. For instance, in a recent study applying computer vision to automatically count the number of modules installed in modular construction projects, Zheng et al. (2020) simulated the modular construction process to capture virtual images. The simulation was used to increase the size of the dataset that was to be used to train the object detection algorithm to detect building modules, since the number of real images was limited. In a similar study directed at automatically tracking the hoisting and installation of precast walls, Wang et al. (2021b) had to manually select 580 images containing precast walls from a dataset of surveillance images and videos from four different construction sites. They further adjusted the resolution of the images and labelled the precast objects before training the computer-vision model. In a related work, Wang et al. (2021c) used 600 similar images, representing 6.5 hours of GPU computational requirements and equating to more than 24 CPU-mode computational hours. In another example, Ahmadian Fard Fini et al. (2021) endeavored to develop an automated method for measuring the onsite installation rate (cm^2/minute) of prefabricated panels in panelized construction. They used a wide-angle lens to capture the entire scene on a construction site, but this resulted in non-uniform zoom levels across each image. As a result, they had to further pre-process the images to remove distortions in order to minimize the effects on the model's performance. In another example, Martinez et al. (2021) collected and manually labelled 1,069 images in order to detect workers and equipment (i.e., a multi-function bridge and a crane) in a wood panel manufacturing facility, focusing on the floor panel station. There is a clear value to these productivity- and efficiency-related custom vision applications, and this value could be increased if more consideration is given to the "timeliness" dimension of data quality. Therefore, there is a need for a timely and efficient method for deploying computer-vision technology for automated data acquisition.

1.2.2 Dimension II: Factors influencing cycle times

In light of the impact of variability on production operations as discussed above, researchers have employed machine-learning models for estimating process time variables considering a variety of related influencing factors. Examples of the considered factors include (1) product-related (or design-related) factors such as the number of single studs, double studs, doors, windows, cutting zones, drill holes, nails, screws, etc., for estimation in wood wall panel production (Shafai, 2012), the number of fittings and cut-outs for steel fitting (Song et al., 2008), the number of bolts, length of weld, length of wide flange beams, etc., for estimation in structural steel manufacturing (Hu et al., 2015), and the nominal height, weight, width, concrete volume, finishing area, reinforcement weight, concrete strength, etc., for estimation in precast concrete production (Benjaoran et al., 2004); (2) worker-related factors such as the number of workers (Benjaoran et al., 2004), and skill level (Song et al., 2008); (3) material-related factors such as length and weight (Hu et al., 2015; Song et al., 2008); (4) machine-related factors such as machine breakdowns (Song et al., 2008); (5) factory operations-related factors such as work shift (Song et al., 2008); and (6) production line-related factors such as activity precedence relationships, queuing, and rework (Song et al., 2008).

The benefit of investing effort on identifying these factors is twofold; it not only helps to identify factors that could hold significant information with regards to the estimated process time variable, but it also deepens the estimator's understanding of the process under study. This, in turn, allows the estimator to follow a prescriptive approach for selecting and representing the influencing factors (also known as predictor variables) used in the machine-learning estimation models (Kuhn et al., 2019). In other words, it enables the estimator to carry out knowledge-driven modelling alongside data-driven modelling, reducing the risk of overfitting to erroneous data patterns or of generating models that cannot be rationally interpreted (compared to an approach that relies solely on empirical data) (Kuhn et al., 2019). As such, there is significant value in giving consideration to a range of influencing factors and studying their impact on the performance of models used in estimating process time variables—including CTs—in offsite construction factories. However, although the reliability of a machine-learning model is a function of the exhaustiveness of the influencing factors considered in the model (Benjaoran et al., 2004), the identified studies either have not taken a systematic approach or have not thoroughly discussed the approach followed for identifying the list of factors that may have an effect on the process time variables under study. As such, there is a need for a structured approach to identifying the factors influencing CTs in offsite construction.

1.2.3 Dimension III: Data- and knowledge-driven cycle time estimation

As mentioned above, machine-learning-based methods—or, in other words, data-driven methods—have been deployed in a number of studies for estimating process time variables in offsite construction. For example, neural networks (NNs) and multivariable linear regression (LR) models were applied to precast concrete production to estimate productivity in consideration of influencing factors related to product shape, material, and manpower (Benjaoran et al., 2004, 2006). Mohsen et al. (2022) conducted research on productivity in wood panelized construction, training several machine-learning algorithms to estimate the time required to complete processes on a wall production line as a function of both design-related factors (e.g., length, width, number of studs) and factors related to work in progress (e.g., the count of wall panels being processed on the production line). In another study in panelized construction, this one targeting the transportation phase, Ahn et al. (2020) trained support vector regression models using GPS data in order to predict transportation durations for a given project as a function of product-related factors such as the total floor area and total wall area, as well as site-related factors such as location and the maturity of the neighbourhood.

However, there is a lack of research focusing on estimating process CTs at the workstation level in the production phase for offsite construction—a requirement for production time estimation. Shafai built LR models for estimating the durations of specific tasks (e.g., spray foam insulation) as a function of the unique design properties of the given panel that are significant to the given task (e.g., number of studs, number of openings, number of cutting zones). Stochastic factors (e.g., triangular distribution) were also incorporated in the regression models to account for uncertainties such as worker performance or machine breakdown. Similar estimation methods have been used in other studies to estimate the processing times at workstations (Altaf et al., 2014; Bhatia et al., 2019; H. Liu et al., 2015). Although such approaches constitute an improvement compared to using a single average value to model the duration of an entire process at a workstation, further investigation is warranted with respect to the following three research areas:

(1) As discussed in the Motivation section of this chapter, continuously acquired data is needed for estimation purposes in order to capture variations in CTs over time. Hence, there is value in testing the feasibility of training machine-learning-based CT estimation models using automatically acquired data. Given the growing momentum of computer-vision applications as discussed above, it is worth testing the validity of using computer-vision data for training estimation models.
(2) In relation to Dimension II, it is worthwhile to consider a range of influencing factors and to study their impact on the performance of estimation models in predicting CTs in offsite construction factories. A deep understanding of these factors and their impact on CT variability allows for the integration of knowledge-driven machine-learning modelling in addition to data-driven modelling, ultimately improving the predictive accuracy of the machine-learning models.
(3) Another area that merits further exploration has to do with the performance of the machine-learning algorithms used in the estimation approaches. While LR models are commonly considered, their potential misuse when the assumptions that underlie them do not hold has been noted in previous research (Lu, 2000). Meanwhile, NNs are capable of modelling complex problems, which can be difficult to model using traditional classical mathematical methods (Adeli, 2001). Hence, NNs have been long deemed a suitable tool for modelling problems in construction research (Moselhi et al., 1991), and they have been used for various applications, e.g., (Benjaoran et al., 2004, 2006; Lu et al., 2000; Song et al., 2008). However, in the aforementioned study by Mohsen et al. (2022), among the models considered, including random forest (RF), LR, *k*-nearest neighbour, and NN, the LR model was found to perform slightly better than the NN model when trained on an engineered dataset to predict the production time of wall panels, and the best performing model was the RF model (Mohsen et al., 2022). These findings suggest that further examination of the performance of NN, LR, and RF models in predicting CTs at the workstation level in wood offsite construction is warranted.

Therefore, further investigation into CT estimation using machine-learning techniques and in consideration of these three research areas is warranted. More detailed descriptions of these research areas are provided in Chapter 5 of this thesis.

1.2.4 Dimension IV: Dynamic production scheduling and digital twins

With respect to production scheduling, the literature has tended to focus on optimizing the production sequence of building components or jobs. There have been numerous studies on such topics seeking to minimize the total production time, makespan, or related parameters (Ko et al., 2011; Leu et al., 2002; Z. Xu et al., 2020). These studies have addressed an essential aspect of production scheduling (i.e., sequencing), but these methods still leave a gap in terms of determining how much time is actually spent by each building component at each workstation, instead relying on average estimates. For this reason, other studies have integrated simulation methods with their sequence optimization methods in order to account for a certain degree of production variability. For instance, Du et al. (2021) integrated a multi-objective genetic algorithm that generates near-optimal schedules with a multiagent system that simulates production according to the near-optimal schedules and includes a risk agent that triggers uncertain events such as machine failure. However, their study still relied on average estimates in modelling the processing times. Alternatively, Altaf et al. (2014, 2018) integrated production sequence optimization models with discrete-event simulation, employing regression equations to model processing times as a function of panel design properties and using statistical distributions to model four types of delays. While this approach constitutes an advancement compared to the practice of relying on single average values to estimate the durations of entire processes at workstation, it can further benefit from machine-learning-based CT estimation models developed in consideration of the research areas discussed under Dimension III. Nevertheless, as noted in the Motivation section above, machine-learning models cannot guarantee the complete capture of all factors contributing to variability. Hence, for dynamic scheduling of operations, there is also need for real-time production monitoring so that production can be tracked and production schedules regularly updated in response to variability. In this regard, the concept of digital twins is a promising solution. Given that digital twins mimic real operations and continuously acquire production data, they make it possible to monitor operations in real time and simulate actual operations for the purpose of estimating production times and developing (and updating) production schedules accordingly.

The concept of digital twins is increasingly being applied across a wide range of industries (Attaran et al., 2023). However, although the controlled factory setting in offsite construction provides a suitable setting for leveraging the benefits of digital twins, relatively few studies have targeted the application of digital twins in offsite construction. In fact, in searching the Scopus database, the Compendex database, and the Web of Science platform for publications on digital twins in offsite construction following the review steps detailed in Alsakka et al. (2023a) fewer than twenty publications were identified as addressing the use of digital twins in offsite construction (as of May 2023). Previous digital twin applications in offsite construction have focused on addressing diverse aspects within the industry, including on-site assembly planning and scheduling, management of hoisting operations, management of transportation risks, quality assessment, and the management of hoisting-related safety risks (Jiang et al., 2022; D. Lee et al., 2021; Z.-S. Liu et al., 2022; Rausch et al., 2021; Tran et al., 2021; Y. Zhao et al., 2022). Overall, these studies have demonstrated the versatility and promising potential of digital twin applications in digitizing various management tasks in offsite construction. However, digital twin applications in offsite construction are still in their infancy. Moreover, no digital twins have been developed for the purpose of production time estimation, scheduling, and real-time monitoring in offsite construction. Therefore, there is a need for a digital twin for production time estimation, scheduling, and real-time monitoring in offsite construction factories.

1.3 Problem statement and research objectives

In light of the above, the high-level problem addressed in this thesis has to do with the use of average production rates for production estimation and scheduling in offsite construction, given that, in reality, there is a high degree of CT variability in production operations due to a multitude of influencing factors (related to building components, workers, machines, materials, workstation setup, production line, factory operations, and external circumstances).

Given its ability to closely replicate real-world production operations in a virtual environment, digital twin has the potential to enable more dynamic and accurate estimation and scheduling of production operations amidst variability in offsite construction. Specifically, a digital twin that is capable of (1) continuously acquiring data from offsite construction factories (satisfying Dimension I), (2) predicting process CTs at workstations considering various CT-influencing factors (satisfying Dimensions II and III), (3) simulating operations and, accordingly, generating production schedules, and (4) mirroring and tracking operations in real time (satisfying Dimension IV) can be expected to improve the performance of production estimation and scheduling tasks in offsite construction compared to the current practice of relying on experience-based average production rates.

Hence, the overarching goal of this thesis is to develop a digital twin for dynamic production time estimation and scheduling and real-time monitoring in offsite construction factories. In this regard, as discussed in the Background section above, there are four main research areas that demand further investigation and development to enable the creation of a practical digital twin capable of fulfilling the above-mentioned functions. These areas and the corresponding objectives are summarized as follows:

- Automated data acquisition: A digital twin requires continuous acquisition of CT data in order to regularly tune the CT estimation models. Computer-vision technology is promising in this respect, but the typical process of building computer-vision models often requires extensive up-front set-up efforts. Moreover, the burden of model-training increases as building components change in shape and size while progressing through production lines.

  Hence, the first research objective is to design and examine a system that deploys computer-vision technology for the automated and accurate acquisition of CT data in a timely and cost-effective manner.

- Factors influencing CTs: There are a variety of factors that exert a continuous influence on CTs at workstations in offsite construction factories. Several previous studies have considered such factors in the development of estimation models for process time-related variables. However, these studies have lacked a systematic approach for, or comprehensive discussion on, the identification of the relevant factors.

  Hence, the second objective is to devise a methodical approach for the identification and understanding of influencing factors as well as evaluating how this identification process improves the accuracy of CT estimation.

- Data- and knowledge-driven CT estimation: There has been relatively little research on the estimation of process CTs at the workstation level during the production phase of offsite construction. Moreover, prior studies in this regard have focused primarily on developing prediction models, particularly using LR models, which rely mainly on unique design properties of building components. As such, several research areas in this domain require further investigation: (1) there is a need to investigate the effect of considering various types of CT-influencing factors on the performance of estimation models; (2) there is a need to examine the feasibility of training the estimation models using automatically acquired data; and (3) there is no consensus in the literature regarding the performance of machine-learning models used for estimation in offsite construction.

  Hence, the third objective is to develop and evaluate a data- and knowledge-driven system that estimates CTs considering various influencing factors and using automatically collected data, to increase the estimation accuracy compared to traditional estimation methods. This objective comprises three sub-objectives:

    - Examine the effect of considering a variety of influencing factors on the estimation performance of the aforementioned models. This subobjective considers the influencing factors identified under the second objective.
    - Explore the reliability of using data collected automatically (via computer vision) to train the estimation models. This subobjective uses the automated computer-vision-based data acquisition system developed under the first objective.
    - Examine the use of different machine-learning algorithms, including the feed-forward ANN, LR, and RF algorithms, for CT estimation considering various influencing factors.
- Digital twin design: The application of digital twins in offsite construction is currently in its nascent phase, and there are no existing digital twins specifically designed for the purpose of production time estimation and scheduling in offsite construction factories.

Hence, the fourth and final objective of the research presented in this thesis is to devise and test a feasible design of a digital twin that enables dynamic and more accurate production scheduling and real-time monitoring in offsite construction factories. The outcomes of the first three objectives are integrated into the development of this digital twin.

The high-level problem addressed in this thesis, the hypothesis underlying the research, the corresponding research needs, the research objectives, and the overarching research goal are summarized in Fig. 1-1.

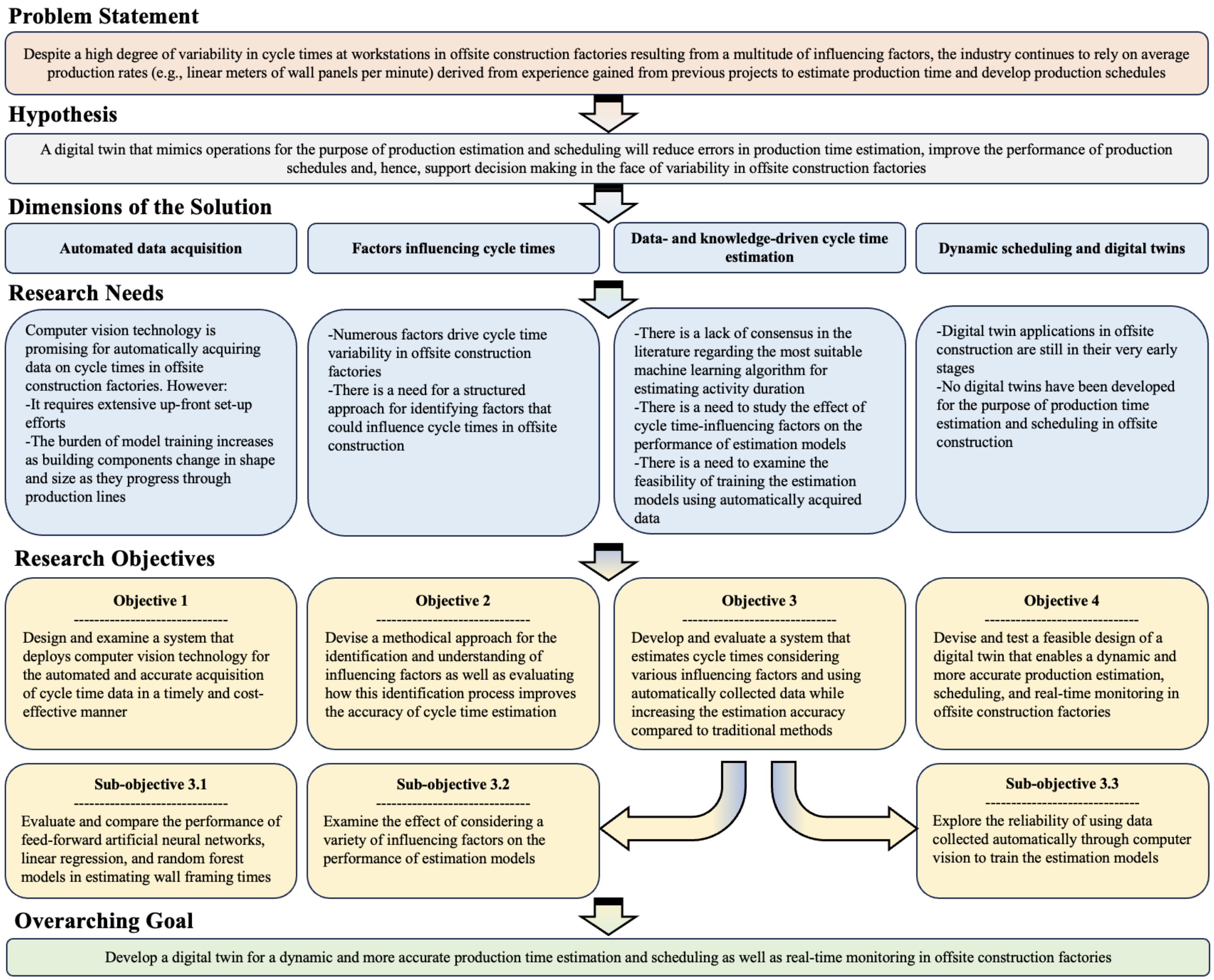


Fig. 1-1. Problem statement, hypothesis, research needs, objectives, and overarching goal.

1.4 Brief overview of research methods

The objectives undertaken in this research are interrelated and had to be pursued in the order listed above (from Objective 1 to Objective 4). This order was imperative because the outcomes of each objective laid the foundation for the subsequent ones, culminating in the development of the digital twin architecture shown in Figure 1-2.

The first step taken in devising this architecture was to address the first objective and develop an automated data acquisition system (highlighted in blue in Fig. 1-2). A computer-vision-based data acquisition system was designed and developed in such a manner as to enable the use of object detection algorithms pre-trained to detect objects commonly encountered in everyday life for studying other custom objects, thereby eliminating the need to retrain the models on the custom objects. The system was developed and evaluated through its application to a semi-automated wood-wall framing workstation at an offsite construction factory and using a YOLOv4 object detection algorithm pretrained on the COCO dataset.

The subsequent step entailed addressing the second objective, which revolved around formulating a qualitative approach for identifying factors influencing CTs at the workstation level in offsite construction factories in order to gain insights on the effects of variability prior to performing any numerical analysis. This approach integrated qualitative methods such as direct observation, process mapping, literature review, and semi-structured interviews with workers. The approach was also applied to the wall framing workstation in order to identify factors influencing framing CTs, as this information was needed for the subsequent objectives (highlighted in orange in Fig. 1-2). The significance of identifying these CT-influencing factors was demonstrated by evaluating their impact on the performance of an NN model developed to estimate framing times. Data collected during the execution of the first objective was used in this evaluation.

The third step entailed addressing the third objective, i.e., designing and evaluating a CT estimation system (highlighted in green in Fig. 1-2) integrating the computer-vision-based data acquisition system developed under the first objective, the influencing factors identified under the second objective, machine-learning-based prediction models, statistical models, and 3D simulation. The system evaluation was also carried out with reference to the framing workstation.

The final step was to address the fourth objective, and this consisted of designing a digital twin architecture capable of fulfilling all essential functions as discussed in the previous section for dynamic and accurate production scheduling. In addition to the methods employed in the first three objectives, ultrasonic sensors were incorporated for real-time tracking of factory operations.

Comprehensive explanations of the methods employed for each objective can be found in the corresponding chapters. The contents of each of the chapters that follow are outlined in the following subsection.

Objective 1

- Chapter 2: Computer Vision Applications in Offsite Construction (Paper 1)
- Chapter 3: Computer Vision-based Process Time Data Acquisition for Offsite Construction (Paper 2)

Main Methods

- YOLOv4 object detection algorithm
- COCO dataset

Objective 2

- Chapter 4: Factors Influencing Cycle Times in Offsite Construction (Paper 3)

Main Methods

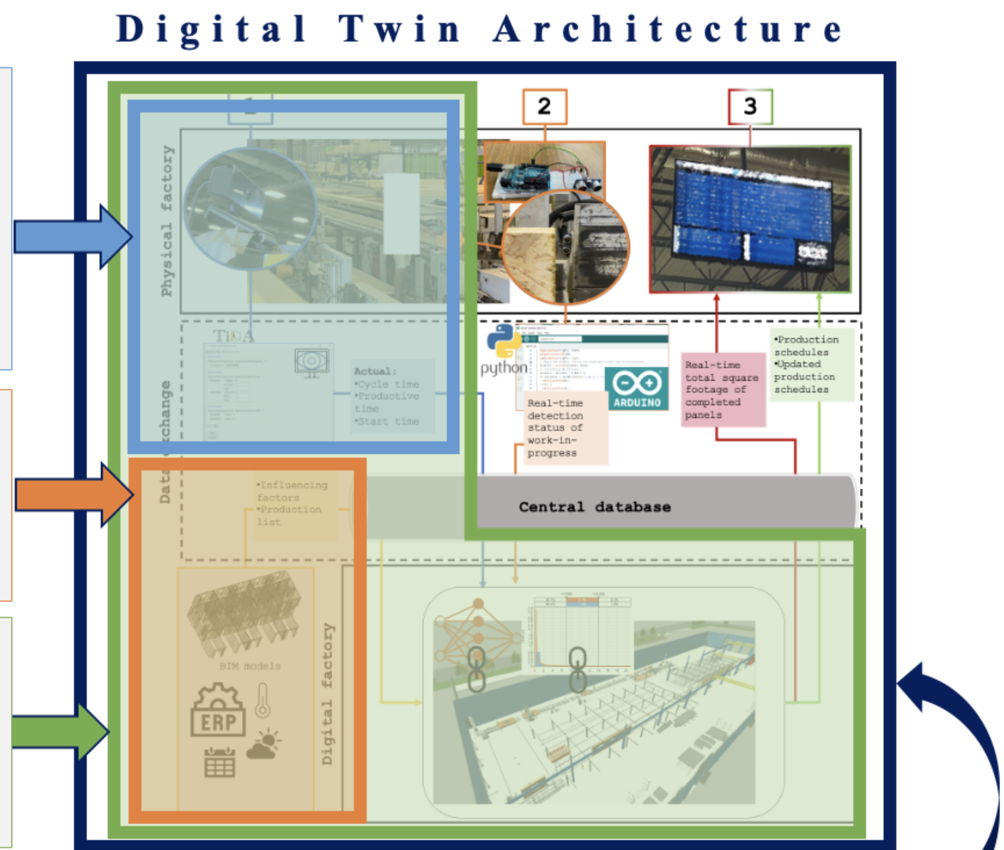
- Direct observation
- Process mapping
- Literature reviews
- Semi-structured interviews

Objective 3

- Chapter 5: Data- and Knowledge-driven Cycle Time Estimation In Offsite Construction (Paper 4)

Main Methods

- Computer vision
- Machine learning-based prediction
- Statistical modelling
- 3D simulation



Objective 4

- Chapter 6: Digital Twin for Production Estimation, Scheduling and Real-Time Monitoring in Offsite Construction (Paper 5)

Main Methods

Methods used in objectives 1 to 3 in addition to ultrasonic sensors

Fig. 1-2. Summary of methods.

1.5 Thesis structure

This thesis follows a paper-based format and consists of seven chapters, organized as follows:

- Chapter 1 serves as an introductory chapter, describing the motivation for the research, offering background information, and presenting the research hypothesis, gaps, goal, and objectives.
- Chapter 2 is an independent paper that presents a scoping review of computer-vision applications in offsite construction. It aims to deepen the reader's understanding of computer-vision technology and its value for data acquisition in the context of offsite construction. It provides a comprehensive foundation on the topic of computer vision, setting the stage for the subsequent exploration in Chapter 3.
- Chapter 3 is an independent paper that introduces a computer-vision-based system developed as part of this research that ultimately serves as a data acquisition system for the digital twin. Chapter 3 addresses Objective 1.
- Chapter 4 is an independent paper that presents a qualitative approach for identifying factors influencing CTs at workstations in offsite construction factories. It includes a case study on a semi-automated wood wall framing workstation and discusses its significance in enhancing the performance of CT estimation models. Chapter 4 addresses Objective 2.
- Chapter 5 is an independent paper that introduces an estimation system trained on data collected using the computer-vision-based system described in Chapter 3. It uses machine learning and statistical modelling to estimate CTs based on the relevant influencing factors, with an application to the wall framing workstation. Chapter 5 addresses Objective 3.
- Chapter 6 is an independent paper that presents the proposed digital twin, integrating the findings from Chapters 3 to 5. Chapter 6 addresses Objective 4.
- Chapter 7 serves as the concluding chapter, summarizing the key research findings, highlighting the contributions of the research, and discussing potential avenues of future research.

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 2

Computer-vision Applications in Offsite Construction

2.1 Introduction

2.1.1 Computer vision definition, tasks, and approaches

The field of artificial intelligence (AI) is considered a “game-changer” with the potential to contribute up to $15.7 trillion to the global economy by 2030 (Anand S. Rao et al., 2017). AI is a multidisciplinary field involving intelligent systems capable of performing tasks that would ordinarily require human intelligence (Shapiro, 1992). These AI tasks include machine learning (which involves making predictions/decisions based on data), natural language processing (which involves speech and text recognition), robotics, expert systems (which involves making decisions or recommendations based on a set of rules or knowledge), and computer vision among others (Shapiro, 1992). Computer vision, in turn, is a subfield of AI that focuses on the development of autonomous systems to mimic certain tasks performed by the human visual system (Huang, 1996). It incorporates the extraction of meaningful information from visual components, such as digital images, videos, cameras, and closed-circuit television (CCTV), allowing for informed data-driven decisions and recommendations (IBM, 2022). The computer vision field has also seen rapid growth in recent years and is projected to continue growing in the future (Data Bridge Market Research, 2022; KBV research, 2020; Verified Market Research, 2021)

The growth of computer vision is largely attributable to its ability to perform various visual tasks, such as object detection (Abbas et al., 2018; Sudharsan et al., 2019), image classification (Nath et al., 2014), object or motion tracking (Azhar et al., 2020; Host et al., 2020), action recognition (Das et al., 2018; Tan et al., 2021), human pose estimation (J. Wang et al., 2019; M. Zhao et al., 2018), semantic segmentation (Orsic et al., 2019; Siam et al., 2018), instance segmentation (Hafiz et al., 2020) optical character recognition (Chaudhuri et al., 2017), facial recognition (Tolba et al., 2006), and scene or 3D reconstruction (Kolmogorov et al., 2002). These tasks can be carried out using a variety of approaches, such as (1) template matching, which consists of comparing a predetermined template image with portions of a larger image to find any matches (Brunelli, 2009); (2) geometric-based approaches, which rely on the geometric properties of objects and scenes and mathematical models to extract information pertaining to the objects and their relationships in the scene (Szeliski, 2010); (3) rule-based approaches, in which a set of rules is predefined in order to detect and recognize objects in images (Szeliski, 2010); (4) physics-based approaches, which rely on mathematical models and the physical properties of light and its interactions with objects in a scene to analyze images (Szeliski, 2010); and (5) machine-learning approaches, which involve training algorithms on large amounts of data so that they can identify patterns and relationships between objects and corresponding image features (Szeliski, 2010). It is worth noting that, while a number of different machine-learning methods are used in computer vision, deep learning approaches have been particularly successful in the field (Chai et al., 2021). Deep learning refers to a subset of machine-learning applications that integrate multiple processing layers of interconnected neural networks within a variety of unsupervised and supervised feature learning algorithms. The role of the network in such applications is to mimic the function of the brain in perceiving and understanding multimodal information, allowing the network to learn from large amounts of data (Géron, 2019). The versatile nature of computer vision tasks and the success of computer vision approaches have led to a wide range of applications across various industries, as discussed in the following section, showcasing the power of these techniques.

2.1.2 Versatility of computer-vision applications and research motivation

The computer vision field has seen growth not only in overall uptake, but also in the breadth of applications in various industries (Szeliski, 2010). For example, in healthcare, computer vision has been used in medical imaging applications to help professionals to better visualize certain organs and make more sound diagnostic decisions accordingly (Szeliski, 2010). In retail, it has been used to detect objects at self-checkout lanes as a way of facilitating the self-checkout process and limiting fraud (Szeliski, 2010; B.-F. Wu et al., 2016). In the automotive industry, it has been employed in safety mechanisms in vehicles for detecting pedestrians on the roadway (Szeliski, 2010) and in self-driving applications to detect roads, pedestrians, and vehicles (Tseng et al., 2018). The technology has also been increasingly deployed in traditional onsite construction, where it is applied for a variety of purposes, including safety monitoring, resource tracking, activity monitoring, productivity analysis, quality control, and infrastructure inspection, to name a few (Martinez et al., 2019c; Paneru et al., 2021; S. Xu et al., 2021). To provide some specific examples, it has been used to detect workers' compliance with wearing personal protective equipment (PPE) (M.-W. Park et al., 2015), to detect the presence of structural supports on construction sites in order to reduce exposure to fall hazards (Fang et al., 2018, 2019), to evaluate construction project progress (Asadi et al., 2018; Hamledari et al., 2017; Roh et al., 2011), to analyze workers' movements or monitor their activities (Gong et al., 2011; Luo et al., 2018), to track the locations of different resources on site (Teizer, 2015), to recognize actions carried out by construction earthwork equipment (e.g., digging, dumping) (Golparvar-Fard et al., 2013), and to recognize dimensional discrepancies in structural components (Maalek et al., 2019).

As the technology proves increasingly promising in various industries and in light of the fourth industrial revolution (Industry 4.0), which has been driving advancements in manufacturing practices, the offsite construction industry (also known as "construction manufacturing" or "prefabricated construction") stands to benefit significantly from the incorporation of various computer-vision applications. In offsite construction, building components, systems, or structures are produced or prefabricated in a controlled factory setting and then transported to the construction site for installation. While the adoption of offsite practices has drawn growing interest and the corresponding research has spiked in recent years (Bosche et al., 2009; Martinez et al., 2019c), computer-vision applications in offsite construction remain under-researched. Indeed, Martinez et al. (2019c) reported that a quick search of publications discussing the incorporation of computer-vision applications in offsite construction yielded only two studies as of 2019. As such, there is a need to evaluate whether the industry has caught up with the pace of other industries in their deployment of computer-vision technology and to better understand the most recent developments in the domain.

2.1.3 Study objective

In this context, this chapter presents a scoping review of computer-vision applications in offsite construction. The purpose of a scoping review, it should be noted, is to "provide a narrative or descriptive account" of research available in a particular area of study (Arksey et al., 2005). In other words, it addresses the question "What evidence exists?" with respect to a particular research area (Munn et al., 2022). It provides "an opportunity to identify key concepts; gaps in the research; and types and sources of evidence to inform practice, policymaking, and research" (Daudt et al., 2013). Scoping reviews do not, however, need to be limited to mapping the literature (Munn et al., 2018); among the purposes of scoping reviews are to "clarify key concepts/definitions in the literature", "examine how research is conducted on a certain topic or field", and "identify key characteristics or factors related to a concept". It should be noted that scoping reviews differ from systematic reviews in that they do not necessarily include a formal quality assessment of the studies reviewed and may not include a synthesis of the results (Munn et al., 2018). The overall aim of the scoping review presented in this chapter is to provide an overview of the status and applications of computer-vision technology in offsite construction. Specifically, it provides the following: (1) summaries of, and discussions on, the research areas in which computer vision is used in offsite construction, the computer vision tasks undertaken, the algorithms or approaches used, and the corresponding performance evaluation results and limitations, (2) a tabulated summary of performance-related terms commonly used in computer-vision applications (to facilitate understanding of the performance evaluation results reported in the review), and (3) potential avenues of future research. Such a review provides a useful point of reference for practitioners and researchers in the offsite construction industry, aiding their understanding of current practices, limitations, possible gaps, and potential opportunities to apply computer vision, and lays the foundation for a future systematic review on the topic.

2.2 Review methodology

The framework followed in the present study (Fig. 2-1) generally aligns with the framework outlined by Arksey and O'Malley (2005). The steps undertaken in each of the framework's five stages are delineated in the following subsections.

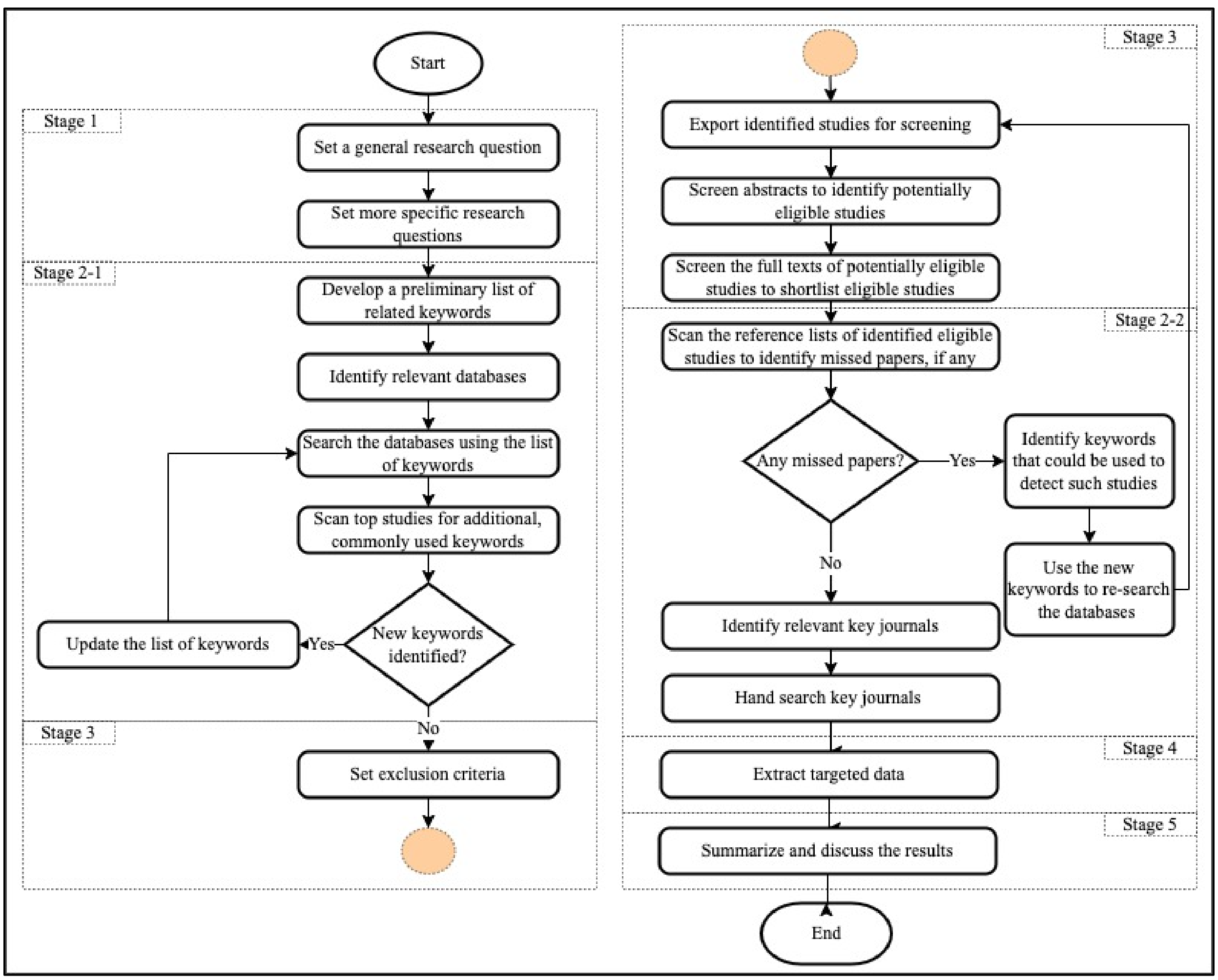


Fig. 2-1. Review framework.

2.2.1 Stage 1: Identification of the research questions

Arksey and O'Malley (2005) suggest starting with a broad research question to mitigate the risk of missing relevant references. Following this recommendation, the general research question targeted in the present study was: "*what is the current state of computer-vision applications in offsite construction?*". Moreover, as suggested by Levac et al. (2010), a well-defined scope of inquiry is helpful in deriving an effective search strategy. As such, the general research question was broken down into more specific questions in order to identify which types of studies would be considered. The specific questions included: (1) What was the overall goal of the study using computer vision? (2) What was computer vision specifically used for in the study? (3) Which computer vision algorithm(s) or method(s) was(were) used? (4) How did the algorithm(s)/method(s) perform? (5) What were the encountered limitations?

2.2.2 Stage 2: Collection of relevant studies

The identification of relevant studies was conducted in an iterative manner, since familiarity with the literature increased as more papers were identified and reviewed. The steps undertaken in this stage were as follows:

Step 1. A preliminary list of keywords related to the use of computer vision in offsite construction was created. The list was initially developed accounting for terms used synonymously (e.g., "offsite construction" versus "prefabricated construction"), alternate spellings (e.g., panelized versus paneli*s*ed), alternate stylizations (e.g., "offsite" versus "off-site"), alternate suffix use, which was accounted for using an asterisk (e.g., "detecting" versus "detection"), and subfields of offsite construction (e.g., panelized construction) and of computer vision (e.g., machine vision).

Step 2. The databases mostly relevant to offsite construction were identified with the help of librarians at the University of Alberta. The Scopus database, the Compendex database, and the Web of Science platform were selected as the sources of relevant literature.

Step 3. The databases were searched using the preliminary list of keywords. In Scopus, only the title, abstract, and keywords were searched, as looking for keywords in all fields (including the reference list of each publication) resulted in a much larger number of publications, many of which were not relevant. On the other hand, in the cases of Compendex and Web of Science, all fields were searched, as doing so did not significantly increase the number of publications identified. As the initial searches did not yield a large number of relevant studies, no additional restrictions (e.g., limited range of publication years, source types, etc.) were imposed.

Step 4. The top studies identified during the database searches were quickly scanned for any additionally keywords used that were not included in the preliminary list of keywords for the search, and the list was updated accordingly.

Step 5. The updated list was used to search the databases again, and Step 4 was repeated. This process was repeated until no more keywords could be identified following this strategy. The resulting references were then exported for screening.

Step 6. After filtering eligible studies as explained in the following subsection, the reference list of each eligible study was scanned for any additional relevant studies that may have been missed. Additional publications were identified following this strategy, and the reviewers looked for additional keywords in these newly identified publications that could be used to detect other additional studies in case others were still missing. The list of keywords was revised accordingly to include "image processing" and "detection", as it was discovered that some studies used these terms to refer to computer-vision applications without mentioning any of the initially identified keywords in the searched fields (e.g., title, abstract, and keywords in the case of Scopus, and all fields in the case of the other two databases). The keyword "detection" was used instead of "detect*", as the latter captured a large number of publications that were not relevant. The list of keywords was once again updated accordingly.

Step 7. The final list of keywords was used to search the databases on November 10, 2022. The final list of keywords was as follows: *("modular construction" OR "construction manufacturing" OR "off-site construction" OR "prefabricated construction" OR "offsite construction" OR "building manufacturing" OR "home prefabrication" OR "modular building" OR "modular home" OR "industrialized construction" OR "industrialized building" OR "prefabricated building" OR "precast construction" OR "pre-cast construction" OR "off-site manufacturing" OR "offsite manufacturing" OR "prefab construction" OR "precast building" OR "panelized construction" OR "panel prefabrication" OR "panelised construction") AND ("computer vision" OR "machine vision" OR "custom vision" OR "object detect*" OR "object recognition" OR "object track*" OR "edge detect*" OR "image recognition" OR* *"feature recognition" OR detection OR "3d reconstruction" OR "scene reconstruction" OR "photogrammetry" OR "image-based" OR "vision-based" OR "motion analysis" OR "segmentation" OR "motion track*" OR "image processing").*

Step 8. Arksey and O'Malley (2005) recommend conducting a hand-search of studies published in key journals, as the databases may not include all publications. The list of journals to be hand-searched was formulated by (1) checking the journals in which the eligible studies appear and selecting those that are likely to include studies pertinent to this scoping review, and (2) screening the top one hundred journals in the list of top journals in the building and construction field found in Resurchify portal (Resurchify, 2022). The following journals were searched, and no additional relevant studies were identified: *Automation in Construction*; *Journal of Information Technology in Construction*; *Journal of Computing in Civil Engineering*; *Engineering, Construction and Architectural Management*; and *The International Journal of Advanced Manufacturing Technology.*

Since this stage was completed in an iterative manner, there was a large number of duplicates found in the final list of imported studies. This is due to the fact that some of the studies captured during the initial database search were captured again during the successive searches, as they included multiple relevant keywords. A total of 2,355 studies were imported for screening, and 1,468 duplicates were removed, leaving 887 remaining (Fig. 2-2).

2.2.3 Stage 3: Filtering of eligible studies

Before screening the 877 studies identified in the previous stage, a list of exclusion criteria was established. The following studies were excluded: studies that were not available in English; studies from which the targeted data could not be extracted; studies that were deemed not exclusive to offsite construction; and studies that only present theory (without practical applications). The PRISMA diagram is shown in Fig. 2-2. The screening was completed in two stages; in the first stage, the 887 abstracts were screened to identify potentially eligible studies, and this resulted in 839 irrelevant studies being removed. Then, the full text of each of the 48 remaining studies was assessed in light of the exclusion criteria, resulting in another 24 studies being removed.

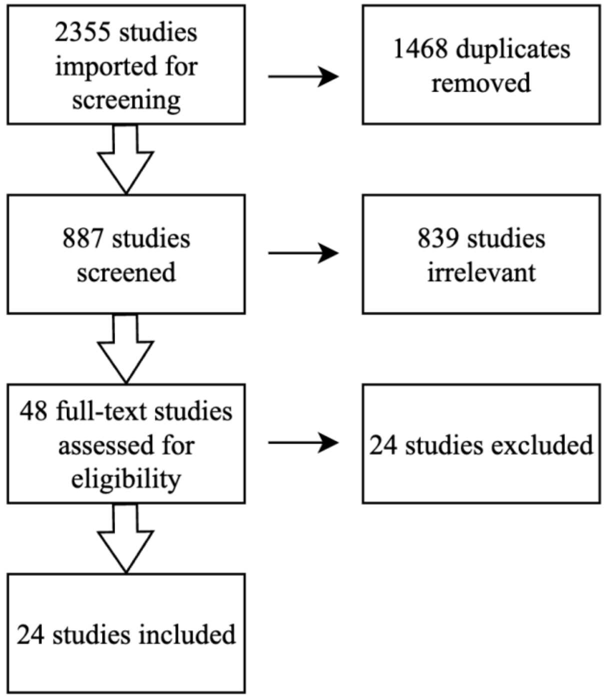


Fig. 2-2. Screening results.

2.2.4 Stage 4: Extraction of targeted data

A template was designed for extracting the data needed to answer the research questions established at the outset of the review process. The targeted data included the (1) publication year, (2) research area (e.g., productivity measurement, quality control, ergonomic analysis...), (3) goal of the study, (4) task(s) for which computer vision was implemented, (5) computer vision algorithm(s) or tool(s) used, (6) algorithm performance, and (7) additional notes, including study limitations reported by the authors. Two reviewers independently filled out the template for each eligible study, and the extraction results were compared and checked for any discrepancies. In the case of there being discrepancies, the extraction process was repeated for the corresponding studies until the review team reached consensus.

2.2.5 Stage 5: Summary and discussion of the results

This stage comprised two steps. In one step, the extracted data was numerically analyzed. The analysis involved (1) the number of studies published each year, (2) the number of studies per research area, and (3) the number of times each computer vision algorithm was used. In the other step, the review team aggregated, summarized, and discussed the various research areas in which computer vision was used, the algorithms used and related performance, the reported limitations, and the research growth trends.

2.3 Results and discussion

A total of 24 studies were identified during the review undertaken in this study. The review results are summarized in Table 2-2, presented at the end of this section, and discussed in the following subsections. The studies were grouped into seven research areas including (1) "progress monitoring and productivity measurement", (2) "quality assurance and control", (3) "ergonomic analysis", (4) "process guidance", (5) "safety management", (6) "disruption management", and (7) "general CV applications". It should be noted that studies tackling multiple research areas are classified in Table 2-2 based on the research area they mainly focus on, but are addressed under each of the relevant research areas in the following subsection.

2.3.1 Research areas of applications

The distribution of the studies across different research areas is shown in Fig. 2-3. As shown in the figure, the research areas with the highest number of publications were "progress monitoring and productivity measurement" and "quality assurance and control", while relatively few studies were found that focused on the other research areas. In the subsections that follow, both a background on the issues encountered in relation to each of these research areas in general and a description of the corresponding applications of computer vision in offsite construction are provided.

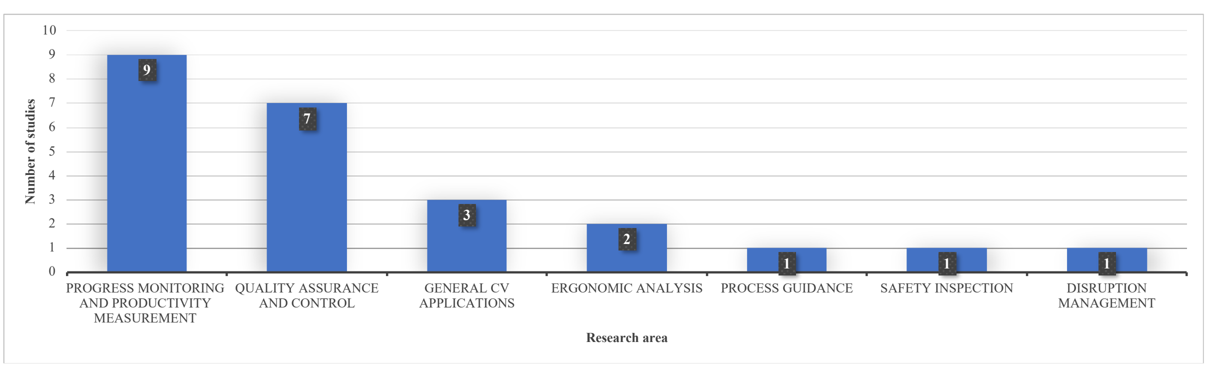


Fig. 2-3. Research areas of application.

2.3.1.1 Progress monitoring and productivity measurement

Construction productivity is considered to be one of the core elements of performance in construction, as it is closely associated with a project’s budget and duration/schedule (El-Gohary et al., 2017). With traditional construction having lagged behind other industries in terms of productivity for decades (Bertram et al., 2019a), offsite construction is seen as having the potential to improve productivity in the construction sector (Bertram et al., 2019a; Hogarth, 2020). In order for productivity to be improved, however, it must be continuously monitored and measured. Meanwhile, some offsite construction companies continue to rely on fixed ratios such as square footage or linear feet per day to measure overall productivity (Alsakka et al., 2023b). Computer vision provides a solution to this issue by automatically collecting data on progress and productivity, saving significant time and effort that would otherwise be spent on manual time studies.

Indeed, computer vision has been successfully applied in offsite construction for progress monitoring and productivity measurement, with nine of the studies identified in the current review having involved research on this topic. In fact, this research area had the highest number of related studies, as shown in Fig. 2-3. Of the nine studies, six targeted the installation phase in offsite construction projects. For instance, Zheng et al. (2020) developed a computer-vision-based model to automatically count the number of installed modules as a way of monitoring the progress. Zhang et al. (2020) developed a model that allows for real-time tracking of the modules. Additionally, computer vision tools have been used to detect precast concrete walls and track their trajectory during the hoisting and installation processes (Z. Wang et al., 2021; Z. C. Wang et al., 2021). Another study developed a model by which to detect prefabricated walls and slabs, in addition to workers and onsite activities, in order to measure the corresponding progress as a proactive control mechanism for monitoring and evaluating the installation schedule (Yan et al., 2022). Another of the identified studies proposed a model to automatically measure the installation rate of prefabricated panels in order to overcome the limitations of manual methods (which are considered to be time-consuming and highly prone to errors) (Ahmadian Fard Fini et al., 2021). The three remaining studies identified in this section targeted the fabrication phase in offsite construction. Panahi et al. (2022b) developed a computer-vision-based model to track volumetric modules through the various stations along the production line, with the resulting information used to automatically measure the station’s cycle time. Park et al. (2021), meanwhile, proposed a model to recognize steps in the module's assembly process in a virtual factory environment, with the output used to identify the next step in the assembly process. Martinez et al. (2021), finally, developed a model to track the progress of floor panel fabrication in order to capture productivity-related metrics (i.e., duration and man-hours) in an accurate and timely manner.

In short, it is evident that the use of computer-vision techniques can effectively support progress monitoring and productivity measurement in offsite construction. With the increased feasibility and convenience of automated progress and productivity measurements, it is now much easier to investigate strategies to improve productivity. It is thus incumbent upon researchers to continue adopting computer vision tools in this research area and using computer vision data to test the implementation of strategies for productivity improvement.

2.3.1.2 Quality assurance and control

Ensuring adequate quality has been long considered one of the pillars defining the success of construction projects (Cox et al., 2003). This is because quality-related issues can lead to significant cost and schedule overruns, given the need to identify and implement the proper mitigation measures to address them (Marasini et al., 2010). Offsite construction projects have been shown to have a lower defect rate compared to traditional construction (Johnsson et al., 2009), likely due to the controlled factory setting in which production takes place. Manual quality inspection, however, can be tedious and time-consuming. For instance, a recent case study in the area of precast construction reported that it took a worker a total of 60 min to manually inspect 142 rebar spacings installed in two precast concrete panels (Q. Wang et al., 2017). On the other hand, advancements in technology such as computer vision have made it possible to automate quality control and assurance tasks. Additionally, the tightly controlled factory environment characteristic of offsite construction is particularly conducive to the adoption of automated quality assurance and control techniques and technologies, as these can be integrated into production operations on a permanent basis in such an environment.

Accordingly, computer vision has recently been deployed for quality management in offsite construction projects, as reported in seven relevant studies identified in the present review. Two of the identified studies developed computer-vision-based models to inspect the quality of precast concrete elements (S. Lee et al., 2022; S. J. Lee et al., 2020). These models were shown to be capable of detecting cracks (S. Lee et al., 2022; S. J. Lee et al., 2020) and breakages (S. J. Lee et al., 2020), and of extracting the corresponding features of the cracks, such as the length and width (S. Lee et al., 2022; S. J. Lee et al., 2020). Additionally, three of the identified studies tackled the quality inspection of screw-fastening operations during the fabrication of light-gauge steel frames as a means of helping to avoid potential damage to the frame and possible failure of the screw-fastening operations (Martinez et al., 2019a, 2019b, 2020). The developed models were shown to be capable of: (1) detecting pre-drilled screw holes in the frames (Martinez et al., 2019b), (2) detecting manually assembled studs to validate the assembly (in order to ensure proper screw-fastening operations) (Martinez et al., 2019a), and (3) detecting stud edges and screws in order to evaluate the framing quality during screw-fastening operations (Martinez et al., 2020). These models were also shown to be helpful in identifying the necessary corrective measures, such as adjusting the location of the screw driving manipulators (Martinez et al., 2019b) and suggesting corrections to the manually assembled frame (Martinez et al., 2019a). Additionally, computer vision tools were used as part of a procedure that ensures proper fabrication of panel elements through identification of potential errors, where an error is defined as a mismatch between the fabricated panel and the corresponding design, detectable by superimposing the as-designed model onto the working area using a tool that incorporates computer vision (S. Ahn et al., 2019). In another study, a computer-vision-based model was developed to inspect the quality of prefabricated components by re-constructing a 3D model for precast columns and aluminum pipes (D. Lee et al., 2020).

Notably, the case studies conducted on screw-fastening operations and panel fabrication showcase the potential of computer vision in enabling continuous quality assurance across the various production processes, extending beyond the mere identification of defects in final products (e.g., cracks in precast concrete elements). By detecting errors in real time, computer vision can help to prevent the occurrence of defects in the first place, translating to time and cost savings. Given this, in addition to the feasibility of integrating computer-vision-based quality assurance and control systems into operations, offsite construction researchers should investigate the application of computer vision to other factory-based production processes to ensure that all fabricated elements are defect-free before they are shipped to the construction site for installation.

2.3.1.3 Ergonomic analysis

The transition of the construction industry towards offsite construction involves the implementation of principles and methods that underpin manufacturing systems. Put simply, a manufacturing system can be defined as a "combination of humans, machinery, and equipment that are bound by a common material and information flow" (Caggiano, 2014). In this regard, in offsite construction factories, workers, machinery, and equipment are typically positioned in fixed locations, while production processes are distributed across multiple workstations. (A detailed description of a sample design of offsite construction operations can be found in a recent study by Alsakka et al. (2020).) This implies that workers at each workstation are assigned a well-defined set of tasks, meaning that their work is highly repetitive. Meanwhile, manual construction tasks carried out by workers involve frequent motion of different body parts, including the neck, knees, wrists, and shoulders. Depending on the amount of time being subject to these tasks and the amount of physical strain involved, workers are likely to experience fatigue and to develop injuries and even suffer from permanent musculoskeletal damage (Ray et al., 2012). As such, analyzing manual construction tasks from an ergonomics perspective helps to identify appropriate remedial measures to improve the health and safety of workers (Ray et al., 2012) and reduce delays related to lost-time claims and disabling injury claims (Hinze et al., 1991).

Computer-vision techniques could be employed to perform ergonomics analysis. For instance, computer-vision-based approaches could be employed to identify and analyze worker posture, which can be used to estimate the body's joint angles for various tasks, such as lifting boxes (Gonsalves et al., 2009), or to determine whether or not a worker is maintaining proper posture based on predefined posing rules (Ray et al., 2012). While the value of such approaches in the context of offsite construction is significant, as remedial measures can be conveniently implemented (e.g., adjusting the height of working tables, using a vacuum lifter to load heavy items), the present review identified only two studies that address the use of computer-vision applications for ergonomic analysis. In both of these studies, computer vision tools were used in modular construction factories to detect the worker of interest and then detect the corresponding body parts and joints (Chu et al., 2019, 2020). The output was then used to reconstruct the 3D body model and perform the necessary posture analysis (Chu et al., 2020). In another study, a model was developed to detect workers' ergonomic postures as part of a broader effort to identify and address schedule delays (Yan et al., 2021).

In summary, while the fixed layouts of offsite construction factories facilitate the implementation of measures to improve worker ergonomics, the application of computer vision for ergonomic analysis in the offsite construction industry has been rather limited. Given that most processes in offsite construction factories still heavily rely on manual labour due to low levels of automation, this research area warrants further study (particularly considering its potential to benefit both workers' physical health and project performance).

2.3.1.4 Process guidance

The fact that the different building components in offsite construction (e.g., walls, roofs, and floors) have varying designs forces workers in offsite construction factories to frequently refer to shop drawings in order to determine the necessary steps involved in the fabrication of a given element (e.g., nailing an L-shaped stud at a specific location), despite the repetitive nature of the work (e.g., generally nailing an element). Hence, automated process guidance systems that provide guidance to the worker as to how to complete the task may increase productivity. Computer vision has already been proposed in the manufacturing sector as a means of guiding the assembly of spacecraft cabins for example (Y. Liu et al., 2015), and for the manufacture of products in general (Hercog et al., 2022), and it stands to reason that it could be used in a similar manner in the context of offsite construction. In this regard, two studies identified in the review demonstrate the potential of computer vision in providing real-time guidance for certain activities not only in the factory production phase but also in the site installation phase. In one of these studies, computer-vision techniques were used to guide workers in the process of installing precast concrete columns on site (K. Zhang et al., 2019). The model predicts the trajectory in order to align rebar installed in the ground with the holes of the precast columns to be installed. The approach is intended to expedite the installation of precast columns, as it assists workers in quickly identifying the final installation position of the columns. In the other study, computer-vision techniques were used to guide workers in identifying the correct positions of the different panel elements during panel fabrication (S. Ahn et al., 2019). This was achieved by superimposing an as-designed model (i.e., an augmented reality model) onto the working area. Such a strategy can eliminate the wasted time associated with having to frequently refer to paper drawings to identify the next steps required in the process, in addition to serving as a quality assurance strategy as discussed above.

The successful implementation of computer-vision-based automated process guidance systems for these offsite construction activities—along with demonstrating its potential to improve productivity and reduce errors—underscores the potential for further research opportunities exploring the development of such systems for other activities.

2.3.1.5 Safety management

Offsite construction has been shown to reduce safety risks compared to traditional construction due to the reduced occurrence of work-at-height tasks, of falling objects-related hazards, and of weather-related incidents, to name a few (S. Ahn et al., 2020). However, risks such as being struck by a moving object (e.g., forklift) or becoming trapped between stationary objects and moving objects still exist in offsite construction (S. Ahn et al., 2020). Moreover, Gibb et al. have contended that, as fabrication activities are shifted to the factory, the numerous onsite risks that have a high probability of occurring, but low consequences are replaced by risks with a lower probability of occurring (as they are generally easier to identify and control) but more significant consequences when they do occur (Gibb et al., 2004). The higher impact of onsite risks in offsite construction is attributable in part to the increased use of cranes (to handle large and heavy prefabricated elements) (Gibb et al., 2004). While computer vision can be employed to track moving objects and detect proximity between different objects in order to identify hazards in real time, the only safety-related study identified in the review focused on the automated detection of workers' PPE compliance on site (S. Liu et al., 2021). As PPE remains among the most important safety measures to reduce the risk of injuries and fatalities on construction sites (Barro-Torres et al., 2012), and since workers often elect not to wear the required PPE or wear it in an incorrect manner, automated PPE detection systems are of value. However, the potential of computer vision in detecting hazardous proximities between workers and moving objects should be researched as an additional safety-related application of computer vision in offsite construction, especially given the considerable safety-related risks associated with hazardous proximities in offsite construction compared to in traditional construction (in the case of crane-related hazards, for example).

It is noteworthy that the present review identified only one study on safety management, compared to nine studies on progress monitoring and productivity measurement and seven studies on quality assurance and control, this despite the high potential of computer vision to significantly assist in mitigating safety risks in offsite construction. Therefore, more studies should be conducted to explore the potential of computer vision in detecting and preventing various hazardous situations in offsite construction.

2.3.1.6 Disruption management

While factory production in offsite construction can be tightly controlled, the shipment of prefabricated elements to the site, as well as the site installation itself, may be subject to more uncertainties and disruptions. Managing uncertainties and disruptive events is crucial, as these may lead to deviations from the project's budget and schedule, resulting in cost and schedule overruns (Yu et al., 2004). Since most of the activities in offsite construction take place in the factory, where production can be tightly controlled and monitored, and relatively few activities are carried out on site enables, it is easier to identify disruptions and detect schedule delays in offsite construction compared to in traditional construction. Moreover, the principal sources of disruptions and delays on site (e.g., delay in delivery of prefabricated elements, installation delay) can be visually identified, presenting an opportunity for the use of computer-vision technology. In this regard, the two studies on disruption management identified in the review focused on the site installation phase of offsite construction projects. In one of these studies, a computer-vision-based model was developed to manage four types of schedule disruptions: (1) delay in delivery of prefabricated components and concrete mix, (2) traffic obstruction on the jobsite, (3) delay in installation of prefabricated components, and (4) ergonomics-related delays (Yan et al., 2021). In the model they developed, once the disruption is detected, its impact on the original schedule is evaluated against the set deviation tolerance from the original schedule in order to determine whether the schedule needs to be adjusted to get the project back on track. The other study proposed a computer-vision-based system that measures and evaluates the progress of prefabricated slab/wall installation and other manual work on site, allowing for the detection and timely handling of schedule disruptions (Yan et al., 2022).

Since greater predictability of time and cost is one of the main advantages of offsite construction (Bertram et al., 2019b), the use of computer vision to automatically detect and manage disruptions is of high potential benefit to the offsite construction industry. As such, the initiatives pursued in the identified studies are commendable, but future research studies should also target the factory production phase. While the risk of disruptive events may be lower in a controlled factory setting, their impact on production may be higher due to operations being designed in a manner conducive to manufacturing. Specifically, workstations on the same production line affect one another because they are linked by a common workflow. This means that disruptive events that impede work at one workstation could result in the stoppage of the entire production line. As such, research should be pursued that examines the applicability of computer vision to automatic identification of disruptions and the evaluation of schedule delays in factory production.

2.3.2 Algorithms used and reported performance

Despite the relatively small number of studies identified in the review, the number of computer vision algorithms and approaches investigated in these studies was relatively large. More than 20 algorithms and approaches were deployed and evaluated among these studies for various computer vision tasks, including object detection and classification (e.g., CNN, R-CNN, Blob detection, YOLOv2, YOLOv3, Extended YOLOv3, SlimYOLOv3, Speeded Up Robust Features, MV-CNN, Single Shot MultiBox Detector, Mask R-CNN, Faster R-CNN), object tracking (e.g., DeepSORT), feature extraction (e.g., CNN, Scale Invariant Feature Transform), post estimation (e.g., DeeperCut), segmentation (e.g., RGB segmentation, DeepLabv2, Mask R-CNN) and various combinations of the tasks. The number of times each algorithm was tested or deployed in the studies is plotted in Fig. 2-4. The Faster R-CNN algorithm, followed by the Canny edge detector and the Mask R-CNN algorithm, were the algorithms most frequently used in the identified computer-vision applications in offsite construction for object detection, edge detection, and object detection and segmentation, respectively. Given their frequency of use, these algorithms are described in greater detail in the following paragraphs. For details on the other algorithms, refer to Table 2-2.

The Faster R-CNN algorithm is a deep convolutional network used primarily for near-real-time and accurate object detection (Ren et al., 2015). It was used in the identified studies to detect various kinds of objects, including cracks and breakages in precast concrete members (S. J. Lee et al., 2020), workers (Martinez et al., 2021; Yan et al., 2022, 2021), machines (Martinez et al., 2021), cranes (Martinez et al., 2021), walls (Yan et al., 2022, 2021), slabs (Yan et al., 2022, 2021), panels (C. Liu et al., 2021), safety barricades (C. Liu et al., 2021), and fences (C. Liu et al., 2021). Regarding the performance of the identified applications, it should be first noted that (1) some of these studies evaluated the performance of their overall methodology (which included computer vision tasks in addition to other tasks) as well as the performance of the computer vision algorithms themselves, whereas others reported only the performance of their overall methodology, and (2) some of the studies did not report on the performance either of the developed algorithms or of the overall methodology. Table 2-3 aids understanding of the performance evaluation results reported in these studies, as well as providing a list of performance-related metrics commonly used in computer-vision applications in general. Sample performance metrics reported for the Faster R-CNN algorithm in the identified studies include an F1-score of 93.2% (Martinez et al., 2021) and an $mAP_{0.50}$ of 90.7% (Yan et al., 2022) for worker detection, an F1-score of 99.7% for machine detection (Martinez et al., 2021), an F1-score of 72.2% for crane detection (Martinez et al., 2021), an $mAP_{0.50}$ of 82.0% (Yan et al., 2022) and an $mAP_{0.50}$ of 82.1% (Yan et al., 2021) for wall detection, an $mAP_{0.50}$ of 82.1% (Yan et al., 2022) and an $mAP_{0.50}$ of 82.0% (Yan et al., 2021) for slab detection, an $mAP_{0.50}$ ranging from 80.0% to 82.4% for truck detection (Yan et al., 2021), an AR@10 and AR@100 of 92.08% for detecting barricades (C. Liu et al., 2021), and an mAP0.5 and mAP0.75 of 99.96% for detecting barricades (C. Liu et al., 2021).

Edge detectors are an essential element of many computer vision systems, as they significantly reduce the amount of data that needs to be processed while retaining the dimensional information on object boundaries (Canny, 1986). The Canny edge detector (Canny, 1986) is among the more prominent edge detection algorithms. It was used in the identified studies to extract the length and width of cracks in precast concrete members (S. J. Lee et al., 2020), to detect the edges of pre-drilled screw holes in light-gauge steel frames (Martinez et al., 2019b), to detect the edges of light-gauge steel studs (Martinez et al., 2020), and to detect edges in an image of a building during the installation phase (Ahmadian Fard Fini et al., 2021). In terms of performance, these studies all focused on the overall methodology, which in each case combined the Canny edge detector with other methods. As such, the performance is not summarized in this section.

The Mask R-CNN algorithm, meanwhile, as an object segmentation algorithm, is an extension of the Faster R-CNN. It efficiently detects and classifies objects while concurrently generating segmentation masks for each instance of the objects (Kaiming et al., 2017). These masks encode the spatial structures of given objects' instances (Kaiming et al., 2017). Mask R-CNN was used in the identified studies to detect and segment modules (Zheng et al., 2020), to detect and segment precast walls (Z. Wang et al., 2021; Z. C. Wang et al., 2021), and to segment and track workers (Xiao et al., 2022). Sample performance metrics reported for the Mask R-CNN algorithm in the identified studies include a precision of 96%, a recall of 92%, and an AP of 91% for module detection (Zheng et al., 2020), an $AP_{0.50}$ of 88% and a recall of 89% for precast wall detection (Z. C. Wang et al., 2021), an $AP_{0.50}$ of 93% and $AP_{0.75}$ of 0.85 for precast wall detection (Z. Wang et al., 2021), and a precision of 98.7%, a recall of 96.7%, and an F1-score of 97.6% for worker detection (Xiao et al., 2022).

It is important to note that the reported performance metrics do not necessarily represent the robustness of the algorithms used, since performance is also a function of other factors, such as the size of the training dataset, the quality of the images/videos used for training, and the complexity of the problem being addressed, to name a few. In other words, a thorough analysis of the performance of the different algorithms used in the identified studies would necessitate a critical evaluation of the methods followed to train and test the algorithms. This is a potential avenue for future research.

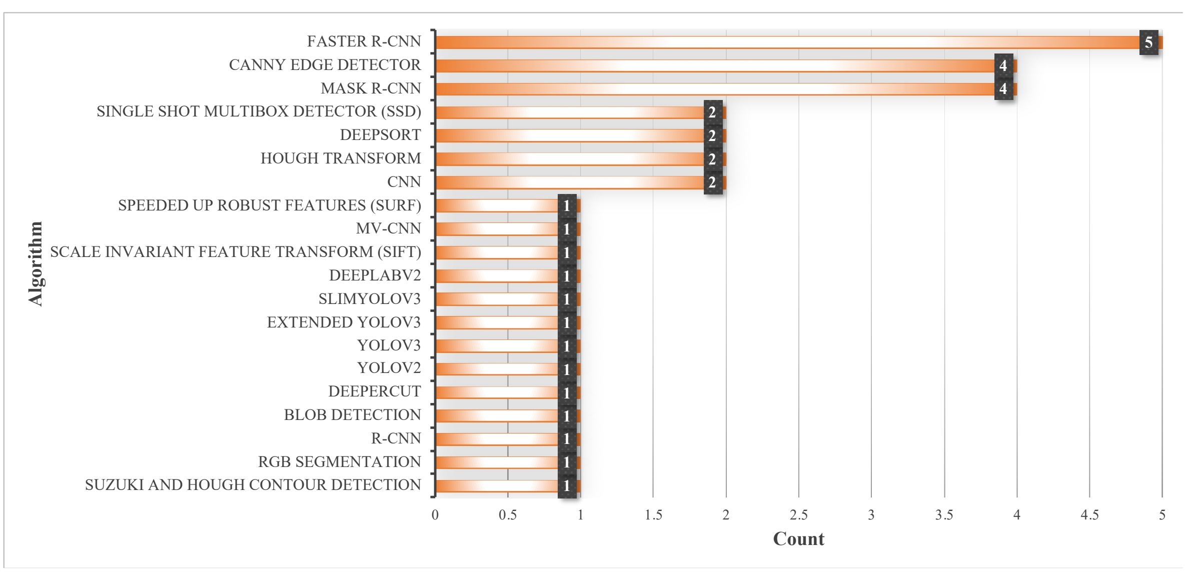


Fig. 2-4. Algorithms used.

2.3.3 Reported limitations

The most prominent limitations reported in the identified studies had to do with the object detection and classification, object tracking, and feature extraction applications, as summarized in Table 2-1. It should be noted that only prominent study limitations that were explicitly stated in the reviewed studies were included in the present scoping review. (In Table 2-2, a dash sign is displayed if the given study did not explicitly note any study limitations). The larger number of limitations reported with respect to the object detection applications is partially attributable to the larger number of object detection applications compared to other applications in offsite construction. The identified limitations are discussed in the following subsections.

2.3.3.1 Occlusions

Generally, occlusions are among the most commonly reported limitations in the identified studies, if we consider the different computer vision tasks as shown in Table 2-1. As this observation confirms, occlusions remain one of the most significant challenges in computer-vision applications, as they reduce the amount of useful visual information that can be extracted from images/videos of objects of interest. Among the reviewed studies, occlusions were most commonly reported with respect to the indoor applications of computer vision, since the confined space typical of a factory setting may result in other objects obstructing the camera's view of the object of interest (S. Ahn et al., 2019; Panahi et al., 2022a; Xiao et al., 2022). Such occlusions can be minimized, however, through careful planning of camera shot framing. For example, placing the camera at a high vantage point, placing the camera at an angle, and using multiple cameras are all strategies that can increase and improve the field of view, thereby minimizing the risk of occlusions. On the other hand, one of the occlusion-related limitations reported in an indoor application in one of the studies had to do with the nature of the particular computer vision application (K. Park et al., 2021). Specifically, computer vision was employed to automatically identify the next step in the assembly process of 3D modules. In the case of the process being investigated in that study, it was observed that, as modules are assembled, the completed portions of the modules occlude incomplete portions, making it more challenging to identify (using computer vision) the work yet to be performed on the incomplete portions of the modules. In other words, although typical occlusions resulting from other objects obstructing the camera's view of the object of interest can be mitigated through better positioning of the cameras (as discussed above), occlusions resulting from the object of interest itself, as in that case, may be more challenging. Hence, for applications such as tracking progress on module assembly, it should be examined whether other technologies, such as a location tracking system, could be better suited than computer vision.

2.3.3.2 Illumination

Another prominent challenge in object detection is illumination issues—the most frequently reported limitation with respect to the object detection applications in the identified studies. While a study involving object detection in an outdoor environment (studying onsite installation of precast walls) noted the typical illumination-related challenges resulting from changes in natural lighting conditions (Z. Wang et al., 2021), a frequently reported illumination issue in other studies was the glare from reflective surfaces of the objects of interest themselves (e.g., steel studs and panels) in both indoor and outdoor applications (Ahmadian Fard Fini et al., 2021; Martinez et al., 2019a, 2020). Moreover, although indoor lighting conditions are usually more controlled and stable than outdoor lighting conditions, natural light coming from open doors and windows was identified in one of the studies as potentially adversely affecting performance (Panahi et al., 2022a). Nevertheless, there are several solutions available to mitigate such limitations. For instance, thermal cameras can be used for object detection in low-light and dark conditions (Ippalapally et al., 2020). As for eliminating glare in the case of objects with reflective surfaces, using polarizers on the camera lens, or even on the light sources in the case of factory applications, can help (Walker, 2012). With regards to the effect of light from open doors and windows at the factory, additional lighting fixtures can be installed to balance the lighting. In general, it is important to carefully consider lighting conditions when designing and implementing computer vision applications, whether off site (indoor/factory setting) or on site (outdoor setting).

2.3.3.3 Limited view ranges of cameras

In addition to occlusions and illumination issues, another limitation that is likely to be encountered in computer vision applications in offsite construction, despite being mentioned in only one study, is the limited field of view of cameras (Z. Wang et al., 2021). This is particularly problematic when the objects of interest are prefabricated elements that are large in size (e.g., a long wall panel), and can be even more challenging in the case of indoor (factory) applications, since the confined space further imposes constraints on the positioning of cameras. One potential solution for detecting and tracking large, prefabricated elements is to use Pan-tilt-zoom (PTZ) cameras, which allow for the camera's direction and zooming capabilities to be controlled. Hsia et al. (2022) showcased the application of object tracking using a camera with a 360° horizontal and 90° vertical movement range. Another potential solution identified in the literature is image-stitching, which consists of combining multiple images captured from multiple viewpoints with overlapping fields of view into a single, larger image that captures a wider field of view (e.g., (Brown et al., 2007)).

2.3.3.4 Camera lens blockage

Another limitation that may arise in computer vision applications in offsite construction factories is the obscuring of camera lenses with dust resulting from fabrication activities such as cutting and drilling, especially in the case of wood-frame construction. While this limitation was mentioned in only one of the studies in the present review (Panahi et al., 2022a), it is likely to occur if proper measures are not taken, since working with wood generates a considerable amount of sawdust. The quality of captured images degrades as the lens becomes clogged with dust, adversely affecting the performance of the given computer vision application. As such, it is important to take measures to mitigate this effect by positioning cameras at high points, by using industrial cameras, or by furnishing the cameras with dust protection, to name a few examples.

2.3.3.5 Object complexity

A notable limitation in object detection reported in two of the studies is the design complexity (or irregularity of shape/size of the objects (Ahmadian Fard Fini et al., 2021; Martinez et al., 2019a). This limitation becomes particularly problematic when computer vision is employed to track building elements, such as walls, during production, since they change in size, shape or, more generally, features as they progress through production lines. For example, during exterior wood wall production, the wall frame increases in length and changes shape as framing elements are added to the frame, and, once the wall frame is completed, sheathing is installed, hence hiding the frame elements from view. Given that the features that could be used to recognize the walls during framing may be eliminated after sheathing installation, it may be difficult to train the computer vision algorithms to track walls during production. To overcome this challenge, the use of deformable models, a group of computer vision algorithms capable of modelling object variability, could be explored (Albrecht et al., 2009).

2.3.3.6 Other limitations

Other limitations reported in the identified studies that could be mitigated through careful planning include: (1) camera vibration or blurry images (Martinez et al., 2019a; Z. Wang et al., 2021; Zheng et al., 2020), which could be addressed by using stabilizing devices to keep the camera steady or by using cameras with anti-vibration systems, for example; (2) image resolution issues (S. Ahn et al., 2019), which could be addressed by using high-quality cameras, for example; (3) objects moving in a direction perpendicular to the camera (in object tracking) (K. Zhang et al., 2019), which could be addressed through proper placement of the camera or by using multiple cameras, for example; (4) scale variation (in object tracking) (Xiao et al., 2022), which could be addressed through the use of scale-invariant computer vision algorithms, such as Scale-Invariant Feature Transform (SIFT) or Speeded Up Robust Feature (SURF), for example; (5) adverse weather conditions (e.g., rain, mist) (Ahmadian Fard Fini et al., 2021), which could be addressed through the use of a hydrophobic camera lens coating, for example; and (6) high computational requirements (Z. Wang et al., 2021) and camera memory capacity and battery life limitations (Ahmadian Fard Fini et al., 2021), which can be addressed through the use of more powerful hardware.

Table 2-1. Limitations frequently reported in the literature.

Computer Vision Task	Limitations	Studies
Object detection	Illumination (or lighting) / Reflective surface of objects	(Ahmadian Fard Fini et al., 2021; Martinez et al., 2019a; Martinez et al., 2020; Panahi et al., 2022a; Z. Wang et al., 2021)
	Occlusions	(Ahmadian Fard Fini et al., 2021; Ahn et al., 2019; Panahi et al., 2022a)
	Camera vibration / Blurry images	(Martinez et al., 2019a; Z. Wang et al., 2021; Zheng et al., 2020)
	Design complexity / Irregular shapes / Size of objects	(Ahmadian Fard Fini et al., 2021; Martinez et al., 2019a)

Table 2 1. Limitations frequently reported in the literature (continued).

Computer Vision Task	Limitations	Studies
Object detection (continued)	Adverse weather conditions (e.g., rain, mist) / Dust	(Ahmadian Fard Fini et al., 2021; Panahi et al., 2022a)
	Image resolution	(S. Ahn et al., 2019)
	High computational requirements	(Z. Wang et al., 2021)
	Camera set up issues (e.g., memory, battery)	(Ahmadian Fard Fini et al., 2021)
Feature extraction	Occlusions / Appearance of random objects in the regions of interest	(Panahi et al., 2022b; K. Park et al., 2021)
Object tracking	Object moving in a direction perpendicular to the camera	(K. Zhang et al., 2019)
	Occlusions	(Xiao et al., 2022)
	Sudden variations in the scale of tracked objects (i.e., scale variation)	(Xiao et al., 2022)
	View range (and movement) of cameras	(Z. Wang et al., 2021)

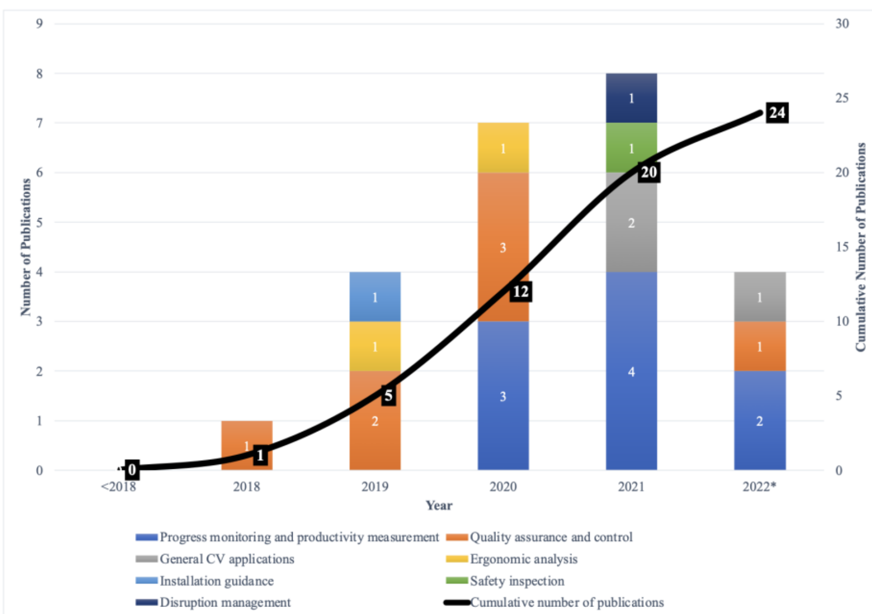


Fig. 2-5. Research growth trend[1].

[1] The review covers studies published as of November 10, 2022.

Table 2-2. Summary of extracted data.

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Quality assurance and control				
Automatically inspect the quality of precast concrete members (S. J. Lee et al., 2020) 2020	•Detect cracks and breakages (whether there is a damage or not) •Extract features (length and width) of cracks	•Faster R-CNN •Canny edge detector	Performance not reported	-
Automatically inspect the quality of precast concrete members (S. Lee et al., 2022) 2022	•Detect cracks •Extract features (length and width) of cracks	CNN (The detection performance of other algorithms is reported)	•Recall = 75% •Precision = 71% •The error obtained for measuring crack width ranged from 0.01 mm to 0.02 mm •An error of 0.1 mm was obtained for measuring crack length	-
Enable real-time automatic correction of screw driving operations in light-gauge steel frame fabrication to ensure screw driving operations are accurately performed (Martinez et al., 2019b) 2019	Detect pre-drilled screw holes	•Canny edge detector for edge detection •Suzuki and Abe contour detection algorithm (Suzuki, 1985)	The average error was 3.14 mm	-
Automatically inspect manually assembled light gauge steel frames and propose corrections if needed before screw-fastening operations are performed by the machine (Martinez et al., 2019a) 2018	Detect studs	Hough transform	The authors did not present the performance of the model. However, they reported that the numbers of detected studs were correct	•Light-gauge steel reflects light to the camera. Hence, high intensity lighting may result in an incorrect definition of studs •Vibration during image capturing results in blurry edges in the image frame and, hence, inaccurate metrics •In case of more complex designs (e.g., having multiple studs of varying lengths), some studs may be considered noise, and hence not detected

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Quality assurance and control (continued)				
Automatically inspect screw-fastening operations (squareness of stud connections and quality of fastened screws) in light-gauge steel frame manufacturing (Martinez et al., 2020) 2020	•Detect stud edges •Detect and classify screws	•Canny edge detector •Hough transform algorithm for edge detection •R-CNN for screw detection	•The average of mean errors for squareness estimation was 1.63 degrees •Screw-fastening detection had an overall accuracy of 91.67%	•Light-gauge steel is highly reflective and contain superficial marks and dents •The proposed system does not account for actual practices taken in response to such inspection results
Automatically guide workers to correctly position the different panel elements during panel fabrication and to look for any errors after the task is complete by projecting/superimposing an as-designed model (an augmented reality model) onto the working area in panelized construction. (S. Ahn et al., 2019) 2019	•Segment markers (i.e., colored stickers attached to the panel to delineate the area onto which the drawing is projected, known by projection area) and projection area •Detect markers and projection area	•RGB (red-green-blue) object segmentation •Blob analysis (blob detection algorithm) for object detection	•The average offset distances measured between the centers of the markers and the corners of the projected area were less than 6.35 mm •The accuracy was found to be inversely related to the projection distance (i.e., distance between the projector and projection area). As the projection distance was increased from 5 m to 8 m, the mean offset distances increased by 3.95 mm and 4.12 mm for high and low illumination, respectively	•The resolutions of the camera and projector affect performance •In practice, limited space and the tasks performed in working areas may result in possible occlusions (e.g., workers and other moving objects) and force having longer projection distances •High illumination may wash out the drawings projected on the panel surface. •The color and size of markers affect the accuracy of their segmentation depending on the color of the background
Automatically inspect the quality of prefabricated components (concrete columns and aluminum pipes) (D. Lee et al., 2020) 2020	3D re-construct precast columns and aluminum pipes	•VisualSFM tool (C. Wu, 2022) (used as GUI application for 3D reconstruction) •Multi-View Environment (MVE) (NavLab, 2015) (used for dense reconstruction)	The average error measured between points selected in base images and the corresponding points in the 3D reconstruction models was 6 pixels (6.61 mm) in the case of concrete columns and 167 pixels (13.74 mm) in the case of aluminum pipes	•The points on the images were selected manually, potentially resulting in human errors •The surface of the aluminum pipe's 3D reconstruction model had noise points

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Progress monitoring and productivity measurement				
Automatically monitor progress on module installation in modular construction (Z. Zhang et al., 2020) 2020	Detect and track modules	The algorithm was not specified.	Performance was not reported.	-
Automatically count the number of installed modules in modular integrated construction (Zheng et al., 2020) 2020	Detect and segment modules	Mask R-CNN	•They investigated the effect of using virtual images in addition to real-life images on the model's performance •Without virtual images, precision = 89%, recall = 83%, AP = 81% •With 1000 virtual images, precision = 96%, recall = 92%, AP = 91% •They applied the algorithm in two case studies; the AP reached 90% and 92%	The model cannot successfully detect modules in blurry images
Automatically track the hoist and installation process of precast walls (Z. C. Wang et al., 2021) 2021	•Detect and segment precast walls •Track (the trajectory of) precast walls over time	Mask R-CNN for detection DeepSORT for tracking	For IoU of 0.5, $AP_{0.50}$ = 88% and recall = 89%	-
Automatically monitor progress on hoisting and installation processes of precast walls (automatically collect data on the installation time of installed walls (timestamps) and the location of walls) (Z. Wang et al., 2021) 2020	•Detect and segment precast walls •Track (the trajectory of) precast walls over time	•Mask R-CNN for detection •DeepSORT for tracking	For mask R-CNN: •$AP_{0.50}$ = 0.93 and $AP_{0.75}$ = 0.85 •The area under the ROC curve for IoU of 0.5 and 0.75 is 0.92 and 0.89, respectively •Considering different construction sites, illumination conditions, and occlusions, precision ranged from 88.1% to 93.9%, recall ranged from 79.1% to 83.5%, and miss rate ranged from 16.5% to 20.9% •The performance of DeepSORT was not reported, but the complete framework successfully timestamped 10 out of 12 walls	•Poor illumination and shaking of the camera resulted in failure to timestamp the missed two walls •The movements and view range of the surveillance cameras in the construction sites can impact the performance of the proposed framework •The algorithms, especially Mask R-CNN, used in the study require high-performance hardware in practical application

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Progress monitoring and productivity measurement (continued)				
Automatically measure the onsite installation rate (square centimetres per minute) of prefabricated panels in panelized construction (Ahmadian Fard Fini et al., 2021) 2021	Detect installed panels	•Canny edge detector •Speeded up robust features (SURF)	•The algorithm resulted in a 95% correct panel detection rate •It produced dimensional panel information accurate to the centimetre •The accuracy of time-lapse data extracted from images was in interval of 30 seconds	•The proposed method may not be efficient in detecting steel panels due to their reflective texture •Capturing images is subject to limitations related to (1) photography apparatus whereby the camera might run out of memory and the battery might drain during data collection, (2) blockage of lens, and (3) adverse weather conditions such as rain, storms, and mist •The algorithm's performance is limited when detecting small panels with irregular shapes and/or when they are double handled (i.e., installed, removed and re-installed)
Automatically measure cycle time at workstations in modular construction factories (Panahi et al., 2022b) 2022	Extract features of regions of interest at the workstations	•Scale Invariant •Feature Transform (SIFT)	Accuracy after denoising images with the median filter = 100%	The appearance of random objects in the regions of interest for long durations may result in misclassifications
Automatically identify the next step in the module assembly process in modular construction using computer-vision models trained with virtual images (K. Park et al., 2021) 2021	Extract geometric features from 2D images of the modules and classify their 3D shapes	MV-CNN	Accuracy, precision, recall, and F1-score = 97%	The completed portions of the module occlude incomplete portions, thereby resulting in wrong predictions later in the assembly process
Automatically track progress and measure productivity (duration and man-hours) for floor panel fabrication in panelized construction (Martinez et al., 2021) 2021	Detect workers, machine, and crane	Faster R-CNN	•For worker detection, F1-score = 0.932 •For machine detection, the F1-score = 0.997 •For crane detection, the F1-score = 0.722 •The accuracy of estimating duration and man-hours required per task reached more than 92%	The proposed system is dependent on the standardized order of operations. Accounting for any potential variations to the typical order of operations requires considerable modelling effort and increases its complexity

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Progress monitoring and productivity measurement (continued)				
Automatically measure progress on and evaluate the status of the installation schedule of prefabricated components (walls and superimposed slabs) and other onsite activities (Yan et al., 2022) 2022	Detect walls, slabs, and workers	•Faster R-CNN with ResNet 101 architecture	For wall detection, $mAP_{0.50}$ = 82.0% For slab detection, $mAP_{0.50}$ =82.1% For worker detection, $mAP_{0.50}$ = 90.7% For predicting durations of activities, the error ranged from 6% to 18% for different activities For estimating the duration needed to complete activities at the floor level, MAPE = 9.3%	•The proposed system does not monitor prefabricated components other than walls and slabs •The system assumes that workers' physical conditions are recovered at the start of each activity
Ergonomic analysis				
Automatically perform ergonomic posture assessment of workers (Chu et al., 2019) 2019	Detect and identify a specific worker (whose body postures are to be analyzed) from other workers •Detect 2D body joints (e.g., ankle, knee, elbow, and wrist) Segment 2D body parts (e.g., head, torso, arm, and leg) •Reconstruct a 3D human body from 2D joints	DeeperCut algorithm (Insafutdinov et al., 2016) •DeepLabv2 algorithm (L.-C. Chen et al., 2017) •MDNet architecture (Multi-domain CNN (Nam et al., 2016))	They did not provide the performance of each algorithm. Rather, they evaluated the performance of their whole framework by measuring the accuracy of joint angles. The average error was 17.5 degrees	-
Automatically perform biomechanical analysis or ergonomic posture assessment (Chu et al., 2020) 2020			They did not report the performance of each algorithm. Rather, they evaluated the performance of their whole framework by measuring the accuracy of joint angles. The absolute average error was 11.7 degrees. The distribution of the angle errors has a mean value of 0.94 degrees and standard deviation of 17.55 degrees	-
Process guidance				
Develop a visual guiding technology to predict the trajectory for aligning holes of precast concrete columns with rebar installed in the ground (peg-in-hole assembly problem) (K. Zhang et al., 2019) 2019	Detect reinforcement (target position) and hole at the bottom of the precast member (moving target)	Improved Single Shot MultiBox Detector (SSD)	Only the confusion matrix was reported: the number of TP is highest (TP=1129), and the number of FN was more than number of FP (FN=232 > FP=119) (Refer to the paper for another reported performance metric (Hausdorff distance))	•The prediction accuracy decreases when the moving target (hole) moves in a direction perpendicular to the camera •The increase in wind speed decreases the success rate (fraction of successful assembly) of the method

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
Safety management				
Automatically identify workers and their safety characteristics (e.g., safety helmet) in prefabricated building construction (S. Liu et al., 2021) 2021	Detect workers and extract/recognize features (safety helmet, protective clothing, and other visible objects)	•Extended-YOLOv3 algorithm •YOLOv2 •YOLOv3 •SlimYOLOv3	•Object detection performance: The extended-YOLOv3 algorithm resulted in the highest F1-score of 70.1% while slimYOLOv3 resulted in the lowest F1-score of 66.3% •YOLOv3 resulted in the highest mAP of 60.1% while slimYOLOv3 resulted in the lowest mAP of 49.6% •A precision value of 94.6% was achieved using Extended-YOLOv3, but the recall value only reached 58.1% (Refer to paper for more performance measures)	-
Disruption management				
Automatically detect, evaluate, and respond to four types of schedule disruptions (delay in delivery of prefabricated components and concrete mix, jobsite traffic block, prefabricated components installation delay, ergonomics-related delays) on prefabricated building projects (Yan et al., 2021) 2021	•Detect and track mixer truck and prefabricated components-loaded truck for material arrival delay disruption •Detect trucks and road for jobsite traffic block disruption •Detect prefabricated slabs and walls for installation delay disruption •Detect worker's posture for ergonomic-related delay disruption	Faster R-CNN with ResNet 101 architecture	•Material arrival delay: -For mixer truck detection, $mAP_{0.5}$ = 82.4% -For PC-loaded truck detection, $mAP_{0.5}$ = 80% -For truck tracking, $mAP_{0.5}$ = 81.0%; ID switch = 0; MOTA = 93.6%; MOTP = 86.2% Jobsite traffic block: -For truck detection, $mAP_{0.5}$ = 80.36% -For recognizing the spatial relationship between a truck and the road, the mean error for estimating proximity between two points was 0.83m •Prefabricated components installation delay: -For detecting prefabricated walls, $mAP_{0.5}$ = 82.1% -For detecting slabs, $mAP_{0.5}$ = 82.0% •Ergonomics-related delays: -They monitored ergonomics-related delays based on work-rest cycles and, hence, based on the estimated duration to complete ergonomics-related work. They did not provide performance measures for worker detection. Rather, they provided MAPE for predicting durations which amounted to 11% (Refer to the paper for additional performance measures)	-

Table 2-2. Summary of extracted data (continued).

Study Goal / Year	CV Application	Algorithm(s)/ Approaches(s) Used	Performance	Limitations
General CV application				
Automatically track workers in offsite construction (Xiao et al., 2022) 2022	Segment and track workers	Mask R-CNN	•The algorithm resulted in average: 1) precision of 98.7%, 2) recall of 96.7%, 3) F1-score of 97.6%, 4) MOTA of 96.4%, 5) MOTP of 86.2% •The success rate of tracking worker trajectories was 95.1% (39 out of 41 were successfully tracked)	•Heavy occlusions (when over 80% of object area is occluded) result in the algorithm temporarily losing the object •Sudden variations in the scale of workers within a few frames may result in tracking failure
Automatically identify whether workers are performing nailing tasks or other activities in modular construction (Panahi et al., 2022a) 2021	Detect workers and classify their activities	CNN with ResNet-50 architecture	•Recall = 92% •Precision = 95% •F1-score = 93.5%	•Workers constantly move which results in a large number of occlusions •Changes in ambient lighting coming from windows and open doors may adversely affect results •Dust generated during drilling and cutting activities may hinder the detection of workers •Locating cameras close to workstations reduces the information that can be extracted from captured frames
Evaluate the performance of two object detection algorithms in detecting modular construction objects (C. Liu et al., 2021) 2021	Detect modular construction objects including panels, safety barricades, and fences	•Faster R-CNN •Single shot multi-box detector (SSD)	•Performance measured based on pixels: The highest AR corresponded to detecting barricades using faster R-CNN (AR@10 = AR@100 = 0.9208), while the lowest AR corresponded to detecting panels using SSD (AR@1 =0.1588). The highest mAP corresponded to detecting barricades using faster R-CNN ($mAP_{0.5}$ = $mAP_{0.75}$= 0.9996), while the lowest mAP corresponded to detecting panels using SSD (mAP = 0.5388). In short, faster R-CNN showed better performance •Performance measured based on counting bounding boxes: The precision of faster R-CNN ranged from 0.67 to 0.9 while its recall value ranged from 0.59 to 0.98. As for SSD, precision ranged from 0.79 to 0.94 and recall ranged from 0.53 to 0.64. RCNN's recall was higher, but SSD's precision was higher (Refer to paper for more performance results)	•The size of used images (990) was relatively small compared to disciplines other than construction (over 1000) •Both algorithms' detection of panels was poor •The backgrounds of the images could decrease recall and precision values of detecting objects

Table 2-3. Glossary of performance-related terms.

Metric	Symbol	Description
True positive	TP	TP occurs when the model identifies a positive class when the truth is positive. e.g., a model detects a person in an image, and the detected object is actually a person.
True negative	TN	TN occurs when the model identifies a negative class when the truth is negative. e.g., a model does not detect a person in an image that does not display a person.
False positive	FP	FP occurs when the model identifies a positive class when the truth is negative. e.g., a model detects a person in an image, and the detected object is not a person.
False negative	FN	FN occurs when the model identifies a negative class when the truth is positive. e.g., a model fails to detect a person in an image that displays a person.
Intersection-over-union	IoU	In object detection, a model borders the detected object with a bounding box that is aimed at predicting a specific object's position (or localizing the object) in the image. To evaluate the performance of the model in doing so, the actual bounding box (called ground-truth bounding box) is manually drawn and compared to the predicted bounding box. The metric intersection-over-union (IoU) is then computed as the ratio of the area of the predicted bounding box that overlaps with the ground-truth bounding box (intersection, ∩) to the combined area of both boxes (union, ∪) (Hymel, 2022). An IoU value of zero indicates a complete miss (i.e., zero overlap), while a value of one indicates a perfect overlap (Hymel, 2022). An IoU threshold can be set to filter out predictions will low IoU value. A bounding box having an IoU value lower than the threshold is considered a FP (Hymel, 2022).
Accuracy	-	It is the ratio of correct predictions (TP+TN) to total predictions (TP+TN+FP+FN).
Precision	-	It is the ratio of correct positive predictions (TP) to the total positive predictions (TP+FP).
Recall	-	Recall is the ratio of correct positive predictions (TP) to the total number of positive cases including those correctly predicted and missed by the model (TP+FN).
F1-score	-	F1-score combines the precision and recall metrics and is equal to their harmonic mean ($2 \times \frac{precision \times recall}{precision+recall}$).
Output confidence score	s	It represents the confidence of the model in correctly predicting the output.
Confidence score threshold	t	A threshold value is set to ensure that predictions have a minimum confidence score value and filter out FP (Wenkel et al., 2021). All positive predictions having a confidence score lower than the threshold (s<t) are considered negative. This helps reduce potential FP predictions. Nevertheless, since there is a probability that a positive class (P) has a score lower than the threshold (i.e., P(s<t/P)) (Krzanowski et al., 2009), setting a threshold may also eliminate some TP predictions, thereby increasing the number of FN. In mathematical terms: $Threshold\ (\uparrow) \rightarrow FP\ may\ (\downarrow)\ but\ TP\ may\ (\downarrow)\ and\ FN\ (\uparrow)$. As such, precision and recall values are contingent upon the selection of the confidence score threshold.

Table 2 3. Glossary of performance-related terms (continued).

Metric	Symbol	Description
Average precision	AP AP_{IoU}	For different values of the confidence score threshold, precision and recall are computed, and a precision-recall curve can be plotted. The AP is equal to the area under the precision-recall curve ($AP = \int_0^1 p(r)dr$) (Hymel, 2022). In object detection, the precision-recall curve varies depending on the selected IoU threshold and hence, there is an AP value for each IoU threshold (denoted as AP^{IoU} or AP_{IoU}) (Hymel, 2022).
Mean average precision	mAP mAP_{IoU}	The average precision metric is computed for each class separately. To evaluate the overall performance of a model, the mAP metric is computed as the mean of the AP metrics computed for all classes (Hymel, 2022). In object detection, mAP is traditionally computed for a single IoU threshold (denoted as mAP_{IoU}) (Hymel, 2022). However, there are variations to this definition. For example, in the evaluation metrics used by COCO, mAP is averaged over multiple IoU thresholds and all classes (denoted as just AP or mAP) (*COCO - Common Objects in Context*, 2022).
Receiver operating characteristic curve	ROC curve	As previously explained, there is a probability that a positive class (P) has a confidence score lower than the confidence score threshold (i.e., $P(s<t/P)$). Similarly, there is a probability that a negative class (N) has a score greater than the threshold (i.e., $P(s>t/N)$); this probability is known as the false positive rate (Krzanowski et al., 2009). Following the same nomenclature, the recall metric defined above is equivalent to the probability that a positive class has a score greater than the threshold (i.e., $P(s>t/P)$). This is also known as the true positive rate. The ROC curve plots the true positive rate (or recall) against the false positive rate for different threshold values (Krzanowski et al., 2009)
Average recall	AR	In object detection, the recall value varies depending on the selected IoU threshold. AR is the average of recall values computed across IoU thresholds (Hosang et al., 2016). For IoU ranging between 0.5 and 1, the AR metric is equal to two times the area under the recall-IoU curve ($AR = 2\int_{0.5}^1 recall(IoU)d(IoU)$) (Hosang et al., 2016). In the evaluation metrics used by COCO, AR is also averaged across classes (*COCO - Common Objects in Context*, 2022). Also, the AR is sometimes calculated for a fixed number of detections per image (denoted as $AR^{max=10}$ or AR@10 in case of ten detections per image for instance), as used by COCO (*COCO - Common Objects in Context*, 2022).
Identity switch (also called mismatch errors)	IDS	In multi-object tracking, a model should assign a unique tracking identity (ID) to each object; this ID should remain constant for each object to ensure consistent tracking over time (Bernardin et al., 2006). If different objects get close to each other, the model may switch their IDs (Bernardin et al., 2006). Moreover, the model may assign a new identity to an object it previously lost track of (due to an occlusion for example) and then recaptured (Bernardin et al., 2006).

Table 2 3. Glossary of performance-related terms (continued).

Metric	Symbol	Description
Multiple Object Tracking Precision	MOTP	In multi-object tracking, MOTP measures how precise the estimates of objects' positions are (Bernardin et al., 2006). It is equal to the average of the distances measured between the ground-truth object's position and the corresponding tracker's output's position (Bernardin et al., 2006).
Multiple Object Tracking Accuracy	MOTA	In multi-object tracking, MOTA is a function of the number of FP, FN, and mismatch errors or ID switches (Bernardin et al., 2006).

2.4 Conclusions

2.4.1 Summary of findings and avenues of future research

This chapter presented a scoping review of computer vision applications in offsite construction. A total of 887 studies were screened, 48 studies were manually assessed for eligibility, and 24 studies were ultimately included in the review based on a pre-defined set of exclusion criteria (with the review reflecting studies published as of November 2022). Six main research areas were identified in the review: (1) "progress monitoring and productivity measurement", (2) "quality assurance and control", (3) "ergonomic analysis", (4) "process guidance", (5) "safety management", and (6) "disruption management". Of the 24 studies, 21 were classified under these research areas, while the remaining three were classified as "general CV applications". Moreover, more than 20 computer vision algorithms and approaches were deployed among the identified studies to perform various types of computer vision tasks, including object detection, classification, object tracking, feature extraction, segmentation, and 3D reconstruction. The principal findings of the present review and potential avenues of future research are summarized as follows:

a) With regard to the research areas of application:

(1) The most frequently targeted research area was "progress monitoring and productivity measurement". The studies showcase the promising opportunity in offsite construction to automatically acquire productivity-related data quickly and efficiently. With the

increased feasibility and convenience of automated progress and productivity tracking, offsite construction companies can more easily identify areas of inefficiency or bottlenecks in their operations and make the necessary adjustments. This can greatly enhance the productivity of offsite construction, which is especially notable considering the construction industry's longstanding productivity challenges compared to other industries. However, none of the reviewed studies provided a comprehensive comparative analysis of production or installation activities performed with and without computer-vision-based progress and productivity tracking. It would be of value for the extent to which automated tracking of progress and productivity could reduce activity duration or improve productivity in offsite construction to be investigated.

(2) The second most targeted research area in the reviewed studies was "Quality assurance and control". These studies highlighted the potential of computer vision in enabling continuous quality assurance throughout production processes, extending beyond the mere identification of defects in final products. Real-time error detection by computer vision, ideally, will help to prevent defects from occurring in the first place, thereby reducing the time and money spent on correcting errors. Such a system would be of significant value to offsite construction companies if it could help to eliminate, or at least reduce, instances of defective prefabricated elements being shipped to site for installation, especially considering that correcting errors on site necessitates having the requisite tools and materials available on site, thereby adding to the time loss and cost associated with quality issues. The existing studies on this topic showcase the feasibility of developing computer-vision-based quality assurance and control systems, but there is a need to test such systems on real projects and evaluate their reliability and efficiency compared to manual quality inspection procedures. Ultimately, since all quality issues on a project should be resolved, it is essential to determine to what extent we can rely on fully automated computer-vision-based quality assurance and control systems.

(3) While the fixed layouts of offsite construction factories make it easier to make adjustments to operations that improve worker ergonomics, there has been relatively little exploration of the use of computer vision for ergonomic analysis in offsite construction. Since offsite construction still heavily relies on manual labour due to low

levels of automation, this research area requires further investigation, as the use of computer vision for ergonomic analysis has the potential to improve both workers' physical health and project performance.

(4) Despite the repetitive nature of their work, workers in offsite construction factories must frequently refer to shop drawings to determine the necessary steps in their work due to the design variability of different building components. In this regard, one of the studies identified in the reviews superimposed as-designed models onto the working area to help workers identify the correct positions of the different panel elements during panel fabrication. An automated process guidance system, by enabling workers to continuously compare the built panel with the as-designed panel during fabrication, has the potential not only to improve productivity by reducing the time wasted by workers having to frequently consult the paper drawings, but also to reduce errors. Therefore, further research should be undertaken that follows this line of thinking, using computer vision to develop automated guidance systems for other processes.

(5) Although offsite construction generally reduces the safety risks compared to traditional construction, workers in offsite construction are still at risk of being struck by a moving object (e.g., forklift) or becoming trapped between stationary and moving objects. While computer vision can be deployed to detect and signal such hazards, safety-related research on the use of computer vision in offsite construction has been quite limited. In fact, our review identified only one such study as of November 2022, that one focusing on the automated detection of PPE. Future research could be undertaken to investigate the use of computer vision for automatic detection and signaling of hazardous proximity between workers and objects and between moving objects in an effort to further reduce safety risks in offsite construction.

(6) Since greater predictability of time and cost is one of the main benefits of offsite construction, using computer vision to automatically detect and manage disruptions is of high potential benefit to the offsite construction industry. Our review identified two studies that developed computer-vision-based disruption management systems for the site installation phase in offsite construction. While the risk of disruptive events may be lower in a controlled factory setting, their impact on production may be higher, particularly given the manufacturing-based design of operations typical of offsite

construction. Specifically, because workstations are linked by a common workflow, a disruption at one workstation could affect the entire production line. As such, there is a need to develop automated disruption management systems for factory production.

b) With regard to the algorithms used and the studied objects:

(1) The Faster R-CNN algorithm, followed by the Canny edge detector and the Mask R-CNN algorithm, were the algorithms most frequently used in the identified computer vision applications. Future research could explore the reasons underlying their frequent use in the studies reviewed. Future work could also investigate whether offsite construction research is harnessing the full potential of these algorithms, but this would require a critical evaluation of the methods used in the studies in order to provide insights into the application of these algorithms. Finally, the deployment of other emerging algorithms, such as YOLOv7, could be examined.

(2) A variety of different objects have been targeted in the identified computer vision applications in offsite construction, including cracks in precast concrete members, workers, machines, cranes, precast walls, slabs, wood panels, safety barricades, fences, light-gauge steel studs, screws, modules, reinforcement bars, PPE, and trucks. In this regard, it would be worthwhile to build a large-scale, open-source dataset containing labelled images of the objects commonly targeted in computer vision applications in offsite construction. Such a dataset would be of great value in supporting the training of computer vision algorithms in offsite construction applications.

c) With regard to the prominent limitations encountered:

(1) Occlusions were among the most commonly reported limitations in the reviewed studies. Occlusions are likely to be encountered in computer vision applications in offsite construction factories, as the confined space within a factory setting may result in other objects obstructing the camera's view of the object under study. As such, future work should be undertaken to carefully plan the placement of cameras, investigating strategies such as placing the camera at a high vantage point, placing it at an angle, or using multiple cameras to increase and improve the field of view. As for occlusions

caused by the object under study itself, such as the completed portions of the module obstructing incomplete portions during module assembly, research should be undertaken to assess whether alternative technologies, such as a location tracking system, could be better suited than computer vision.

(2) Illumination issues were the most frequently reported limitation of object detection applications in the identified studies, with these issues encompassing changes in the natural lighting conditions in the site installation phase, glare from the reflective surfaces of steel studs and panels in both indoor and outdoor applications, and natural light coming from open doors and windows in factory settings. There are several solutions available to mitigate such limitations, including using thermal cameras for nighttime applications, furnishing camera lenses and light sources with polarizers for glare, and installing light fixtures to balance the lighting in factories, to name a few.

(3) Given that the target objects of most applications of object detection in offsite construction are large, prefabricated elements, such as long walls, floors, or roofs, the limited range of view of most cameras may pose a challenge, especially in confined factory spaces, where the camera positioning may be restricted. To address this issue, future research could explore such solutions as using cameras with flexible vertical and horizontal rotation capabilities, or employing the image-stitching approach, which combines images captured from multiple viewpoints in order to expand the field of view and accommodate large, prefabricated elements.

(4) The obscuring of camera lenses with dust resulting from fabrication activities such as cutting and drilling is likely to occur in factory applications, especially when working with wood. As such, failure to take proper measures can adversely affect the performance of computer vision applications. Therefore, future research should take the necessary measures to address this issue, such as positioning cameras at high points, using industrial cameras, or providing cameras with dust protection, to name a few.

(5) Two of the studies identified reported a notable limitation in object detection related to the complexity of the object's design or irregular shapes/sizes. This limitation could be particularly problematic when using computer vision to track building components such as walls during production, as their size and shape change throughout the production process. In such cases, training computer vision algorithms to track building

components during production could be challenging, since the features used to recognize components at one workstation may be eliminated (or otherwise be undetectable) at a subsequent workstation. To address this challenge, research could be undertaken to explore the use of computer-vision models capable of modelling object variability, such as deformable models.

d) With regard to the growth of applications:

(1) The earliest publication identified in the present review on computer vision in offsite construction was published in 2018. In other words, research in this area began relatively recently, considering that the use of computer vision in traditional construction dates back to as early as 1994. Future research could investigate why the offsite construction industry tends to lag behind the traditional construction industry and other industries in the application of digital technologies such as computer vision.

(2) The yearly rate of publications has been generally growing in recent years, but this growth has been quite modest, ranging from one to eight publications per year between 2018 and 2021. Still, the cumulative number of studies has been growing over the past five years (2018 to 2022). Given the projected growth in global market size of computer vision and the momentum around its adoption, it would be worthwhile for another scoping review similar to the present one to be undertaken in the coming years to examine the extent to which the offsite construction has advanced in its adoption of digital technologies.

2.4.2 Review limitations and challenges

The present scoping review was subject to several limitations and challenges, summarized as follows:

- The Scopus database, the Compendex database, the Web of Science platform, as well as five leading journals in the building and construction field (*Automation in Construction*; *Journal of Information Technology in Construction*; *Journal of Computing in Civil Engineering*; *Engineering, Construction and Architectural Management*; *The*

International Journal of Advanced Manufacturing Technology) were searched to extract studies on the use of computer vision in offsite construction. These sources were selected due to their relevance to the use of digital technologies in construction. However, there may be studies using computer vision in offsite construction that are not available in any of the databases or journals searched.

- Capturing all the studies related to a given topic based on keywords is inherently challenging given that differing, and sometimes, uncommon nomenclatures may have been used depending on the study. Future work in this area could include a review of the nomenclature used to refer to various computer vision tasks as the basis for developing a comprehensive list of keywords for use in future literature reviews in this area.
- The authors' knowledge plays a key role in the accuracy of the results of any review study. There is a risk of potential bias and misinterpretation in the review process. Measures were taken to minimize this risk by (1) having two individuals review the studies independently, (2) comparing the extraction results from the two reviewers for any discrepancies, and (3), in the case of discrepancies, repeating the extraction process for the corresponding studies until reaching consensus.

Acknowledgements

This work was supported by Alberta Innovates. The authors would also like to acknowledge the support of Ms. Alison Henry and the technical review of Mr. Jonathan Tomalty.

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 3

Computer-vision-based Process Time Data Acquisition for Offsite Construction

3.1 Introduction

3.1.1 The need for time data and the significance of data quality

With traditional construction having lagged behind other industries in terms of productivity for decades (Bertram et al., 2019), it has been argued that offsite construction, also known as construction manufacturing or prefabricated construction, gives the construction industry a "productivity boost" (Bertram et al., 2019; Hogarth, 2020). Additionally, as the fourth Industrial Revolution ("Industry 4.0") precipitates an evolution of manufacturing practices more broadly (Lasi et al., 2014), various digital technologies are showing great promise in terms of bringing about improved efficiency and productivity in the offsite construction industry (M. Wang et al., 2020).

The use of digital technologies for efficiency- and productivity-related matters, however, involves the acquisition and processing of large volumes of time data. In this regard, data quality and accessibility issues constitute a significant barrier-to-entry, inhibiting the incorporation of data science projects into industry standards, with only 13% of these projects actually becoming part of production (Venture Beat, 2019). In addition to the challenge of data accessibility, data quality issues encountered may include accuracy, objectivity, relevancy, timeliness, completeness, size, interpretability, and consistency, to name a few (R. Y. Wang et al., 1996). Data quality, it should be noted, has been defined as "the capability of data to be used effectively, economically, and rapidly to inform and evaluate decisions" (Karr et al., 2006). Thus, the notion of quality of data is also contingent upon the purpose for which the data is to be used. For instance, we can consider the lightweight wood panelized construction factory used for the case study presented herein. Panelized construction is a type of offsite construction in which wall panels, floor panels, and roofs are fabricated in the factory and shipped to the construction site for installation. A radio frequency identification (RFID) system is used at the case factory to track wall panels as they flow through the wall production line. An RFID tag is attached to each panel so that it can be tracked by the system, and timestamps indicate a panel's detected location at a given time. Although such data is suitable for tracking manufacturing progress, multiple deficiencies of this system have been identified when used for process time measurement, including high noise-to-signal ratios, extensive data storage requirements, and rigorous data featuring and engineering when used for process time measurements (Altaf, 2016; Mohsen, 2021). The system necessitates that RFID tags be installed, sometimes manually, onto individual components, introducing substantive manual input into otherwise semi- or fully-automated processes. For these reasons, the case company elected to suspend the operation of its RFID system after having invested a significant amount of capital and effort in developing and implementing it.

The company currently measures manufacturing efficiency and productivity based on ratios of total square footage (of wall, floor, and roof panels) produced per day to the corresponding total labour hours. However, such ratios combine efficiency and productivity measures corresponding to different workstations at which various levels of automation, different number of workers, varying number of tasks, different team leads, and generally variable conditions are at play. As such, such ratios cannot be relied upon when evaluating work at the workstation level or identifying opportunities for improvement. Offsite construction companies need a reliable means of collecting process time data, and the selection of a system that acquires data of sufficient quality for the intended purpose is crucial. A notable technology currently in use within the offsite construction industry for automated collection of time-related data is computer vision, as described in the following subsection.

3.1.2 Use of computer vision for time data acquisition in offsite construction

Computer vision is a field of artificial intelligence in which a computing system is used to extract meaningful information from visual components (IBM, 2022), such as digital images, videos, and other visual inputs, including cameras and closed-circuit television (CCTV), in order to mimic the way in which the human brain perceives and understands visual information (Huang, 1996). Such approaches have enabled the accurate identification and classification of objects, leading to data-driven recommendations or actionable insights within various systems. Examples of computer vision applications include object detection (Papageorgiou et al., 1998), motion tracking (Lowe, 1992), action recognition (Ji et al., 2012; H. Wang et al., 2013), human pose estimation (Sun et al., 2019; Toshev et al., 2014), and semantic segmentation (L. C. Chen et al., 2018; Long et al., 2015). The technology has been successfully deployed in the offsite construction industry for acquiring productivity/efficiency-related data. The value of this technology was demonstrated in a recent study in which it was applied to automatically track and monitor the installation of modules in real time in a critical, fast-tracked modular construction project in Hong Kong (Zhang et al., 2020). In this study, the real-time monitoring was deemed to have been a critical factor in the rapid delivery of the project. Indeed, computer vision has been used for progress monitoring in a number of recent studies. For instance, Zheng et al. (2020) deployed the technology to track the number of modules installed on modular construction projects and calculate the durations during which modules appear in a region of interest, achieving an accuracy of 97.7%. Wang et al. (2021), meanwhile, implemented computer vision to automatically track the trajectory of the hoisting and installation of precast walls, achieving precision and recall of 88% and 89%, respectively. In a similar study, computer vision was used to automatically collect data on the installation time of precast walls (timestamps) and the location of walls (Z. Wang et al., 2021a). In that study, the researchers were able to successfully timestamp 10 out of 12 walls using their computer vision method (Z. Wang et al., 2021a). Computer vision has also been deployed to automatically measure the onsite installation rate (cm^2/minute) of prefabricated panels in panelized construction, with an error of less than 5% having been achieved (Ahmadian Fard Fini et al., 2021). Finally, computer vision has been used to automatically track progress on, and measure the duration and man-hours spent on, floor panel fabrication in panelized construction, with the developed method achieving an overall accuracy of over 92% (Martinez et al., 2021).

Notwithstanding these examples of the successful application of computer vision for acquiring productivity/efficiency-related data, there are still technical challenges with respect to this technology that need to be addressed, as detailed in the following subsection.

3.1.3 Research motivation: Set-up effort in custom vision applications

Computer vision applications in offsite construction have shown promising performance as described in the previous section. However, depending on the design of the computer-vision-based method used in these applications, the typical process of building computer-vision models often requires extensive up-front set-up efforts. Typically, a dataset comprising images or videos of the targeted object must first be created. The dataset must then be annotated, and this could entail classifying images or drawing bounding boxes around the objects of interest and classifying them. The annotated dataset is then used to train a computer-vision model, and this task also requires significant set-up efforts. The annotated dataset must also be of sufficient size and quality to ensure adequate performance. In other words, the task of developing the computer-vision models is an iterative process that requires significant time, effort, and resources. In the context of the present study, computer-vision models built following the steps described above are referred to as custom vision models. As per Microsoft's definitions, computer-vision models are pre-trained, and they process images based on users' input (Microsoft, 2022a), while custom vision models are trained by users (Microsoft, 2022b). The set-up effort required in order for custom vision to be successfully deployed is evident in a number of productivity- and efficiency-related applications described in the literature. The following tasks had to be carried out in the studies noted below as prerequisites for training the custom vision models:

- Zheng et al. (2020): Virtually simulate the modular construction process in order to capture virtual images of the process. (This task was necessary in order to increase the size of the dataset used for training the object detection algorithm to detect building modules, since the number of real images was limited.)

- Wang et al. (2021): Manually select 580 images containing precast walls after obtaining surveillance images and videos from four different construction sites, adjust the resolution of images, and label the precast objects before training the computer-vision model.
- Wang et al. (2021b): Create a training dataset of 600 images and train the model for 6.5 hours of GPU computational time (equating to more than 24 CPU-mode computational hours).
- Ahmadian et al. (2021): Pre-process the training dataset prior to training in order to remove image distortion (resulting from the non-uniform zoom levels caused by the wide-angle lens used to capture the whole scene on the construction site), stabilize images, and correct them.
- Martinez et al. (2021): Collect and manually label 1,069 images to detect workers and equipment.

These examples demonstrate the range of potential issues (e.g., insufficient size of training dataset, distorted images, high computational requirements) that may be encountered when training custom vision models. Moreover, the model-training burden may be compounded in the case of studying the offsite construction manufacturing process, since building objects change in shape and size as they progress through production lines. For instance, exterior wall panel production begins with framing, where the wall frame increases in length and changes shape as framing elements are added to the frame. Once the wall frame is completed, sheathing is installed, thus hiding the frame elements from view. This process makes it inherently challenging to train an object detection model to track the wall during the manufacturing process. Tackling this challenge is at the core of the present study, as explained in the following subsection.

3.1.4 Research objectives and contribution

Despite the significant potential that custom vision holds for the offsite construction industry, a review of computer vision (custom vision) applications in construction identified only two related publications in offsite construction as of 2019 (Martinez, Al-Hussein, et al., 2019). Although more studies have been published since the time of that review study (as noted above), the assertion on the part of Martinez et al. (2019), that the use of computer vision (custom vision) in offsite construction is under-researched, still holds. Meanwhile, the field of computer vision in general has seen rapid growth. Despite potential discrepancies between the global computer-vision market size figures forecast by different market researchers (e.g., valued at USD 10.5 billion in 2019 (KBV research, 2020), USD 14.82 billion in 2020 (Verified Market Research, 2021), and USD 12.78 billion in 2021 (Data Bridge Market Research, 2022) and estimated to reach USD 17.9 billion by 2026 (KBV research, 2020), USD 27.02 billion by 2028 (Verified Market Research, 2021), USD 24.19 billion by 2029 (Data Bridge Market Research, 2022)), there is a consensus on a projected upward trend. Given the momentum around its adoption in other fields, any effort to simplify its application will increase its attractiveness and perceived value to offsite construction professionals, in turn spurring its adoption in the offsite construction industry.

In this context, and given the need for time data in offsite construction, this chapter presents a low-cost, low-setup requirement computer-vision-based time data acquisition system (TiDA) that automatically collects productivity-related data (productive time, cycle time, and process start-time, as detailed in Section 3.2 below) at the workstation level on standardized production lines in panelized construction factories. TiDA makes use of an open-source object detection model that has been extensively pre-trained to detect a set of objects commonly encountered in everyday life. In other words, the logic underlying the tool proposed in the present study, TiDA, enables the user to employ object detection models for custom objects without having to create annotated datasets and train models, thereby significantly reducing the set-up requirements.

The following criteria underlie the design and development of TiDA:

- *Comprehensibility*: In order for practitioners to have confidence in the solutions generated by the system, the underlying logic must be easy to understand. Although complex methods may be needed to build a system that offers the targeted functionalities, the logic should be represented in layman's terms.
- *Ease-of-use:* Practitioners should be able to quickly learn how to use the system without assistance and adapt it to their needs. Since the familiar process of making rough estimates

is the default alternative to any new system, the ease-of-use criterion is vital for increasing adoption of the system.

- *Adaptability:* Panelized construction factories vary from one another in terms of the level of automation of the operations, ranging from primarily labour-based to semi- or fully-automated. Moreover, factories with comparable levels of automation may still differ significantly from one another in terms of the manufacturing practices employed. Hence, adaptability is a key factor in the success of any data acquisition system.
- *Timeliness:* The system should be capable of fulfilling the given requirements in a timely manner with minimal human input as an attractive alternative to the current practice of making rough estimates. Ensuring that the system is time-efficient will reduce the likelihood that the burden of implementing and operating the system is perceived to outweigh its potential benefits.
- *Cost-effectiveness*: Even when a system is deemed to be value-adding and practical, cost may be a barrier to implementation. Since the current practice of relying on rough estimates does not entail any direct costs, any alternative solution proposed must be low-cost in order to establish a competitive advantage.

These criteria are revisited in Section 3.6.2 of this chapter to demonstrate the value and contribution of TiDA. The rest of the chapter is organized as follows. Section 3.2 describes the logic underlying TiDA's design and its architecture. The procedure for applying TiDA to workstations in panelized construction factories is described in Section 3.3. Section 3.4 presents the user interface and provides a general overview of how the system is operated. Section 3.5 provides an assessment of TiDA's performance based on a case implementation. Finally, Section 3.6, in addition to describing the value and contributions of TiDA as noted above, provides a summary of the results obtained, outlines the limitations of TiDA, and proposes potential avenues of future work in this area.

3.2 TiDA design and architecture

This chapter describes the logic of TiDA in reference to workstations on the wall production line at a lightweight wood panelized construction factory. A virtual representation of these workstations is displayed in Fig. 3-1.

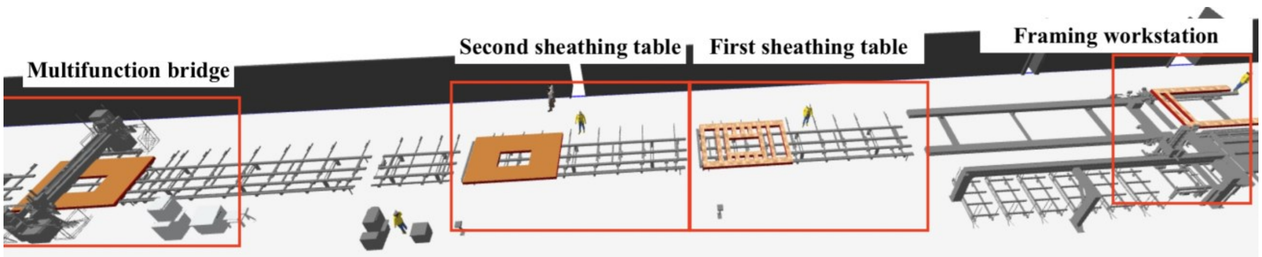


Fig. 3-1. Virtual representation of workstations.

A number of different tasks are performed on these workstations. First, wall panels are framed at the framing workstation, where a semi-automated wood-framing machine is used to automatically perform nailing, drilling, and cutting operations. One operator is needed to load the machine with framing elements (e.g., studs, subassemblies for doors and windows) when prompted by the machine to do so. The wall frame is then transferred to the first sheathing table using a conveyor system, where workers manually inspect the frame for any errors, perform the necessary adjustments, and attach missing elements. Next, the panel is pulled to the second sheathing table using the conveyor, where sheathing is manually placed and temporarily stapled to the frame. Next, workers transfer the wall panel to the multifunction bridge, which is a fully automated machine that fastens sheathing and cuts openings. Different wall panel types, it should be noted, have different designs, and therefore vary in terms of both the types of elements used and the quantity of each type. In other words, the amount of time needed to complete tasks at different workstations varies depending on the design of the given panel.

3.2.1 System functionality

TiDA measures three process time variables for each panel at a workstation, as outlined below:

- Start-time (*ST*): *ST* is the time at which work on a panel is initiated at the workstation.

- Productive time (PT): PT is the time during which work is being performed either by the machine or by the worker. In other words, it is the time actually spent by resources working on a panel at the workstation.
- Cycle time (CT): CT is the time spanning from the start of the process undertaken at the workstation to the end of the process. Under ideal conditions where time efficiency is 100% (i.e., zero waste of time), CT is equal to PT. However, 100% efficiency is not possible in reality, as various forms of delay (D) inevitably arise (e.g., machine breakdown, workers off task, material shortage, etc.). CT is equal to the sum of PT and D $(i.e., CT = PT + D)$.

3.2.2 System logic and assumptions

3.2.2.1 Object detection model

In the application of TiDA as described in the present study, YOLOv4 (Bochkovskiy et al., 2020) is the model used for object detection. YOLOv4 is the fourth version in the "You Only Look Once" (YOLO) family of object detection models. Object detection models are trained to look at an image and identify the probability that a detected object is a subset of object classes, thereby classifying the object (Redmon et al., 2016). However, as mentioned above, creating datasets for training custom object detection models for this purpose requires extensive time and effort, especially in the context of panelized construction, where panels come in different shapes and dimensions. To facilitate the model-training process, there are a number of large open-source datasets available that can be used for training models to detect a variety of objects, a notable one being the COCO dataset (COCO Consortium, 2022). The COCO dataset contains more than 200,000 labelled images under 91 object classes of objects commonly found in everyday scenes (e.g., people, stop signs, vehicles) (Lin et al., 2014). Given its large library of images of everyday objects, COCO has been used to train and evaluate a number of different open-source object detection models. TiDA leverages such models pre-trained on common objects to study custom objects (e.g., wall frame) that are outside the scope of the COCO dataset and in many cases particular to panelized construction. Specifically, TiDA uses the detection of any common object the user may choose from the COCO dataset's object classes, in addition to the detection of human subjects, to measure process time variables as delineated in the following subsections. In other words, the user does not have to train an object detection model to detect and classify custom objects (such as wall frames and panels) in order to acquire data concerning these custom objects. Instead, the user simply downloads publicly available pre-trained weights (e.g., (Alexey, 2021)) and uses them to run detection of common object. As such, TiDA significantly reduces the set-up requirements (by eliminating the need to train custom object detection models).

3.2.2.2 Start-time and cycle time measurements

For some workstations, such as the two sheathing tables and the multifunction bridge (Type I workstations), the process at that station begins with work-in-process (e.g., a partially completed wall panel) being pulled into the workstation from the upstream station and ends with the work-in-process being pulled by the downstream workstation. For other workstations, such as the framing workstation (i.e., the first station on the wall production line) (Type II workstations), the process involves loading elements (e.g., studs for panel frames) onto the workstation for assembly rather than beginning with work-in-process being pulled from upstream. Either way, work-in-process/material flows into and out of workstations; consequently, certain points along the production line cyclically become blocked with work-in-process/material. As such, if we place common objects (e.g., stop signs) in the field of view at these locations that can be detected by the object detection model, these objects cyclically become blocked with work-in-process/material and then unblocked when work-in-process is pulled to the downstream workstation. (The procedure followed to locate object(s) at a workstation for this purpose is described in Section 3.3.1.)

To demonstrate TiDA's logic, let us first consider the example of the second sheathing table. In Fig. 3-2a, Panel 1 has been completed and is leaving the workstation as Panel 2 is entering the workstation. Two stop signs located at the workstation (Fig. 3-2b) were blocked when Panel 1 was still at the workstation (Fig. 3-2a) and became unblocked when it was pulled downstream. As Panel 2 enters the workstation, the smaller stop sign becomes completely blocked and the larger stop sign becomes partially blocked, as shown in Fig. 3-2c. Hence, the object detection model will only detect both stop signs when there is no work-in-process at the workstation. In this respect, whether there is work-in-process/material at a workstation is linked to whether the object detection model detects or does not detect the objects at the workstation.

This means that, for Type I workstations such as the sheathing tables and the multifunction bridge, where the flow of work-in-process is linked to the start-times and end-times of the processes undertaken at the workstation, a link can be established between the start-times and end-times of the processes and the detection status (i.e., detected versus not detected) of objects at these workstations.

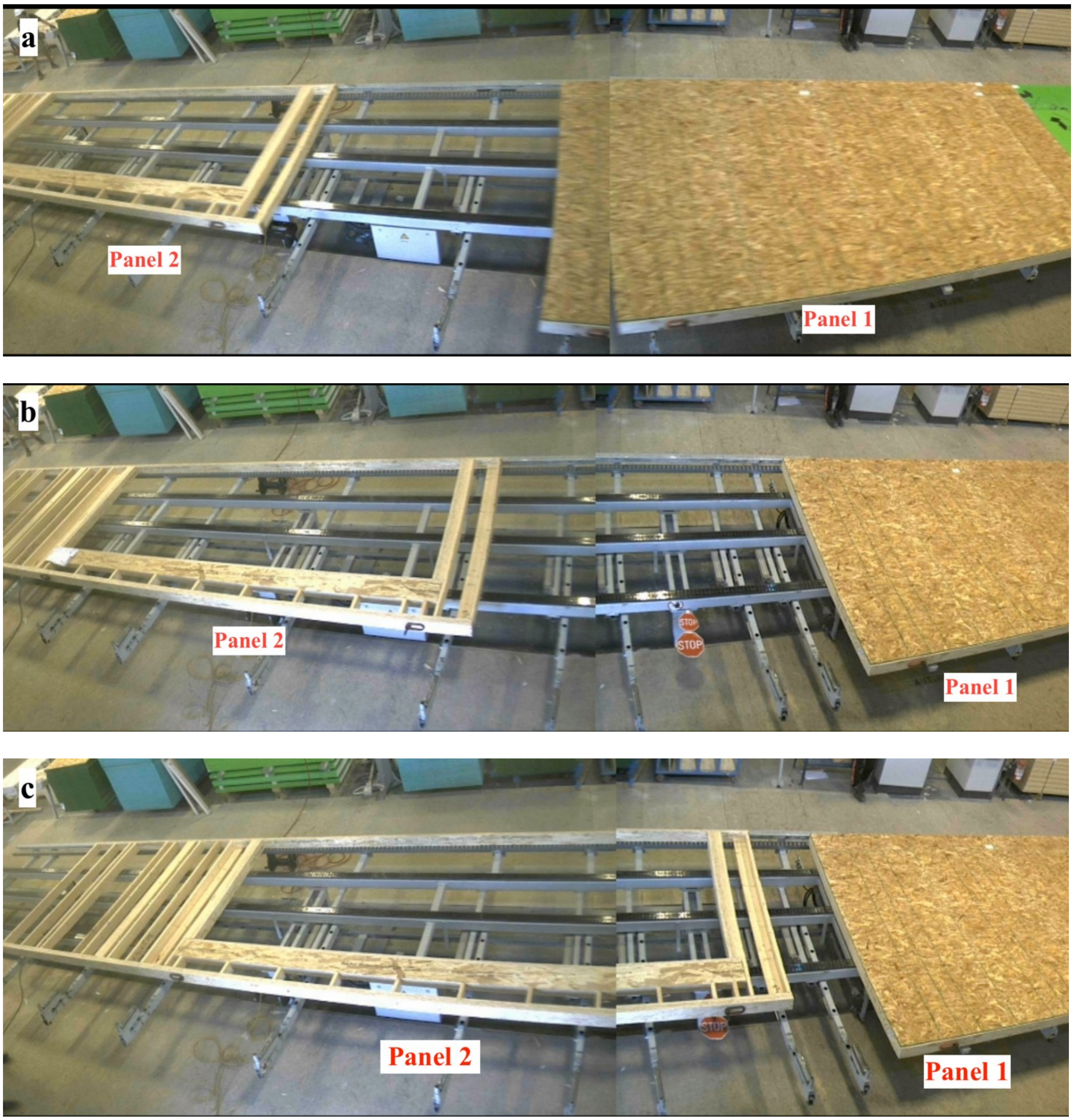


Fig. 3-2. Example of second sheathing table.

As for Type II workstations such as the framing workstation, it may not be possible to directly link the start of the process to the detection status of objects placed at these workstations (although the end of the process is almost always linked to the detection status as work-in-process or completed work flows out of the workstation). Nevertheless, the first task that involves loading material onto the workstation can be linked to the detection status. To clarify the difference between Type I and Type II workstations with respect to detection status, let us consider the example of the framing workstation: the operator starts the framing process by setting up the computer and then loading a top-plate. If stops signs are located where the top-plate is to be loaded, the presence of the top-plate in this location will block the stop signs from view, as shown in Fig. 3-3, and the object detection model will not detect the stop signs until the panel frame is completed and pushed downstream. Hence, the subprocess spanning from the time at which the top-plate is loaded until the end of the framing process can be linked to the detection status of objects at the workstation. Nevertheless, as mentioned above, the operation at this workstation starts with the operator setting up the computer prior to loading the top-plate, and this first subprocess comprising these two tasks of setting up the computer and loading the top-plate cannot be directly linked to the detection status as in the case of the second subprocess. However, if subprocesses that cannot be automatically linked to the detection status are repetitive in nature and show little variability in terms of the time needed to complete them, their completion time can be represented using statistical modelling. This is the case for the framing workstation investigated in this chapter. The cases in which the subprocesses that cannot be automatically linked to the detection status show high variability in terms of completion time, meanwhile, are discussed in Section 3.6.

Fig. 3-3. Framing workstation example.

In general, a process may be divided into multiple subprocesses, if necessary, that are mutually exclusive and collectively exhaustive. Let us consider the following definitions to translate TiDA's logic into mathematical terms:

- ct_i is the cycle time of subprocess i, and $i = 1$ for the first subprocess.
- $DO_{i,j}$ is the number of detected object(s) j linked to subprocess i at the workstation, and $j = 1$ for the first subprocess that can be linked to detection status.
- $t_{DO_{i,j} > LB}$ is the time at which $DO_{i,j}$ is greater than a lower bound, LB.
- $t_{DO_{i,j} < UB}$ is the time at which $DO_{i,j}$ is smaller than an upper bound, UB.
- e_{ct_i} is the statistical estimate of the time needed to complete a subprocess that is statistically modelled as explained above.

As such, *CT* can be computed satisfying Eq. (3.1) and Eq. (3.2):

$$CT = \sum_{1}^{n} ct_i \tag{3.1}$$

$$ct_i = \begin{cases} t_{DO_{i,j} > LB} - t_{DO_{i,j} < UB}, & if\ subprocess\ i\ is\ linked\ to\ detection\ status \\ e_{ct_i}, & if\ subprocess\ i\ is\ statistically\ modelled \end{cases} \tag{3.2}$$

where n is the total number of subprocesses. For Type I workstations, n is equal to 1. As for *ST,* it can be computed satisfying Eq. (3.3).

$$ST = \begin{cases} t_{DO_{1,1} < UB}, & if\ the\ first\ subprocess\ can\ be\ linked\ to\ detection\ status \\ t_{DO_{2,1} < UB} - e_{ct_1}, & if\ the\ first\ subprocess\ is\ statistically\ modelled \end{cases} \tag{3.3}$$

Examples of the use of these equations are presented in Section 3.3.1.

3.2.2.3 Measurement of productive time

For labour-based and semi-automated workstations, productive time is measured based on the detection of workers at the workstation. TiDA assumes that work is being performed whenever a certain number of workers are detected within the workspace allocated to the workstation. The minimum number of workers that need to be detected within the workspace in order for the system to consider that work is being performed is contingent upon the nature of tasks performed at the given workstation. For example, if a specific task cannot possibly be performed by a single worker and necessitates at least two workers, the lower bound can be set at 1 (i.e., when more than one worker is detected at the workstation, work is being performed). However, it is important to note that there are specific steps to consider in relation to camera shot framing when considering this assumption, as described in Section 3.3.2. Based on this assumption, over the span of a subprocess, the total time during which a minimum number of workers are present and detected at the workstation is the productive time, pt_i, of the subprocess. This assumption holds for most cases, but results in error when workers are present at the workstation but not performing any work (interacting casually with co-workers, for example). The resulting error, however, is often offset by the detection model's occasional failure to detect workers performing work (as in the case application presented in the chapter).

On the other hand, this assumption does not hold for fully automated workstations (e.g., the multifunction bridge), as productive time in these workstations is related to machine operation. Alternatively, the logic described for measuring *ST* and *CT* could be adapted to measure *PT* in such cases. For instance, the fully automated machines typically used in panelized construction are computer numerical control (CNC) machines. CNC machines typically start from a "home position" to perform specific tasks and return to the home position after completing the tasks. As such, whether the machine is performing tasks can be linked to the detection status of an object located in the home position of the machine. However, measuring productive time, or machining time, at fully automated workstations is a straightforward task since it is a direct function of the machine's speed and the coded tasks, and as such it can be calculated rather than measured based on image data. As such, this chapter focuses on more complex cases in which the tasks performed involve human input.

3.2.3 Architecture of the TiDA system

TiDA's development is based on the logic delineated in the previous section. It integrates four modules for measuring *PT*, *CT*, and *ST*. The architecture of TiDA is displayed in Fig. 3-4, while the four modules it comprises are described in the subsections that follow. For the purpose of the case application, stop signs are selected as the common object that is detectable at the workstations by the pre-trained object detection model.

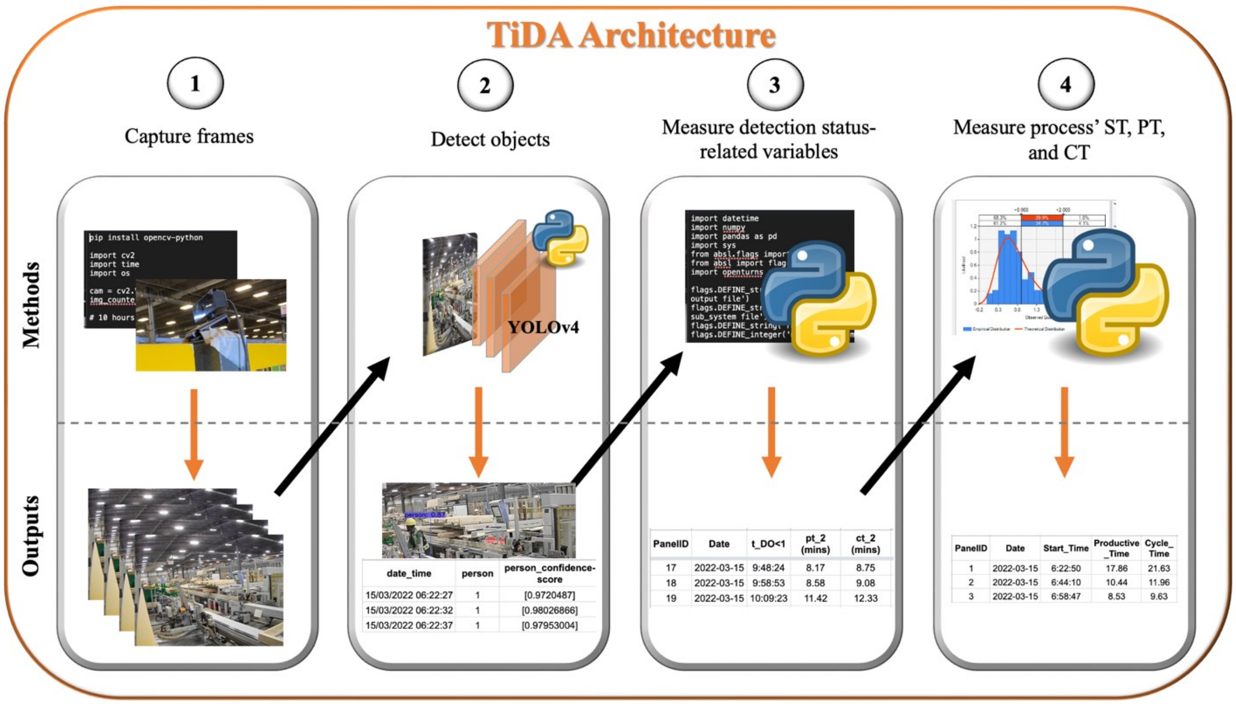


Fig. 3-4. TiDA architecture.

3.2.3.1 Module 1: Capture frames

The function of, input to, structure of, and output of the first module are as follows:

- **Function:** It captures frames (images) of the workstation, from which the time data is extracted.
- **Input:** The input is the real-world characterization of the studied process.
- **Structure:** This module comprises a camera set to capture frames of the workstation and a Python program that captures frames at a time interval or a "frames per second" (fps) rate that is set by the user.
- **Output:** The output of Module 1 is the frames.

3.2.3.2 Module 2: Detect objects

The function of, input to, structure of, and output of the second module are as follows:

- **Function:** It detects workers and stop signs located at the workstation.
- **Input:** The input to Module 2 is the frames from Module 1.
- **Structure:** This module uses the YOLOv4 object detection model to detect workers and stop signs.
- **Output:** The output of Module 2 is a list of the objects detected in the frames, along with their count and the confidence score, date, and time of each detection.

3.2.3.3 Module 3: Measure detection status-related variables

The function of, input to, structure of, and outputs of the third module are as follows:

- **Function:** This module measures $t_{DO_{i,j} < UB}$ and ct_i for each subprocess that can be linked to detection status and the corresponding productive time, pt_i.
- **Input:** The input is the list of stop-sign and worker detections from Module 2.
- **Structure:** This module comprises a Python program that (1) identifies patterns in the stop-sign detections used to measure $t_{DO_{i,j} < UB}$, $t_{DO_{i,j} > LB}$, and, accordingly, ct_i and (2) computes the total time during which the number of workers detected by Module 2 over the span of each subprocess is greater than the lower bound selected by the user for the purpose of defining and measuring pt_i for the given case.
- **Outputs:** The outputs are the measures of $t_{DO_{i,j} < UB}$, ct_i, and pt_i for each subprocess that is linked to detection status.

3.2.3.4 Module 4: Measure the CT, ST, and PT of the process

The function of, inputs to, structure of, and outputs of the fourth module are as follows:

- **Function:** This module measures *CT*, *ST,* and *PT* for the process at the workstation.
- **Inputs:** The inputs are the measures of $t_{DO_{i,j} < UB}$, ct_i, and pt_i from Module 3.

- **Structure:** This module comprises a Python program that approximates (1) the cycle times of subprocesses that cannot be linked to the detection status (e_i), if any, using statistical modelling, (2) *CT* and *ST* using Eq. (3.1), Eq. (3.2), and Eq. (3.3), and (3) PT based on the sum of pt_i. It should be noted that there are a number of different statistical models that can be used to model subprocesses that cannot be linked to detection status. The user can test different methods to identify the one that results in the least modelling error in approximating *CT*, *ST,* and *PT*. Examples of such statistical models are presented in Section 3.3.7.
- **Outputs:** The outputs of this module are the measures of *CT*, *ST*, and *PT*.

3.3 Procedure to deploy TiDA

This section presents the procedure used to apply TiDA to workstations in panelized construction factories. The steps are described mainly in reference to their application to the framing process (since the workstation involved in this process is of Type II workstations, which are more complex than Type I workstations). Occasional references are made to the second sheathing table, a Type I workstation, only to further demonstrate the adaptability of TiDA.

3.3.1 Step 1: Identify the location(s) of stop sign(s) and apply equations

The first step in applying TiDA is to identify (1) the first task that involves loading material/work-in-process onto the workstation and that can be linked to the detection status of the stop sign(s) located at the workstation and (2) the location(s) where the stop sign(s) should be placed in order to establish the link. It should be noted that there may be multiple locations where the stop sign(s) could be placed, and that these locations may be a function of the camera's position. For Type I workstations, this step may be very straightforward. For instance, for the second sheathing table, the process starts with work-in-process being pulled into the workstation and ends with work-in-process being pushed to the downstream workstation. Hence, a stop sign could be placed horizontally anywhere under the panel if the camera captures frames from above the workstation (as shown in Fig. 3-2 above) or vertically on one side of the table if the camera is positioned to capture frames horizontally from the opposite side. Moreover, the default number of stop signs to be located at the workstation is one, but more than one stop sign may be needed at certain workstations. The need for more than one stop sign can be easily determined by placing one stop sign and observing the process to confirm whether a link between the process and status detection has been established. For the second sheathing table, for instance, two stop signs are needed because of the variations in panel height. When the process starts, work-in-process blocks at least one of these two signs (i.e., $DO_{1,1} < 2$), whereas, when the process ends, both stop signs become unblocked (i.e., $DO_{1,1} > 1$). As such, TiDA's logic is applied to this case satisfying Eq. (3.4) and Eq. (3.5):

$$CT = t_{DO_{1,1}>1} - t_{DO_{1,1}<2} \tag{3.4}$$

$$ST = t_{DO_{1,1}<2} \tag{3.5}$$

As for productive time, it is equal to the total time during which at least one worker is detected at the workstation between $t_{DO_{1,1}>1}$ and $t_{DO_{1,1}<2}$. For Type II workstations, this first step in applying TiDA may be less straightforward, but it is important to note that it is the only ad hoc step although the logic underlying the step is generic. For the framing workstation, the first task that can be linked to detection status is the task of loading the top-plate (as explained above). Locations A1 and A2, highlighted in Fig. 3-5a, become blocked when the top-plate is properly loaded (Fig. 3-5b). The reason for using two locations instead of just one in this case has to do with the end-time of the process. Locations A1 and A2 become unblocked when the operator completes the framing process, cuts and removes the excess portion of the top-plate (if any), and pushes the framed panel downstream, as shown in Fig. 3-5c. However, the operator may in some cases complete the framing process but not directly push the framed panel downstream, such that Location A1 becomes unblocked but Location A2 remains blocked, as shown in Fig. 3-5d. In a similar manner, the operator may sometimes push the framed panel downstream but not immediately remove the excess portion of the top-plate from the table, in which case Location A2 becomes unblocked but Location A1 remains blocked. Nevertheless, since the specific times at which the two locations become unblocked are only seconds apart under normal operations, the end-time of the framing process can be considered to be approximately the same as the time at which Locations A1 or A2 become unblocked. Given this, and as explained above, the framing process can thus be divided into two subprocesses, where the first subprocess is statistically modelled and the second subprocess is linked to the detection of the stop signs. In such cases, TiDA's logic is applied satisfying Eq. (3.6) and Eq. (3.7):

$$CT = ct_1 + ct_2 = e_{ct_1} + (t_{DO_{2,1}>0} - t_{DO_{2,1}<1}) \tag{3.6}$$

$$ST = t_{DO_{2,1}<1} - e_{ct_1} \tag{3.7}$$

As for productive time, the user may set it to be equal to the time during which at least one worker is detected at the workstation between ST and $t_{DO_{2,1}>0}$ or, alternatively, the user may approximate the productive time associated with the first subprocess using statistical modelling and add it to the time during which at least one worker is detected at the workstation during the second subprocess (i.e., between $t_{DO_{2,1}<1}$ and $t_{DO_{2,1}>0}$).

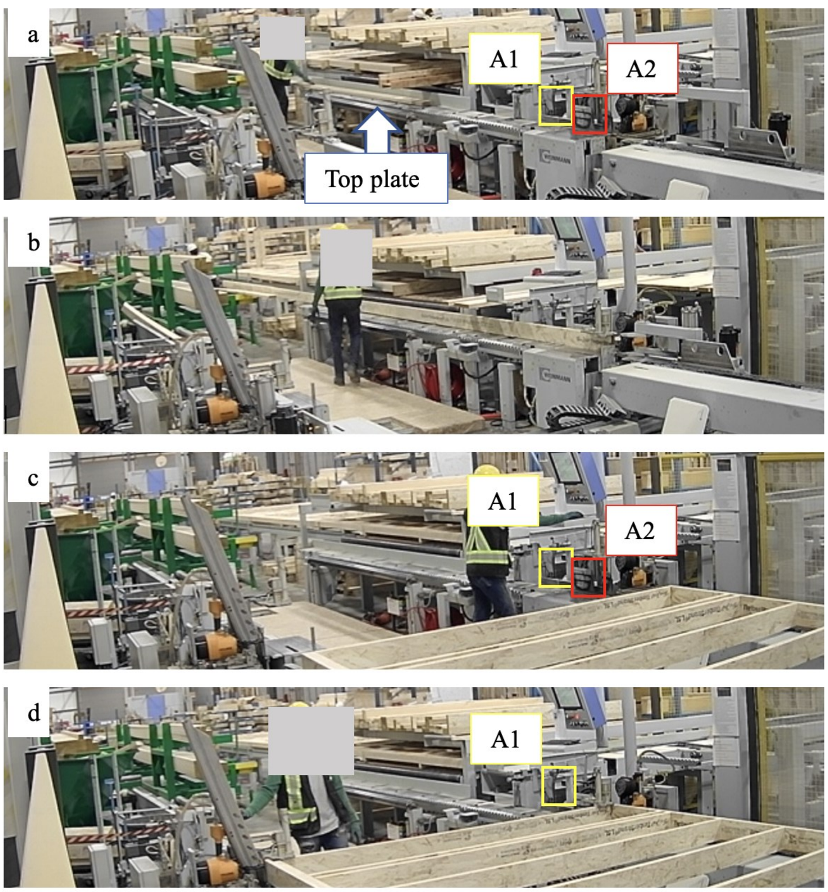


Fig. 3-5. (a) Top-plate loading; (b) Locations blocked; (c) Locations unblocked; (d) Location A1 unblocked and Location A2 blocked.

3.3.2 Step 2: Position and set up the camera

Two elements must be captured in the frames for object detection: (1) the installed stop signs, and (2) the particular workspace within the workstation where workers perform tasks. The user may choose to use one camera to capture both elements or to use one camera for each element, depending on the nature of the workstation and user preference. It may be simpler in the case of a large workstation to use two cameras, although one camera should be sufficient if it has a wide lens and can be installed above the workstation. If a single camera is used in the case of a large workstation and the stop signs are small relative to the size of the workstation, the Python program in Module 1 can be adjusted to pre-process the captured frames and create cropped copies that focus on the stop signs. The camera(s) must be firmly attached to avoid any vibrations that could result in blurry images and thereby adversely affect the performance of the object detection model—a limitation that has been reported in a number of studies applying computer vision to offsite construction (Martinez, Ahmad, et al., 2019; Wang et al., 2021b; Zheng et al., 2020). Given that tasks in offsite construction are performed indoors under artificial lighting (Fig. 3-6) and that reflective surfaces are not typically found at workstations, the lighting issues reported in some computer vision applications (e.g., (Liu et al., 2019; Wang et al., 2021a)) are not likely to be encountered at workstations in wood panelized construction factories.

For the workspace, the camera should be positioned in such a way that it only captures the workers assigned to the workstation and does not detect other workers who may appear in the background. In this regard, it may be advisable to position the camera directly above the workstation under study. It may also be feasible to use existing CCTV cameras, as in the case of the second sheathing table at the case company. Regardless, if background workers do appear in the frames, it should be determined whether the object detection model is capable of detecting them. If the model fails to detect them, no further action is required (as in the case application described herein). If, on the other hand, it is determined that the model is capable of detecting background workers, the Python program in Module 1 can be adjusted to automatically mark regions of interest, that exclude irrelevant spaces appearing in the frames, for object detection.

Once the camera is positioned, the user sets the fps rate. Typically, the higher the fps rate is, the more information can be extracted from the frames. However, to minimize storage requirements and reduce the time needed to run Module 2, the fps rate should be minimized while ensuring that the time gap between frames is short enough that the system does not miss steps being performed by the worker. A similar logic applies to frame resolution: the resolution should be just high enough not to adversely affect the performance of the object detection model, since high frame resolution is correlated with excessive data storage requirements.

For the framing workstation at the case company, since there was no existing CCTV camera above the workstation, a full HD 1080P webcam set to capture frames of the workstation was installed as shown in Fig. 3-6. This camera features a lens that can capture sharp frames at fixed distances and that is best suited for distances ranging from 0.5 m to approximately 4 m. Moreover, the fps rate was set to 0.2 fps, meaning that frames are captured at five-second intervals. Finally, the resolution of the frames was set to 640 × 480 pixels.

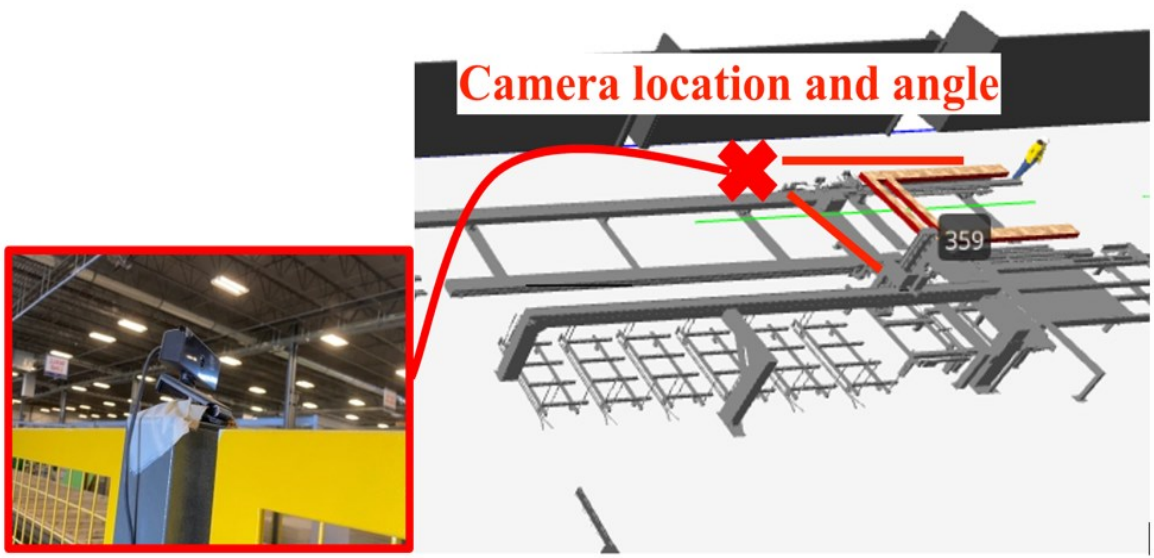


Fig. 3-6. Camera installed for the framing workstation.

3.3.3 Step 3: Select a confidence score threshold

When the model detects and classifies an object, it reports a confidence score that signifies the confidence of the model in predicting the correct output for the given detection. In other words, the higher the score is, the higher the probability that the given detection is correct. Consider for example the detection of the operator in Fig. 3-7. The model correctly classifies the detected operator as a person, and the confidence score is high (87%). On the other hand, the model incorrectly detects a portion of the machine and classifies it as a person. The corresponding confidence score is low (42%). A threshold value is typically set in order to filter out any detections that have low confidence scores and therefore are more likely to be incorrect (e.g., Fig. 3-7). Nevertheless, some correct detections may have relatively low confidence scores and therefore could be unwittingly filtered out if the threshold is set too high. This trade-off must be accounted for when setting the threshold value. For the framing workstation, a threshold value (of 0.4 as later explained) was set by trial and error with the aim of minimizing the combined effect of missed detections and false detections on the performance of TiDA.

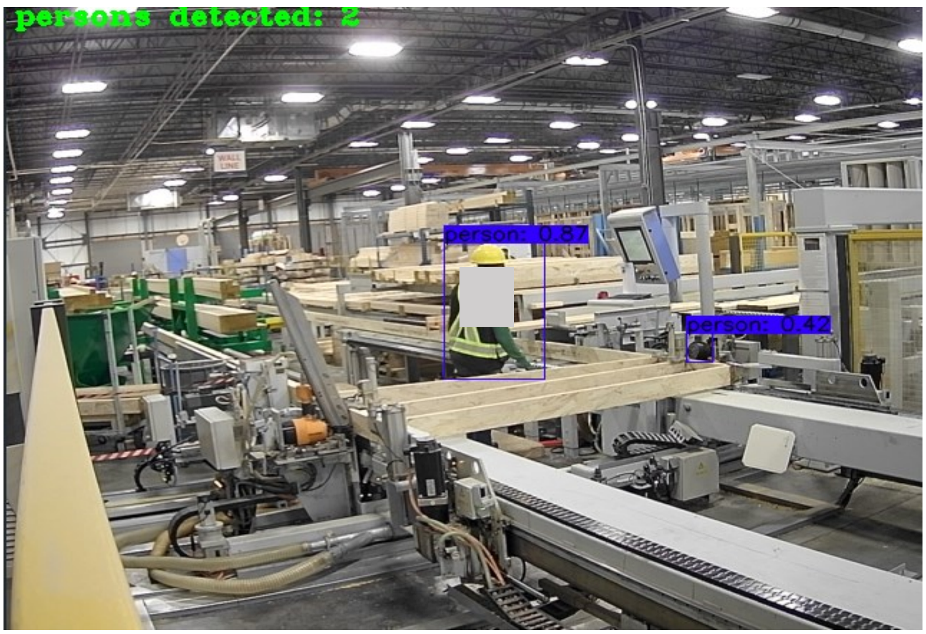


Fig. 3-7. Correctly detecting the operator as a person but incorrectly identifying a machine part as a person.

3.3.4 Step 4: Select the size of the stop sign(s)

The stop signs must be large enough to be detected by the model when they are unblocked but small enough to be sufficiently covered by work-in-process/material. For example, in the case of the framing workstation, the stop signs should be sufficiently covered by the smallest size of the top-plates (which come in different sizes). Generally, for a given frame, the smaller the object we are interested in detecting is, the harder the detection becomes, and this may result in a lower confidence score. Consequently, reducing the size of the stop signs may increase the risk of the corresponding detections being rejected due to low confidence scores. Using a small enough size of stop sign that it can be completely blocked by work-in-process/material is not necessary, however. As long as work-in-process/material hides a certain proportion of the stop sign, the model will not detect the visible portion. Given these considerations, a trial-and-error method can be used to select the optimal size of stop signs digitally. In other words, stop signs of different sizes can be digitally added to the frames and experimented with, thereby eliminating the workflow interruptions that would result from physical experimentation at the workstation. It should be noted that “optimal size” refers to the size or range of sizes that cannot be detected when there is work-in-process/material but that can be detected (with a confidence score that always satisfies the threshold value) when there is no work-in-process/material. As shown in Fig. 3-8, the model did not detect the partially hidden digital stop signs even in the case of small top-plates. Similarly, when stops signs were installed at the framing workstation, the model detected them when they were unblocked and did not detect them when they were partially blocked as shown in Fig. 3-9.

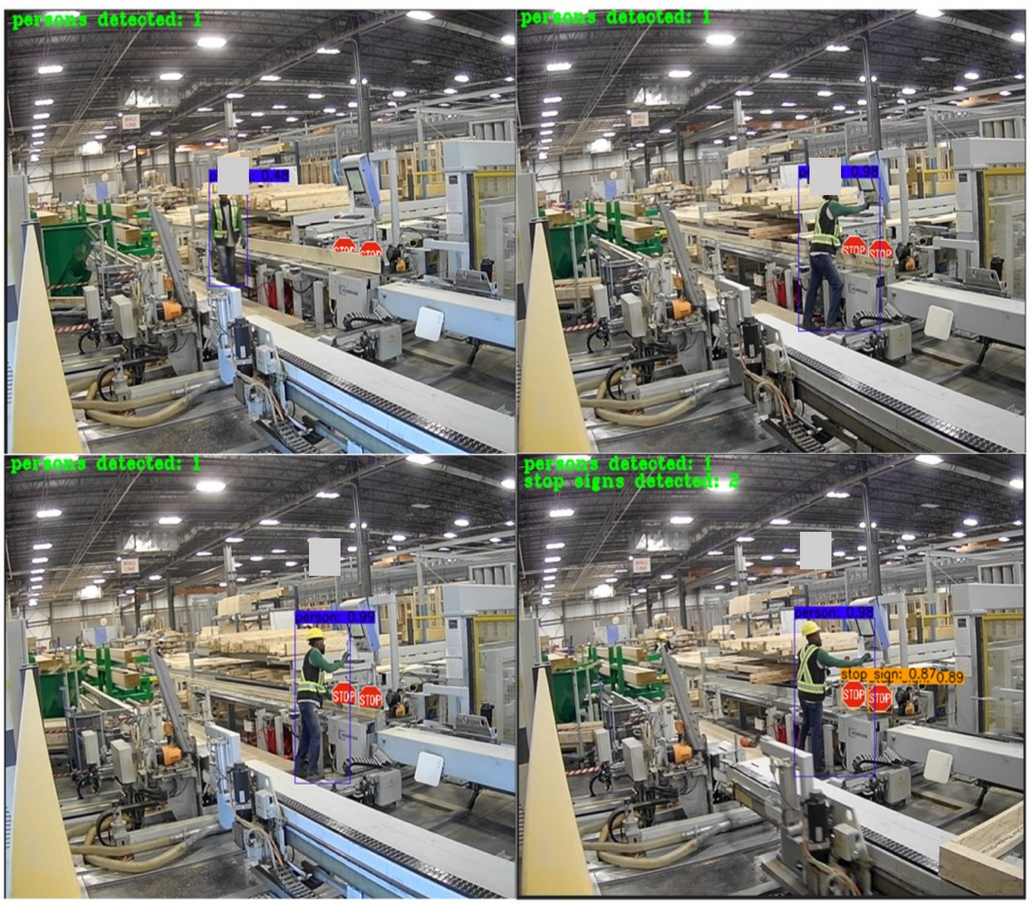


Fig. 3-8. Stop sign detection digital testing.

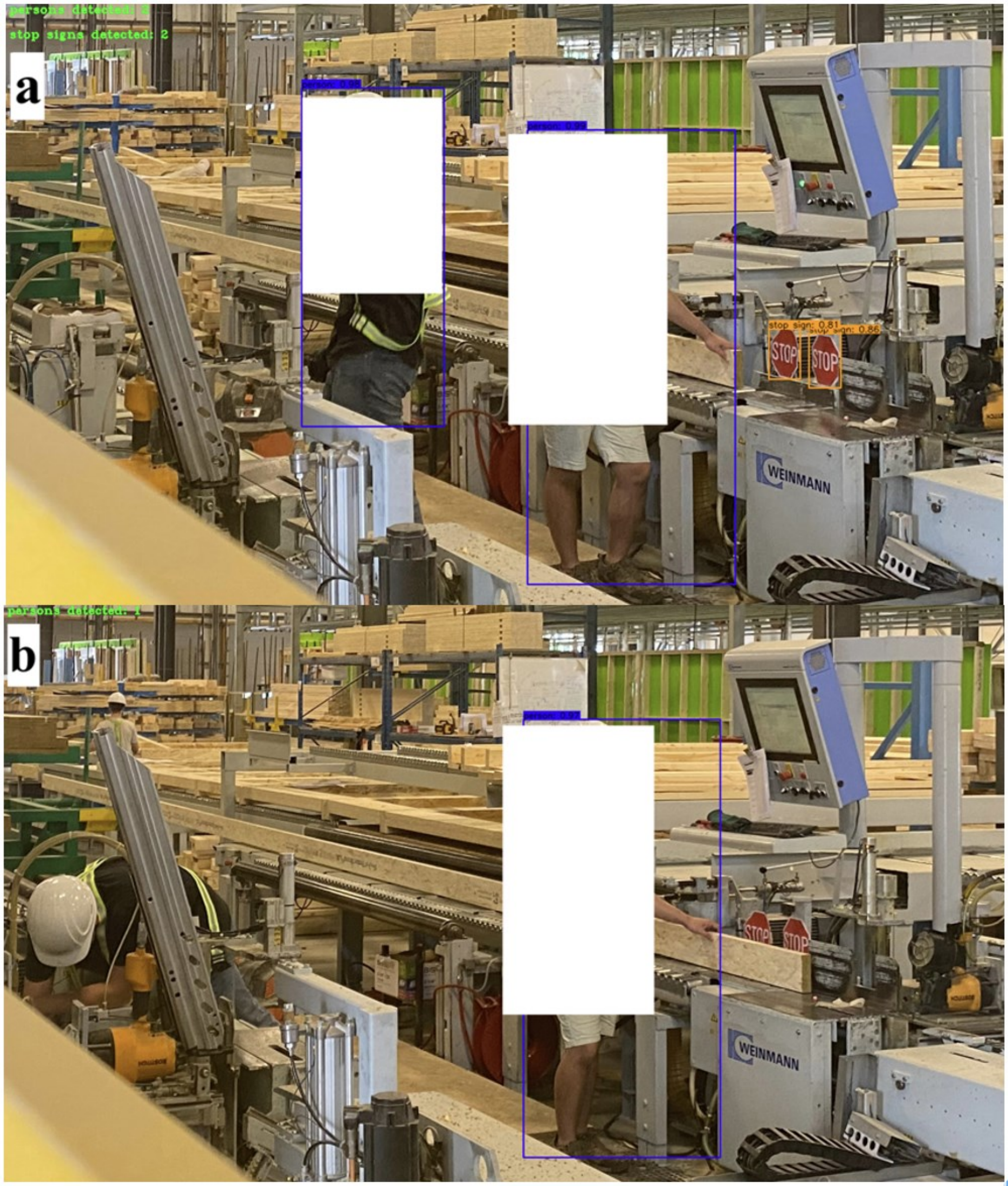


Fig. 3-9. (a) Stop signs unblocked; (b) stop signs blocked.

3.3.5 Step 5: Select performance evaluation metrics

The overall performance of TiDA is governed by the respective performance of the four modules it comprises. Selecting the proper performance metrics based on the intended purpose of the system is essential for a sound evaluation of the system. The metrics used for evaluating the performance of the modules for the framing workstation are described below. The user may select any other performance metrics depending on the intended use of TiDA.

3.3.5.1 Detection/classification-related performance (Module 2) and quality of frames (Module 1)

The quality of detections and classifications is a function of the performance of the object detection model and the quality of the dataset used. As such, to evaluate the performance of the object detection model embedded in Module 2 and the quality of the frames captured using Module 1, the following metrics were used to assess the quality of detections and classifications:

- True positive (TP): TP occurs when the model identifies a positive class when the truth is positive, e.g., the model detects a worker (or stop signs) in frames in which the worker (or stop signs) is (are) visible.
- True negative (TN): TN occurs when the model identifies a negative class when the truth is negative, e.g., the model does not detect a worker (or stop signs) in frames in which the worker (or stop signs) is (are) not visible.
- False positive (FP): FP occurs when the model identifies a positive class when the truth is negative, e.g., when the model classifies a non-person as a person.
- False negative (FN): FN occurs when the model identifies a negative class when the truth is positive, e.g., a model fails to detect a worker (or stop signs) in frames in which the worker (or stop signs) is (are) visible.
- Accuracy: Accuracy is the ratio of correct predictions (TP+TN) to total predictions (TP+TN+FP+FN).
- Precision: Precision is the ratio of correct positive predictions (TP) to total positive predictions (TP+FP).
- Recall: Recall is the ratio of correct positive predictions (TP) to the total number of positive cases, including those correctly predicted and missed by the model (TP+FN).
- F1-score: F1-score combines the precision and recall metrics and is equal to their harmonic mean, calculated satisfying Eq. (3.8):

$$2 \times (precision \times recall)/(precision + recall) \tag{3.8}$$

3.3.5.2 Evaluation of detection-based measures (Module 3) and overall performance (Module 4)

To evaluate the performance of Modules 3 and 4, the following metrics were used:

- Measurement error (ME) calculated satisfying Eq. (3.9). This metric was used for assessing whether Modules 3 and 4 tended to overapproximate (positive error value) or underapproximate (negative error value) true values and for determining the degree of variance of the measures from their true values for each panel.

$$ME_i = M_i - A_i \tag{3.9}$$

where ME_i is the error corresponding to panel i, M_i is the measure generated by Module 3 or Module 4 for panel i, and A_i is the actual corresponding to panel i (collected by manually checking the frames).

- The sum of measurement errors (SME) calculated satisfying Eq. (3.10). Since Modules 3 and 4 may overapproximate or underapproximate true values, some of the errors may cancel one another out when the total durations needed to frame a batch of panels are being measured. As such, the sum of these errors allows determining whether the total durations corresponding to a batch of panels are overapproximated or underapproximated.

$$SME = \sum_{1}^{n} ME_i \tag{3.10}$$

where n is the total number of panels used for evaluation.

- The mean absolute error (MAE) calculated satisfying Eq. (3.11). This metric allows determining the average degree of variance of the measures generated by TiDA for each panel from their true values, regardless of whether they had been overapproximated or underapproximated.

$$MAE = \frac{\sum_1^n |ME_i|}{n} \tag{3.11}$$

3.3.6 Step 6: Select sample size for the distribution-fitting dataset (if applicable) and testing dataset

3.3.6.1 Distribution-fitting dataset (if applicable)

This step is only required for Type II workstations when a subprocess is statistically modelled. The size of the dataset used for fitting statistical distributions to model a subprocess affects the choice of the distributions that would best represent reality. This is because a large sample may include certain ranges of data points that are not captured in a smaller sample. Consider, for example, one day of operations during which forty panels are scheduled for manufacturing. If data collection starts in the morning and stops when data on twenty panels is collected, the sample will not include data on panels framed in the afternoon. If, for some reason, the operator takes more time to complete the tasks in the afternoon, the collected sample will not be representative of reality for the entire day of operations. In technical terms, a representative sample reflects the true characteristics of the population from which it originates.

Statistical modelling was used to approximate the cycle time (ct_1) and productive time (pt_1) corresponding to the first subprocess at the framing workstation. Accordingly, statistical distributions were fitted using Simphony.NET software (Engineering at Alberta, 2022). The (1) least squares (Wolfram MathWorld, 2022), (2) maximum likelihood (Pearson, 1936), and (3) moment matching (Pearson, 1936) estimation methods were considered, as shown in Fig. 3-10. The Kolmogorov-Smirnov (K-S) test was conducted in Simphony.NET to evaluate the goodness-of-fit of each statistical distribution (i.e., to determine how well the distribution fits the data). The K-S test, it should be noted, is based on the maximum difference between the empirical and theoretical cumulative distributions (i.e., the K-S statistic shown in Fig. 3-10) (Massey, 1951). Based on the results of the three fitting methods, the distribution resulting in the least K-S statistic was deemed to be the distribution having the best goodness-of-fit.

Distribution	K-S
Gumbel	0.08375
Laplace	0.09938
Logistic	0.10009
Cauchy	0.11205
Pearson5	0.11529
Gamma	0.11777
Beta	0.14436
Normal	0.17537
Exponential	0.26554
Weibull	0.41524
Triangular	0.45462
Uniform	0.60287
ChiSquare	0.61854

Least Squares | **Maximum Likelihood** | Moment Matching

Fig. 3-10. Fitting results using Simphony.NET.

In order to select a reasonably representative sample size, experiments were conducted to test the effect of increasing the sample size on the type of distribution having the best goodness-of-fit. In the first experiment, a sample size of 50 was selected, and the distributions with the best goodness-of-fit for ct_1 and pt_1 were identified accordingly. The sample size was then increased by an increment of 20 data points in each successive experiment.

The results corresponding to ct_1 (see Table 3-1) demonstrate clearly the significance of the sample size. For 50 data points, the distribution with the best goodness-of-fit was found to be the LogLogistic distribution. However, adding only 20 additional data points resulted in the Gamma distribution having the best goodness-of-fit, while adding another 20 data points resulted in the Pearson Type V (Pearson5) distribution having the best goodness-of-fit. As shown in Fig. 3-11, the probability density functions (PDFs) of these distributions have different shapes; consequently, datasets sampled from these distributions would have different characteristics.

Table 3-1. Sample size testing for distribution-fitting.

Sample Size	Actual pt_1				Actual ct_1		
	Best Fit Distribution	Parameters			Best Fit Distribution	Parameters	
		Shape	Scale	Location		Shape	Scale
50	LogLogistic	2.44	0.79	-	LogLogistic	1.97	1.00
70	Gumbel	-	0.40	0.61	Gamma	1.84	0.63
90	Gamma	2.21	0.36	-	Pearson5	1.60	0.99
110	Gumbel	-	0.39	0.56	Gamma	1.79	0.58
130	Gumbel	-	0.37	0.56	Pearson5	1.79	1.25
150	Gumbel	-	0.35	0.57	Pearson5	1.95	1.37

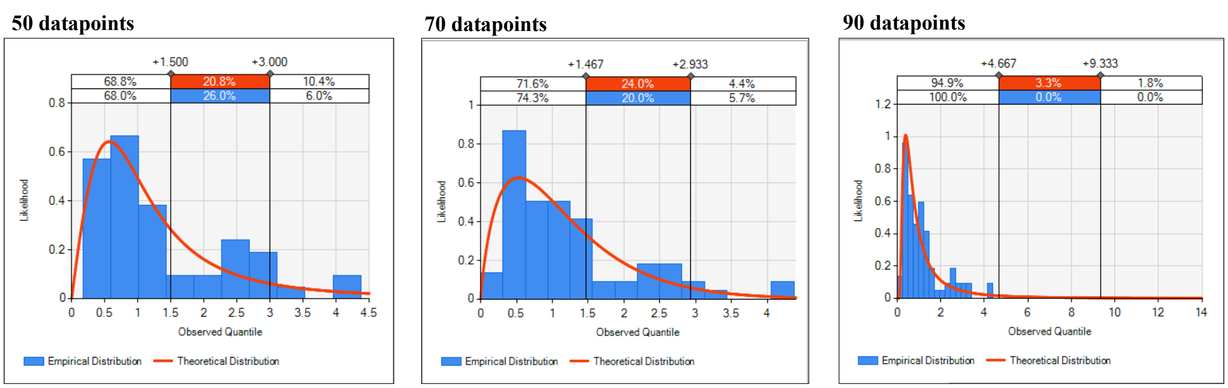


Fig. 3-11. Effect of additional data on distribution-fitting.

After reaching a sample size of 130, the addition of data points no longer had a significant effect on the shape of the datasets or on the results of distribution-fitting for either pt_1 (Fig. 3-12) or ct_1 (Fig. 3-13), notwithstanding some minor variations in the parameters of the distributions with the best goodness-of-fit. As such, a sample size of 150 data points was selected for the statistical modelling of pt_1 and ct_1.

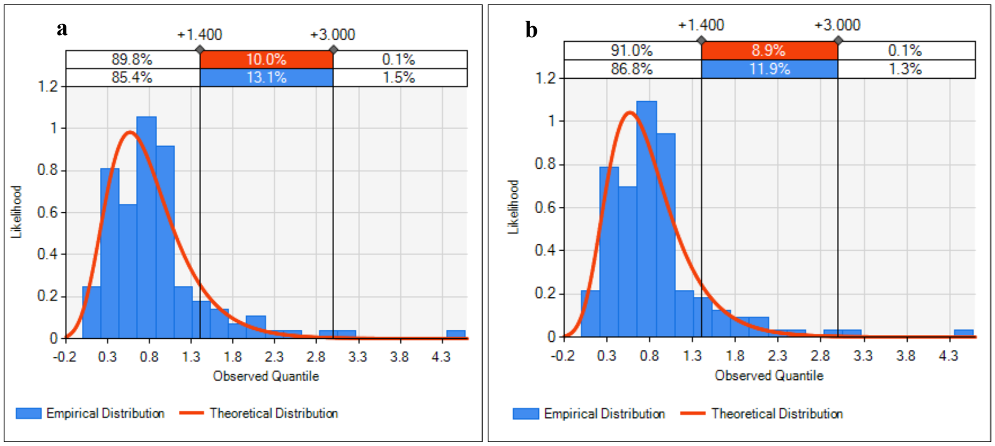


Fig. 3-12. (a) pt_1 Distribution for a sample size of 130; (b) pt_1 Distribution for a sample size of 150.

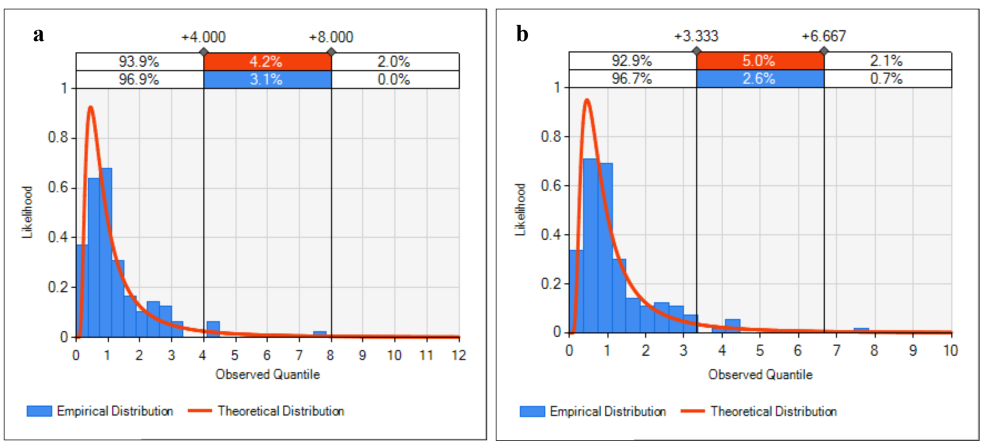


Fig. 3-13. (a) ct_1. Distribution for a sample size of 130; (b) ct_1. Distribution for a sample size of 150.

3.3.6.2 Testing dataset

Following the same rationale discussed with respect to the distribution-fitting dataset, it is important to select a sample representative of a wide variety of panel designs and corresponding process durations for testing the performance of the system. Two factors were considered in determining the sample size for testing the performance of the system:

1. Statistical characteristics of the samples: To determine this factor, based on consultation with workers at the case company, a range of data points could be pre-set representing the minimum, maximum, and most likely values of *PT* and *CT*. Boxplots of the collected *PT* and *CT* (samples provided in Section 3.5) could be developed and compared to the workers'

estimates in order to determine which samples could be tentatively considered to be representative of reality.

2. Variability in the size of errors: Error metrics could be measured and plotted against sample sizes in order to find the sample size beyond which the effect of increasing its value on the error metrics becomes negligible. For the framing workstation for instance, MAE and SME were measured and plotted, along with the moving average (which plots the average of each pair of successive points), based on different sample sizes of testing datasets, as shown in Fig. 3-14 and Fig. 3-15. As shown in Fig. 3-14, the MAE was found to decrease with an increase in sample size before stabilizing at approximately 90 data points. Generally, the effect on the MAE of additional data points was negligible. On the other hand, the SME for *PT* continued to increase, while that for *CT* continued to decrease, with an increase in sample size. The change in the SME was deemed to be normal, as this metric is highly sensitive to outliers. For instance, an ME of −20 minutes for one panel increases the size of the system's underapproximation by additional 20 minutes. Given the sensitivity of the SME metric, the selection of the sample size was based on the MAE metric, and a sample size containing 121 panels framed on four different days was deemed to be sufficient for testing the performance of TiDA.

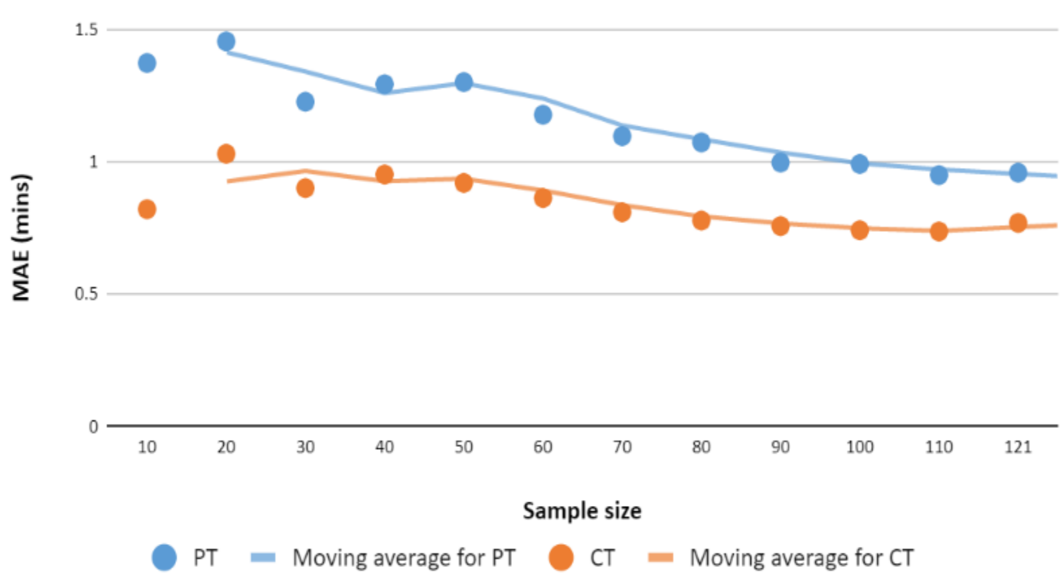


Fig. 3-14. MAE versus sample size.

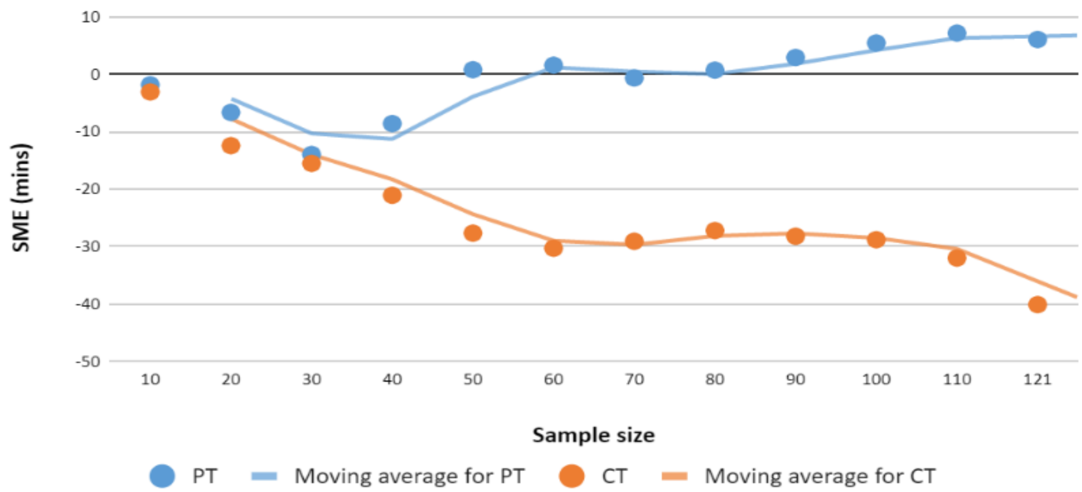


Fig. 3-15. SME versus sample size.

3.3.7 Step 7 (if applicable): Select of the best-performing statistical modelling method

As mentioned above, different statistical models could be used to model subprocesses that cannot be linked to the detection status if there are any. The first step in selecting a statistical model was to select the performance metric(s) based on which the selection would be made. As noted above, the selection metrics should be determined in consideration of the intended purpose of the system under study. For instance, MAE is a suitable metric if the aim is to measure *PT* and *CT* for each panel individually, whereas SME is more suitable if the target is cumulative measures for batches of panels, and a combination of both MAE and SME is appropriate if both individual and cumulative measures are of interest. For the framing workstation, pt_1 was measured for 120 different panels, and the collected data points were used for fitting distributions for the pt_1. The same was done for ct_1. Then, pt_1 and ct_1 were measured for 30 additional panels for testing and used to compare the performance of five statistical models. The manner in which each of the methods was implemented is described below:

- Method 1: Use the averages of the datasets to model pt_1 and ct_1. The mean values of the 120 data points were computed, and approximations of pt_1 and ct_1 were set to be equal to these mean values (0.85 minutes for pt_1 and 1.19 minutes for ct_1) for all 30 panels.

- Method 2: Find the statistical distributions that best fit the datasets and use their mean values to model ct_1 and pt_1. The statistical distributions having the best goodness-of-fit for pt_1 and ct_1 were found to be the Gumbel distribution and the Pearson5 distribution, respectively. Fig. 3-16 and Fig. 3-17 show the corresponding PDFs and empirical distributions of the actual values of pt_1 and ct_1, respectively. Approximations of pt_1 and ct_1 were then set to be equal to the means of these distributions (0.76 minutes for pt_1 and 1.58 minutes for ct_1) for all 30 panels.

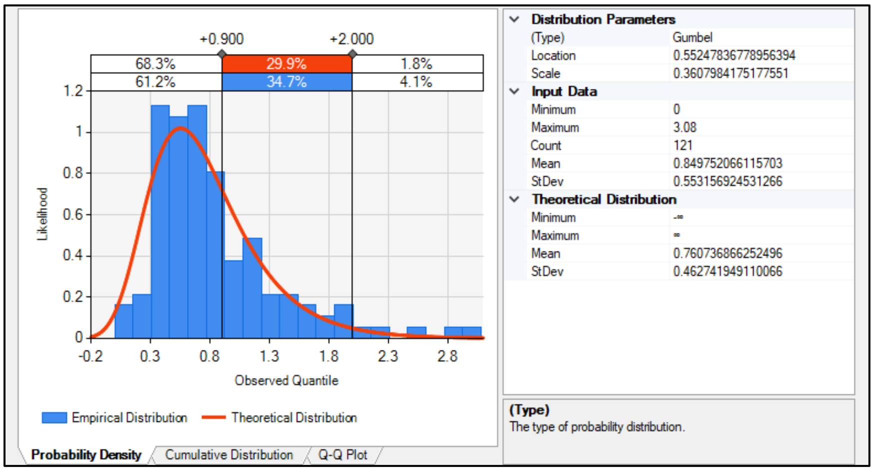


Fig. 3-16. Best fit distribution for pt_1.

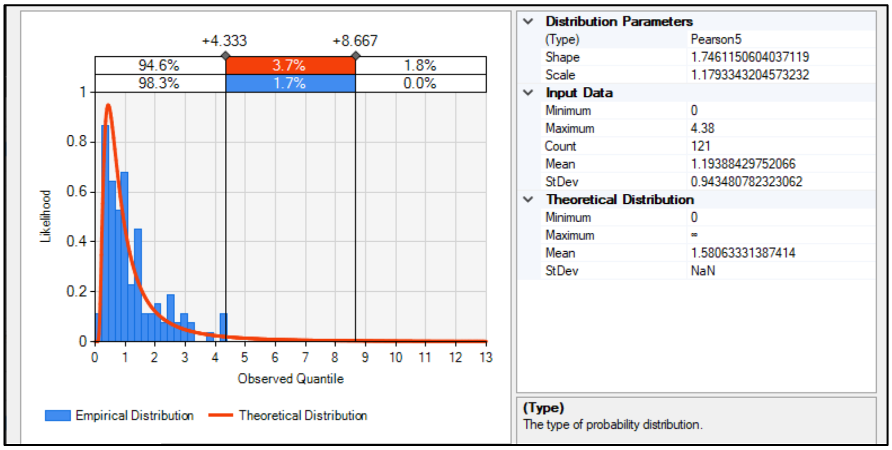


Fig. 3-17. Best fit distribution for ct_1.

- Method 3: Use the statistical distributions that best fit the datasets to randomly sample only positive instances from these distributions to model pt_1 and ct_1. The Gumbel and Pearson5 distributions were used to randomly sample instances of pt_1 and ct_1 for each of the 30 panels. Since the Gumbel distribution having the best fit could take negative values, as shown in Fig. 3-16, only positive instances were sampled from this distribution. This was done by rejecting negative instances and re-sampling until a positive instance was obtained.

- Truncate the statistical distributions that best fit the datasets (used in Method 2) and randomly sample instances from these truncated distributions to model pt_1 and ct_1. As explained above, the Gumbel distribution could assume negative values. Moreover, the Pearson5 distribution shown in Fig. 3-17 is skewed to the right and unbounded, and this could result in some extreme instances of ct_1 being sampled, in turn increasing the magnitude of error. The Gumbel distribution is also skewed to the right in this application, but, since it quickly approaches zero, extreme positive instances are less likely. As such, the distributions were truncated to avoid sampling extreme positive values in the case of ct_1 and negative values in the case of pt_1 while covering the range of collected data. The truncated PDFs for pt_1 and ct_1 are displayed in Fig. 3-18 and Fig. 3-19, respectively.

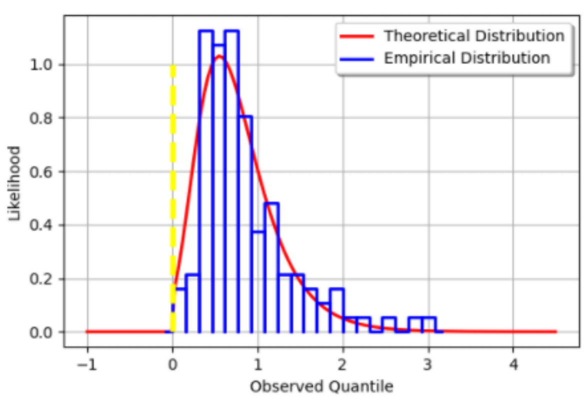


Fig. 3-18. Truncated Gumbel for pt_1.

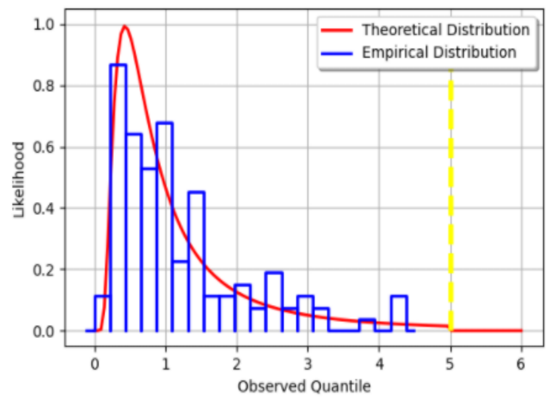


Fig. 3-19. Truncated Pearson5 for ct_1.

- Method 5: Use bounded distributions with good fit and randomly sample instances from these distributions to model pt_1 and ct_1. The goodness-of-fit measures were found to be acceptable for the bounded Beta and Triangular distributions used to model pt_1 and ct_1, respectively. The corresponding PDFs are displayed in Fig. 3-20 and Fig. 3-21. These distributions were used to randomly sample instances of pt_1 and ct_1 for each of the 30 panels.

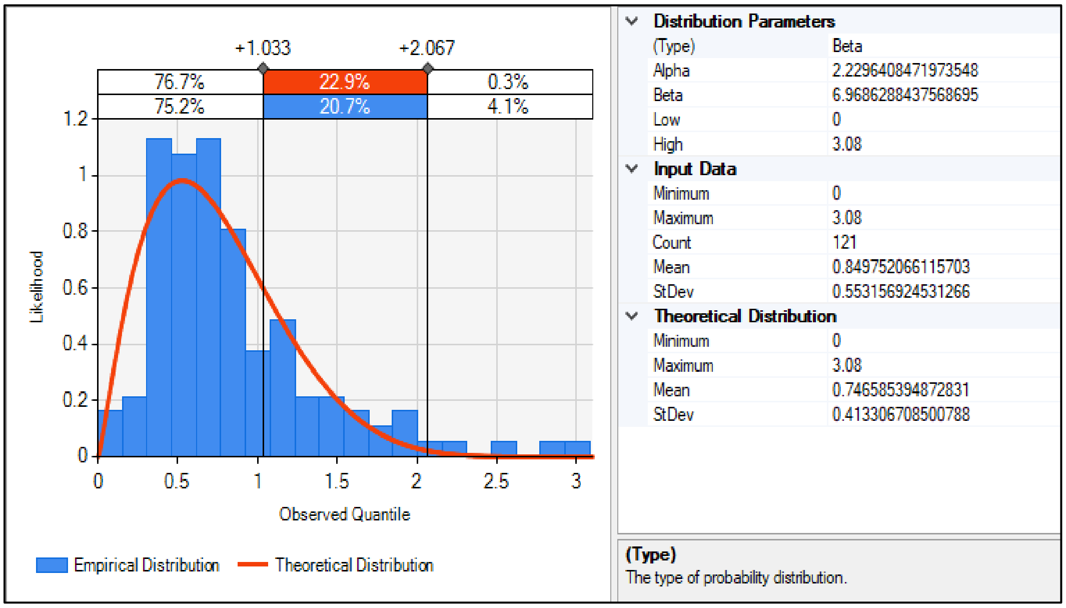


Fig. 3-20. Beta distribution for pt_1.

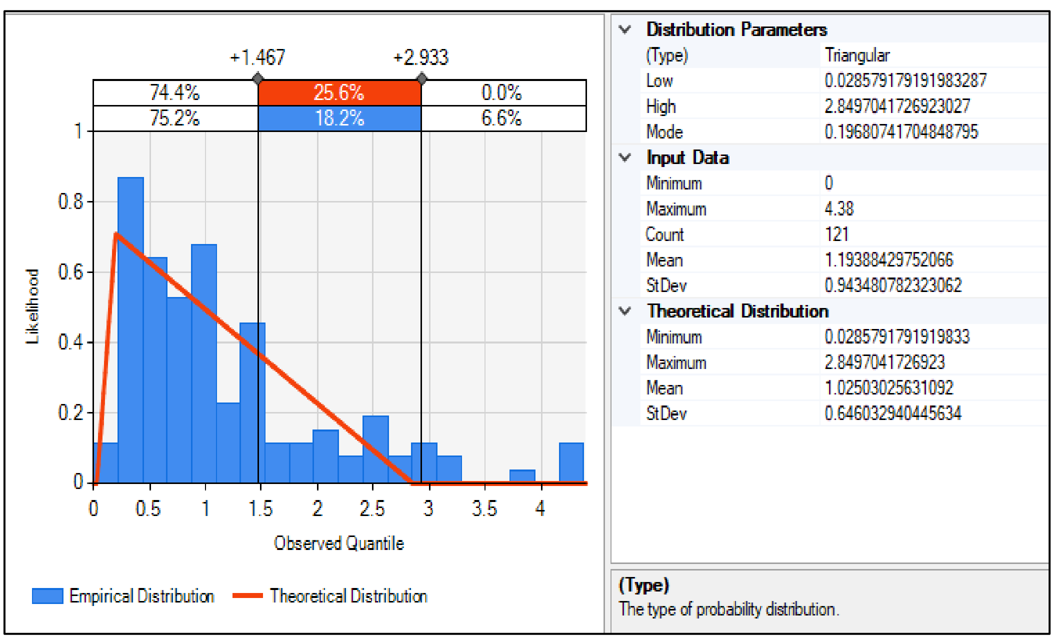


Fig. 3-21. Triangular distribution for ct_1.

The effects of each method on the errors measured for the *PT* and *CT* were also examined given that the overall performance depends on the interaction between Module 3 and Module 4; specifically, for the framing workstation, if Module 3 underapproximates ct_2 (i.e., the cycle time of the second subprocess which can be linked to object detection) while the statistical model overapproximates ct_1, the error corresponding to CT will be smaller (since the errors offset one another). Hence, it was essential to assess the global effect of each method on the overall system performance. The MEA and SME measured based on 30 testing data points and considering the different methods are summarized in Table 3-2, with the smallest error values for each variable highlighted in the table. This selection process was repeated using different fitting and testing datasets (drawn from the same set of 150 data points) to ensure that the selection of the best method would not be sensitive to the samples used for fitting and testing. The results are summarized in Table 3-3. As can be seen, changing the fitting and testing datasets did affect some of the methods, as the changes in the highlighted cells indicate. (This effect would be expected to decrease if larger sample sizes were to be used for the testing datasets.)

Table 3-2. Performance summary for the five modelling methods – First run.

Method	**MAE (minutes)**				**SME (minutes)**			
	pt_1	PT	ct_1	CT	pt_1	PT	ct_1	CT
1	0.39	0.80	0.76	0.82	−1.93	0.66	−2.58	−11.98
2	0.39	0.80	0.99	0.90	−4.60	−2.01	9.02	−0.38
3	0.57	0.94	0.87	1.10	−3.51	−0.93	−16.78	−26.18
4	0.54	0.99	0.87	1.15	−3.02	−0.44	−13.12	−22.52
5	0.59	0.91	0.87	1.00	−10.03	−7.45	−7.35	−16.75

Table 3-3. Performance summary for the five modelling methods – Second run.

Method	**MAE (minutes)**				**SME (minutes)**			
	pt_1	PT	ct_1	CT	pt_1	PT	ct_1	CT
1	0.49	1.23	0.82	0.90	−2.78	−15.53	−4.83	−15.89
2	0.49	1.27	0.90	0.87	−5.24	−17.99	1.31	−9.75
3	0.62	1.23	1.02	1.20	−4.18	−16.93	−19.15	−30.22
4	0.60	1.33	1.03	1.23	−3.72	−16.47	−15.88	−26.94
5	0.62	1.33	0.97	1.22	−5.69	−18.44	−16.92	−27.99

The results show how the selection of the most suitable method depends on the performance metric used for evaluation. For instance, in the case of PT, Method 1 and Method 2 performed best based on MAE, whereas Method 4 had the best performance based on SME. The importance of selecting the appropriate statistical model is clearly demonstrated in the notable differences in SME values corresponding to CT approximations made for 30 panels. Specifically, with respect to these CT approximations, there is about a 25-minute difference between the SME values corresponding to Method 2 and those corresponding to Method 3 in the first run (Table 3-2), and about a 20-minute difference in the second run (Table 3-3). The disparity can be expected to further increase with increasing sample size of the testing datasets. Hence, it is important to select a suitable statistical model for subprocesses that cannot be linked to detection status.

For the framing workstation, the selection of a statistical model was based on the MAE metric. As such, Method 1 was selected for approximating pt_1 and ct_1. Approximations of 0.86 minutes and 1.21 minutes were made for pt_1 and ct_1, respectively, based on 150 datapoints when sample $actual\ pt_1 = [1.08, 0.58, 0.67, 0.92, 2.00 \dots]$ and sample $actual\ ct_1 = [1.35, 1.43, 0.85, 1.17, 2.43, \dots]$.

3.4 TiDA user interface

The user interface (UI) of the program developed for TiDA is displayed in Fig. 3-22. Following the logic described above, in this UI the user can tap the (+) button to add as many subprocesses as needed (e.g., one subprocess for the sheathing process and two subprocesses for the framing process as described above). For each subprocess, the user selects whether its time will be measured using object detection (referred to as imaged-based measurement in the UI) or statistical modelling (referred to as numerical measurement in the UI). For the image-based measurement, the user uploads a folder of images of the corresponding subprocess. As for the numerical measurement, the user selects a statistical modelling method (e.g., Gamma distribution) and inputs its parameters. The user then runs the program to obtain the *ST*, *PT* and *CT* measurements for the overall process.

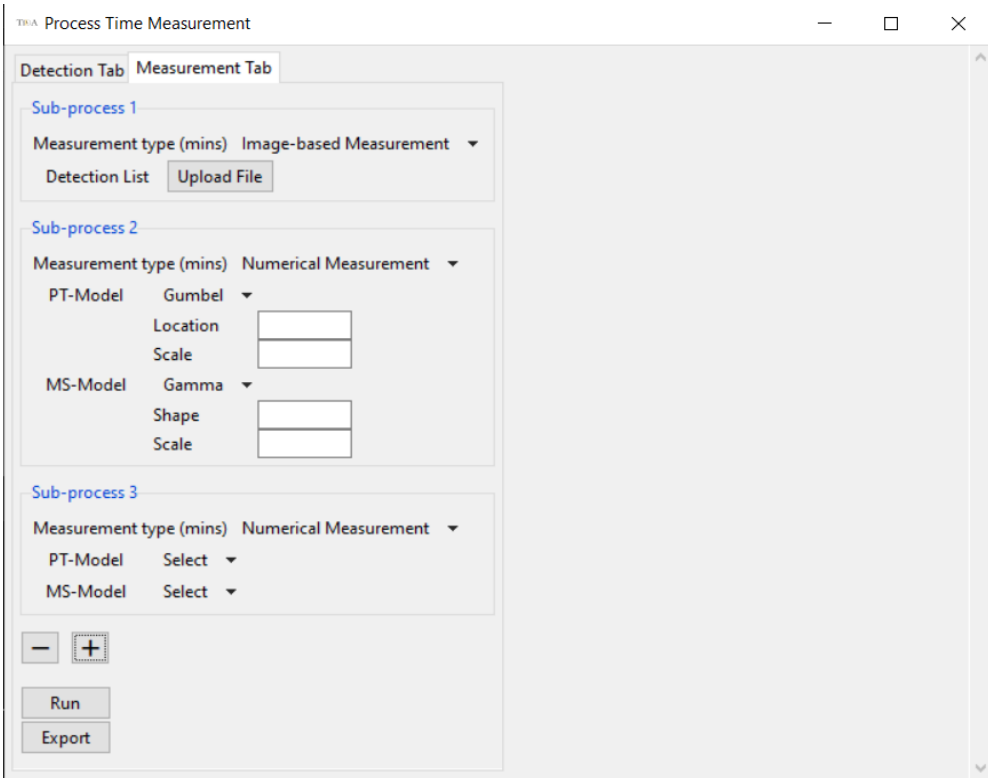


Fig. 3-22. TiDA UI.

3.5 TiDA application: Results and discussion

This section presents the results obtained in the application of TiDA to the framing workstation at the case company.

3.5.1 Quality of frames (Module 1) and detection/classification-related performance (Module 2)

Given the view angle of the webcam installed at the framing workstation, there was a large number of frames in which the operator was partially or almost fully hidden from view by components of the machine (Fig. 3-23) when either standing behind the components or bending down. These occlusions resulted in low confidence scores for many correct detections, and resulted in the object detection model failing to detect the operator in many other frames. It should be noted that most of the detection misses occurred when the operator was bending down close to the waste receptacle and therefore was almost entirely hidden from view of the camera (Fig. 3-23). It should also be noted that increasing the resolution of the frames did not enable the detection of the operator when they were almost completely hidden. As such, the confidence score threshold was lowered to reduce filtering out of correct detections of the operator (i.e., to reduce FN), although lowering the threshold also resulted in increasing the number of false detections (i.e., FP). The results obtained for operator detection based on a confidence score threshold of 0.4 are summarized in Table 3-4. Despite the low threshold value, the number of FPs was deemed acceptable, as the precision was found to be high (99.6%). However, the number of FNs was high, hence the lower recall (89.4%) compared to precision. Had the camera been positioned above the workstation, the number of occlusions would have been greatly reduced. Nevertheless, even though the accuracy (90.7%) and F1-score (94.2%) could have been increased by reducing the number of FNs, the overall performance of TiDA was still found to be satisfactory, as shown in the subsections that follow.

Table 3-4. Operator detection and classification performance.

TP	TN	FP	FN	Precision	Recall	Accuracy	F1-score
13880	2753	50	1652	0.996	0.894	0.907	0.942

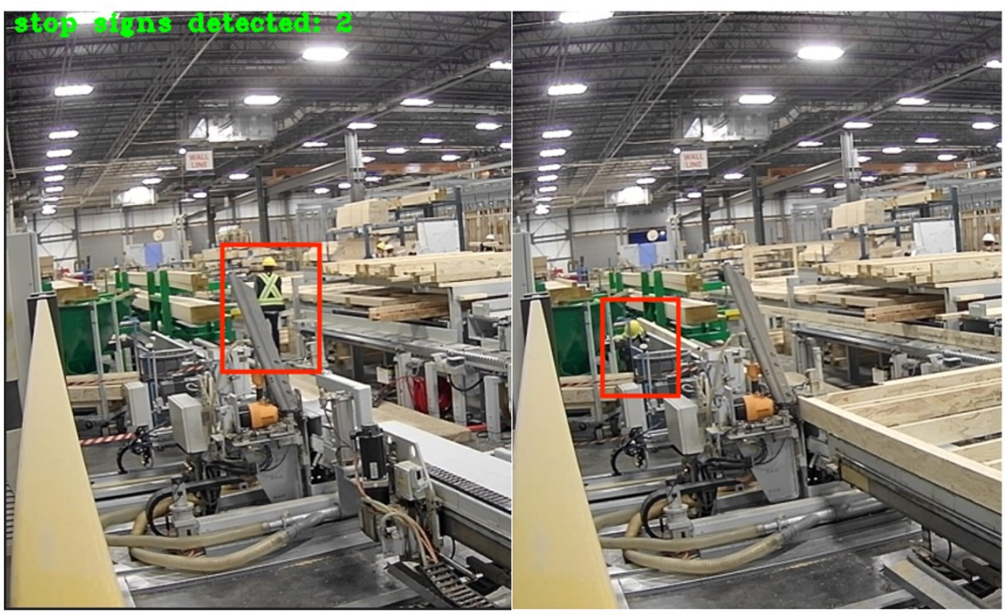


Fig. 3-23. Sample occlusions.

3.5.2 Evaluation of detection-based measurements (Module 3)

3.5.2.1 pt_2 measurement

The performance evaluation results for pt_2 are summarized in Table 3-5.

Table 3-5. Performance evaluation results for pt_2.

Date	**Panel #**	pt_2			
		Measured (minutes)	**Actual (minutes)**	ME **(minutes)**	|ME| **(minutes)**
3/15/2022	1	17.00	16.42	0.58	0.58
3/15/2022	2	9.58	10.00	−0.42	0.42
3/15/2022	3	7.67	8.25	−0.58	0.58
3/15/2022	4	8.58	9.33	−0.75	0.75
3/15/2022	5	8.25	10.25	−2.00	2.00
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
3/18/2022	119	7.17	6.83	0.33	0.33
3/18/2022	120	11.50	12.42	−0.92	0.92
3/18/2022	121	8.83	8.58	0.25	0.25
		∑(measured)	∑(actual)	SME (minutes)	MAE (minutes)
		973.75	966.75	7.00	0.90

The actual pt_2 measured for the 121 panels ranged from approximately 1 minute to approximately 28 minutes, with 75% of the panels having values greater than 5.5 minutes, as per the boxplot shown in Fig. 3-24. As shown in the boxplot in Fig. 3-25, meanwhile, 75% of the ME values were found to fall between −2.25 minutes and 0.42 minutes and 50% of them between −0.75 minutes and 0.42 minutes.

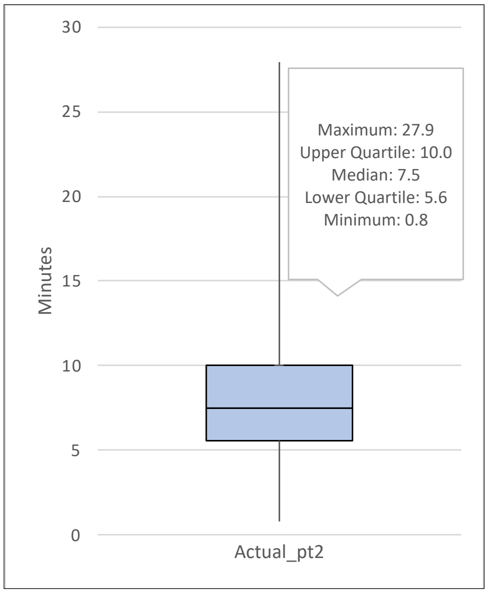


Fig. 3-24. Boxplot for actual values of pt_2.

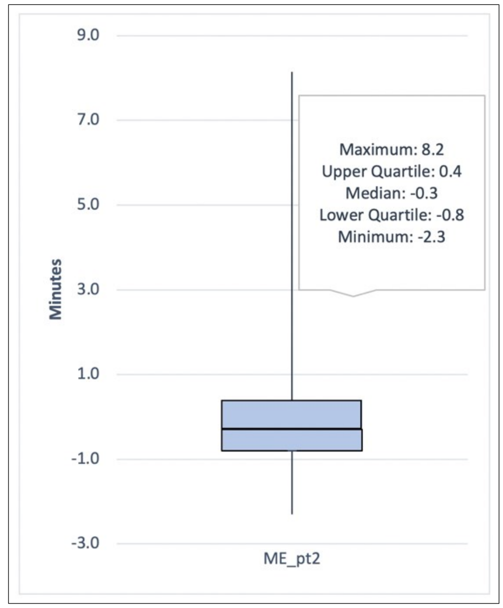


Fig. 3-25. Boxplot for ME of pt_2.

As shown in the streamgraph in Fig. 3-26, Module 3 overapproximated pt_2 for some panels and underapproximated it for other panels, with less than ten overapproximations/underapproximations exceeding 2 minutes and only two values exceeding 3 minutes. Despite these extreme values of ME, the MAE was just 0.91 minutes. Omitting the two values that exceeded 3 minutes, meanwhile, the MAE decreases to 0.82 minutes (based on 119 panels).

There were found to be mainly two scenarios in which Module 3 may underapproximate or overapproximate pt_2 as well as the size of the ME:

- Scenario 1: Module 2 fails to detect an operator who is in fact working at that time (i.e., is on task). This scenario contributes to underapproximation.
- Scenario 2: Module 2 detects an operator when in fact the operator is off task (Fig. 3-27) or incorrectly identifies an object (e.g., a machine part) as a person when in fact no work is being performed. This scenario contributes to overapproximation.

If the number of occurrences of Scenario 1 is greater than that of Scenario 2 for a given panel, Module 3 underapproximates pt_2 for that panel; whereas, if the number of occurrences of Scenario

1 is less than that of Scenario 2 for a given panel, Module 3 overapproximates pt_2 for that panel. The size of the overapproximation/underapproximation depends on how frequently one scenario occurs compared to the other, where greater error values (e.g., the two values exceeding +3 minutes as described above) are attributable to greater differences between the frequencies of occurrence of the two scenarios. For instance, in the case of the +8.17-minute error, the operator was at the workstation, but was frequently off task (Scenario 2). Hence, the assumption that work is being performed anytime workers are detected was frequently violated for this panel. Nevertheless, the violation was only extreme for one of the 121 panels under study.

If the two scenarios occur at a similar rate, or if neither scenario occurs at all (although this is not likely to be the case), Module 3 will achieve perfect performance in measuring pt_2 for a given panel. In the application to the framing workstation, Module 3 overapproximated pt_2 for 47 of the 121 panels under study with an MAE of 1.03 minutes (excluding the two outlier values exceeding 3 minutes). Meanwhile, it underapproximated pt_2 for the remaining 71 panels, with an MAE of 0.72 minutes. However, in reference to Fig. 3-26, in can be seen that, in aggregate, an approximate balance was achieved between overapproximation and underapproximation cases, as the SME was only 7 minutes for the set of 121 panels.

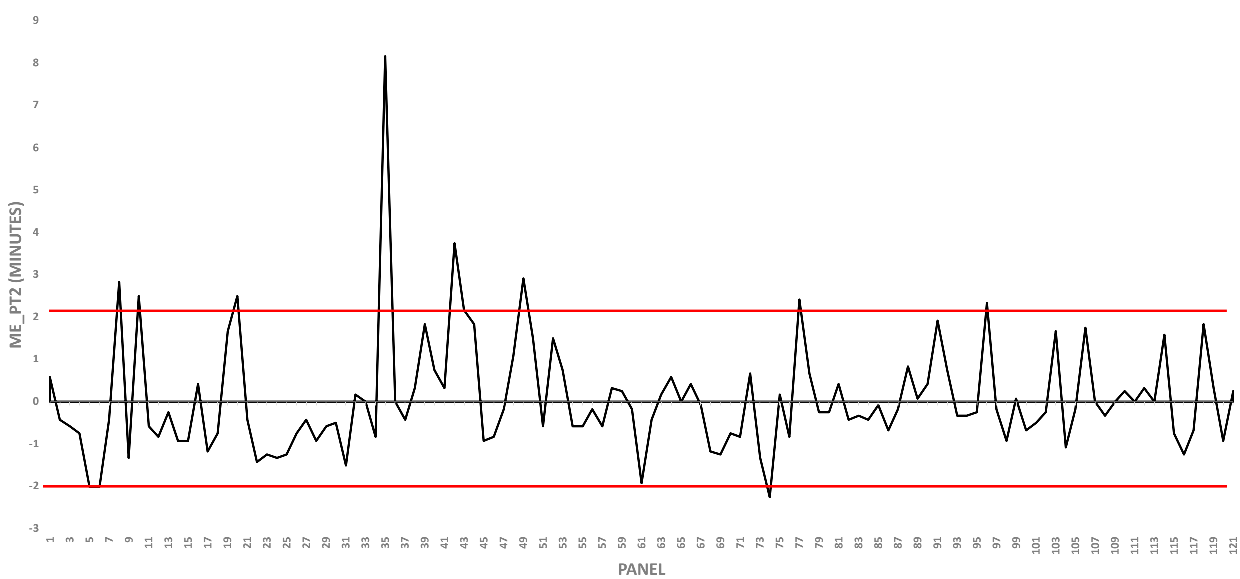


Fig. 3-26. Streamgraph for ME of pt_2.

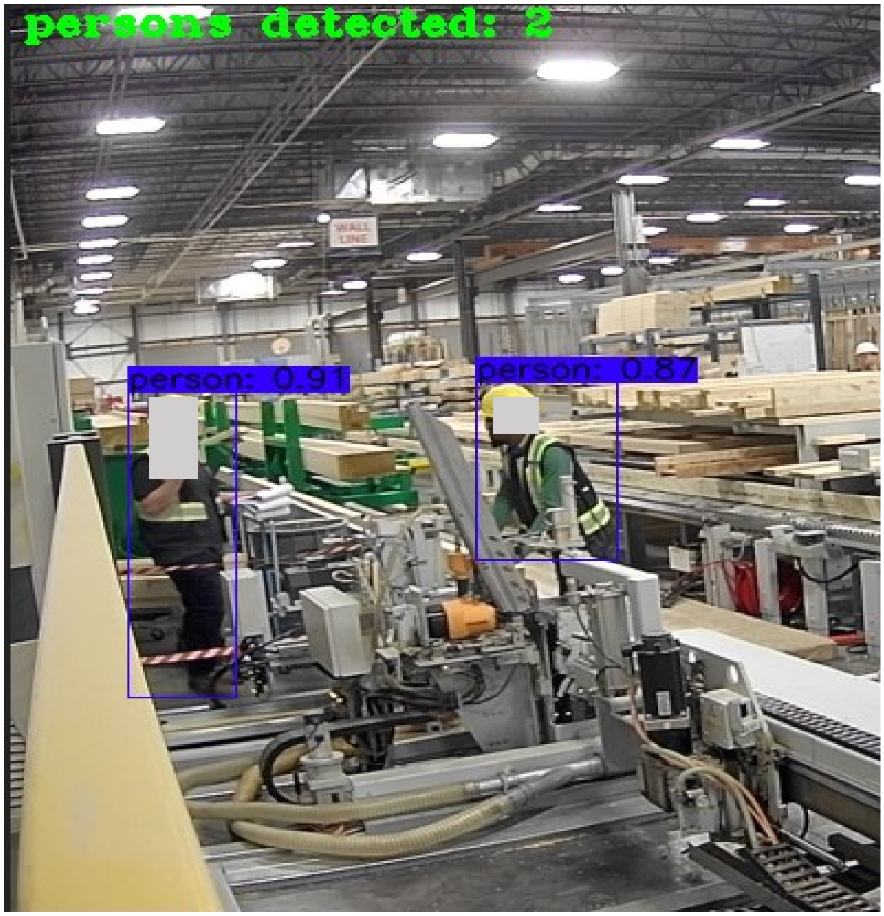


Fig. 3-27. Worker detected but not on task.

As the results show, it was found that, for a batch of panels, Module 3 may overapproximate or underapproximate the total productive time needed to complete tasks depending on the extent of overapproximation/underapproximation of individual panels. In the example presented in Table 3-5, Module 3 overapproximated by 7 minutes the productive time needed to complete the tasks for the 121 panels. However, the extent of the overapproximation can be considered insignificant given that the sum of the pt_2 for the 121 panels was approximately 974 minutes (meaning that it overapproximated by a margin of less than 1%).

As such, the performance of Module 3 was deemed to be highly satisfactory for measuring the pt_2 for individual panels, and to be outstanding for measuring the total pt_2 for a batch of panels.

3.5.2.2 ct_2 measurement

The performance evaluation results for ct_2 are summarized in Table 3-6.

Table 3-6. Performance evaluation results for ct_2.

Date	Panel #	$\mathbf{ct_2}$			
		Est (minutes)	**Act (minutes)**	ME **(minutes)**	|ME| **(minutes)**
3/15/2022	1	20.42	20.67	−0.25	0.25
3/15/2022	2	10.75	10.92	−0.17	0.17
3/15/2022	3	8.42	8.30	0.12	0.12
3/15/2022	4	9.42	9.40	0.02	0.02
3/15/2022	5	9.42	10.77	−1.35	1.35
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
3/18/2022	119	7.50	7.92	−0.42	0.42
3/18/2022	120	11.92	12.53	−0.62	0.62
3/18/2022	121	9.25	9.42	−0.17	0.17
		∑(measured)	∑(actual)	SME (minutes)	MAE (minutes)
		1,160.92	1,205.02	−44.10	0.40

The actual ct_2 measured for the 121 panels ranged from approximately 1 minute to approximately 48 minutes, with 75% of the panels having values greater than 6.5 minutes, as shown in the boxplot in Fig. 3-28. According to the boxplot in Fig. 3-29, 75% of the ME values were found to fall between −0.52 minutes and 0.82 minutes and 50% of them between −0.53 minutes and −0.10 minutes.

As shown in the streamgraph in Fig. 3-30, only a small portion of the data exceeded −1 minute, with only one value exceeding −2 minutes. The MAE was just 0.42 minutes, which can be considered negligible compared to the likely range of actual values. Module 3 generally underapproximated ct_2 with only a few exceptions. This underapproximation was due to the fact that the stop signs were located upstream from the cutting saw and the worker was cutting the top-plate, hence revealing the stop signs, before loading the last stud. This issue could be addressed by locating the stops signs downstream from the saw. Now, since this occurred for the majority of the panels, and given that the time needed to complete the rest of the work shows little variability (less than 1 minute in most cases), simply adding a correction factor of 0.4 minutes (approximately equal to MAE) would significantly reduce both MAE and SME. This correction was not implemented, though, since the purpose of the TiDA is to measure *CT*, which is also a function of ct_1. (In other words, it would be advisable in such a case to examine the combined effect of ct_2 and ct_1 before tuning the modules independently.)

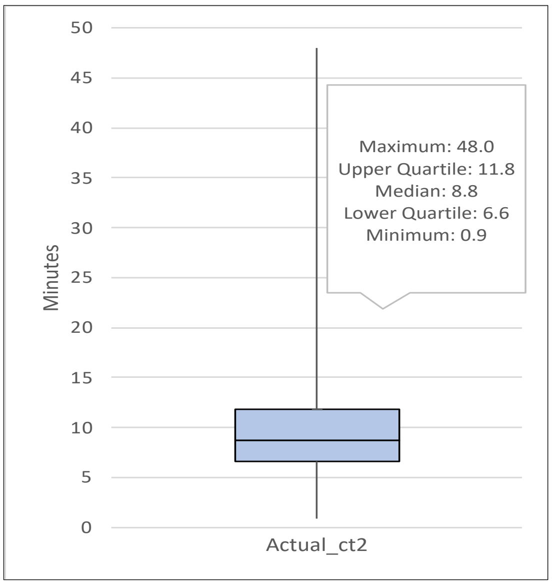


Fig. 3-28. Boxplot for actual values of ct_2.

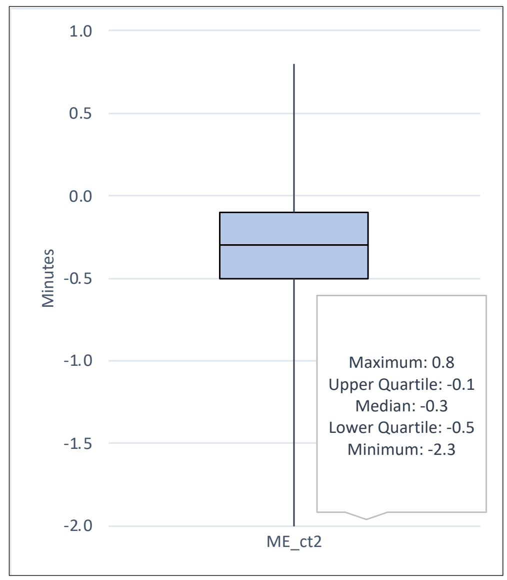


Fig. 3-29. Boxplot for ME of ct_2.

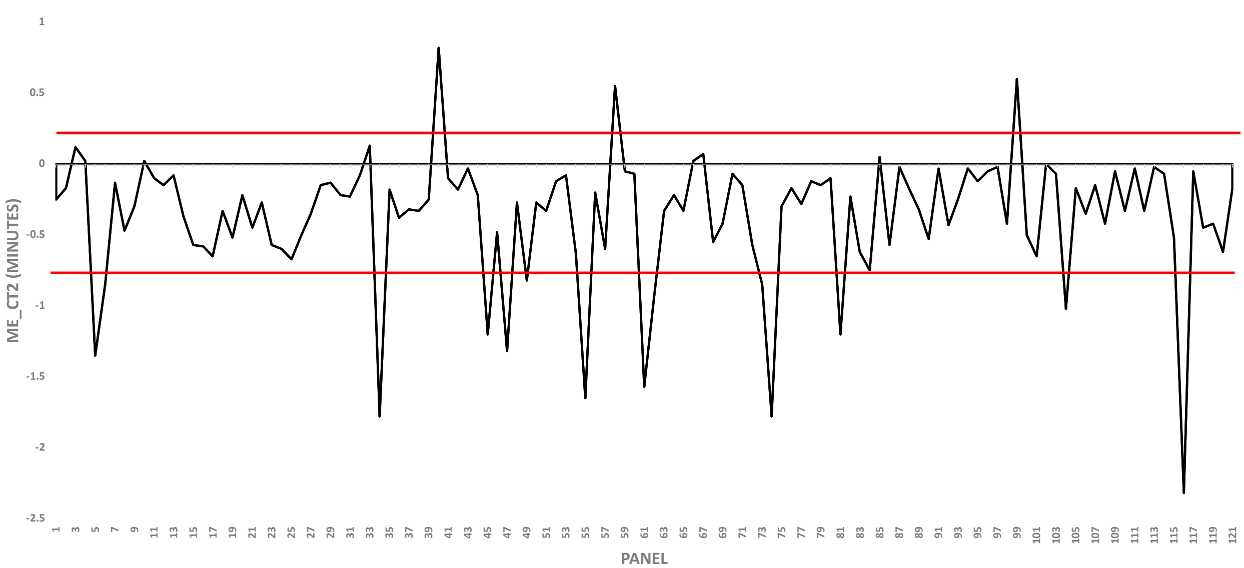


Fig. 3-30. Streamgraph for ME of ct_2.

Given that the ME was found to be negative for almost every panel, the ME values of the various individual panels did not cancel each other out when a batch of panels was considered in aggregate the way that they did in the case of pt_2. This implies that SME increases as the batch size increases. Nevertheless, the size of this error was found to be relatively small. The SME was -44.10 minutes for the 121 panels, whereas the sum of ct_2 for the 121 panels was found to be approximately 1,205 minutes.

3.5.3 Overall performance (Module 4)

3.5.3.1 PT measurement

The performance evaluation results for *PT* are summarized in Table 3-7.

Table 3-7. Performance evaluation results for *PT*.

Date	**Panel #**	***PT***			
		Est (minutes)	**Act (minutes)**	ME **(minutes)**	|ME|**(minutes)**
3/15/2022	1	17.86	17.50	0.36	0.36
3/15/2022	2	10.44	10.58	−0.14	0.14
3/15/2022	3	8.53	8.83	−0.31	0.31
3/15/2022	4	9.44	9.92	−0.47	0.47
3/15/2022	5	9.11	10.92	−1.81	1.81
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
3/18/2022	119	8.03	7.50	0.53	0.53
3/18/2022	120	12.36	17.00	−4.64	4.64
3/18/2022	121	9.69	9.42	0.28	0.28
		∑(measured)	∑(actual)	SME (minutes)	MAE (minutes)
		1,077.99	1,072.00	5.99	0.96

The actual *PT* measured for the 121 panels ranged from approximately 1 minute to approximately 30 minutes, with 75% of the panels having values greater than 6.3 minutes, as shown in the boxplot in Fig. 3-31. According to the boxplot in Fig. 3-32, the ME ranged between −4.64 minutes and 6.11 minutes, with 50% of the values falling between −0.64 minutes and 0.65 minutes. Based on the streamgraph in Fig. 3-33, only a small portion of the data was overapproximated/underapproximated by more than 2 minutes. The ME value exceeding 6 minutes corresponded to the panel whose measured error for pt_2 exceeded +8 minutes. As for the ME value that exceeded −4 minutes, the corresponding error of pt_2 was less than −1 minute (refer to Fig. 3-26), which implies that the error mainly resulted from setting pt_1 to be equal to an average value. The error of pt_1 was high in the case of that particular panel (i.e., panel #120 in Table 3-7) since the tasks performed within the corresponding timeframe took longer than usual, as the operator started framing the panel and then stopped to pick up and dispose of material waste. This was an unusual occurrence, as the operator typically performs such tasks before initiating the framing process for a new panel. As for MAE, it was found to be just 0.96 minutes.

Generally, the errors corresponding to PT measurement were higher than those recorded for pt_2 as a result of setting pt_1 to be equal to an average value as mentioned above (Method 1).

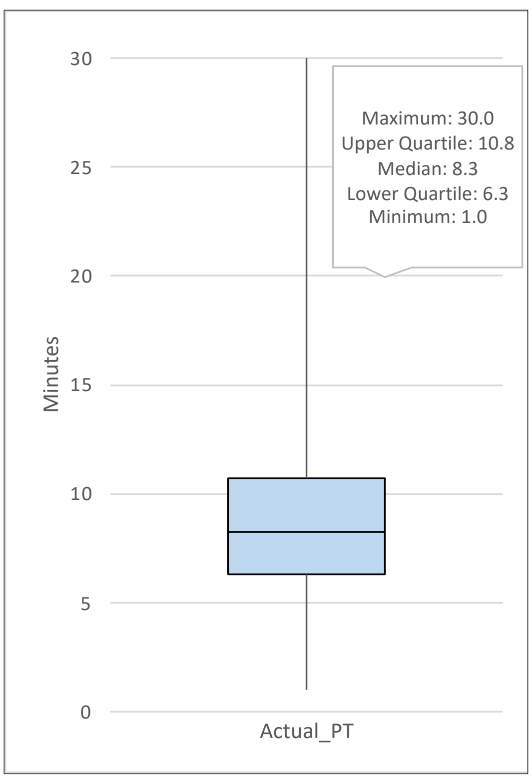


Fig. 3-31. Boxplot for actual values of *PT*.

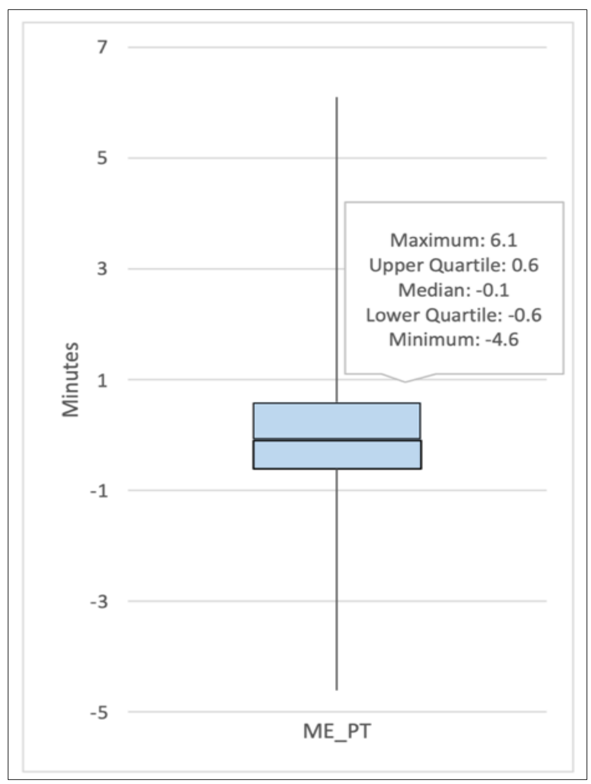


Fig. 3-32. Boxplot for ME of *PT*.

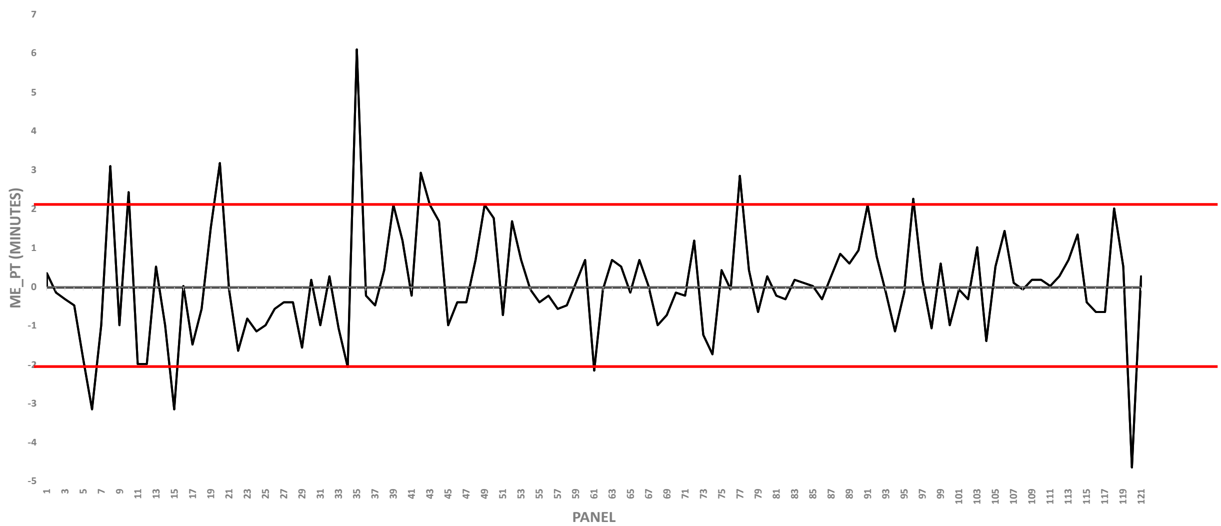


Fig. 3-33. Streamgraph for ME of *PT*.

As shown in Fig. 3-33, like Module 3, TiDA overapproximated *PT* for some panels and underapproximated it for other panels. There was also a relative balance between overapproximation and underapproximation, as the SME was small, amounting to 5.99 minutes for the 121 panels. As such, it could be concluded that Module 4, like Module 3, may overapproximate or underapproximate the total productive time necessary to frame a batch of panels depending on the sizes of overapproximations and underapproximations corresponding to individual panels. In the application to the framing workstation, Module 4 overapproximated the productive time needed to frame 121 panels by approximately 6 minutes. The size of this overapproximation can be considered negligible, however, considering that the sum of actual *PT* for the 121 panels was 1,072 minutes.

In short, the overall performance of TiDA was deemed to be highly satisfactory for measuring *PT* for individual panels, and to be outstanding for measuring the total *PT* for a batch of panels.

3.5.3.2 CT measurement

The performance evaluation results for *CT* are summarized in Table 3-8.

Table 3-8. Performance evaluation results for *CT*.

Date	Panel #	***CT***			
		Est (minutes)	**Act (minutes)**	ME **(minutes)**	|ME|**(minutes)**
3/15/2022	1	21.63	22.02	−0.39	0.39
3/15/2022	2	11.96	11.50	0.46	0.46
3/15/2022	3	9.63	8.90	0.73	0.73
3/15/2022	4	10.63	10.00	0.63	0.63
3/15/2022	5	10.63	11.43	−0.81	0.81
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
3/18/2022	119	8.71	8.67	0.04	0.04
3/18/2022	120	13.13	20.37	−7.24	7.24
3/18/2022	121	10.46	11.27	−0.81	−0.81
		∑(measured)	∑(actual)	SME (minutes)	MAE (minutes)
		1,307.47	1,347.57	−40.10	0.77

The actual *CT* measured for the 121 panels ranged from 1 minute to 48.5 minutes, with 75% of the panels having values greater than 7.7 minutes, as shown in the boxplot in Fig. 3-34. According to the boxplot in Fig. 3-35, meanwhile, 75% of the ME values were found to fall between −0.70 minutes and +1.61 minutes, and 50% of them between −0.70 minutes and +0.33 minutes. Based on the streamgraph in Fig. 3-36, only a small portion of the data fell outside the range bounded by +1 minute and −2 minutes, with only one value exceeding −4 minutes. This error value, which reached −7.24 minutes, occurred in the case of panel #120, i.e., the case mentioned above in which the operator paused working on the panel to dispose of material waste. The MAE, meanwhile, was found to be just 0.77 minutes—negligible compared to the likely range of actual values.

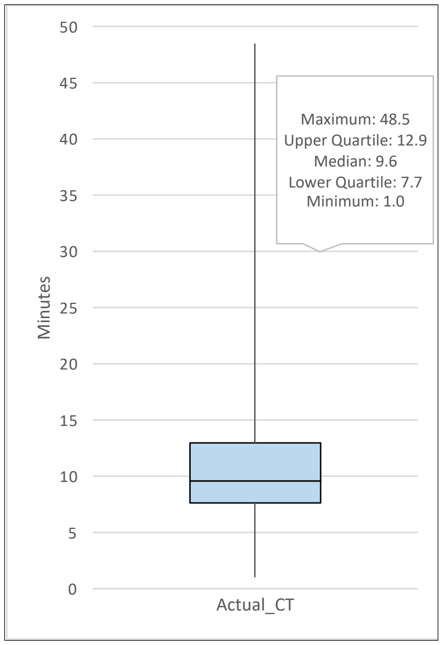


Fig. 3-34. Boxplot for actual values of *CT*.

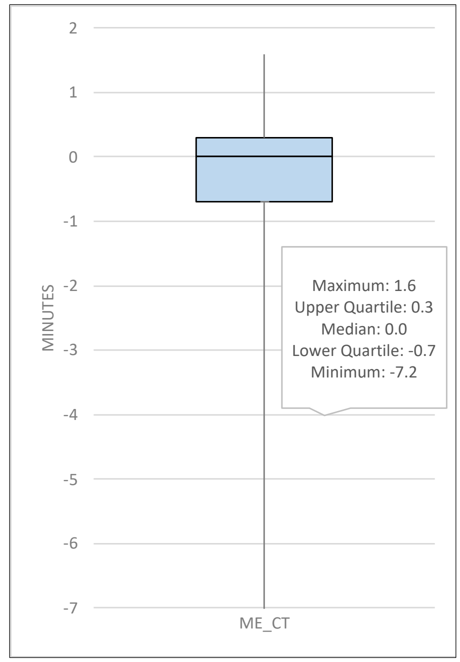


Fig. 3-35. Boxplot for ME of *CT*.

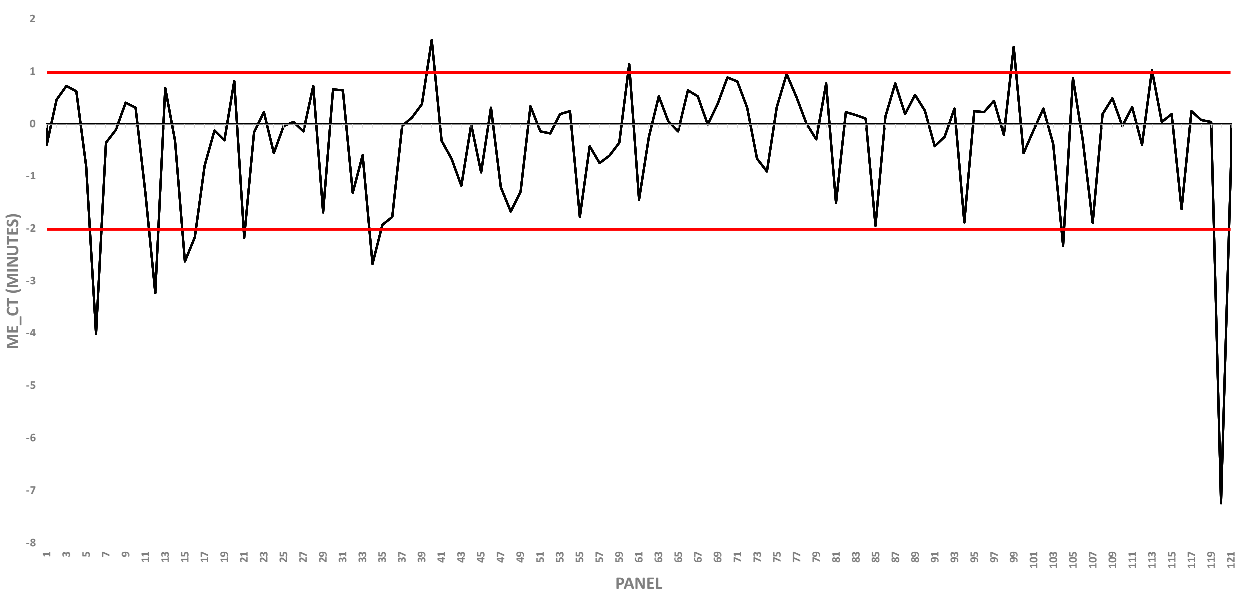


Fig. 3-36. Streamgraph for ME of *CT*.

As shown in Fig. 3-36, TiDA overapproximated *CT* for some panels and underapproximated it for other panels. The error range of underapproximations was greater than that of overapproximations, however. Such results are reasonable since Module 3 mostly underapproximated ct_2, meaning that CT will be further underapproximated if ct_1 is underapproximated but will be less underapproximated or overapproximated if ct_1 is overapproximated. Since the size of the underapproximations was found to be greater than that of the overapproximations, it can be inferred that TiDA generally underapproximated the total cycle time for a batch of panels for the case framing workstation, with the SME amounting to approximately −40 minutes for the 121 panels. As explained above, this underapproximation was mainly due to the fact that the stop signs were located upstream from the cutting saw and the worker was cutting the top-plate, hence revealing the stop signs before loading the last stud. Nevertheless, the size of this error was found to be relatively small compared to the sum of actual *CT* for the 121 panels, which was approximately 1,348 minutes.

As such, the overall performance of TiDA was also deemed to be highly satisfactory for measuring *CT* for individual panels, and to be outstanding for measuring the total *CT* for a batch of panels.

3.5.3.3 ST measurement

The performance evaluation results for *ST* (measured for the 121 panels) are summarized in Table 3-9. The results show whether the measured timestamps corresponding to *ST* are earlier or later than the actual timestamps, as well as how much the measured timestamps deviate from the actual timestamps. As can be seen, the MAE was found to be just 0.72 minutes, and 92.7% of $|ME|$ were less than 2 minutes. Moreover, as per the boxplot in Fig. 3-37, 50% of the ME values were between −0.7 and 0.1 minutes. Only Panel #120 showed a high error value (6.62 minutes), this being for the case mentioned above in which the operator paused working on the panel to dispose of material waste since *ST* is measured based on ct_1.

Hence, the results also demonstrate the promising performance of TiDA in measuring the start-time of framing operations.

Table 3-9. Performance evaluation results for *ST*.

Date	**Panel #**	***ST***			
		Est	**Act**		|ME|**(minutes)**
3/15/2022	1	7:22:50	7:22:42	Later	0.13
3/15/2022	2	7:44:10	7:44:48	Earlier	0.63
3/15/2022	3	7:58:47	7:59:24	Earlier	0.62
3/15/2022	4	8:07:46	8:08:23	Earlier	0.62
3/15/2022	5	8:18:05	8:18:38	Earlier	0.55
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
3/18/2022	119	13:57:50	13:58:18	Earlier	0.47
3/18/2022	120	14:13:35	14:06:58	Later	6.62
3/18/2022	121	14:28:03	14:27:25	Later	0.63
					MAE (minutes)
					0.72

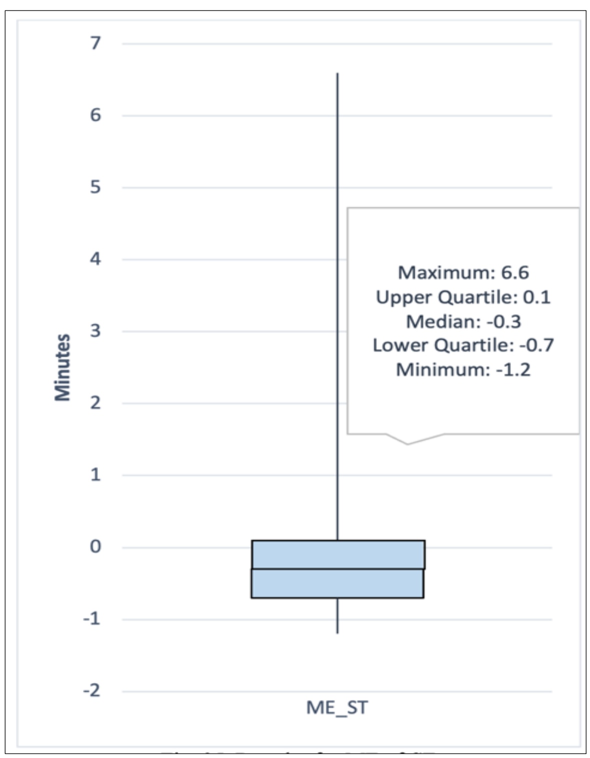


Fig. 3-37. Boxplot for ME of *ST*.

3.6 Conclusions

3.6.1 Summary of results

This chapter presented a computer-vision-based system, TiDA, that allows the user to measure the start-time of a process performed at a workstation in a panelized construction factory, the productive time needed to complete the process, and the actual cycle time (i.e., from the start of the process to the end) spent on the process. The logic of TiDA was demonstrated through its application to two different workstations in a light-wood panelized construction factory, and its performance was evaluated based on its application to the framing workstation at this factory.

TiDA showed promising performance in measuring the productive time, cycle time, and start-time. With respect to productive time, some values were overapproximated while others were underapproximated, but the degree of variance from true values was deemed to be acceptable. Specifically, the MAE computed for productive time measured for 121 panels was found to be just 0.96 minutes, while the SME revealed an overapproximation of approximately 6 minutes. The extent of this overapproximation can be considered negligible, given that the sum of actual productive times measured for the 121 panels was approximately 1,072 minutes. TiDA similarly overapproximated cycle time for some panels while underapproximating it for other panels, with the error range of underapproximations being greater than that of overapproximations, although both ranges were also deemed acceptable. The MAE computed for the 121 panels was just 0.77 minutes, while the SME showed an underapproximation amounting to approximately -40 minutes. Again, the size of this underapproximation can be considered negligible when we consider the sum of actual cycle times for the 121 panels (i.e., approximately 1,348 minutes). With respect to start-time, the MAE for the 121 panels was 0.72 minutes, and 92.7% of the absolute values of ME were less than 2 minutes.

The main sources of error originating in the object detection were occlusions and awkward body postures, as well as the assumption that the presence of the operator indicates productive time. However, for each panel, errors resulting from occlusions and body posture (underapproximations) were approximately offset by errors resulting from the assumption of productive time (overapproximations). Given this, Module 3 overapproximated pt_2 for 47 of the 121 panels under study with an MAE of 1.03 minutes (excluding the two outlier values exceeding 3 minutes). Meanwhile, it underapproximated pt_2 for the remaining 71 panels, with an MAE of 0.72 minutes. Moreover, using mean values to model the durations of some tasks also resulted in errors, but these errors were generally minor, with the one exception being the case of Panel #120. The underapproximation of CT was a result of locating the stops signs upstream from the cutting saw when the worker was cutting the top-plate, which resulted in the stop signs being revealed prior to loading of the last stud. Nevertheless, all the errors encountered did not have a significant impact on the performance of TiDA, which was found to be promising.

3.6.2 Benefits of the developed system

In principle, we can see that TiDA was found to satisfy the development criteria defined for this research (presented earlier in this chapter). Nevertheless, a thorough evaluation of these criteria that compares the use of TiDA to building a training dataset and training a custom object detection model is needed. (More details on this follow in the next section). A general discussion on how TiDA was found to satisfy the development criteria is provided as follows:

- *Comprehensibility*: The logic underlying TiDA is simple. Productive time is linked to the presence of workers at the workstation for labour-based and semi-automated workstations, and cycle time is mainly linked to the detection state of an object(s) located at the workstation.
- *Ease-of-use:* Once the user understands how the system functions and the system is set up for a workstation, the only inputs required are images of the workstation.
- *Timeliness:* Setting up the system for Type I workstations is straightforward and can be completed within hours. For Type II workstations, setup of the system can be completed within a day. Once the system is set up, the user only needs to run the code.
- *Adaptability:* The logic underlying TiDA is generic and can be applied to various processes involving varying levels of automation in panelized construction. TiDA's measurements of start-time and cycle time are mainly based on the flow of work-in-process/material at a workstation, regardless of whether there are workers or machines at the workstation. This is demonstrated in the examples above (i.e., the framing

workstation at the case company is semi-automated, whereas the sheathing workstation is labour based). As for productive time, its measurement is related to the detection of workers for labour-based and semi-automated workstations. TiDA can also be deployed to measure productive time at fully automated workstations, as described in Section 3.2.2.3, although productive time can also be evaluated in such cases using simple calculations as explained above.

- *Cost-effectiveness*: The time requirement for setting up TiDA is considerably lower than the time requirement for training and using custom object detection, resulting in time and cost savings. Moreover, using a pre-trained object detection model significantly reduces computational requirements compared to an approach that involves model-training, so the use of Google Colab, open-source Python libraries, and a personal computer (rather than powerful GPU resources) is sufficient for implementing TiDA. Finally, the system only requires the following hardware: (1) a camera to capture images of the workstation (a CAD 40 webcam was used in the application presented in this study—further cost savings can be achieved if existing CCTV systems can be used for image capture); (2) stickers of stop signs to be installed at the workstation; and (3) a computer to run the code for performing the measurements. It should be noted that it is also possible to extract the object detection model into a Yolov4 tiny model and run TiDA on a Raspberry Pi with a TPU dongle (e.g., Google Coral USB TPU dongle).

As shown in the summary above, in addition to exhibiting promising performance, TiDA generally meets the pre-defined development criteria and, thus, can be considered an attractive solution for acquiring time-related data in panelized construction. Meanwhile, the process time variables measured by TiDA can be used for (1) evaluating the efficiency of manufacturing processes at workstations (by comparing productive times to cycle times), (2) identifying bottlenecks on production lines (by analyzing cycle times), (3) monitoring progress at workstations (if the images are processed in real time), (4) training machine-learning models to predict productive times and cycle times at different workstations, and (5) virtually modelling operations in simulation models and digital twins, to name a few examples. In fact, the frequent need for such data at the case company and the lengthy time studies typically conducted to fulfill this need were the primary motivation for developing TiDA.

3.6.3 Limitations and future work

In applying TiDA at certain workstations, there are two limitations the user may encounter. First, the use of the system is not advisable when the process of interest includes a subprocess that cannot be linked to detection status and yet shows high variability. Applying statistical modelling to such a subprocess would result in high error, in turn resulting in increased errors in the measurements of process time variables. However, such processes are not common in panelized construction factories, where most tasks are of a repetitive nature. Still, more work needs to be done to address this issue. The second limitation is that, if workers at the workstation are frequently off task, productive times will be greatly overestimated. Again, this is not frequently encountered in panelized construction factories, since operations are typically consistently monitored and controlled by team leads and supervisors. In the case application presented in this study, only one instance of this issue was encountered over the course of four days. Nonetheless, an outlier analysis method could be integrated with TiDA in the future to identify outlier productive time values, given the design properties of panels (e.g., a 40′ wall panel should take longer to frame than a 20′ wall panel).

Although the detection misses encountered in the application presented in this chapter did not have a significant impact on the performance of TiDA, the system can be improved in future work to make it better able to handle the occlusions that may be encountered. Although camera-positioning can, as explained above, play a significant role in reducing occlusions, it would be of value to integrate occlusion handling techniques with TiDA and/or experiment with alternative object detection models such as YOLOv7 (Wang et al., 2022).

It should also be noted that two related studies are underway. In one study, object detection is being combined with low-cost ultrasonic sensors to measure the same process time variables considered in the present study (i.e., start-time, productive time, and cycle time). This alternative is being examined as an approach that may be more effective at locating sensors in tight spaces at certain workstations. In the other study, we are creating and annotating a dataset to train a YOLOv7 object detection model on wall frames and panels to provide a thorough comparison of the differences between operating TiDA with minimal set-up requirements versus building a custom object detection model from scratch. In that comparative study, we delve into the specifics of the development criteria described in the present study.

Acknowledgements

This work was supported by Alberta Innovates. The authors would also like to acknowledge the technical support of Mr. Thomas Huynh and the technical review of Mr. Jonathan Tomalty.

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 4

Factors Influencing Cycle Times in Offsite Construction

4.1 Introduction

4.1.1 Motivation: The hare approach

The offsite construction industry, also known as the construction manufacturing industry, is rooted in the broad shift of construction practice from traditional in-situ methods to manufacturing methods. One may intuit that moving towards manufacturing methods will inevitably pave the way for comprehensive and streamlined implementation of lean philosophy in construction. There is a degree of truth to this, as many offsite construction companies have sought to leverage the benefits of lean principles such as standardisation, waste reduction, continuous flow, production line balancing and others with the notable case of a panelized construction enterprise in Edmonton, Canada, applying these principles described in a recent study (Alsakka et al., 2022a). Several studies have evaluated the benefits of implementing lean principles in offsite construction, including waste minimisation and workload and workforce density balancing in modular construction (Moghadam et al., 2013; Y. Zhang, 2017), and batch and inventory size reduction in precast construction (El Sakka et al., 2016), to name a few. In reality, however, the variable nature and other unique characteristics of construction make the implementation of lean manufacturing practices such as *Heijunka* (i.e., levelling out the work schedule (Liker, 2004)) inherently challenging. As argued by Ohno (1988) decades ago and reiterated by Liker (2004), "*the slower* *but consistent tortoise causes less waste and is much more desirable than the speedy hare that races ahead and then stops occasionally to doze. The Toyota Production System can be realized only when all the workers become tortoises*.” The point here is not to advocate for slow production, but rather for steady production that reduces the likelihood of over- and under-utilisation of resources. The variability inherent in construction projects, however, forces workers and machines in offsite construction factories to follow the so-called “hare” approach. Let us consider, for example, a production line for fabricating wood house walls with one workstation dedicated to framing wall panels and another workstation dedicated to installing sheathing. Since walls are of different types, dimensions, and designs, the time it takes to frame a wall panel or install sheathing, if any, will vary depending on the wall type/design. As a result, if, for a given batch of panels, wall framing takes less time than sheathing installation, the workers at the sheathing workstation will be pressured to speed up their work to keep pace with the framers and keep the production line moving. If this batch is followed by a batch of interior walls for which no sheathing is installed, then the workers at the sheathing workstation will be under-utilized while the framers will be pressured to speed up their work to avoid starving the downstream workstations. In other words, the workers are pressured to function as “hares”. Operators in offsite construction do endeavor to reallocate tasks among workstations in order to mitigate this effect, as described in (Alsakka et al., 2022a), but effective levelling of a production line requires detailed knowledge of the variable cycle times at workstations that is not readily available in current practice.

4.1.2 The need for influencing factors

Given this, researchers have employed machine-learning models trained to consider relevant influencing factors (or prediction variables) in order to estimate process time-related variables. For instance, Shafai (2012)—who argued that average task times should not be used to estimate the durations of highly variable tasks performed for manufacturing wall panels of different designs since the manufacturing time is contingent upon the unique design properties of each panel—built linear regression models for estimating the duration of each task (e.g., spray form insulation) as a function of the given panel’s design properties relevant to the task at hand (e.g., number of studs, cut zone area, etc.). In another study, (Benjaoran et al., 2004) used multivariable linear regression and neural networks to estimate the duration of production processes in a precast factory as a function of twenty influencing factors such as material weight and concrete strength. However, few studies are available in the literature that have followed this line of thinking for estimating process times or cycle times at the workstation level in offsite construction, although there are studies that have followed this paradigm for estimating other related variables, such as a study estimating man-hour requirements for structural steel fabrication jobs using linear regression (Hu et al., 2015), and another estimating the productivity of steel fitting activities in steel fabrication using artificial neural networks and simulation (Song et al., 2008). Moreover, despite the critical role prediction variables play in determining the performance of machine-learning models, the identified studies either have not taken a systematic approach or have not thoroughly discussed the approach followed for identifying the factors that may have an effect on the time variables under study (i.e., task time, cycle time, man-hours, and productivity). The value of expending effort on such an approach is not only that it allows for the relevant influencing factors to be identified; it also helps modellers to gain knowledge about the process under study, in turn allowing them to follow a prescriptive approach for selecting and representing predictor variables (Kuhn et al., 2019). In this manner, they can perform experience-driven modelling alongside empirically-driven modelling, thereby reducing the risk of overfitting to erroneous data patterns or of generating models that cannot be rationally interpreted, compared to an approach that relies solely on empirical data (Kuhn et al., 2019).

4.1.3 Study objective and contributions

In this context, there is a need for a structured approach for identifying factors that could influence cycle times in offsite construction. This study thus presents a qualitative approach for identifying factors that influence cycle times at the workstation level in offsite construction factories. The identification of these factors, it should be noted, is an important preliminary step when deploying machine-learning techniques to develop cycle time prediction models as part of the lean practice of production line levelling. The approach is demonstrated through its application to a semi-automated, wood-wall framing workstation in a panelized manufacturing factory in Edmonton, Canada. The study contributions are as follows: (1) shedding light on the significance of analysing cycle times at the workstation level in offsite construction factories; (2) presenting the implementation of a generic approach as a way of encouraging researchers and practitioners to expend effort on identifying the factors influencing cycle time, which are significant for the performance and interpretability of machine-learning models developed to predict cycle times or related process time variables for the purpose of optimising production lines and production schedules (in order to ensure more balanced and efficient production); and (3) providing a preliminary list of factors that could influence cycle times at semi-automated wood framing workstations in offsite construction—a list that could serve as a starting point for researchers or practitioners studying other types of framing workstations.

4.2 Approach and methods

The study followed a three-stage qualitative approach that leverages the benefits of process mapping and semi-structured interviews to identify the factors exerting an influence on cycle times at a wood framing station. The approach is presented in a generic manner in this section, while the next section describes its application to the case framing workstation.

4.2.1 Stage I: Understand the process

An adequate understanding of the tasks involved in a process, the resources allocated to it, the manner in which the tasks are carried out, the process inputs, and of the process outputs enables rapid identification of a number of factors influencing cycle time. Process mapping of the current state, in turn, is an effective means of gaining a thorough understanding of a given process. Process mapping generates an abstraction of the process, allowing for it to be better understood and demonstrated and its performance assessed (Giachetti, 2011). The steps followed in building a process map are described in detail in a previous study by Alsakka et al. (2022). Validating the accuracy of the process map with input from the workers actually assigned to the workstation under study is crucial. The case application described in the present study demonstrates the significance of this validation task. The framing machine is equipped with a cutting saw that is used to cut through the top- and bottom-plates of the panels. Throughout the period of observation that formed the basis of the process mapping, the operator at the framing workstation was manually operating the cutting saw for every wall panel. As a result, moving the cutting saw was recorded as a step in the framing process. However, consultations with the operator revealed that in fact the framing machine was in disrepair, and hence, the operator was manually performing a step that would normally be performed automatically by the machine. In other words, what seemed to the analyst to be a normal part of the process (based on observation alone) was in fact the result of equipment breakdown (i.e., a factor affecting cycle time at the workstation). This example underscores the importance of validating the process map based on consultation with workers on the production line as a crucial step in identifying the factors influencing cycle time.

4.2.2 Stage II: Compile a list of potential factors

Based on the results of the first stage, the analyst may identify a variety of factors that influence cycle time at the workstation with regards to various elements involved in the process. The analyst may start by specifying high-level classes that could encompass the different types of factors to be identified, since doing so helps to structure and, hence, facilitate the process analysis task. In this respect, a set of eight major classes is proposed in the presented approach— “product”, “worker”, “machine”, “material”, “workstation setup”, “production line”, “factory operations”, and “external factors”—these classes having been preliminarily selected based on the authors’ understanding gained during the first stage, then confirmed based on a review of the relevant literature (refer to the Case Application section). In relation to each of these classes, the analyst may identify factors that influence cycle time at the workstation. (Examples of factors that belong to different classes are described in the case application section of this chapter.) In addition to the process map, a review of previous research that analyses cycle times, productivity, or related aspects of the process under study, or of similar processes in offsite construction factories, could help to identify additional factors and, possibly, additional classes (over and above the eight classes proposed). At that juncture, the analyst would have a profound knowledge of the process under study and would be well positioned to extract relevant factors from the literature. It is advisable to extract all factors that could potentially have an impact on cycle time at this stage as doing so can further bolster the understanding of the process, even if some of the factors are ultimately excluded at a later stage. The outcome of this stage is a group of classes comprising factors that may impact cycle time at the workstation under study.

4.2.3 Stage III: Solicit workers' input on the factors

As the cutting saw example described above demonstrates, the input of workers regarding cycle time-influencing factors is critical, since they are the most knowledgeable about the process. The workers' input may help the analyst to better understand certain factors, highlight significant factors, determine which factors are less important, identify additional factors, or identify relationships between different factors. Hence, upon compiling a preliminary list of factors in the second stage, semi-structured interviews can be conducted to solicit workers' input on the factors in the list. Semi-structured interviews, it should be noted, involve a mixture of close-ended and open-ended questions that are often followed with "why" or "how" questions (Adams, 2015). Semi-structured interviews are valuable when the interviewer (i.e., the analyst, in the context of this study) is interested in the independent thoughts of the interviewee (i.e., the worker) or when there are unknown but potential issues and the interviewer needs to pinpoint beneficial leads and pursue them (Adams, 2015). For each of the identified factors, the analyst may start by asking the worker if the factor affects or does not affect cycle time (i.e., a Yes/No question) and then asking follow-up questions such as "why it affects (or does not affect) cycle time", "how it affects cycle time (i.e., positively/negatively)", and "to what extent it affects cycle time (i.e., significance)". In the case application presented in this study, this approach was found to trigger valuable discussions that yielded useful insights.

Given that a fixed and limited number of workers are typically assigned to each workstation in offsite construction factories, it is possible that some workstations will only have a single worker. This means that there may be just a single worker who is deeply knowledgeable about the current state of the process under study in some cases. However, this would not be critical, as the factors would have been previously identified based on a detailed analysis of the process and previous research work and will be further analysed during the machine-learning process in which the factors will be used. In other words, there are multiple input sources for the factors.

4.3 Case application

This section presents the implementation of the described approach on a semi-automated wood-wall framing workstation located in a panelized construction factory. In a recent case study on this workstation, cycle times were found to vary significantly, ranging from approximately 1 minute to about 48.5 minutes (Alsakka et al., 2023b). This wide range of cycle times underscores the importance of determining the factors that influence cycle times at such workstations.

4.3.1 Stage I: Understand the process

The case framing workstation has a semi-automated wood-wall framing machine that performs three operations: nailing, drilling, and cutting. An operator loads the machine with framing elements when prompted by the machine to do so, and the machine performs the required operations. An automated material feeding system moves studs from their inventory location to a location at the framing workstation from which the operator can directly pull them. The components are made ready half a shift or one shift before they are needed, and are placed on a rack located at the framing workstation in the same order in which they will be required by the framer. Moreover, the top and bottom plates of wall panels are stored on a rack located next to the workstation in such a manner that the operator can directly pull the plates to their loading locations on the framing machine. Fig. 4-1 shows the locations of the different elements.

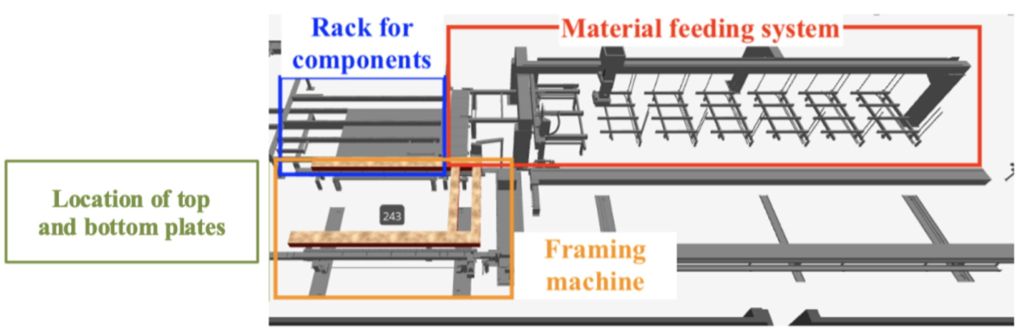


Fig. 4-1. Virtual model of the framing workstation.

Given that there are multiple resources (i.e., machine, operator, feeding system) interacting at the framing workstation to frame wall panels, cross-functional diagrams, also known as “swim-lanes”, were developed to aid understanding as to which tasks are performed by each resource. Cross-functional diagrams, it should be noted, are used to map the workflow of interrelated activities and resources that transform inputs into outputs, as well as to portray the relationships among the various resources performing actions (Damelio, 2011; Giachetti, 2011). A portion of the mapped diagram is displayed in Fig. 4-2. The diagram was first mapped based on observation, and then verified and adjusted based on the operator's feedback.

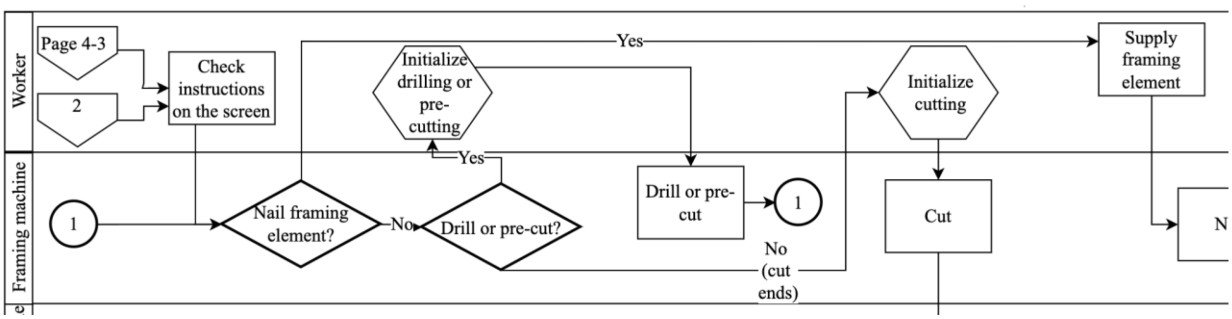


Fig. 4-2. Portion of the framing workstation's cross-functional diagram.

4.3.2 Stage II: Compile a list of potential factors

For each of the eight classes mentioned above, the factors understood to affect cycle times at the framing workstation were identified based on the authors' understanding gained during the first stage. This was followed by a review of the relevant literature to confirm the comprehensiveness of the classes identified. Because, as previously mentioned, only a limited number of directly related studies were identified, studies examining related metrics such as man-hour requirements and productivity were also reviewed. The factors identified in the relevant literature included, to name a few representative examples, (1) product-related (or design-related) factors such as [length, width, height, surface area,…] for the production of steel panels (Ayinla et al., 2019), [number of single studs, double studs, doors, windows, cut zones, drill holes, nails, screws,…] for the production of wood wall panels (Shafai, 2012), [number of fittings, cut-outs] for steel fitting (Song et al., 2008), [number of bolts, length of weld, length of wide flange beams,…] for structural steel manufacturing (Hu et al., 2015), and [nominal height, weight, and width, concrete volume, finishing area, reinforcement weight, concrete strength,…] for precast concrete production (Benjaoran et al., 2004); (2) worker-related factors such as the number of workers (Benjaoran et al., 2004), and skill level (Song et al., 2008); (3) material-related factors such as length and weight (Hu et al., 2015; Song et al., 2008); (4) machine-related factors such as breakdowns and interactions of material handling systems (Song et al., 2008); (5) factory operations-related factors such as work shift (Song et al., 2008); and (6) production line-related factors such as activity precedence relationships, queuing, and rework (Song et al., 2008). Finally, a list was compiled for 36 classified factors of which the cycle time at the case framing workstation may be a function. These factors are listed in Table 4-1 below. It should be noted that certain factors that, although may influence framing cycle time, are highly complex and may require a comprehensive analysis of their own (e.g., worker morale, work environment, worker wellness, pay, etc.) were excluded from the case study. It should also be emphasized that, while there may be factors that influence cycle times at framing workstations in other companies, or in other workstations at the case company, only factors influencing cycle time at the framing workstation under study were considered. For instance, the availability of tools and machines is a commonly encountered factor that influences cycle time, but these resources at the case workstation are not shared with other workstations and, hence, are always available. The input of the operator on these factors (presented in the following section of this chapter) helps to further clarify meaning, and provide a preliminary justification for inclusion, for the listed factors.

4.3.3 Stage III: Solicit workers' input on the factors

At this stage, the operator's input was solicited (via semi-structured interview) concerning the list of potential factors. The operator consulted, it should be noted, has more than ten years of experience working at the framing workstation at the case company, making him highly knowledgeable about the process. The operator was asked whether or not, why (if applicable), and in what manner (if applicable) each of the listed factors affects cycle time. The operator indicated that some of the listed factors are correlated with other factors, which means that they hold information also held by other factors with regards to cycle time. The interview results are summarized in Table 4-1 (found in the following subsection), where (✓) indicates that the given factor was considered by the operator to influence framing cycle time, (X) indicates that the given factor was not considered by the operator to influence framing cycle time, and (C) indicates that the given factor was considered by the operator to influence framing cycle time, but that the factor is correlated with another one. The operator's comments included in the table are based on written notes taken during the interview. (It should be noted that the comments as represented are a mix of the exact words of the operator and reformulations of some of the operator's input.) Factors for which no specific comments were made during the interview are denoted by a dash symbol in the "operator's comments" cell in the table.

4.4 Results and discussion

Based on the interview results, the majority of the factors identified in the first two stages were deemed to be relevant based on the operator's input. Accordingly, it was determined that these factors (highlighted in green in Table 4-1) should be left for the machine-learning process. During the semi-structured interview, the operator provided information that directly resulted in the exclusion of previously included factors, as it became evident based on this information that these factors (highlighted in red in Table 4-1) do not influence framing cycle time. Removing these factors would help to avoid unnecessary effort expended collecting data on factors that would have been removed during the machine-learning process anyway, as well as reducing the complexity of the machine-learning process. However, factors with respect to which workers may make subjective judgements were not excluded (even when flagged as candidates for exclusion) unless the machine-learning process confirms their irrelevance. For instance, the hypothesis underlying the wall panel design complexity factor is that it may take the operator more time to interpret the shop drawings and load the elements accordingly for more complex wall panels (since they typically require more framing tasks compared to less complex wall panels). Even though the operator indicated that this factor does not affect the time it takes to frame a panel, relying solely on their experience-based input may introduce bias, as it is difficult to assess how long it takes to interpret a shop drawing or load elements from different locations without a quantitative analysis. As such, these factors (highlighted in orange in the table) should be examined in the machine-learning process. Moreover, the operator identified two factors as being correlated with other factors. One of these was panel length (highlighted in yellow), which was indeed found to be correlated with the number of cuts. However, the panel length factor may hold additional information that is not captured by the number of cuts factor or by other factors. In fact, many of the previous studies in this area have used panel length as a factor (as discussed above), further supporting the hypothesis that it is an influencing factor. Additionally, panel length is correlated with the number of holes used for lifting, a consideration that the operator did not mention. This justifies the consideration of panel length as a potential influencing factor, as well as its inclusion in the final list of factors. The other factor identified by the operator as being correlated with other factors was the distance between the nail inventory location and the workstation, this factor being correlated with the nail gun refill factor. Reaching the nail inventory during the process of framing a wall panel was found to be 100% correlated with the nail gun refill factor, and for this reason the former factor can be excluded. Finally, the operator noted that adjusting the machine's opening to accommodate panels of different heights adds an extra step to the framing process for certain panels. Hence, the height difference between a panel and its preceding panel should be examined as an influencing factor. The framing sequence of panels should be also included in the final list of factors to account for any other correlations between cycle times of subsequent panels. Additionally, the operator mentioned that events occurring on certain days may affect productivity (Factor 21). Thus, the framing date should also be considered to better understand cycle times.

Table 4-1. Results of semi-structured interview.

Class	Factors and operator's comments	Effect?
Product	**1. No. of single studs**: - **2. No. of double studs**: "They take more time to nail than single studs as they require more nails." **3. No. of L-shaped studs**: "They also take more time to nail than single studs as they require more nails." **4. No. of multi-ply studs**: "They take more time to nail than the previous three types of studs, and they take more time to nail with every additional ply." **5. No. of regular doors**: "They could take about 6 times longer to nail single studs." **6. No. of large doors**: "They could take about 10 times longer to nail single studs." **7. No. of garage doors**: "I need to do some manual work for garage doors, so they take much longer than large doors." **8. No. of regular windows**: "They could about 6 times longer to nail single studs." **9. No. of large windows**: "They could take 10 times longer to nail single studs." **10. No. of cuts**: "Cutting takes about as long as nailing single studs." **11. No. of drill holes**: "The time needed to drill a hole is close to the time needed to nail single studs." **12. No. of blocks**: - **13. No. of components**: -	✓

Table 4 1. Results of semi-structured interview (continued).

Class	Factors and operator's comments	Effect?
Product (continued)	**14. Panel length**: "Longer panels typically comprise multiple wall panels that are grouped together. As such, they necessitate a higher number of cuts, but this effect is correlated with the number of cuts per panel."	C
	15. Panel height: "This factor may affect cycle time in two ways. First, for higher wall panels, all panel elements (e.g., stud) are heavier. Whether this factor affects or does not affect cycle time depends on each worker. Some workers may find it harder to lift and load longer elements while other workers may not be affected. Second, I should adjust the machine's width between panels of different heights. This task is not required when a batch of panels of equal height are framed sequentially. Moreover, before I can adjust the machine, I must be able to push the completed panel downstream which means that the downstream station must be available. This task adds additional time to the cycle time for certain panels."	✓
	16. Panel thickness: "A thicker panel is composed of thicker elements (e.g., 2×6 studs versus 2×4 studs). First, thicker elements are heavier and may be more difficult to lift and load. Second, thicker elements require a larger number of nails."	✓
	17. Availability of shop drawings: "Shop drawings are always made readily available before they are needed."	X
	18. Wall panel design complexity (It reflects the variety of framing tasks that the operator must complete for a wall panel): "Aside from the varying time required by each type of element (e.g., single stud versus door), having a panel composed of single studs only versus a panel with a mix of various elements does not affect cycle time as the same steps are followed to load each element and run the machine."	X
	19. Quality of shop drawings (i.e., dictates the frequency of errors + delay + rework time if any): "This factor has a high impact on cycle time. The framing machine cannot read drawings with errors. As a result, I have to stop the work, inform the drafter, and wait for the revised draft before work can be resumed."	✓
Worker	**20. Work shift (i.e., morning vs. afternoon) (which could relate to fatigue)**: "This may have an effect, but it depends on each worker and the workstation. For instance, younger workers may work faster at the beginning of the day and start slowing down throughout the day. Meanwhile, older workers may be more consistent in their speed throughout the day. Moreover, when the workstation is semi-automated, the worker's pace may be dictated by the machine's pace, which increases consistency. Sometimes, however, random events may happen throughout the day, and workers could become mentally drained in the afternoon."	✓

Table 4 1. Results of semi-structured interview (continued).

Class	Factors and operator's comments	Effect?
Worker (continued)	**21. The day of the week (which could relate to work motivation or fatigue accumulation)**: "Monday mornings may be less productive as workers return from weekends, which may involve disrupted sleep schedules, alcohol, etc. Tuesdays are more productive as workers become dialled in. Thursdays (given that the company has a four-day work week) may be also productive because workers are motivated to finish their work earlier and start their weekend. Regarding fatigue accumulation, this factor may be more critical in the summer as workers get tired more quickly in higher temperatures and may get less rest after work. This means that their bodies may recover less between workdays, which may lead to fatigue accumulation."	√
	(Note: the operator's comment on this factor was generic and is not applicable to the case workstation given the operator's long years of experience. **22. Learning curve**: "This factor has a high impact on cycle time, but it varies among workers. Some workers are fast learners and retain knowledge, while others constantly seek help from others, thereby increasing cycle times."	X
Machine	**23. Breakdowns**: "Some breakdowns result in complete work stoppages while others may only cause minor interruptions. For instance, the nail gun may occasionally shoot double nails, requiring extra work to cut the defective nails each time it occurs. Although this extends the framing process, it does not entirely halt production. These issues may occur approximately once every two weeks. In contrast, machine failures that require complete shutdown may last anywhere from 15 minutes to an entire day, and may occur approximately once every six months.	√
	24. Errors: "The machine may result in errors (e.g., nailing defect); a couple of minutes may be spared per incident."	√
	25. Nail gun refills: "The machine's nail gun was replaced with a new one of a different brand, but the new one must be refilled more frequently. Nail refills add more time to the cycle time for certain panels."	√
	26. Motion speed: "The machine has a constant motion speed."	X
Material	**27. Material type**: "Different types of materials (e.g., Laminated strand lumber (LSL) versus Spruce wood) vary in weight (e.g., LSL is heavier than Spruce), and heavier elements may be more difficult to lift and load."	√
	28. Delays in raw material supply: "There are no delays related to raw material supply."	X
	29. Delays in material preparation activities (e.g., sub-assembling door openings): "Material preparation activities are completed one shift or half a shift before the material is needed."	X

Table 4 1. Results of semi-structured interview (continued).

Class	Factors and operator's comments	Effect?
Workstation setup	**30. Distance between material inventory location and installation location**: "I must reach the nail inventory location every time a nail gun refill is needed, but this factor is correlated with the "nail gun refills" factor. All the other materials are reachable from my work location."	C/X
	31. Distance between tools location and workstation: "All the tools are located in a way that I can reach them without travelling."	X
Production line	(Note: the framing workstation is the first workstation on the wall production line) **32. Delay at downstream workstations**: "While waiting for the downstream workstation, I could start setting up the machine for the following panel instead of standing idle."	✓
Factory operations	**33. Workload – Sq. Ft. per day**: "If the workload is low, the workers may become slower. Meanwhile, high workload may have two outcomes depending on the worker; while some workers may become faster trying to finish the scheduled work during working hours, other workers may become overwhelmed with the increased workload which, in turn, adversely affect their productivity."	✓
	34. Overtime shift: "It depends on each worker. My speed during overtime shifts and regular shifts is consistent if overtime shifts are occasional."	X
	35. Weekly cumulative overtime: "In case of multiple overtime shifts during a week, workers do not have enough time to recover and become less productive."	✓
External factors	**36. Ambient temperature**: "When the temperature exceeds 20 °C, workers get tired more quickly and become slower since there is no air conditioning in the factory."	✓

Following this approach, the modeller will have a set of factors that are highly likely to influence cycle time at a given workstation (highlighted in green), another set of factors that are likely to influence cycle time (highlighted in orange and yellow), and a third set of factors that show minimal or zero likelihood of influencing cycle time (highlighted in red). The modeller will also have a good understanding of how these factors could influence cycle time and, hence, will be better positioned to rationally interpret the performance and the results of a machine-learning model developed to predict cycle times at the workstation under study. This facilitates the development of prediction models that more accurately capture the complexity of the process under study. It is important to note that not all influencing factors will become part of the prediction model. Some factors may be excluded due to various reasons such as data unavailability, an insufficient sample size, or weak correlations with cycle time compared to other factors. The modeller will nevertheless have an awareness of the potential effect of the excluded factors on the results. To further demonstrate the importance of following such a systematic approach for identifying influencing factors, let us consider a brief overview of the results obtained for building a model that predicts processing times (excluding delay times) at the case framing workstation. A multi-layer feedforward artificial neural network model was trained and cross-validated (using a 10-fold cross-validation) using data collected on 172 wall panels framed at the case framing workstation. The case company estimates the capacity of the workstation in linear metres per minute (m/min), so only panel length was used as a predictor variable in the first model. Based on cross-validation results, the mean absolute error was found to be 2.18 min. Adding the geometric properties of the given panel (i.e., Factors 1–16 in Table 4-1) reduced the error to 1.94 min, resulting in an 11% reduction in the error. Moreover, considering the complexity, day, shift, temperature, height difference, framing sequence, and date factors further reduced the error to 1.80 min, resulting in a total error reduction of 17%. The details of this neural network model are presented in Chapter 5 of this thesis. Nevertheless, this brief overview of the results serves to highlight the value of dedicating time and effort to identifying and understanding the factors that influence process cycle time; Having a comprehensive pool of influencing factors is vital for the development of more accurate prediction models. As such, following the same approach for identifying influencing factors and building prediction models for different workstations, the modeller gains a deeper understanding of what factors drive cycle time variability and becomes well positioned to analyse cycle times across workstations. This, in turn, can facilitate workload balancing across workstations to ensure leaner operations.

4.5 Conclusions

This chapter presented a structured approach for identifying and understanding the factors influencing cycle times at workstations in offsite construction factories, an essential step toward more accurate analysis of cycle times across workstations for the purpose of balancing production lines. The application of the approach was demonstrated in reference to a semi-automated, wood-wall framing workstation in a panelized manufacturing factory in Edmonton, Alberta, Canada. A total of 36 potential factors categorized into eight classes were identified based on observation, a cross-functional diagram of the process, and a literature review. These factors were further investigated based on the input of the workstation operator solicited in a semi-structured interview, and the factors were further discussed in light of the interview results. A brief demonstration of their effect on the performance of an artificial neural network model was presented, where using more factors as prediction variables in the model reduced the mean absolute error by 17%. In short, this study demonstrated the value of expending effort on the identification and understanding of the factors influencing cycle times at workstations in offsite construction. Doing so can be expected to aid in streamlining and improving the accuracy of cycle time analysis for the purpose of applying *Heijunka* and balancing production lines, thereby minimising instances in which workers find themselves playing the role of the "hare".

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 5

Data- and Knowledge-Driven Cycle Time Estimation in Offsite Construction Factories

5.1 Introduction

5.1.1 Cycle time variability in offsite construction factories

The construction sector has exhibited a trend towards increasing adoption of offsite construction (also known as “prefabricated construction” or “construction manufacturing”), which involves fabricating building components in a controlled factory setting and then transporting them to the construction site for assembly. Shifting towards manufacturing entails adopting principles and methods from manufacturing. In essence, a manufacturing system refers to a "combination of humans, machinery, and equipment that are bound by a common material and information flow" (Caggiano, 2014). As such, workers, machinery, and equipment in offsite construction factories are typically positioned in fixed locations at workstations, while production processes (e.g., wall framing, windows/doors installation) are distributed across these workstations. However, in contrast to traditional mass production, offsite construction factories experience significant variability in the cycle times of their processes at these workstations, with cycle time referring to the time spanning from the start to the end of a process cycle. For instance, a recent case study of a panelized construction factory (in which wall panels, floor panels, roof components, and staircases are prefabricated for shipment to the site for installation) reported cycle times for wood wall framing operations ranging from approximately 1 min to as much as 48.5 min (Alsakka et al., 2023b). This high variability arises from a variety of influencing factors related to the components, workers, machines, materials, workstation setup, production line, factory operations, and external circumstances (Alsakka et al., 2023c).

Considering the wide range of cycle times and the multitude of factors influencing them, the current practice in offsite construction of relying on average production rates (such as square footage of a building component per minute) to estimate production time (i.e., the total time required for producing building components) and create production schedules poses certain difficulties. To elaborate, the total production time is contingent upon cycle times at various workstations constituting the entirety of production operations. Consequently, overly optimistic cycle time estimates can lead to an underestimation of the time necessary for producing building components, thereby yielding overly optimistic and impractical production schedules. Therefore, an estimation approach is needed whereby process cycle time is estimated for each workstation as a function of the factors influencing it. An overview of related estimation methods explored in previous research is provided in the subsequent subsection.

5.1.2 Process time estimation methods

Numerous data-driven methods of estimating the durations of construction processes (e.g., machine learning, simulation) have been proposed in the literature, and a variety of different factors influencing the productivity of these processes have been incorporated into the estimation models to improve accuracy. For example, Chao (2001) used neural networks (NNs) combined with a simulation model to estimate cycle times of earthmoving operations based on a variety of factors related to the equipment used (e.g., weight of truck), the site conditions (e.g., soil type), and the nature of the operation (e.g., swing angle from truck to cutting location). Zayed and Halpin (2004) built simulation models of the piling process to estimate the corresponding productivity, considering equipment-, site-, and operations-related factors. Zayed and Halpin (2005) later conducted a similar study to estimate piling process productivity using artificial neural networks (ANNs), while Tam et al. (2002) used NNs and regression models to estimate hoisting times of tower cranes considering similar types of factors. In a later study, Tam et al. (2010) estimated hook times of mobile cranes as a function of operations and load weight. Many other studies that followed the same line of thinking can be found in the literature (e.g., (Chao et al., 1994; Lu et al., 2000)).

Similar estimation methods have been used in industrial construction and offsite construction. For example, Hu et al. (2015) followed a similar approach for estimating man-hour requirements for steel fabrication. They trained regression models to estimate man-hours mainly based on design properties, such as the lengths of beams, the quantity of bolts, and the weight of the structure under fabrication. Song and AbouRizk (2008) trained an ANN model for estimating steel-fitting activity duration in steel fabrication as a function of design-related factors, such as weight and length, worker skill, as well as work shift. In another study, NNs and multivariable linear regression (LR) models were applied to precast concrete production to estimate productivity in consideration of influencing factors related to product shape, material, and manpower (Benjaoran et al., 2004, 2006). Conducting research on productivity in wood panelized construction, Mohsen et al. (2022) trained a number of different machine-learning algorithms to estimate the time required to complete processes on a wall production line as a function of design-related factors (e.g., length, width, number of studs) and factors related to work in progress (e.g., the count of wall panels being processed on the production line). In another recent study in panelized construction, this one targeting the transportation phase, Ahn et al. (2020) trained support vector regression models using GPS data in order to predict transportation time for a given project as a function of product-related factors such as the total floor area and total wall area, as well as site-related factors such as location and the maturity of the neighbourhood.

However, there is a lack of research that focuses on estimating process cycle time at the workstation level in the production phase for offsite wood construction. A relevant study is the one carried out by Shafai (2012) in wood panelized construction. In this study, LR models were developed to estimate the durations of specific tasks (such as spray foam insulation) as a function of the unique design properties of the given panel that are significant to the given task (e.g., number of studs, number of openings, number of cutting zones). Stochastic factors (e.g., triangular distribution) were also incorporated in the regression models to account for uncertainties such as worker performance or machine breakdown. Similar estimation methods have been used in other studies to estimate the processing times at workstations (Altaf et al., 2014; Bhatia et al., 2019; H.

Liu et al., 2015). Such approaches constitute an improvement upon using a single average value to model the duration of an entire process at a workstation. Nevertheless, further research with a specific focus on cycle time estimation in offsite construction factories is warranted. This is due to the vital role cycle times play in various aspects of production planning and scheduling (as discussed above), as well as in efficiency and optimization in manufacturing settings. This rationale aligns with the availability of research addressing cycle time estimation in other manufacturing settings (e.g., Chen, 2013; Sun et al., 2022; Tai et al., 2012). Moreover, three related research areas warrant further exploration in the context of cycle time estimation in offsite construction, as explained in the following subsection.

5.1.3 Research areas for further investigation

5.1.3.1 Cycle time-influencing factors

Based on the above, the approaches devised for estimating various process time variables mostly rely on machine-learning techniques. However, as argued by Benjaoran et al. (2004), the reliability of machine-learning models is a function of the exhaustiveness of the influencing factors considered in the models. In this regard, a variety of different factors have been considered in the above mentioned studies, including (1) product-related (or design-related) factors such as number of single studs, double studs, doors, windows, cutting zones, drill holes, nails, screws, etc., for the production of wood wall panels (Shafai, 2012), number of fittings and cut-outs for steel fitting (Song et al., 2008), number of bolts, length of weld, length of wide flange beams, etc., for structural steel manufacturing (Hu et al., 2015), and nominal height, weight, and width, concrete volume, finishing area, reinforcement weight, concrete strength, etc., for precast concrete production (Benjaoran et al., 2004); (2) worker-related factors such as the number of workers (Benjaoran et al., 2004), and skill level (Song et al., 2008); (3) material-related factors such as length and weight (Hu et al., 2015; Song et al., 2008); (4) machine-related factors such as breakdowns (Song et al., 2008); (5) factory operations-related factors such as work shift (Song et al., 2008); and (6) production line-related factors such as activity precedence relationships, queuing, and rework (Song et al., 2008).

The benefit of investing effort on identifying these factors is twofold; it not only helps to identify factors that could hold significant information with regards to the estimated process time variable, but also deepens the estimator's understanding of the process under study. This, in turn, allows the estimator to follow a prescriptive approach for selecting and representing influencing factors used in the machine-learning estimation models (Kuhn et al., 2019). In other words, it enables the estimator to carry out knowledge-driven modelling alongside data-driven modelling, reducing the risk of overfitting to erroneous data patterns or of generating models that cannot be rationally interpreted (compared to an approach that relies solely on empirical data) (Kuhn et al., 2019). Therefore, it is valuable to consider a range of influencing factors and to study their impact on the performance of machine-learning models in estimating cycle times in offsite construction factories.

5.1.3.2 The need for continuous system tuning

As discussed above, various factors pertinent to the components, workers, machines, materials, workstation setup, production line, factory operations, and external factors influence cycle times. However, if a particular factor remains constant during the timeframe covered in the training dataset, even if it has a significant impact on cycle time, machine-learning models will fail to capture its correlation with cycle time. For instance, suppose an expert worker at a particular workstation is substituted with a less-experienced, less-efficient new employee. In such a scenario, if the dataset used to train machine-learning models for predicting cycle time at that workstation only includes data collected during the tenure of the skilled worker, the models are likely to underestimate the time required to finish a cycle performed by the new employee. In other words, substituting the workers will introduce a new source of variability in cycle time that the models were not trained to capture. In this case, the models will need to be retrained to account for this new variance associated with the employee performing the work. Generally, in order to sustain and enhance the performance of the machine-learning models used to predict cycle times, regular training of the models is necessary. However, consistent training necessitates the ongoing acquisition of training data or, in other words, automated data acquisition. This can be facilitated through the use of computer-vision technology, which has shown promise in the context of offsite construction (Alsakka et al., 2023a).

Computer vision is a branch of artificial intelligence focused on creating autonomous systems capable of imitating specific tasks executed by the human visual system (Huang, 1996). In computer vision, useful information is extracted from visual components (i.e., digital images, videos, cameras, and closed-circuit television (CCTV)) and is analyzed to facilitate informed, data-driven decisions and recommendations (IBM, 2022). The field of computer vision has experienced substantial expansion in recent years and is expected to continue growing in the future (Data Bridge Market Research, 2022; KBV research, 2020; Verified Market Research, 2021). The technology has been successfully implemented in the offsite construction industry for automatically collecting productivity- and progress-related data. For instance, Alsakka et al. (2023b) deployed the technology to automatically measure the start time, productive time, and cycle time of a wood-wall framing process at a panelized construction factory, achieving a mean absolute error (MAE) of less than 1 min. Zheng et al. (2020) used the technology to monitor and track the installation of modules in modular construction projects, as well as to track the duration for which modules are detected in a designated region of interest, achieving a 97.7% accuracy. Wang et al. (2021) deployed computer vision for automatically tracking the hoisting and placement of precast concrete wall panels, achieving a precision and recall of 88% and 89%, respectively. Similarly, another study used computer vision to gather timestamp data for precast concrete wall panel installation operations (Wang et al., 2021). They succeeded in correctly capturing timestamp data for 10 out of 12 walls using this method (Wang et al., 2021). Moreover, computer vision has been employed for measuring the installation rate (cm^2/min) of prefabricated panels in panelized construction, achieving an error rate of less than 5% (Ahmadian Fard Fini et al., 2021). Finally, Martinez et al. (2021) implemented computer vision to track the progress, measure the duration, and calculate the man-hours expended on floor panel fabrication in a panelized construction facility, achieving an overall accuracy of over 92%. Given that computer-vision technology has achieved promising performance as exemplified in the aforementioned studies, a further investigation of the performance of machine-learning models trained on computer vision data for estimating cycle times is warranted.

5.1.3.3 Machine-learning algorithms employed

Another area that merits further exploration relates to the performance of the machine-learning algorithms used in the estimation approaches. While existing methods use LR for building the estimation models, the potential misuse of regression and correlation analysis when the assumptions that underlie them do not hold has been considered in earlier related research (Lu, 2000). Examples of assumptions that may be critical for cycle time estimation in offsite construction given the amount of variability present in operations include the linearity between the predictor variables (i.e., the influencing factors) and the response variable (e.g., cycle time), the independency of observations, and the constancy of the standard deviation and variance of the residuals (i.e., the difference between an observed value and the corresponding predicted value on the regression line) for all values of the predictor variables (Casson et al., 2014). Alternatively, NNs are known for their capability to model complex problems, which can be difficult to model using traditional classical mathematical methods (Adeli, 2001). Hence, NNs have been long recognized as a suitable tool for modelling problems in construction research (Moselhi et al., 1991), and have found extensive use across various applications, as discussed in the previous section (e.g., (Benjaoran et al., 2004, 2006; Lu et al., 2000; Song et al., 2008)). An NN, it should be noted, is defined as "an interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal neuron" (Gurney, 1997). Various NNs have been developed to imitate desirable characteristics of the human brain such as its learning ability, generalization capability, and adaptivity (Jain et al., 1996). In the aforementioned study by Mohsen et al. (2022), however, among the models considered—i.e., random forest (RF), LR, *k*-nearest neighbour, and NN—the LR model was found to perform slightly better than the NN model when trained on an engineered dataset to predict the production time of wall panels, and the best performing model was the RF model (Mohsen et al., 2022).

In light of this and given the variety of influencing factors to be considered in the development of estimation models in this study, further examination of the performance of multiple machine-learning models in estimating cycle times at the workstation level in wood offsite construction, while considering different influencing factors, is warranted. Specifically, based on the results obtained in previous works as explained above, the use of NN, LR, and RF models for cycle time estimation will be explored.

5.1.4 Study aim, objectives and contribution

The aim of the research presented in this chapter, then, was to develop a data- and knowledge-driven system that estimates cycle times considering various influencing factors and using automatically collected data while increasing the estimation accuracy compared to traditional estimation methods. The system was designed to be trained using data collected through a computer vision system and to use machine learning, statistical modelling, and 3D simulation techniques for estimating cycle times at workstations considering a set of cycle-time-influencing factors. The system's performance was examined through its application to a semi-automated wood-wall framing workstation in a panelized construction factory. In relation to the research needs highlighted above, the secondary objectives underlying this research and the corresponding contributions are as follows:

(1) Examine the effect of considering a variety of influencing factors on the performance of cycle time-estimation models: Doing so helps to demonstrate the importance of expending time and effort on the identification and understanding of influencing factors prior to building prediction models.

(2) Explore the reliability of using data collected automatically through computer vision to train the estimation models: The findings of this task can shed light on the extent to which we can rely on automatically acquired data for building estimation models.

(3) Examine the use of different machine-learning algorithms, including the feed-forward ANN, LR, and RF algorithms, for cycle time estimation considering various influencing factors: As discussed in the literature review section above, there is no consensus regarding the behaviour of these machine-learning algorithms in the context of process time estimation applications. Therefore, there is merit to gaining a better understanding of how these models perform given various influencing factors in the context of such applications.

The rest of this chapter is organized as follows: Section 5.2 provides definitions of the set of process time variables relevant to the developed estimation system, as well as a description of the system and its architecture. Section 5.3 outlines the procedure followed and methods used to develop the system for the case framing workstation. Section 5.4 presents the evaluation results of the system's performance. Section 5.5 presents an additional analysis of the results with regards to the use of computer vision data, the influencing factors used, the performance of the machine-learning algorithms, and the effect of unpredictable delays on estimation systems. Finally, Section 5.6 summarizes the findings of the study, discusses implications for the industry, lists the limitations, and suggests avenues of future research.

5.2 System description and architecture

5.2.1 Process time variables

For the purpose of this study, the following definitions of process time variables relevant to the estimation system to be developed were considered. (It should be noted that variations in the definitions of these variables may be found in the literature.)

- Cycle time (CT): CT refers to the total time spanning from the start of the process undertaken at a workstation for a given component until the end of the process, where a "cycle" refers to the set of tasks assigned to a workstation for a single component (e.g., one wall panel). CT is a function of two variables:
- Processing time (PT): PT is the time spent by resources processing a component during a cycle at a workstation. Under ideal conditions, CT is equal to PT.
- Cycle delay (CD): CD is the time during which work is not performed on the component during a cycle at a workstation. In other words, it is the amount of time it takes a cycle to be completed beyond the expected completion time, which is PT. We can further differentiate between two types of delays: predictable cycle delays (PCD) and unpredictable cycle delays (UCD). PCD refers to interruptions to a cycle that can be anticipated and estimated to a certain extent. Examples of PCD include scheduled breaks, meetings, training sessions, predictable unavailability of resources, scheduled maintenance, and waiting for a slow material preparation process. UCD, on the other hand, arises from random events such as machine breakdowns, machine malfunctions, errors in shop drawings, power outages, worker injuries, phone calls, conversing with co-workers, and bathroom breaks.

Given these definitions, CT is calculated satisfying Eq. (5.1).

$$CT = PT + CD = PT + (UCD + PCD) \tag{5.1}$$

- Inter-cycle total delay (ITD): ITD is the total time spanning from the end of a cycle at a workstation to the start of the subsequent cycle. ITD is a function of the following variables:
- Downstream-related waiting time (DW): DW is the time spent waiting for the completed component to be transferred to the downstream workstation. Specifically, it is the time spanning from the end of a process undertaken at a workstation to the time at which the component is transferred downstream. Various scenarios could result in DW. One such scenario is when the downstream workstation is busy and there is no inventory between the two workstations or there is inventory that is already at full capacity. Another example scenario that could result in DW is when the resources responsible for transferring the component between workstations are busy with other tasks. Although DW is not factored into the CT for a given cycle, it affects the start time of the subsequent cycle.
- Upstream-related waiting time (UW): UW is the time spent waiting (after a completed component is transferred to the downstream workstation) for the upstream workstation to complete work before a new cycle can be started. This occurs when a given workstation is faster than the upstream workstation(s). Since it occurs before a new cycle is started, UW, like DW, is not factored into CT.
- Inter-cycle additional delay (ID): ID is the time by which the start of a new cycle is delayed beyond DW and UW due to any of the aforementioned reasons that cause CD. Note that the total duration of the related delay may be longer than ID, but it may overlap with DW and UW, which is why ID, as defined herein, specifically refers to the additional delay that exceeds the durations of DW and UW. Like CD, ID can arise from both predictable and random events, generally rendering it a random occurrence.

Given these definitions, ITD is calculated satisfying Eq. (5.2).

$$ITD = DW + UW + ID \tag{5.2}$$

5.2.2 System architecture and components

The system was designed to be capable of continuously learning from actual process time data and of predicting the process time variables for each component processed at a given workstation as a function of relevant influencing factors. The system's architecture is displayed in Fig. 5-1, and the major components it comprises are the following (more details are provided in Section 5.3):

(1) A computer vision system for actual data collection at each workstation: As discussed above, various influencing factors contribute to CT variability in production factories. For this reason, the estimation system should be trained regularly in order to better capture the variability arising from various influencing factors, thus a continuous stream of data from the production factory is required. Therefore, the estimation system uses the computer-vision-based time data acquisition system developed in a previous work (Alsakka et al., 2023b) for automated data acquisition. The system automatically acquires data on the cycle start time, productive time (i.e., the time actually spent by resources working on a component at a workstation—equivalent to PT in the context of this study), and CT for each component processed for a given operation at a workstation.

(2) A prediction model for PT at each workstation: PT is predictable to a certain degree when the factors that influence its value for a given component at a workstation are known. The degree of predictability, however, may fluctuate at the workstation across different time frames. This is because most of the operations in offsite construction factories are still labour based. Labour-based tasks, even if they are well-defined and standardized, are still subject to high variability because of the inconsistency of human productivity. In fact, PT can even vary for the exact same task depending on the worker's physical health, mental health, work environment, motivation, and other factors influencing their pace of work. Due to this variability, PT prediction can be highly complex at certain workstations. Nevertheless, machine-learning models can be leveraged to model such complexity. As such, the system developed in the present study uses machine-learning models to predict PT as a function of influencing factors at workstations.

(3) Estimation models for UCD and ID at each workstation: Given the random nature of the events causing UCD and ID, probability distributions were used to model these variables.

(4) A 3D simulation model for PCD, DW, and UW: Workstations on a production line affect one another, as the above discussions on DW and UW serve to demonstrate. A simulation model can be used to model the interdependencies among different workstations that determine the durations of UW and DW (which, in turn, affect the start times of cycles). A simulation model can also model the dependencies of workstations on materials/resources, scheduled events, and other forms of PCD. In the present study, a 3D simulation model was used in the estimation system to leverage the benefits associated with its realistic visual representation of the real factory. (Specifically, the 3D visual representation allows the user to determine whether the model is error-free and to validate its representativeness of reality. A 3D model that is developed to be representative of reality also helps users to better understand and analyze the real manufacturing operations.)

The system was developed to include a database serving as a data hub for storing (1) actual process time data measured using the computer vision system, (2) daily production lists of scheduled jobs, (3) panel design properties (e.g., panel length, number of studs) extracted from BIM models, and (4) data concerning other influencing factors.

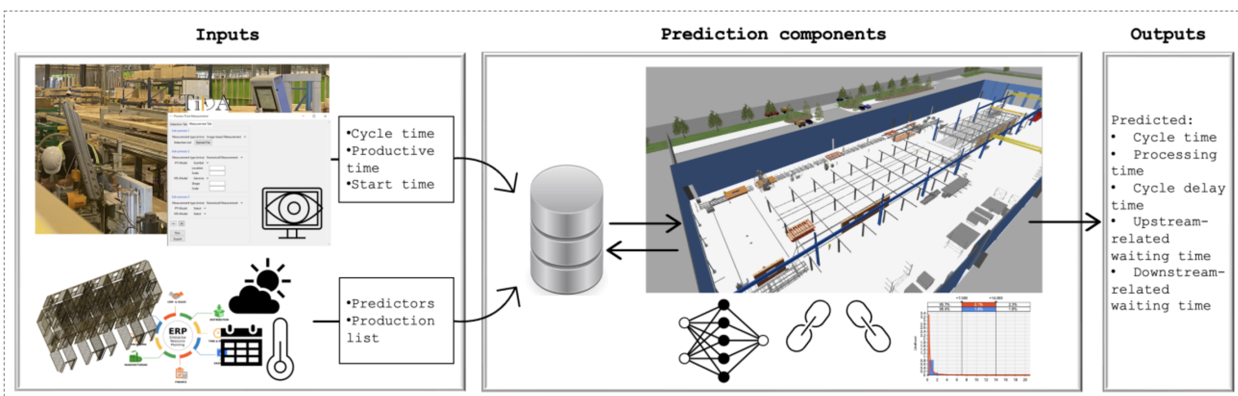


Fig. 5-1. System architecture.

5.3 Procedure and methods for system deployment

The procedure followed to deploy the system is described in reference to a wall framing workstation at a lightweight wood panelized construction factory. The framing workstation, it should be noted, is the first workstation on a wall production line that comprises several workstations. A portion of this production line is displayed in Fig. 5-2. First, a semi-automated wood-framing machine is used to frame wall panels at the framing workstation, automatically performing nailing, drilling, and cutting operations. An operator loads the machine with framing elements (e.g., single studs, double studs, subassemblies for large doors) when prompted to do so by the machine, which then performs the operations. The wall frame is then transferred to the downstream workstations for further processing. This section describes the steps followed and methods used to deploy the above-described estimation system for the case framing workstation.

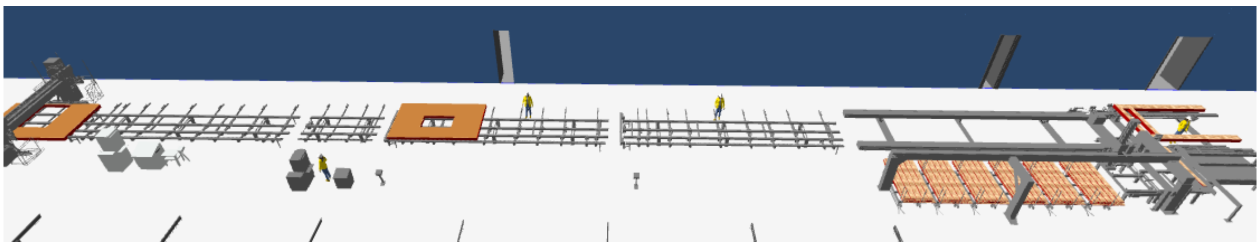


Fig. 5-2. A portion of the wall production line.

5.3.1 Step 1: Deploy computer vision for automated data acquisition and manually collect data for testing

The computer vision system developed in a previous work (Alsakka et al., 2023b) to automatically measure a process' start time (ST), PT, and CT was deployed for automated data acquisition in the CT estimation system. Data is automatically acquired using computer-vision technology through the following approach: In offsite construction factories, work-in-process (WIP) or material flows into and out of workstations in a cyclic manner. This causes specific points along the workstation to cyclically become blocked and unblocked by WIP or material. By strategically placing objects, such as stop signs, at these points, these objects alternate between being blocked and unblocked with each cycle as shown in Fig. 5-3. Hence, the detection status of these objects can be correlated with the start and end of a cycle, and thus, with ST and CT. This logic was employed to enable the use of object detection algorithms that have been extensively trained to detect objects commonly found in our everyday life using one of the open-source datasets containing a large volume of annotated images of such objects (e.g., (COCO Consortium, 2022)). This was aimed to reduce the significant amount of time and effort needed to train object detection algorithms to detect building elements, which may be challenging as they change in shape and size while progressing through the production line. As for productive time, the system assumes that when the worker is detected during a cycle at the workstation, the workstation is actively in use.

The system's performance was evaluated with reference to its application to the case framing workstation and was found to measure the framing process' ST, PT, and CT with mean absolute errors that are less than 1 min. In that application, the object detection model YOLOv4 (the fourth version of the "You Only Look Once" object detection algorithm) (Bochkovskiy et al., 2020) trained on the COCO dataset (COCO Consortium, 2022) to detect commonplace objects was used. Hence, for consistency, the data used for building estimation models in the present chapter was also collected based on detections made by YOLOv4. Detailed descriptions of the system's logic, application, and performance can be found in the previous work (Alsakka et al., 2023b).

Fig. 5-3. (a) Stop signs unblocked; (b) stop signs blocked (Alsakka et al., 2023b).

The computer vision system was used for automatically collecting data on +200 wall panels framed at the framing workstation. Actual UCD values can be computed by subtracting PT and PCD from CT following Eq. (5.1). Over the course of the study period, PCDs were limited to scheduled breaks for the case framing workstation. This is because there were no scheduled events that interrupted work, the needed resources were consistently and exclusively dedicated to this workstation (i.e., they are not shared with other workstations), and all of the necessary materials were consistently made ready before they were needed. As such, actual UCD values for the wall panels were measured by subtracting PT and scheduled breaks from CT. To compute ID, ITD can be first computed by subtracting the finish time of a given cycle from the start time of the subsequent cycle. Next, DW and UW can be computed, if they are not null, by considering the finish times of cycles at the upstream and downstream workstations. Finally, the actual ID can be determined by subtracting DW and UW from ITD, according to Eq. (5.2). Since the framing workstation is the first workstation on the production line, its UW is null as there is no upstream workstation. Meanwhile, the computer vision system has not been implemented yet for the downstream workstation, so no data could be obtained on DW in the present case study. Due to the unavailability of DW data, it was assumed that ID was equal to ITD for the framing workstation for the purpose of building an estimation model for ID. This assumption is not critical, however, since the role of ID in CT estimation is limited to affecting a cycle's start time which, in turn, may occasionally affect the value of two predictor variables, namely average hourly temperature and work shift, which are used in the prediction of framing CT as described below.

For testing purposes, the variables were also measured manually based on recorded videos of the framing process for additional 40 wall panels.

5.3.2 Step 2: Identify influencing factors

A multi-stage procedure was followed to devise a list of factors influencing framing CT as described in detail in a previous study (Alsakka et al., 2023c). To summarize the procedure, first, a cross-functional diagram of the framing process was mapped in order to gain a deeper understanding of the tasks involved in the process, the resources allocated to it, the manner in which the tasks are carried out, the process inputs, and the process outputs. Eight classes of factors were identified accordingly: "product", "worker", "machine", "material", "workstation setup", "production line", "factory operations", and "external factors". Then, factors were identified in relation to each class based on the understanding of the process and a review of relevant literature. The identified factors were then compiled as a list, and a semi-structured interview was conducted to solicit the framing workstation operator's input on the list. The full list of factors and the interview results can be found in the previous study (Alsakka et al., 2023c). Following the operator's input, some factors, including the availability of shop drawings, learning curve, etc., were eliminated, as they were considered to not affect framing CT at the case workstation. Moreover, since machine learning was only to be used for modelling PT in our system, delay-related factors were also eliminated. Ultimately, the list of influencing factors was narrowed down to 24 features, which are detailed in Table 5-1, to be considered in the development of the machine-learning models.

The framing workstation operator's input on these features, as described in the aforementioned previous study (Alsakka et al., 2023c), helped to clarify their consideration as features influencing PT. In what follows we briefly outline some of the reasons for their inclusion. First, the date feature was considered because certain events that occur on specific dates could affect multiple panels, allowing the machine-learning model to identify any patterns or variability associated with particular dates. Panel length was also considered, as longer panels typically require more work. The panel thickness feature was included because thicker elements are heavier and require a larger number of nails. Panel height, meanwhile, was considered, as the framing elements of higher wall panels are heavier. The height difference feature (or delta height) was included since the operator needs to adjust the machine's opening when panels of different heights are framed sequentially. In relation to this, the framing sequence feature was considered in order to account for any other potential variations in PT resulting from the sequence of the panels. Moreover, the reason why the different types of studs and openings were separated in the list is that they vary in terms of the required numbers of nails, which, in turn, affects the time it takes to nail the framing element. The wall panel design complexity feature was included to reflect the variety of framing tasks that the operator must complete for a wall panel. These tasks could include loading a stud, loading a double stud, loading an L-shaped stud, loading a multi-ply stud, loading a regular door, loading a large door, loading a garage door, loading a regular window, loading a large window, loading a component, performing a cut, drilling a hole, or manually nailing a block (since this task cannot be performed by the framing machine). The wall panel design complexity feature serves to differentiate between a complex wall panel the production of which would necessitate many of these tasks and a simpler wall panel that would necessitate fewer of these tasks. In other words, this feature was included as a way of determining whether it takes the operator more time to interpret the shop drawings and load the requisite elements for more complex wall panels that require more framing tasks compared to less complex wall panels that require fewer types of tasks even if they are of the same size. The work shift feature (morning versus afternoon), meanwhile, was included as a means of determining whether the operator’s work pace changes throughout the course of the day (due to fatigue, for instance). Similarly, the day of the week feature was included as a way of capturing the dynamics of work motivation and fatigue accumulation over the course of the week. The scheduled workload feature was considered as a means of accounting for the possibility that a higher workload puts pressure on the operator to increase their pace of work. The ambient temperature feature, finally, was included by virtue of its potential effect on work pace.

5.3.3 Step 3: Prepare data, perform exploratory data analysis, and pre-process data

5.3.3.1 Data preparation

With the exception of the wall panel design complexity, shift and ambient temperature features, data on the identified features can be extracted from BIM models and from the company's enterprise resource planning system. In fact, the company's system supports the extraction of production lists of wall panels scheduled on given dates along with panel design properties into a Microsoft Access database. As such, SQL (i.e., structured query language) queries were developed to combine the data collected on process time variables with the data extracted on the features. However, for the period of operations during which the data was collected, the exported database was missing data on a number of wall panels, which were thus removed from the dataset. The final training dataset contained data on a total of 172 wall panels and the testing dataset contained data on 40 wall panels.

As for the wall panel complexity feature, it was modelled using Eq. (5.3) as follows.

$$\begin{aligned} Complexity = {} & x_{studs} + x_{double\ studs} + x_{L-shaped\ studs} + x_{multi-ply\ studs} \\ & + x_{regular\ doors} + x_{large\ doors} + x_{garage\ doors} + x_{regular\ windows} \\ & + x_{large\ windows} + x_{cuts} + x_{drill\ holes} + x_{blocks} + x_{components} \end{aligned} \tag{5.3}$$

where x is a binary variable that takes a value of 1 if the wall panel includes the corresponding element and a value of 0 if it does not. For the ambient temperature feature, average hourly temperatures were extracted from the Time and Date AS database (Time and Date AS, 2023).

Finally, for the shift feature, the value corresponding to each panel was determined based on the start time of its framing cycle.

Following data collection, an exploratory data analysis was conducted to gain understanding of the features, as described in the following subsections. It should be noted that the 24 features included 21 numerical variables and three categorical variables.

5.3.3.2 Exploratory data analysis

5.3.3.2.1 Statistical description

Statistical details about these features, as well as about data on PT and CT are provided in Table 5-1. As presented in the table, the count of non-zero values, mean, standard deviation, minimum value, 25th percentile, 50th percentile, 75th percentile, and maximum value were computed for each feature. Among the various conclusions and observations that could be drawn from these statistics, the ones that stand out with respect to the PT prediction model are the following:

- The range of PT values is wide (i.e., from 1.9 min to 26.8 min), with an average value of 9.3 min for wall lengths ranging from 2 ft to 40 ft and an average length of 30.3 ft.
- The number of panels with large doors, garage doors, preassembled components, double studs, multi-ply studs, and blocks is small relative to the total sample size of 212 wall panels. (Having a small amount of data could compromise accuracy with respect to identifying correlations between the features and the response variable, PT.)
- No wall panels in the dataset were framed on Mondays. As such, any potential effect of this day of the week on PT was not explored.

Table 5-1. Dataset description.

	Outcome							
	Non-zero count	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
PT (minutes)	212	9.3	4.0	1.9	6.7	9.3	11.5	26.8
CT (minutes)	212	11.6	7.5	2.3	7.7	10.2	13.4	58.4

Table 5 1. Dataset description (continued).

	Numerical features								
	Feature	**Non-zero count**	**mean**	**std**	**min**	**25%**	**50%**	**75%**	**max**
1	Length (ft)	212	30.3	11.8	2	25.5	36.2	39.2	40
2	Height (ft)	212	8.6	0.8	5	8	9	9	10
3	Delta height (ft)	51	0.3	0.7	0	0	0	0	4.0
4	Thickness (in)	212	5.1	1.1	3.5	3.5	5.5	5.5	7.2
5	Regular windows	37	0.3	0.7	0	0	0	0	3
6	Large windows	49	0.3	0.7	0	0	0	0	4
7	Regular doors	60	0.5	0.9	0	0	0	1	5
8	Large doors	8	0.0	0.2	0	0	0	0	1
9	Garage doors	8	0.0	0.2	0	0	0	0	1
10	Preassembled components	11	0.1	0.3	0	0	0	0	2
11	Cutting zones	150	1.9	1.9	0	0	2	3	9
12	Drilled holes	164	5.1	3.4	0	2.8	6	7	13
13	Studs	208	15.7	7.5	0	10	17	21	34
14	Double studs	14	0.1	0.3	0	0	0	0	2
15	L-shaped studs	179	2.0	1.5	0	1	2	3	7
16	Multi-ply studs	26	0.2	0.5	0	0	0	0	3
17	Blocks	27	1.1	3.5	0	0	0	0	32
18	Avg hourly temp. (°C)	N/A	3.9	3.4	−5.5	2.5	5	6	12
19	Complexity	212	4.4	1.6	1	3.75	5	5	8
20	Scheduled workload (sf)	212	14,654	2,582	11,098	12,614	14,500	16,492	21,715
21	Panel sequence	212	N/A						

	Categorical features				
	Feature	**Non-zero count**	**Categories**	**Top**	**Freq.**
22	Day	212	['Tuesday', 'Wednesday', 'Thursday']	Tuesday	112
23	Shift	212	['Morning', 'Afternoon']	Afternoon	125
24	Date	212	['2022-03-15', '2022-03-16', '2022-03-22', '2022-03-23','2022-03-24', '2022-03-29', '2022-03-30', '2022-04-05']	2022-03-16	41

5.3.3.2.2 Correlation with PT

Moreover, in order to identify key features affecting PT, scatter plots and Spearman's rank correlation coefficients were analyzed (Fig. 5-4 and Table 5-2). Scatter plots were drawn in order to visualize the relationship between each feature and PT, while Spearman's rank correlation coefficients were computed in order to determine the strength and direction of the correlations. By identifying these correlations, we were able to gain insights into which features were strongly associated with PT and, in turn, use this information to guide the machine-learning process. Of the features analyzed, complexity, length, all types of studs, number of drilled holes, number of cutting zones, blocks, garage doors, and windows all showed clear correlations with PT, while weaker correlations were observed with the other features. (However, it is important to note that an NN model, for example, may still be able to identify more complex relationships between these features and PT.) Based on these results, no features were removed from the analysis, as all may provide valuable information to the machine-learning models. Instead, the correlations identified served to guide the decisions on feature selection in order to optimize the accuracy and efficiency of the machine-learning models.

Table 5-2. Spearman's coefficient.

Feature	Spearman's coefficient	Feature	Spearman's coefficient
Complexity	0.63	Multi-ply studs	0.17
Drilled holes	0.54	Large doors	0.10
Length (ft)	0.49	Regular doors	0.10
Studs	0.42	Double studs	0.10
L-shaped studs	0.39	Height (ft)	0.08
Blocks	0.35	Delta height (ft)	0.05
Cutting zones	0.31	Thickness (in)	0.04
Regular windows	0.28	Scheduled workload (SF)	−0.02
Large window	0.20	Preassembled components	−0.06
Garage doors	0.20	Avg hourly temp. (°C)	−0.11

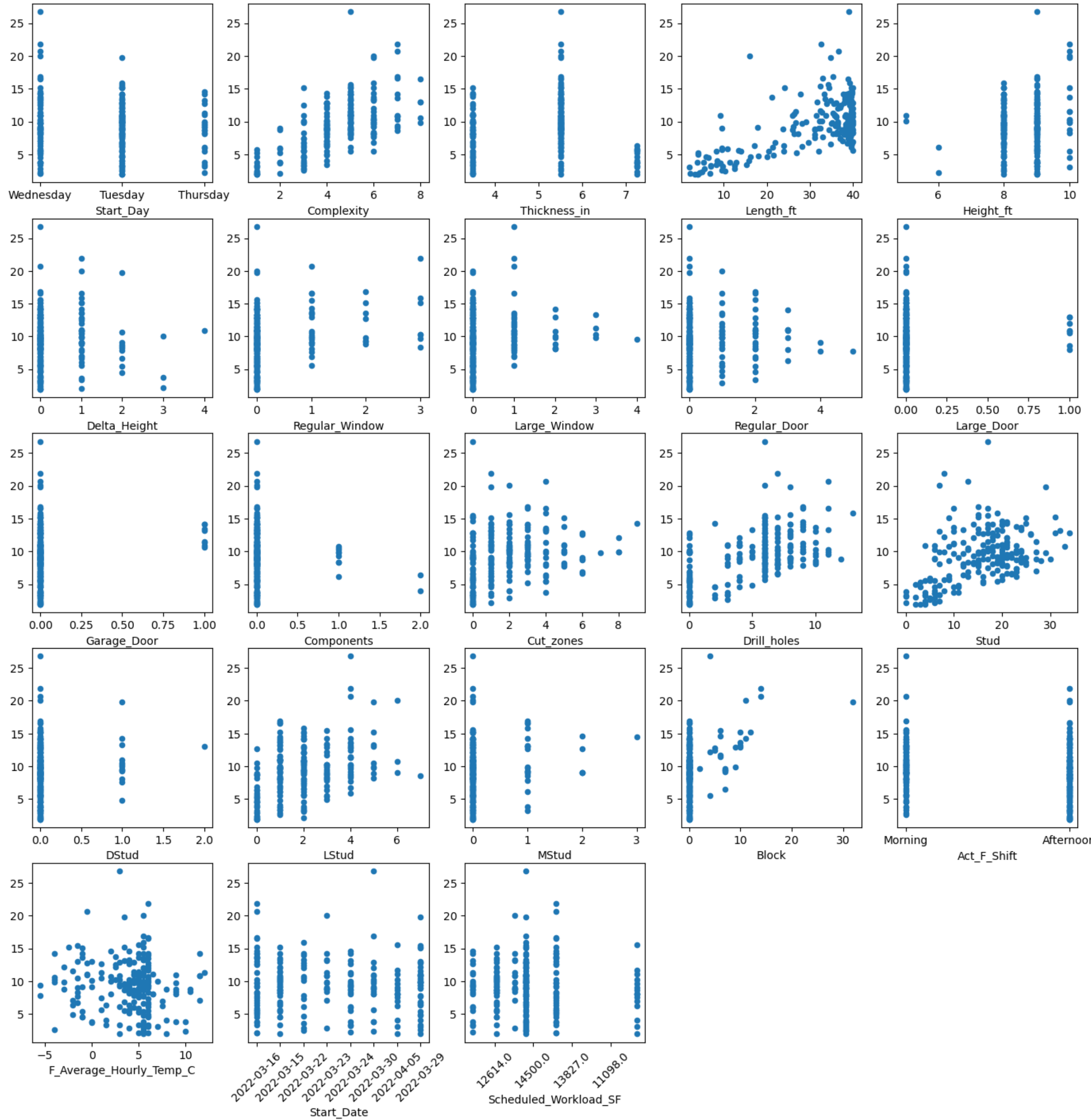


Fig. 5-4. Scatter plots.

5.3.3.2.3 Multicollinearity

The use of LR for predicting PT being one of the prospective solutions to be tested, the variance inflation factor (VIF) was computed for the numerical features in order to measure the degree of multicollinearity between different features. VIF was selected as a diagnostic tool for multicollinearity because it provides interpretable information about the regression coefficients

(O'brien, 2007). For example, a VIF of 10 means that the variance of the coefficient of the i^{th} independent variable is 10 times greater than it would have been if this variable had been linearly independent of the other variables (O'brien, 2007). The results are provided in Table 5-3. As shown in the table, height, length, complexity, scheduled workload, thickness, stud, and number of drilled holes were the features found to have the highest VIF values. This finding can be attributed to the following: (1) longer panels typically have a greater number of studs; (2) longer panels typically require a greater number of holes to be drilled for installing hooks for lifting; (3) the complexity feature is a function of the framing elements, as per Eq. (5.3); (4) the scheduled workload is constant on each date; and (5) longer panels usually consist of multiple interior walls that are framed together, whereas most interior walls are shorter and less thick. No features were removed at this juncture based on VIF results. Rather, the values helped to aid understanding of the results of the LR model.

Table 5-3. VIF for features.

Feature	VIF	Feature	VIF
Height_ft	80.5	Large_Window	2.9
Length_ft	60.7	Regular_Door	2.7
Complexity	48.6	Regular_Window	2.0
Scheduled_Workload_SF	42.2	Start_Day_Thursday	1.6
Thickness_in	36.9	Block	1.6
Stud	31.2	Large_Door	1.4
Drill_holes	11.7	Delta_Height	1.4
Act_F_Shift_Afternoon	6.0	DStud	1.4
F_Average_Hourly_Temp_C	5.8	MStud	1.4
LStud	4.0	Preassembled components	1.3
Cut_zones	4.0	Garage_Door	1.3
Start_Day_Tuesday	3.3		

5.3.3.3 Data pre-processing

To ensure that features with higher scales such as the scheduled workload and length features do not dominate the learning process of the estimation models, numerical features were scaled to have a mean value of 0 and a standard deviation of 1. As for transforming categorical features, the onehot encoding technique was used where each category of a feature was transformed into a binary variable (i.e., separate column) in the dataset. For example, the afternoon shift was added as an additional binary column in which a value of 1 indicates an afternoon shift and a value of 0 indicates a morning shift since there are only two categories for the shift feature. Since a value of 0 in the afternoon shift column indicates a morning shift, including a column for the morning shift category would introduce redundancy in the dataset. Hence, for each categorical feature, one binary variable was dropped to avoid multicollinearity between the introduced binary variables.

5.3.4 Step 4: Select performance evaluation metrics

The following metrics were used to evaluate performance and calculate errors at different stages in the study.

- The prediction error, e, was calculated satisfying Eq. (5.4). This metric was selected to examine whether the predictions tended to overestimate (positive error value) or underestimate (negative error value) true values, as well as to determine the degree of variance of the measures from their true values for each panel.

$$e_i = P_i - A_i \tag{5.4}$$

where e_i is the prediction error corresponding to panel i, P_i is the predicted value for panel i, and A_i is the actual value corresponding to panel i.

- The sum of errors, SE, was calculated satisfying Eq. (5.5). This metric was used to evaluate whether the predictions made for a batch of panels are cumulatively overestimated or underestimated.

$$SE = \sum_{1}^{n} e_i \tag{5.5}$$

where n is the total number of panels used for evaluation.

- The MAE was calculated satisfying Eq. (5.6). This metric was used to determine the average degree of variance of the predictions from their true values, regardless of whether they had been overapproximated or underapproximated.

$$MAE = \frac{\sum_1^n |e_i|}{n} \tag{5.6}$$

- The mean percentage error (MPE) was calculated satisfying Eq. (5.7). This metric was used to evaluate the prediction errors as a percentage of the true values.

$$MPE = \frac{\sum_1^n \frac{e_i}{A_i} \times 100}{n} \tag{5.7}$$

- The root mean squared error (RMSE) was calculated satisfying Eq. (5.8). This metric was used in addition to MAE as it is useful for detecting outlier values since it assigns a higher penalty to larger errors compared to MAE.

$$RMSE = \sqrt{\frac{\sum_1^n {e_i}^2}{n}} \tag{5.8}$$

5.3.5 Step 5: Build PT prediction models

As noted above, the data used for training the machine-learning models was automatically collected using the computer vision system. A 10-fold cross validation was employed to tune the parameters of the models and select features based on RMSE and MAE values. The training results are described in this section, while the testing results are described later in Section 5.4.1.

5.3.5.1 Artificial neural network model

A multi-layer feed-forward ANN model was trained using the open-source machine-learning platform, H2O (H2O.ai, 2023b), accessed through Python, in order to predict PT. The Cartesian grid search method, in which a set of values is specified for each hyperparameter under which to search (H2O.ai, 2023a), together with a trial-and-error approach, was employed to select values for the model's hyperparameters. Two hidden layers with 200 neurons each were ultimately incorporated into the model, and ten epochs were used for training. The activation function that resulted in the best performance based on cross-validation results was the "tanh" function implemented in conjunction with dropout regularization, which helps to reduce overfitting by randomly dropping out neurons during training. Moreover, the "Laplace" distribution was found to achieve the best results and thus was used in the model.

All 24 features were initially included in the model, which yielded an RMSE of 2.67 min and an MAE of 1.91 min based on cross-validation results. The importance of features was calculated following the Gedeon method, which measures the contribution of an input neuron to an output neuron in an NN (Gedeon, 1997). The scaled importance of features was found to range from 0.61 to 1.0, with complexity, length, and block having the highest importance (1.0, 0.901, and 0.896, respectively), and the other features ranging in importance from 0.615 to 0.752. All of the values were relatively close to one another, and the model was found to perform more poorly when testing the removal of the features with the lowest importance. Hence, the correlation results presented in Section 5.3.3 were used in a trial-and-error approach to evaluate whether removing certain features could improve the model's performance. The scheduled workload feature was the first feature to experiment with due to the lack of a clear relationship between this feature and PT, as evidenced by the null Spearman's coefficient and the absence of a clear pattern in the scatter plot. Indeed, the removal of this feature reduced the RMSE and MAE to 2.60 min and 1.80 min, respectively, leading to the decision to remove it from the model. The same experiment was conducted for the thickness, delta height, preassembled components, and height features, given their low Spearman's coefficients of 0.04, 0.05, −0.06, and 0.08, respectively. However, the removal of these features led to increases in RMSE and MAE of 2.69 min and 1.88 min, 2.72 min and 1.90 min, 2.76 min and 1.93 min, and 2.79 min and 1.96 min, respectively. The next feature to experiment with was the shift feature, since no clear relationship could be identified between this feature and PT based on the scatter plot. Nevertheless, removing the shift feature increased RMSE and MAE to 2.77 min and 1.88 min, respectively. Following that, given the small number of panels with large doors (eight panels only), the regular doors and large doors features were combined into one feature, referred to simply as "doors". This resulted in increases in RMSE and MAE of 2.71 min and 1.92 min, respectively. It is worth noting that, when the framing elements (i.e., studs, D-studs, M-studs, large windows, etc.) were combined into a single feature following the approach outlined in the study by Mohsen et al. (2022) and used along with the remaining features in the model, the resulting error was higher compared to when including all 24 features, with the RMSE measured at 3.04 min and the MAE at 2.26 min.

In general, the removal of any of the 23 features other than the scheduled workload feature resulted in a decline in the model's performance, indicating the importance of these features. Therefore, all features other than the scheduled workload feature, were retained in the NN model, resulting in an RMSE of 2.60 min and an MAE of 1.80 min.

5.3.5.2 Linear regression model

An LR model was developed using the RapidMiner software (RapidMiner, 2023). All features with the exception of the date feature were initially included in the model, which yielded an RMSE of 2.83 min and an MAE of 2.18 min based on the cross-validation results. However, the *p*-values corresponding to the regression coefficients were found to be greater than 0.10 for all the features other than the block, length, complexity, number of cutting zones, and day of the week features, indicating that the coefficients may not be statistically significant and that the observed relationship between the feature and PT may be due to chance. (It is interesting to note that length, complexity, and block were the most important features based on the Gedeon method and had the most statistically significant regression coefficients in the LR model.) Next, the effect on the model's performance of removing less significant features was examined in an iterative manner based on cross-validation results. With each iteration, the feature with the highest *p*-value was removed and, if the model's performance improved, the feature was then excluded from the model in subsequent iterations. The features removed, in order of their removal, were large doors, large windows, shift, sequence, regular window, stud, regular door, scheduled workload, height, number of drilled holes, D-stud, ambient temperature, and L-stud. Ultimately, the LR model was reduced to include a limited number of features beyond which removing any feature resulted in increased errors, as expressed in Eq. (5.9). This streamlined LR model yielded an RMSE of 2.61 min and an MAE of 2.00 min based on the cross-validation results.

$$
\begin{aligned}
PT\text{ (min)} = {} & 1.601 * Length\_ft + 0.729 * Complexity + 0.438 \\
& * Thickness\_in + 0.465 * Delta\_Height + 1.430 \\
& * Garage\_Door - 1.453 * Components + 0.297 * Cut\_zones \\
& + 0.561 * MStud + 0.358 * Block - 1.585 \\
& * Start\_Day\_Thursday - 0.949 * Start\_Day\_Tuesday + 3.388
\end{aligned} \tag{5.9}
$$

While the M-stud, garage door, and delta height features all had p-values greater than 0.10, in each case the feature's removal was found to negatively affect the performance of the LR model, indicating its importance in predicting PT. The remaining features all had p-values less than 0.04, indicating their significance based on a 5% significance level. The effect on the LR model's performance of combining the framing features was also examined since following this approach improved the performance of the LR model in (Mohsen et al., 2022). However, no improvement was observed in the model's performance using this approach, where the lowest RMSE and MAE were measured at 2.71 min and 2.10 min, respectively. The best-performing LR model remained the one presented in Eq. (5.9), but it performed more poorly than the NN model based on cross-validation results.

5.3.5.3 Random forest model

An RF model was trained using the open-source machine-learning platform, H2O, accessed through Python, to predict PT. The Cartesian grid search method, together with a trial-and-error approach, was used to select values for the model's hyperparameters. Ultimately, 50 decision trees, each with a maximum depth of 20, a minimum number of rows of 5, and a row sampling rate of 0.8, were selected based on the cross-validation results. The RF model containing all 24 features yielded an RMSE of 2.88 min and an MAE of 2.13 min.

The H2O platform, it should be noted, calculates the importance of each feature as the total improvement in the squared error realized following splits of the trees on the feature (H2O.ai, 2023c). Unlike in the case of the NN model, here the results showed a wide range of importance values for different features. The length feature had the highest importance score, scaled to 1.0, followed by the complexity feature with a scaled importance of 0.59. The importance scores of the number of drilled holes, stud, block, and L-stud features ranged from 0.15 to 0.39, while the number of cutting zones, height, and ambient temperature features had scores ranging from 0.05 to 0.39. The remaining features had importance scores lower than 0.04, with some (i.e., the large window, shift, regular door, D-stud, garage door, preassembled components, and large door features) even below 0.01. However, these scores may not necessarily reflect the true significance of the features in relation to PT, as it was observed that the features with the lowest number of observations—there were only eight large doors, eight garage doors, and 11 preassembled components in the entire dataset—tended to have the lowest importance scores. For example, garage doors had an importance score near 0 when, in reality, garage doors have a significant effect on PT since they necessitate manual work (which cannot be completed using the framing machine), as also evidenced by the statistically significant regression coefficient shown in Eq. (5.9). The RF model was reduced to exclude all the features with low importance. As such, the final model included the length, complexity, number of drilled holes, stud, block, L-stud, and number of cutting zones features and yielded an RMSE of 2.83 min and an MAE of 2.11 min. The RF model performed more poorly than both the NN and LR models based on cross-validation results.

5.3.6 Step 6: Develop UCD and ID estimation models

As previously explained, UCD could be computed by subtracting PT and scheduled breaks from CT for the case framing workstation. As such, UCD was evaluated for the 172 panels included in the training dataset and used to fit statistical distributions in Simphony.NET software (Engineering at Alberta, 2022). The least squares method (Wolfram MathWorld, 2022) was used to fit a set of distributions, and the Kolmogorov-Smirnov (K-S) test was performed in Simphony.NET to evaluate how well the distributions fit the data. The K-S test is based on the maximum difference between the empirical and theoretical cumulative distributions (Massey, 1951). Based on the results, the Pareto distribution with a shape parameter of 0.94 was found to provide a good fit for the data, as also indicated by the Q–Q plot, shown in Fig. 5-5, since most of the points on the Q-Q plot were found to fall close to the diagonal line. However, some points were found to fall well above that line, indicating that the corresponding values in the dataset are smaller than what would be expected under the assumed distribution. This suggests that the actual data has a lighter tail than the Pareto distribution, as also confirmed by the Pareto's probability density function plotted against the empirical distribution in Fig. 5-5. As such, to avoid sampling extreme UCD values from the Pareto distribution, which do not accurately represent reality, a trial-and-error approach was followed to cap random sampling in a manner that would minimize MAE and SE. Ultimately, sampling was capped at 4 min, beyond which the Pareto distribution consistently overestimated the observed values as indicated by the Q-Q plot. Capping the sampling process was accomplished by rejecting any sampled values greater than 4. It should be noted that using a cap of 4 min means that actual UCD values exceeding 4 min will be underestimated by the system. However, the occurrence of such events is relatively minimal based on the training data. Moreover, these occurrences should be addressed through proper control of operations rather than adjusting the system, as discussed later in this chapter.

The Simphony.NET software, the least squares method, and the Kolmogorov-Smirnov (K-S) test were similarly used to build a statistical model for ID. The Chi-squared distribution with two degrees of freedom was found to have the best goodness-of-fit for modelling ID.

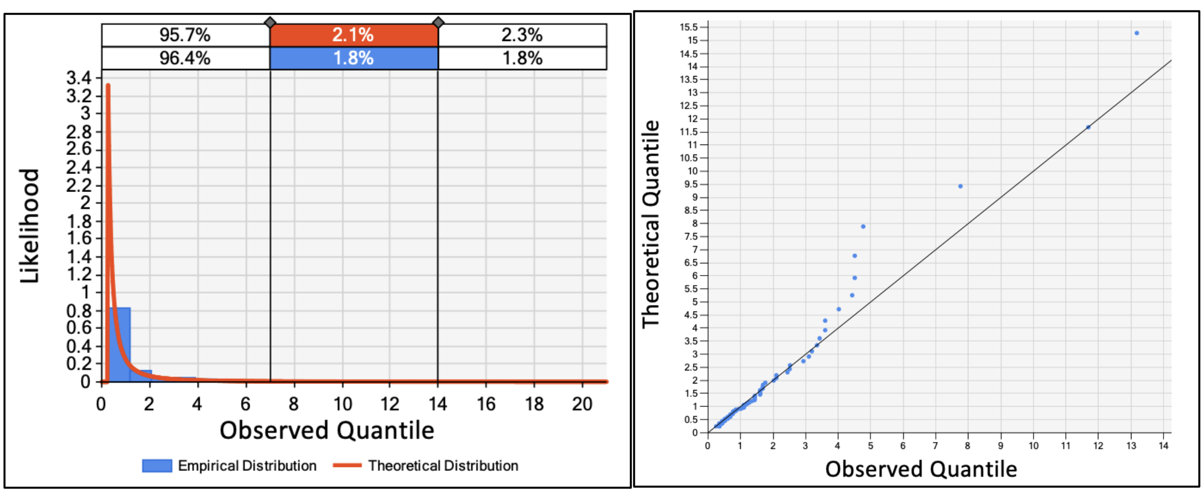


Fig. 5-5. Pareto distribution fitting.

5.3.7 Step 7: Build the simulation model

A 3D simulation model was developed for the entire wall production operation using Simio software (Simio, 2023), as shown in Fig. 5-6. The simulation model was connected to an MS

Access database from which it could retrieve the list of wall panels scheduled for the day as well as data on the corresponding features. Based on a predefined production sequence, the model was developed to mimic actual operations and simulate the production of wood wall panels as they flow from one workstation to another. At the framing workstation, PT, UCD, and ID were modelled using the NN model (selected based on the results presented in Section 5 below) and the statistical models previously developed, respectively, while the simulation model was mainly responsible for (1) determining when a framed wall panel could be transferred to the downstream workstation and, accordingly, when a new framing cycle could be started, (2) halting production during scheduled breaks, and (3) collecting data on cycle start time (ST), finish time (FT), CT, PT, CD, DW, and UW, as shown in Fig. 5-7. The remaining workstations were modelled by incorporating regression and statistical models developed by Shafai (2012) in a prior study conducted on the same case production line. (The regression models predict the durations of various tasks performed at each workstation as a function of wall panel design attributes, and the statistical models estimate delays. However, these models provide only rough estimates of the current cycle times of downstream workstations, and this is a limitation of the present study, as discussed in the Limitations subsection below.)

The simulation model was verified and validated in several ways. First, the model was developed in stages with weekly meetings conducted with the case company's research and development personnel who, at each stage, validated that the model reasonably represented actual operations. Second, the 3D animation functionality of the model significantly facilitated verifying that the model was behaving as intended and validating that it was representative of actual operations as it enabled a direct comparison between virtual operations within the model and actual images captured of real operations. Third, the data collected during simulation, as depicted in Fig. 5-7, provided further support for verifying the model's sound behavior. Finally, CT estimates generated by the model for forty wall panels were compared to actual measurements. The results obtained, which are detailed in the following section, further validated the soundness of the model.

Fig. 5-6. 3D model of the factory.

Panel ID	Framing ST	Framing FT	Framing CT (Minutes)	Framing PT (Minutes)	Framing D (Minutes)	Framing DW (Minutes)	Framing UW (Minutes)
E-16_6002-LH-0119_...	02/03/2023 11:49:29 AM	02/03/2023 11:58:08 AM	8.6580	8.0556	0.6024	0.1467	0.0000
E-17_6002-LH-0119_...	02/03/2023 11:58:17 AM	02/03/2023 12:42:06 PM	43.8121	13.7876	30.0244	0.0015	0.0000

Fig. 5-7. Sample simulation data.

5.4 System testing results and discussion

The performance of the PT and CT predictions was evaluated using the testing dataset manually collected for 40 wall panels, as described in greater detail below. Moreover, the case company uses a linear fixed rate (metres per minute) to estimate the full production capacity of the wall production line under ideal conditions, and was developing production schedules at the time of the present study based on the assumption that the production line was operating at 85% of its full capacity. The performance of the case company's current estimation practice was also evaluated, as described later in this section.

5.4.1 Evaluation of PT predictions

The PT prediction performance of each of the three models and of the current estimation practice at the case company are summarized in Table 5-4. As these results show, the LR model was found to have the best performance, with MAE, MPE, and SE values of 1.52 min, 19%, and 5.62 min, respectively. However, based on MAE and MPE only, the NN and LR models achieved similar levels of performance. Meanwhile, the RF model showed the poorest performance based on MAE and MPE (i.e., 1.81 min and 21%, respectively), although it achieved a lower SE than did the NN model (10.01 min in the case of the former versus 17.68 min in the case of the latter). The SE was positive for each of the three models since the overestimations tended to be more frequent (and in some cases more significant) than the underestimations. Although the NN and LR models were found to be comparable in performance based on MAE and MPE, the NN model had a larger SE since the overestimations it made were larger than the overestimations made by the LR model. Similarly, the size of underestimations made by the LR model was larger than that of the underestimations made by the NN model. This difference is clear in Fig. 5-8. It is also worth noting that errors corresponding to four of the panels framed sequentially skewed the evaluation results, as the prediction error of these panels ranged between 2.66 min and 4.94 min for the NN model, between 2.51 min and 3.86 min for the LR model, and between 2.62 min and 3.86 min for the RF model. Excluding these panels from the evaluation reduced the MAE, MPE, and SE to 1.31 min, 15%, and 1.99 min, respectively, for the NN model, to 1.33 min, 16%, and −7.30 min, respectively, for the LR model, and to 1.67 min, 19%, and −2.37 min, respectively, for the RF model. In other words, with this outlier sequence of four panels removed, the NN model showed the best performance based on the testing results. In examining the recorded framing processes for these panels, no satisfactory explanation for the high error values could be identified aside from the fact that the operator was consistently working at a high speed when framing them. Another potential explanation of the high error values could be the presence of similar panels that took longer to frame or panels for which the computer vision system overestimated the actual PT values in the training dataset. Whatever the cause of the high error values, in future when the system is deployed for continuous data collection, the size of the training dataset will be continuously increasing, and this should improve the performance of the models as they continuously learn from new data.

Table 5-4. Evaluation results for PT predictions.

	Actual PT (min)	NN		LR		RF		Fixed rate	
		Predicted PT (min)	***e* (min)**	**Predicted PT (min)**	***e* (min)**	**Predicted PT (min)**	***e* (min)**	**Predicted PT (min)**	***e* (min)**
1	14.17	11.95	−2.22	10.78	−3.39	11.36	−2.81	7.33	−6.84
2	10.58	11.79	1.21	10.98	0.40	10.82	0.24	8.61	−1.97
3	8.83	10.34	1.51	8.71	−0.12	9.81	0.98	6.83	−2.00
4	9.92	10.00	0.08	9.23	−0.69	11.14	1.22	8.03	−1.89
5	15.17	13.75	−1.42	12.45	−2.72	11.40	−3.77	5.31	−9.86
6	4.58	5.05	0.47	4.42	−0.16	5.95	1.37	4.39	−0.19
	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
38	7.42	8.90	1.48	8.79	1.37	8.78	1.36	8.13	0.71
39	12.67	12.99	0.32	11.77	−0.90	11.08	−1.59	8.71	−3.96
40	11.33	11.00	−0.33	10.46	−0.87	9.83	−1.50	7.77	−3.56
	MAE	1.57		1.52		1.81		2.87	
	MPE	19%		19%		21%		29%	
	SE	17.68		5.62		10.01		−98.64	

To summarize, based on the cross-validation results, among the three models considered it was the NN model that was found to perform best. Although it was the LR model that performed better based on the preliminary analysis of the testing results, the NN model would have outperformed the LR model if it had not significantly overestimated the PT values for the four panels with high errors. In general, the NN model tended to overestimate PT more than did the LR model. However, since capping the UCD values may lead to underestimating long delays, occasional overestimations of PT would reduce SE for the daily scheduled batch of panels. Moreover, an estimation system that tends toward overestimation of PT is preferable to one that tends toward underestimation of PT (since the former gives more conservative estimates). As such, the NN model was selected for use in the estimation system.

Finally, the case company's current estimation practice was found to have an MAE of 2.87 min and an MPE of 29%, while the SE was negative (approximately −99 min). The SE was significantly negative due to the frequency and size of underestimations made based on the fixed rate assumption. In fact, the current estimation practice underestimated PT for 32 of the 40 panels, indicating that the fixed rate assumption is overly optimistic. In other words, all three machine-learning models outperformed the current estimation practice in estimating PT, with the NN model and the LR achieving reductions in MAE of 45% and 47%, respectively, compared to the current practice.

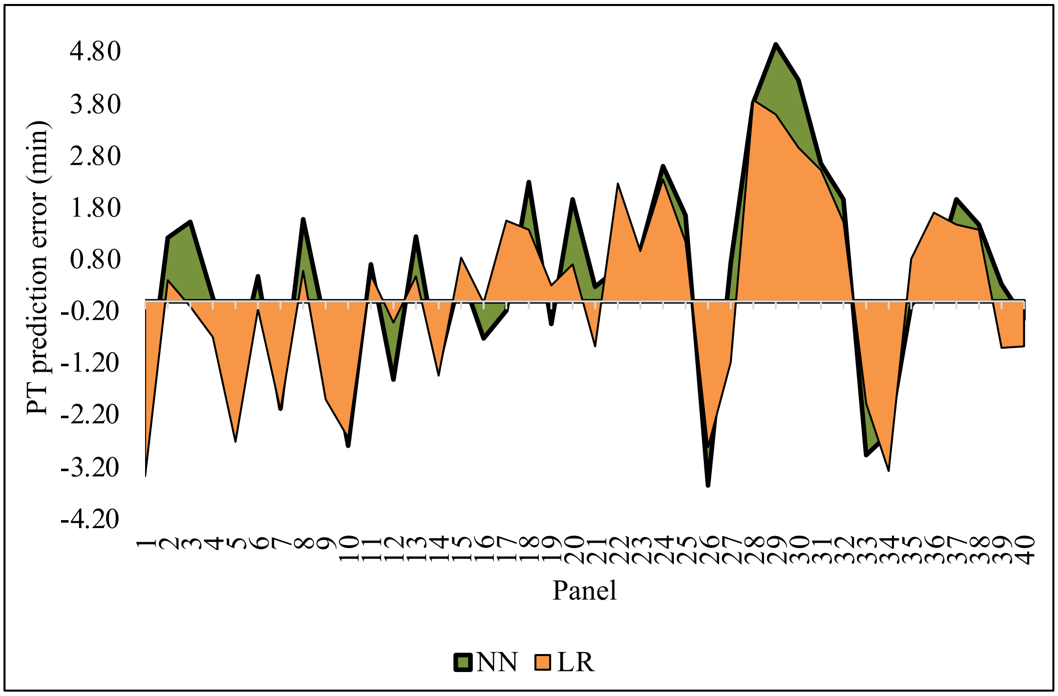


Fig. 5-8. e_i of the NN model versus that of the LR model.

5.4.2 Evaluation of CT predictions

The evaluation of the CT predictions and of the current estimation practice are summarized in Table 5-5. As can be seen, the estimation system achieved an MAE of 3.03 min and an MPE of 23%, while the SE was negative (i.e., −50 min), although the number of panels with overestimated predictions was equal to the number of panels with underestimated predictions (i.e., 20 panels). The negative SE was due to the larger size of underestimations compared to overestimations. This discrepancy is partly attributable to the UCD predictions having been capped at 4 min, resulting in two error values of about 17 min. The actual UCD values for these two panels were around 14 min and 16 min, as the operator had left the workstation during the framing process in both cases.

Omitting these two panels alone reduced the MAE, MPE, and SE to 2.30 min, 21%, and −16.28 min, respectively. Moreover, although the SE of −16.28 min implies that the estimation system slightly underestimated the total time needed to frame a batch of panels, this underestimation is insignificant compared to the total actual time spent on framing 38 panels (i.e., about 420 min).

Nevertheless, notwithstanding the higher CT prediction errors, the developed estimation system performed significantly better than the current estimation practice, which yielded an MAE of 4.72 min, an SE of −156 min, and an MPE of 34%. In other words, the estimation system realized reductions in MAE and SE of about 36% and 68%, respectively, compared to the current estimation practice. The current estimation practice was found to significantly overestimate the actual capacity (or underestimate cycle times) of the framing workstation. Meanwhile, the framing workstation is the first workstation on the production line, so the actual production capacity of the wall production line should be less than or close to that of the framing workstation. This means that the overall production capacity of the entire production line is overly optimistic in current practice. An overly optimistic estimation of production capacity leads to a scenario in which more panels per day are scheduled for production than what can be produced, in turn resulting in production targets being missed and in a physical and mental toll on the well-being of workers. In this regard, it is notable that the developed estimation system was found to significantly outperform the current practice.

Table 5-5. Evaluation results for CT predictions.

	Actual CT (min)	**Prediction system**		**Fixed rate**	
		Predicted CT (min)	***e*** **(min)**	**Predicted CT (min)**	***e*** **(min)**
1	22.02	15.21	−6.81	8.62	−13.40
2	11.5	12.06	0.56	10.12	−1.38
3	8.9	10.66	1.76	8.04	−0.86
4	10	10.68	0.68	9.44	−0.56
5	17.63	14.62	−3.01	6.25	−11.38
6	7.57	7.64	0.07	5.16	−2.41
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Table 5-6. Evaluation results for CT predictions (continued).

	Actual CT (min)	Prediction system	Fixed rate		
		Predicted CT (min)	*e* (min)	Predicted CT (min)	*e* (min)
.	.	.	.	.	.
38	7.82	10.38	2.56	9.56	1.74
39	18.67	14.61	−4.06	10.25	−8.42
40	14.8	12.18	−2.62	9.14	−5.66
MAE (min)			3.03		4.72
MPE			23%		34%
SE (min)			−50.10		−156.25

5.5 Further analysis of the findings

5.5.1 Use of computer vision data

The prediction error of the three models may be partly attributable to measurement errors on the part of the computer vision system. Specifically, the frequency, and, for some panels, the size, of overapproximations made by the models were greater than those of the underapproximations. Meanwhile, the size of the PT overapproximations made by the computer vision system, as reflected in the results of a prior study (Alsakka et al., 2023b), was generally greater than that of the underapproximations. As such, the overapproximations on the part of the prediction models may be partially attributable to the fact that PT was more significantly overapproximated than underestimated by the computer vision system for a number of panels in the training dataset. Nevertheless, the fact that the performance of the three machine-learning models was superior to that of the fixed rate used in current practice points to the suitability of using computer vision data for training the models in applications of this nature. Indeed, computer vision is a promising solution, as it provides a practical and time-efficient supply of data that can be used to continuously tune the models in order to account for future sources of variability.

5.5.2 Influencing factors

As noted above, the features considered in the machine-learning models were initially identified based on a detailed qualitative analysis of the framing workstation as outlined in a previous study (Alsakka et al., 2023c). Interestingly, all of the features with the exception of the scheduled workload feature were found to hold relevant information regarding PT in at least one of the machine-learning models. Notably, the NN model found all of the features, with the exception of the scheduled workload feature, to be significant predictors of PT. However, this does not necessarily imply that scheduled workload does not affect PT. In fact, the data collected for this case study only covered days of operations that resulted in a limited number of observations for this feature. Since any potential effect of workload on PT would be indirect, several months' worth of data may be necessary to ascertain whether or not it has an impact. Moreover, although the majority of the features were found not to have a statistically significant regression coefficient in the initially developed LR model, a number of them were found to have a potentially explainable effect on PT. The LR model trained using all features is expressed in Eq. (5.10).

$$\begin{aligned} PT\ (\text{min}) = {} & 0.117 * Scheduled\_Workload\_SF + 1.544 * Length\_ft - 0.002 \\ & * Daily\_Sequence + 0.646 * Complexity + 0.365 \\ & * Thickness\_in + 0.146 * Height\_ft + 0.501 * Delta\_Height \\ & - 0.153 * Regular\_Window - 0.086 * Large\_Window \\ & + 0.082 * Regular\_Door - 0.145 * Large\_Door + 1.600 \\ & * Garage\_Door - 1.044 * Components + 0.323 * Cut\_zones \\ & + 0.079 * Drill\_holes - 0.029 * Stud - 0.883 * DStud \\ & + 0.196 * LStud + 0.631 * MStud + 0.363 * Block - 0.071 \\ & * F\_Average\_Hourly\_Temp\_C - 1.398 * Start\_Day\_Thursday \\ & - 1.052 * Start\_Day\_Tuesday - 0.193 \\ & * Act\_F\_Shift\_Afternoon + 3.233 \end{aligned} \tag{5.10}$$

As explained in Section 5.3.2, the increase in length, height, and thickness should logically increase PT, and this is reflected in their positive coefficients. The size of these coefficients, however, does not accurately represent the respective independent effect of each on PT, given their high level of multicollinearity with other features (as evidenced by their high VIF values as presented in Table 5-3). For instance, the coefficient of the daily sequence feature is near 0, but its slightly negative value could indicate that the worker becomes more “dialed in” to their work as they frame more panels. Moreover, while it was to be expected that the complexity feature would have a positive coefficient since framing a wall panel with a mix of different elements (e.g., an exterior multi-wall with one large door, one regular door, one large window, two regular windows, etc.) is less straightforward than framing a panel consisting of fewer different types of elements (e.g., an interior wall that consists mainly of studs), given its degree of multicollinearity, the size of its coefficient could not be interpreted independently. As for the height difference feature, its coefficient indicates that it takes an additional half-minute for every foot of difference in height, which is a reasonable amount of time to allow for adjusting the width of the framing machine. The negative coefficients of windows, large doors, and preassembled components are reasonable given the large positive coefficient of the length feature. Windows, doors, and preassembled components are preassembled and only need to be nailed to the frame. Since every linear foot per wall panel adds approximately 1.5 min to PT, such preassembled openings should reduce this duration, as they cover several linear feet of wall panel and only require nailing. It is not clear, however, why the coefficient of regular doors was found to be positive.

With regard to the number of cutting zones and number of drilled holes features, the positive signs of their coefficients are reasonable, as each cutting zone and hole requires that additional steps be performed by the machine. The relatively small coefficient of the number of drilled holes feature, however, may be attributable to its collinearity with the length feature (as previously explained). Similarly, since length is correlated with the number of regular studs, it is not surprising that it was found to have a small coefficient, but the reason for its negative sign is not clear. D-studs, L-studs, and M-studs should logically add more time to PT since they require more nails compared to regular studs. While the coefficients of the M-stud and L-stud features align with this logic, that of the D-stud feature is negative. The D-stud feature had the lowest number of records (it was found in just 14 panels, as per Table 5-1), and this may explain why the LR model was not able to identify a logical relationship between the D-stud feature and PT. The coefficient of the blocks feature is reasonable, as each block must be manually nailed by the operator. As for the ambient temperature feature, its coefficient was found to be small and a negative value. It is important to note, however, that the range of the recorded temperature data was not particularly wide (−5.5 °C to 12 °C); this data range is not sufficient to test the hypothesis proposed by the operator consulted in this study, which is that workers become tired more and their work pace slows when the ambient temperature exceeds 20 °C since there is no air conditioning in the factory (Alsakka et al., 2023c). The negative coefficient of this feature could be attributable to the fact that the temperature typically increases throughout the day, such that it could have a similar effect to that of the sequence feature on PT. In other words, although the NN model found the ambient temperature feature to contribute to PT, the information it identified in this feature is not necessarily related to the temperature itself. In future work, data from the summer season should be collected in order to examine the effect of high temperatures (i.e., > 20 °C) versus low temperatures on PT.

Regarding the day of the week feature, the negative coefficients corresponding to Tuesdays and Thursdays align with the observation of the operator consulted in this study; the operator mentioned that Tuesdays see a spike in production since workers tend to become "dialed in" after a more sluggish start to the week on Mondays, and that Thursdays tend to be more productive because workers are motivated to complete their work early in order to start their weekend (the case plant does not operate on Fridays) (Alsakka et al., 2023c). Finally, although the shift feature was discussed in a previous study (Alsakka et al., 2023c), with fatigue expected to result in an increase in PT in the afternoon, its coefficient was negative in the LR model. Nevertheless, the negative sign of its coefficient is consistent with the negative signs of the sequence and temperature coefficients, the values of which increase throughout the day. These results could also imply that workers are more motivated in the afternoon to finish their work so they can return home, or that they are compelled to increase their pace of work in the afternoon in order to complete the set of wall panels scheduled for the day by the end of the shift.

This discussion demonstrates the significance of expending time and effort on identifying and understanding the factors that influence cycle time prior to developing machine-learning models. When considering only the panel length to predict PT, the MAE obtained for the NN model based on cross-validation results was 2.18 min. Taking into consideration other geometric properties of the panels (i.e., height, width, number of cuts, etc.) was found to reduce this error to 1.94 min, representing an 11% reduction in the error. Moreover, taking into consideration the complexity, day, shift, ambient temperature, height difference, framing sequence, and date features was found to further reduce the error to 1.80 min, representing a total error reduction of 17%. As these findings suggest, gaining understanding as to what factors are influencing cycle time helps to improve the accuracy of process time estimation systems.

5.5.3 The performance of different machine-learning algorithms

The NN model was found to be the most suitable model for the case workstation, as previously explained. However, the LR model performed nearly as well as the NN model, and was able to reach that performance using only 11 features (compared to 23 features in the case of the NN model). As such, for the modeller who favors simplicity and interpretability, the LR model may be a more attractive choice. Nevertheless, the NN model's ability to identify relationships between the various features and PT that were not apparent through scatter plots and Spearman's coefficient and that were not deemed important by the LR and RF models is noteworthy. Moreover, given that the numbers of observations in the training dataset was relatively small for some of these features (e.g., large doors, garage doors, preassembled components), increasing the size of the dataset may further improve the performance of the NN model. It is also noteworthy that the NN model was not sensitive to multicollinearity, unlike the LR model, which frequently changed its regression coefficients and improved with each statistically insignificant and dependent feature removed from the model. As for the RF model, it generally showed inferior performance compared to the NN and LR models. Perhaps the most significant observation concerning the RF model was that it performed better after removing all the features that had low frequencies in the training dataset (with the exception of the block feature, which was found to be important given its actual significant effect on PT, since installation of blocking involves time-consuming manual work). Hence, significantly increasing the size of the dataset to include more panels containing the less frequent framing elements (e.g., preassembled components) may result in more of the relevant features being retained in the RF model and, in turn, in improved performance of the RF model.

It can be concluded that the performance of the models was dependent on the specific workstation studied, given that a good number of features were found to be correlated, and that certain features were found to be significantly less frequent than others. As such, models other the NN model may be more suitable to represent other workstations on the same production line, depending on the factors influencing the corresponding PT and on the dataset used for training. In other words, selecting a machine-learning model for the estimation system in the case of other workstations should not rely solely on the results of this study, whereas an independent analysis is required for each workstation by virtue of each having unique characteristics.

5.5.4 Criticality of unpredictable delays

As the results reveal, the system's predictions of PT were found to be more accurate than its predictions of CT. This was to be expected, given the random nature of UCD. The difference in MAE values between PT predictions (MAE = 1.57 min) and CT predictions (MAE = 3.03 min) underscores the significant effect that random delays can have on the accuracy of CT predictions. Even if PT can be predicted with perfect accuracy, high levels of process randomness can diminish the effectiveness of such PT predictions, leading to high CT prediction errors. As such, the effectiveness of process time estimation systems such as the one proposed in this chapter is contingent upon the level of variability present in the process being estimated. It is worth noting that, since PT was found to range from as little as 2 min to as much as 27 min (Table 5-1), long delays such as those recorded for the two panels discussed in Section 4.4.2 could result in a reduction in output of at least one wall panel per day. As such, it is incumbent upon any offsite construction enterprise considering investment in an estimation system to ensure it has in place a robust program for project control and for mitigating controllable sources of variability prior to investing in such a system.

5.6 Conclusions

5.6.1 Summary

This study proposed an estimation system that predicts cycle time at the workstation level in offsite construction factories. The system uses machine learning, statistical methods, 3D simulation, and computer vision to predict cycle times as a function of various influencing factors. The system's performance was evaluated through its application to a semi-automated wood-wall framing workstation in a panelized construction facility. The study also examined the performance of feed-forward ANN, LR, and RF models for predicting PT (i.e., CT − CD) at the case workstation. The results show that the NN model was the model that performed best in cross-validation, yielding more conservative predictions—and being less sensitive to multicollinearity, which is a notable characteristic of the set of features considered in this study—compared to the other models. On the other hand, the LR model was the model that performed best in testing. Moreover, the LR model was capable of performing well using only 11 features (whereas the NN model used 23). As such, if preference is given to simplicity and interpretability, the LR model may be a more desirable choice. Still, the NN model's ability to identify relationships between features that had low frequencies in the dataset is noteworthy. Increasing the size of the dataset to include more observations of these features may allow the NN model to better characterize their relationships with framing time and thus to further improve its performance. The RF model, meanwhile, had the poorest performance among the three models considered in both cross-validation and testing. However, it was observed that this model's performance was affected significantly by the features that had low frequencies in the dataset, and removing those features served to improve its performance. Thus, increasing the dataset size to increase the frequency of these features may also improve the performance of the RF model. In general, it was concluded that the performance of the models was sensitive to the characteristics of the workstation under study and, thus, the decision regarding the best-performing model should not be generalized to other workstations, which require an independent analysis of their own.

The developed system achieved an MAE and an MPE of 3.03 min and 23%, respectively, and these further decreased to 2.30 min and 21%, respectively, after removing the predictions corresponding to two outlier panels; (during the framing of these panels the operator randomly left the workstation for a significant amount of time, an event that cannot be predicted using the developed estimation system). The developed system was found to perform significantly better than the case company's current estimation system, which achieved an MAE of 4.72 min and an MPE of 34%. Moreover, given the superior performance of the machine-learning models compared to the current estimation practice at the case company, the use of computer vision data for training the models can be considered a valid solution, especially given that it allows for continuous tuning of the prediction model, thereby accounting for potential new sources of variability in the future.

The study results also demonstrate the importance of dedicating time and effort to identifying the factors influencing cycle time prior to collecting data and building machine-learning models.

These factors can be identified using qualitative methods such as process observation, process mapping, reviews of previous related studies, and semi-structured interviews. As evidence of the importance of factor identification, the models performed best when using a different set of influencing factors for each. Moreover, taking into consideration the geometric properties of a given wall panel rather than relying solely on its length to predict framing time, as is the case in the case company's current estimation practice, served to reduce the MAE of the NN model in cross-validation by 11%. Additionally taking into consideration the complexity, day, shift, ambient temperature, height difference, framing sequence, and date features further reduced the error, resulting in a total error reduction of 17%. Finally, it is worth noting that in most cases the contribution of a given factor to the framing time could be rationally interpreted. As these observations demonstrate, having a rich pool of influencing factors whose contributions to PT are known or at least tentatively known facilitates the development of prediction models.

Finally, the study results demonstrate the significant effect that unpredictable or random delays can have on cycle time estimation systems. The difference between the accuracy of the PT predictions (MAE of 1.57 min) and that of the CT predictions (MAE of 3.03 min) demonstrates this effect. In fact, disregarding the prediction results for the two outlier panels with long delays reduced the CT prediction error from 3.03 min to 2.30 min. Such results should play a key role in the decision of using an estimation system, such as the one proposed in this chapter, as elaborated in the following subsection.

5.6.2 Implications for the industry

The utilization of average rates as a means of estimating process time has long been regarded as an inaccurate and ineffective method in the construction industry, according to academia. Surprisingly, even in offsite construction, where the controlled factory production of building components should facilitate estimation, average rates continue to be employed. The findings of this study have demonstrated that such rates can be overly optimistic, leading offsite construction companies to make commitments to optimistic production schedules that cannot be fulfilled due to the inherent variability in operations. These practices not only impose physical and mental strain on workers who are compelled to constantly meet unrealistic deadlines but also burden managers who must grapple with schedule delays.

The widespread availability of artificial intelligence, virtual modelling, and smart data acquisition solutions offers companies the opportunity to move away from their current estimation practices and better accommodate the daily variability they encounter. The findings of this study demonstrate that such solutions can significantly enhance the accuracy of cycle time estimates. Improved estimates contribute to a deeper understanding of operations and enable better control, thereby alleviating pressure on both the workforce and management. On another degree, accurate estimation of cycle times can enable companies to quantitively analyse their operations and, hence, to identify various sources of waste such as excessive waiting and delay times as compared to productive times. In essence, companies would be better equipped to improve the efficiency of their operations and reduce unnecessary costs. For example, one way to leverage this knowledge of cycle time variability would involve the reallocation of tasks from workstations with lengthy cycle times to those with shorter ones. This would enable balancing production lines, thereby reducing waiting times for resources assigned to faster workstations and, in turn, saving the costs associated with idle resources.

Nevertheless, it is essential to recognize that the accuracy of cycle time estimates is vulnerable to unpredictable delays, which are inherently random and cannot be predicted. Consequently, addressing the estimation issue involves more than simply leveraging advanced and robust tools. For example, even the most intelligent technologies cannot anticipate the actions of workers and the resulting delays if their behavior remains uncontrolled. In other words, focusing solely on developing an estimation solution is insufficient. Instead, the successful implementation of cycle time estimation systems, such as the one proposed in this chapter, requires a concerted effort to manage and mitigate various sources of variability at offsite construction factories.

5.6.3 Limitations and future work

The evaluation of the developed estimation system as presented herein was subject to several limitations. First, given that there were no scheduled events (except for scheduled breaks) interrupting the work at the case framing workstation, that none of the labour resources were shared with other workstations, and that all the necessary materials were prepared in advance, predictable delays (as per the definition presented herein) were not a critical factor. Hence, the system should also be evaluated through its application to a workstation that has several sources of predictable delays. Second, the hypothesis underlying the inclusion of the ambient temperature feature was that higher temperatures are associated with higher fatigue levels and slower work pace at the case factory. Although relationships between the ambient temperature factor and PT were identified in the dataset used in this study, the ambient temperature was relatively low during the data acquisition period, so this hypothesis could not be adequately tested. In future work, data collected during the summer season could be used to test this hypothesis and account for any variability resulting from high temperatures. Third, the data collected for this case study only covered a short period of operation, and this limited the opportunity to observe the effect of variations in scheduled workload. Since the suspected effect of workload on processing time is indirect, several months' worth of data may be needed in order to ascertain whether or not this factor has an impact on PT. Finally, cycle times at the other workstations in the production line at the case company's production facility were modelled in the simulation using rough estimates, and ID was assumed to be equal to ITD. These contributed to errors in the estimation of the start times of cycles at the framing workstation. Discrepancies between predicted and actual cycle start times may affect the values of the average hourly temperature and shift features corresponding to the panels. Although the errors resulting from those rough estimates and the assumption are in the order of minutes and, hence, should not have a significant effect on the values of the related features, they still affect the total duration needed to frame a batch of panels. As such, an avenue for future research would be to apply the developed system to all workstations on the production line at the case company's production facility and evaluate its estimation performance for all the workstations collectively.

5.7 Acknowledgements

This work was supported by the Alberta Innovates Graduate Student Scholarship, and by the Natural Sciences and Engineering Research Council of Canada (grant file no. ALLRP 577047-22). The authors gratefully acknowledge the technical review of Mr. Jonathan Tomalty.

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 6

Digital Twin for Production Estimation, Scheduling and Real-Time Monitoring in Offsite Construction

6.1 Introduction and background

6.1.1 Production management practice in the offsite construction industry

The construction sector is witnessing a notable shift towards the widespread adoption of offsite construction, also referred to as construction manufacturing or prefabricated construction. This trend is expected to continue, with the global market projected to expand from around $130 billion in 2020 to a potential $230 billion by 2030 (Khandelwal, 2021). This shift towards construction manufacturing involves the application of theories and practices that form the foundation of manufacturing systems. In simple terms, a manufacturing system is a "combination of humans, machinery, and equipment that are bound by a common material and information flow" (Caggiano, 2014). Accordingly, workers, machinery, and equipment in offsite construction factories are typically positioned in fixed locations at workstations with each workstation assigned a well-defined production process (e.g., wall framing, windows/doors installation). Despite this shift towards manufacturing, the industry still lacks well-established management tools on the basis of which to estimate production time (i.e., the total time necessary to produce building components) and, accordingly, develop production schedules. In fact, in current practice, the tasks of estimating production time and developing production schedules are typically performed using average production rates (e.g., linear metres of wall panels per minute (Alsakka et al., 2023d) derived from experience gained from previous projects. There are two main challenges pertinent to the use of average rates:

(1) Production time is a function of cycle time (CT), which is the time needed to complete process cycles at different workstations that collectively form the overall production operations. Meanwhile, CT may be subject to a high degree of variability. In fact, a recent case study of a panelized construction factory (in which wall panels, floor panels, roof components, and staircases are prefabricated for shipment to the site for installation) reported CTs for wood wall framing operations ranging from approximately 2 min to as much as 1 h (Alsakka et al., 2023d). There are many influencing factors that contribute to this high variability, including aspects related to the building components processed, workers, machines, materials, workstation setup, production line, factory operations, and external circumstances (Alsakka et al., 2023c). Given this, average production rates may not provide an accurate representation of actual production. In the aforementioned case study, the use of average rates resulted in overly optimistic CT estimates (Alsakka et al., n.d.). Optimistic CT estimates lead to optimistic production time estimates and schedules, in turn resulting in workers and management personnel being pressured to meet unrealistic schedules, as well as leading to schedule delays (and associated costs).

(2) Manufacturing systems are designed and operated to satisfy specific business objectives (Chryssolouris, 2006). Principally, they involve product-contingent applications of science and technology via their processes and associated machinery (Parnaby, 1979). In other words, manufacturing systems are tailored to meeting the specific requirements of the products they are designed for. This means that substantive changes in product requirements may disrupt manufacturing operations, rendering previous production experience largely irrelevant for estimating and scheduling production of the product at hand. For example, consider an offsite construction factory that fully relies on a fully-automated multi-function bridge for fastening sheathing boards (e.g., plywood, oriented

strand board (OSB)) to exterior building walls. If the multi-function bridge is configured to nail materials with a specific degree of stiffness (dictating the shooting pressure of the nailing gun), which is a common setting in automated multi-function bridges currently used in offsite construction, a significant change in the softness of the sheathing board material could mean an entire disruption to the use of the bridge and, hence, the factory operations. Hence, as Parnaby argued decades ago in this regard, maintaining stability in operations amidst variability requires a supply of information and continuous data acquisition (Parnaby, 1979).

In this context, there is a need in offsite construction for more dynamic production time estimation and scheduling methods that can handle the variability inherent in operations, as well as the variability that arises from unexpected events and external factors. The following subsection provides further background related to production estimation and scheduling in offsite construction.

6.1.2 Production time estimation and scheduling in offsite construction

With regard to production estimation and scheduling in offsite construction, the literature has tended to focus primarily on optimizing the production sequence of building components or jobs with the objective of minimizing the total production time, makespan, or related parameters (e.g., Ko & Wang, 2011; Leu & Hwang, 2002; Z. Xu et al., 2020). While such studies address an essential aspect of production scheduling (i.e., sequencing), the methods proposed still require knowledge of how much time is spent by each building component at each workstation, and average time estimates are typically used for this purpose, this in spite of the inherent variability as noted above. Du et al. (2021) proposed an alternative approach in which they integrated an optimization algorithm with a multiagent system that simulates production and includes a risk agent that triggers uncertain events such as machine failure. However, despite including a stochastic element in their model to account for unexpected events, they still relied on average estimates to model processing times. A different method was employed by Altaf et al. (2014, 2018), who, in addition to integrating their production sequence optimization model with discrete-event simulation to simulate operations, employed regression equations to model processing times at different workstations as a function of the design properties of building components. They also included statistical distributions to model several types of delays. While this strategy is an improvement upon the practice of relying on single average values to estimate the durations of entire processes at workstations, machine-learning techniques (such as the linear regression models they employed to estimate processing times), while robust, cannot guarantee the complete capture or explanation of all sources of variability. As such, they can generate estimation errors that, when left unaddressed, may gradually accumulate and lead to significant deviations between planned production schedules and actual operations. Moreover, even algorithmically sophisticated scheduling solutions can deteriorate amid uncertain processing times, underscoring the significance of dynamic scheduling in which schedules are updated in real time as uncertainties with respect to processing times are encountered (Lawrence et al., 1997). To mitigate these limitations and establish a mechanism for (1) gaining real-time insights into operational activities, (2) monitoring progress on production schedules in real time, and hence, (3) regularly updating production schedules in response to changes, the concept of digital twins is a promising solution. An overview of digital twins and their applications in offsite construction are provided in the subsequent subsections.

6.1.3 Overview of digital twins

The concept of digital twins was introduced in the early-2000s by Grieves, who described a digital twin as a digital representation of a physical product consisting of three elements: the physical product in real space, the virtual product in virtual space, and the bidirectional data and information connections that link them (Grieves, 2014). In recent years, the concept of digital twins has garnered significant attention from technology companies which, in turn, have endeavored to further promote this concept in various industries. While each company may have its own definition of digital twins, these definitions generally align with Grieves' original concept. For instance, Siemens defines a digital twin as “a virtual model of a physical object or system that can be used to simulate the behavior of that object or system to better understand how it works in real life” (Tronel, 2023). Similarly, IBM defines a digital twin as “a virtual representation of an object or system that spans its lifecycle, is updated from real-time data, and uses simulation, machine learning and reasoning to help decision making” (IBM, 2023). Amazon has also joined this tech wave and is promoting the concept of digital twins, with a definition that closely aligns with that of IBM (Amazon, 2023). Moreover, the digital twin concept is now being applied in sectors ranging from retail, to healthcare, mining, manufacturing, agriculture, automotive, education, aerospace, and construction (Attaran et al., 2023).

The momentum surrounding digital twins is fueled by the multitude of applications and benefits this concept offers. For instance, digital twins can be deployed for remote performance monitoring and management, quality assessment, defect detection, safety monitoring, process or product optimization, process or product design, predictive maintenance, project control, productivity improvement, cost reduction, efficiency enhancement, predictive analytics, data visualization, education, and training purposes (Attaran et al., 2023; Autodesk, 2023; Cooper et al., 2022; Gulewicz, 2022; Pavel Orlov, 2023). However, despite these numerous benefits, the application of digital twins in the offsite construction industry is still in its early stages, as explained further in the following subsection.

6.1.4 Digital twin applications in offsite construction

Although the controlled factory setting typical of offsite construction is well suited to leveraging the benefits of digital twins, relatively few studies have targeted the application of digital twins to offsite construction. In fact, a search of the Scopus database, the Compendex database, and the Web of Science platform for publications on digital twins in offsite construction following the steps detailed in a previous review study (Alsakka et al., 2023a) identified fewer than twenty publications discussing digital twins in offsite construction (as of May 22, 2023). Moreover, among the identified studies, only a subset presented full or partial applications, development, or testing of the proposed digital twin systems (e.g., (Jiang et al., 2022; D. Lee & Lee, 2021; Z.-S. Liu et al., 2022; Rausch et al., 2021; Tran et al., 2021; Y. Zhao et al., 2022)).

These studies have examined a range of different aspects of the offsite construction industry, including on-site assembly planning and scheduling, management of hoisting operations, management of transportation risks, quality assessment, and the management of hoisting-related safety risks (Jiang et al., 2022; D. Lee & Lee, 2021; Z.-S. Liu et al., 2022; Rausch et al., 2021; Tran et al., 2021; Y. Zhao et al., 2022). Moreover, each of these studies presented a unique digital twin architecture tailored to the given requirements. For instance, Lee and Lee (2021) developed a digital twin system aimed at monitoring and simulating logistics operations, focusing on the transportation of prefabricated modules. Given the real-time location of transportation trucks, the authors used the Unity game engine, Bing Maps, and a geographic information system (GIS) for virtual modelling, for the assessment of logistics risks (e.g., accidents, road accessibility) on the routes taken by the trucks, and for the generation of alternative routes if needed. It is worth noting that the authors did not implement data collection systems; instead, hypothetical data was used for the location of the trucks. Moreover, the systems needed for communicating the alternative routes with the logistics personnel in real time were not discussed. Another example is a recent study by Tran et al. (Tran et al., 2021) that developed a digital twin-based framework for evaluating the geometric quality of installed prefabricated façades. They deployed laser scanning technology to capture 3D geometric data and, subsequently, to generate a digital replica of the installed façade. The aim here was to compare the digital as-built façade with the as-designed façade in order to evaluate the accuracy, completeness, and correctness. It is worth noting that the conversion of the point cloud data captured using the laser scanner to the 3D as-built model was performed manually, given the challenges associated with automated conversion as highlighted by the authors.

Overall, these studies demonstrate the versatility and promising potential of digital twin applications in digitizing various management tasks in offsite construction, hence encouraging their adoption for production estimation and scheduling purposes as well.

6.1.5 Research objective and contributions

In light of the above, this study aims to develop a digital twin for dynamic production time estimation, scheduling, and real-time monitoring, DiTES, in offsite construction. The specific research objective is to devise a novel design of a digital twin that improves the accuracy of production estimation and scheduling compared to the current practice and provides means for real-time monitoring in offsite construction. To achieve this, DiTES integrates computer vision, ultrasonic sensors, machine learning-based prediction models, and 3D simulation. The study builds upon and extends previous research efforts that yielded a computer-vision-based process time data acquisition system (Alsakka et al., 2023b) and a system that estimates process CTs at the workstation level in offsite construction factories (Alsakka et al., 2023d). The main contribution of the proposed digital twin, DiTES, is two-fold, as outlined below:

- The synergy and capabilities achieved through the integration of various technologies in the developed design of DiTES. Specifically:
    - Machine learning enables the identification of patterns in operations to make improved predictions of process-related time variables.
    - Computer-vision technology enables automated acquisition of the data needed to train and continuously tune the machine-learning models. This saves significant time and effort that would otherwise be spent on manual time studies and ensures that the machine-learning models constantly account for new sources of variability affecting operations.
    - Ultrasonic sensors enable real-time monitoring and tracking of the actual work progress, enabling dynamic production scheduling.
    - 3D simulation enables realistic visualization of all sources of data in a single virtual replica of the factory and the generation of production schedules based on simulations of the real operations.
- The improved accuracy of production schedules compared to those developed using fixed average production rates (i.e., quantity per unit of time) commonly employed in current practice.

The remainder of this chapter is organized as follows: Section 6.2 describes in detail the challenge of production estimation and scheduling in the context of offsite construction. Section 6.3 provides a description of the DiTES' architecture and components and outlines the methods used to deploy DiTES for a semi-automated wood-wall framing workstation at a lightweight wood panelized construction factory. Section 6.4 presents the results of the evaluation of DiTES' performance. Section 6.5 discusses the results. Finally, Section 6.6 summarizes the findings of the study, discusses implications for the industry, lists the limitations, and suggests avenues of future research.

6.2 More on the challenge of production estimation and scheduling in offsite construction

To demonstrate the production estimation and scheduling challenge, consider the first four workstations on the wall production line at a lightweight wood panelized construction factory displayed in Fig. 6-1, which was also used as a case company in the previous works upon which the present study builds (Alsakka et al., 2023d., 2023b). At these workstations on the wall production line, wood walls undergo framing, sheathing, nailing, and cutting operations. Each workstation is assigned a well-defined set of tasks for producing wall panels. Specifically, wall panels are first framed at the framing workstation using a semi-automated wood-framing machine that automatically performs nailing, drilling, and cutting operations while one loads the machine with framing elements (e.g., studs, subassemblies for doors and windows) when prompted by the machine to do so. The wall frame is then transferred to the first sheathing table using a conveyor system, where workers manually inspect the frame for any errors, perform the necessary adjustments, and attach missing elements. Next, the panel is pulled to the second sheathing table using the conveyor, where sheathing is manually placed and temporarily stapled to the frame. Workers then transfer the wall panel to the multifunction bridge which automatically fastens sheathing and cuts openings. The wall panel is then transferred to downstream workstation for further processing.

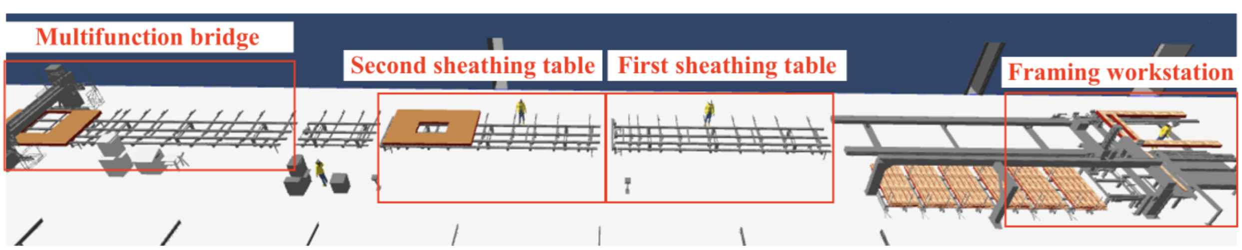
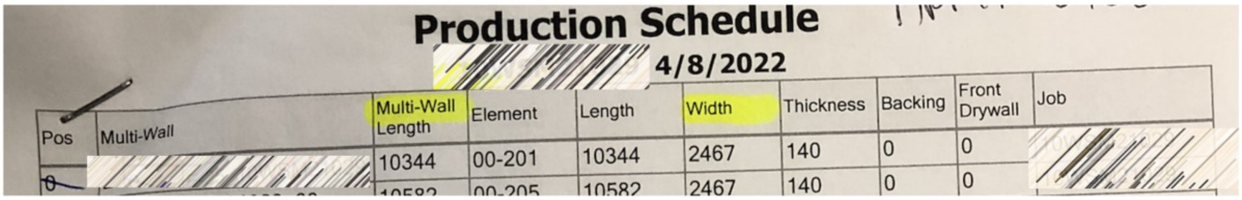


Fig. 6-1. Sample production line.

Given this manufacturing-like production of building components, each workstation receives a daily production schedule listing the components to be processed on that specific day in a given sequence at the workstation as shown in Fig. 6-2. This production schedule will be referred to as "workstation-level production schedule" for the balance of this chapter, distinguishing it from the operations-level production schedule, which provides information on the expected total production time of a component.



Production Schedule

[redacted] 4/8/2022

Pos	Multi-Wall	Multi-Wall Length	Element	Length	Width	Thickness	Backing	Front Drywall	Job
	[redacted]	10344	00-201	10344	2467	140	0	0	[redacted]
		10582	00-205	10582	2467	140	0	0	

Fig. 6-2. Workstation-level production schedule in current practice.

As discussed in the introduction section, variability causes the process CTs at each of these workstations to fluctuate. This results in an imbalanced production line with workstations operating out of sync. Consequently, resources frequently transition between active work and waiting for other resources at different workstations to complete their tasks. Such variability affects the accuracy of production schedules, resulting in a disparity between the scheduled and actual production outputs. The variables provided in Table 6-1, which are relevant to the digital twin proposed in this study, further clarify the dynamics of the operations, which add complexity to the task of developing production schedules. (It should be noted that there may be minor variations in the definitions of these variables offered in other sources in the literature.)

Table 6-1. Variables definitions (Alsakka et al., 2023d).

Variable	Definition
Cycle time (CT)	CT is the total time spanning from the start of the process undertaken at a workstation for a given component until the end of the process, where a “cycle” refers to the set of tasks assigned to a workstation for a single component (e.g., one wall panel).
Cycle start time (ST)	ST is the time at which a cycle starts at a workstation.
Cycle finish time (FT)	FT is the time at which a cycle finishes at a workstation.
Processing or productive time (PT)	PT is the time spent by resources processing a component during a cycle at a workstation. Under ideal conditions, CT is equal to PT.

Table 6 1. Variables definitions (Alsakka et al., 2023d) (continued).

Variable	Definition
Cycle delay (CD)	CD is the time during which work is not performed on the component during a cycle at a workstation or, in other words, the amount of time it takes a cycle to be completed beyond the expected completion time, which is PT. CD can be broken down into two types of delays described below.
Predictable cycle delays (PCD)	PCD is interruptions to a cycle that can be anticipated and estimated to a certain extent. Examples of PCD include scheduled breaks, meetings, training sessions, predictable unavailability of resources, scheduled maintenance, and waiting for a slow material preparation process.
Unpredictable cycle delays (UCD)	UCD arises from random events such as machine breakdowns, machine malfunctions, errors in shop drawings, power outages, worker injuries, phone calls, conversing with coworkers, and bathroom breaks.
Inter-cycle total delay (ITD)	ITD is the total time spanning from the end of a cycle at a workstation to the start of the subsequent cycle.
Downstream-related waiting time (DW)	DW is the time spent waiting for the completed component to be transferred to the downstream workstation. Specifically, it is the time spanning from the end of a process undertaken at a workstation to the time at which the component is transferred downstream. For example, if the workers at the second sheathing table finish their tasks on a wall panel before the multifunction bridge completes its process on the previous panel, they would have to wait for the multifunction bridge to finish before they can transfer the completed panel downstream. Although DW is not factored into the CT for a given cycle, it affects the start time of the subsequent cycle.
Upstream-related waiting time (UW)	UW is the time spent waiting (after a completed component is transferred to the downstream workstation) for the upstream workstation to complete work before a new cycle can be started. For instance, if workers at sheathing workstation complete a panel and transfer it to the multifunction bridge before the framing process is completed for the subsequent panel, they would have to wait for the framing workstation to finish before they can start a new cycle. Since it occurs before a new cycle is started, UW, like DW, is not factored into CT.
Inter-cycle additional delay (ID)	ID is the time by which the start of a new cycle is delayed beyond DW and UW due to any of the aforementioned reasons that cause CD. Note that the total duration of the related delay may be longer than ID, but it may overlap with DW and UW, which is why ID, as defined herein, specifically refers to the additional delay that exceeds the durations of DW and UW. Like CD, ID can arise from both predictable and random events, generally rendering it a random occurrence.

Based on the above definitions, the variables are interrelated as per Eq. (6.1) and Eq. (6.2).

$$CT_i = PT_i + CD_i = PT_i + (UCD_i + PCD_i) \tag{6.1}$$

$$ITD_{[i;i+1]} = ST_{i+1} - FT_i = DW_{i-1} + UW_{i+1} + ID_{[i;i+1]} \tag{6.2}$$

where i refers to the panel of sequence i at a given workstation, $i + 1$ refers to the sequence of the subsequent panel, which is transferred from the upstream workstation, and $i - 1$ refers to the sequence of the previous panel which is processed at the downstream workstation. As expressed in Eq. (6.1) and Eq. (6.2), each CT_i and $ITD_{[i;i+1]}$ is a function of three variables each of which, in turn, is a function of other factors as explained in Table 6-1. This results in a wide range of values for each CT_i and $ITD_{[i;i+1]}$ at each workstation, meaning that determining the list of panels that can be framed or sheathed on a particular day or, in other words, creating workstation-level production schedules for the framing and sheathing workstations is not a simple mathematical calculation. Indeed, using average rates to estimate CT_i and develop a workstation-level production schedule cannot accommodate the inherent complexity of the problem. In other words, such an approach oversimplifies the reality of the situation. Instead, by mimicking the actual operations within a virtual environment, a more accurate representation of the real-world dynamics of the workstations can be obtained.

6.3 DiTES architecture and development methods

DiTES comprises three subsystems that interact with each other to perform the following main functions:

(1) Estimate CTs at each workstation.
(2) Acquire actual data on ST, CT, and PT for each workstation.
(3) Generate production schedules:
    a. Generate workstation-level production schedules which essentially outline the sequence of building components to be produced on a given date at each workstation along with the estimated ST and FT for each component at each workstation.

b. Generate operations-level production schedules, which essentially outline the list of jobs and corresponding building components scheduled for each day along with their estimated total production time.

(4) Track and monitor progress against the production schedules in real time and, when needed, generate updated schedules to accommodate any deviations from the initial schedules.

The architecture of DiTES is displayed in Fig. 6-3, and its subsystems are presented in the following subsections.

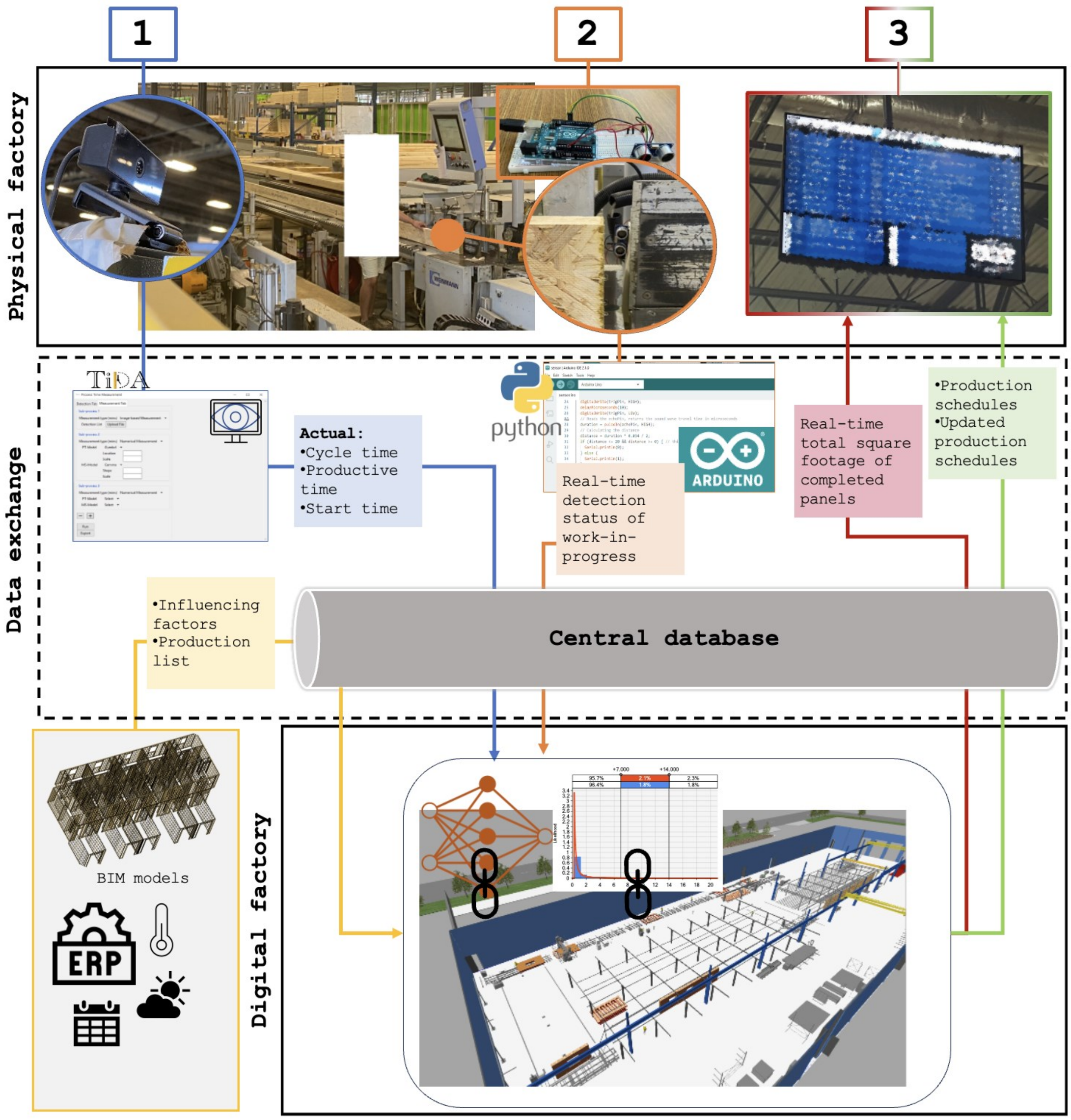


Fig. 6-3. DiTES architecture.

6.3.1 Estimation and scheduling subsystem

6.3.1.1 Overview

This subsystem integrates simulation, machine learning, and statistical modelling to mimic factory operations in a virtual environment as shown in Fig. 6-4. Given a list of jobs and corresponding building components in queue for production, the subsystem simulates their production as they flow from one workstation to another. The subsystem is set to generate the following schedules as it mimics operations:

(a) Daily workstation-level production schedule for each workstation, with estimated framing ST and FT for each panel, as shown in Fig. 6-5(a).

(b) Daily operations-level production schedule for the entire production line, with estimated total production time, as shown in Fig. 6-5(b).

These schedules only include the lists of building components that can be produced within operating hours on a given day, while the production of the remaining components is simulated to occur on subsequent days. These workstation-level production schedules are to be communicated to the relevant workstations through digital signage installed at the workstations (as depicted by the 3rd data connection highlighted in green in Fig. 6-3) at the start of the day. The operations-level production schedule, on the other hand, is intended for the management personnel to incorporate into higher-level (i.e., job-level) schedules.

Fig. 6-4. 3D model of the factory.

a

	Framing ST	Job Nb	Multi_Panel_ID	Length_ft (Feet)	Thickness_in (Feet)	Height_ft (Feet)	Framing FT
1	02/03/2023 7:22:43 AM	10WSM21027	E-6_10WSM21027_00	33.3465	18.0834	8.0000	02/03/2023 7:37:55 AM
2	02/03/2023 7:38:00 AM	10WSM21027	E-25_10WSM21027_00	39.1535	18.0834	9.0000	02/03/2023 7:50:04 AM

b

Production Finish Time	Job Nb	Single_Panel_ID	Production Start Time	Total Production Time (Minutes)
02/03/2023 10:58:14 AM	10WSM21027	186	02/03/2023 9:42:47 AM	75.4435
02/03/2023 11:07:30 AM	10WSM21027	187	02/03/2023 10:10:05 ...	57.4116

Fig. 6-5. (a) Subsystem's workstation-level production schedule; (b) Subsystem's operations-level production schedule.

However, as discussed in the Introduction section of this chapter, and given the multitude of factors that influence operations, estimates generated using this subsystem are likely to contain errors that accumulate over the course of the day and result in deviations between the generated schedules and actual operations. In this regard, if the tracking and monitoring subsystem (described below) detects deviations from the base schedules communicated at the start of the day, updated schedules can be generated and displayed on the signage, ensuring that the shop floor is informed of any changes and can adapt accordingly. However, management personnel may select any of several strategies for rescheduling as they see fit given their factory operations. Specifically, they may follow a periodic (i.e., at regular time intervals), event-driven (e.g., if the schedule deviates by a defined threshold of time from actual production or if there is an unexpected event, such as machine breakdown), or hybrid strategy (i.e., combination of periodic and event-driven) (Ouelhadj et al., 2009).

It should be noted that the production schedules generated using this subsystem follow a predefined production sequence of building components and may not be optimal; the integration of a sequence optimizer into DiTES to obtain optimal schedules is proposed in the "Future work" section of this chapter.

6.3.1.2 Methods

This subsystem extends the CT estimation system developed in a previous related work (Alsakka et al., n.d.). The CT estimation system was developed, evaluated, and then expanded into DiTES' estimation and scheduling subsystem with reference to the case semi-automated wall framing workstation (Figure 6-1). The estimation and scheduling subsystem employs the following methods:

- Machine learning for predicting PT considering various factors that could contribute to its variability: In the application to the framing workstation, a multi-layer feed-forward artificial neural network (ANN) model was used to predict PT as a function of 23 influencing factors: (1) framing date; (2) panel length; (3) panel thickness; (4) panel height; (5) height difference between subsequent panels; (6) panel sequence; (7) number of studs; (8) number of double studs; (9) number of L-shaped studs; (10) number of multi-ply studs; (11) number of regular doors; (12) number of large doors; (13) number of garage doors; (14) number of regular windows; (15) number of large windows; (16) number of cutting zones; (17) number of drilled holes; (18) number of blocks; (19) number of components; (20) wall panel design complexity; (21) work shift (i.e., morning vs. afternoon); (22) the day of the week; and (23) ambient temperature. The model was trained using the open-source machine-learning platform, H2O (H2O.ai, 2023b), accessed through Python.
- Statistical models for estimating UCD and ID given their random nature: In the application to the framing workstation, a Pareto distribution with a shape parameter of 0.94 was used for modelling UCD. The distribution was capped at 4 min (beyond which the sampled UCD values consistently overestimated the observed values). For modelling ID, a Chi-square distribution with two degrees of freedom was used.
- Simulation that (a) models PCD, UW, DW, and ITD as it simulates interactions between workstations, (b) estimates CT in conjunction with the machine-learning and statistical models and, accordingly, (c) generates workstation-level and operations-level production schedules based on a given list of jobs and components in queue for production as it simulates the entire factory operations: In the application to the framing workstation, the Simio simulation software (Simio, 2023) was used for this purpose. Specifically, a simulation model was developed for the entire wall production operations in order to mimic the actual logic of operations (e.g., by determining when a framed wall panel is ready to be transferred to the downstream workstation and, accordingly, when a new framing cycle can be started, or halting production during scheduled events) and, hence, simulate the production of wood wall panels that are stored in a Microsoft Access database linked to the Simio model. For the framing workstation, the ANN and statistical models were used to model PT, UCD, and ID. For the remaining workstations, these

variables were modelled by incorporating the LR models developed by Shafai (2012) in a prior study conducted on the same case production line. (These models are outdated and, hence, only provide rough estimates of the variables, as noted the Limitations subsection below.) The aforementioned study (Alsakka et al., n.d.) provides comprehensive details on the deployment of these methods, with the exception of the scheduling performance, which is discussed in Section 6.5 of this chapter.

Finally, it should be noted that the case factory does not currently have digital signage installed at each workstation. Hence, the green connection shown in Fig. 6-3 has not yet been tested. However, this task is technical in nature and does not affect the estimation and scheduling performance of DiTES.

6.3.2 Training data acquisition subsystem

6.3.2.1 Overview

The machine-learning and statistical models discussed above must undergo regular tuning in order for their performance to be sustained and enhanced. This is due to the fact that machine-learning models may fail to capture the correlation between PT and a given factor if that factor remains constant during the period covered in the training dataset, and the statistical models are likely to become less accurate if new sources of delays arise in operations. For example, we can consider a scenario in which an expert worker is replaced with a less-experienced and less-efficient worker at a particular workstation. If the dataset used for training machine-learning models to predict PT only contains data from when the skilled worker was in place, the models are likely to underestimate the time required for the less experienced/efficient worker to complete a cycle. This substitution of one worker for another introduces a new source of variability in PT that the models were not trained to account for. Consequently, the models would need to be retrained to account for the variability between workers (Alsakka et al., n.d.).

Therefore, continuous acquisition of training data is needed in order to maintain the accuracy of DiTES' estimation and scheduling subsystem. This, in turn, necessitates an automated data acquisition system. As such, this subsystem deploys computer-vision technology to extract training data from images captured using cameras installed at the workstations (as depicted by the 1st data connection highlighted in blue in Fig. 6-3), as explained in the following subsection.

6.3.2.2 Methods

This subsystem uses the computer-vision-based time data acquisition system developed in a previous related work (Alsakka et al., 2023b), which enables the automated measurement of ST, CT, and PT for every component processed at a workstation. This system collects data on these variables based on the following logic. Briefly, for some workstations in offsite construction factories, such as the two sheathing tables and the multifunction bridge in Fig. 6-1, the process begins with work-in-progress (WIP) (e.g., a partially completed wall panel) being pulled into the workstation from upstream and ends with WIP being transferred to the downstream workstation. For other workstations, such as the framing workstation in Fig. 6-1 (i.e., the first station on the wall production line), the process involves loading elements (e.g., studs for panel frames) onto the workstation for assembly rather than beginning with WIP being pulled from upstream. Either way, WIP/material flows into and out of workstations; consequently, certain points along the workstation become blocked and unblocked by WIP/material in a cyclical manner. If we strategically position objects (e.g., stop signs) at these points, these objects become alternately blocked and unblocked with each cycle, as illustrated in Fig. 6-6. Hence, the detection status of these objects can be linked to the cycle's beginning and end and, hence, to ST and CT. This logic was implemented to enable the utilization of object detection algorithms that have undergone extensive training on identifying everyday objects, leveraging open-source datasets containing a substantial volume of annotated images of such objects. The purpose of this was to alleviate the considerable time and effort typically associated with training object detection algorithms to recognize building elements (given their dynamic nature, as they undergo changes in shape and size while progressing through the production line). Regarding productive time, the system operates under the assumption that, when a worker is detected at the workstation during a cycle, the workstation is in use.

Fig. 6-6. (a) Stop signs unblocked; (b) stop signs blocked (Alsakka et al., 2023b).

The computer vision system was also developed and evaluated with reference to the case semi-automated wall framing workstation (Figure 6-1) in the aforementioned previous study (Alsakka et al., 2023b). In that application, the object detection model YOLOv4 (the fourth version of the "You Only Look Once" object detection algorithm) (Bochkovskiy et al., 2020) trained on the COCO dataset (COCO Consortium, 2022) was employed. To ensure consistency, the data collected for building estimation models in this chapter were also based on detections made by YOLOv4. For a more in-depth understanding of the system's logic and application, the interested reader may refer to the previous work (Alsakka et al., 2023b). Moreover, further details concerning the manner in which data collected on ST, CT, and PT is used to train the estimation models can be found in the other previous work on CT estimation (Alsakka et al., n.d.).

6.3.3 Real-time tracking and monitoring subsystem

6.3.3.1 Overview

To track and monitor progress against production schedules in real time, DiTES also contains a subsystem that integrates ultrasonic sensors (as depicted by the 2nd data connection highlighted in orange in Fig. 6-3) and simulation. This subsystem tracks the count of components processed at each workstation, as well as their corresponding total quantity (e.g., square footage for wall panels)—or any other metric the management personnel may be interested in—and mirrors the actual operations in a virtual replica of the factory, as shown in Fig. 6-7.

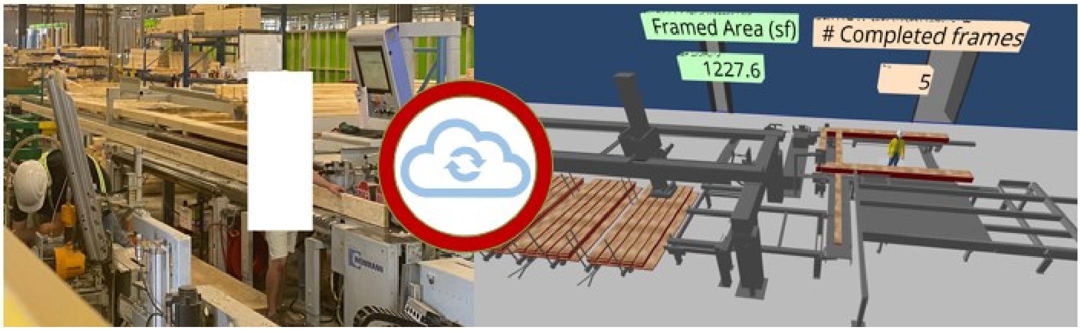


Fig. 6-7. Synced physical and virtual workstations.

The logic underlying this subsystem is similar to that used in the data acquisition subsystem as explained in the previous subsection. By strategically installing an ultrasonic sensor at a workstation in such a manner that it becomes obstructed when (WIP)/material is present and unobstructed when the WIP is transferred out of the workstation, we can track the number of completed cycles.

The tracking subsystem includes a replica of the simulation model used in the estimation and scheduling subsystem, described above. In the remainder of this chapter, the simulation model integrated with the machine-learning and statistical models is referred to as the "first simulation replica" and the simulation model used in the real-time tracking subsystem is referred to as the "second simulation replica". In the second simulation replica, the production of components and their flow across the shop floor are not governed by the estimated CT and ITD. Rather, they are controlled by the statuses of the ultrasonic sensors installed at the workstations in such a manner that the second simulation replica mirrors the real-world operations in real time. By mirroring the actual operations, the second simulation replica tracks the progress of work. This measured progress is communicated with the shop floor in real time via the digital signage (as depicted by the 3rd data connection highlighted in red in Fig. 6-3). The subsequent section further clarifies how this subsystem functions.

6.3.3.2 Methods

The real-time tracking subsystem was also developed and evaluated with reference to the case framing workstation. An infrared sensor (the Robojax E18-D80NK Infrared Obstacle Avoidance sensor) was initially tested at the factory as part of the subsystem's development, but the light coming from open doors and windows was found to interfere with the sensor's detection, causing poor performance. Accordingly, the infrared sensor was replaced with an ultrasonic sensor (the Elegoo HC-SR04 Ultrasonic Module Distance sensor). The ultrasonic sensor was installed at the workstation following the logic described in Section 6.3.3.1 and as shown in Fig. 6-8. It should be noted that the sensor's detection range was adjusted to be smaller than the distance between the sensor and the operator (in order to avoid detecting the operator), as shown in the figure.

Fig. 6-8. Ultrasonic sensor testing.

A simple sensor setup was employed since the sole purpose of this installation was to test the subsystem. The sensor was connected to an Arduino board, and a Python script in conjunction with the Arduino IDE software was used to read the detection status of the sensor and write the status information to a Microsoft Access database. The database, in turn, was connected to the second simulation replica, which was set to check the status of the sensor at one-second intervals. The second simulation replica was configured such that, when the status of the sensor switches from unblocked to blocked, framing of the first wall panel in the sequence of wall panels is started at the virtual framing workstation. Once the sensor status switches back to unblocked, meanwhile, the cycle is ended, and the completed frame is transferred downstream. When the completed frame is transferred downstream, the count of frames completed at the workstation (highlighted in orange in Fig. 6-7) increases by an increment of 1. Moreover, the corresponding surface area of the wall panel is retrieved from a Microsoft Access database and added to the total square footage of framed panels (highlighted in green in Fig. 6-7). When the status switches again to blocked, the framing cycle repeats. In this manner, the virtual framing operations are synced with the real factory framing operations.

As noted above, the workstations at the case factory do not currently feature digital signage. However, the factory does have signage for displaying the total square footage of completed wall panels across the entire wall production line. In current practice, the information displayed on this signage is manually updated by a technician who periodically checks the work progress. The connection between the second simulation replica and the existing digital signage at the factory has not yet been tested, but this is a technical task that would not be expected to affect the performance of the developed system.

6.4 DiTES' performance evaluation results

Since the data acquisition and estimation subsystems were developed in previous works, detailed evaluations of their performance can be found in the corresponding publications (Alsakka et al., 2023d; Alsakka et al., 2023b). All evaluations were carried out with reference to the case framing workstation. In brief, the evaluation of the training data acquisition subsystem demonstrated its ability to measure ST, PT, and CT with mean absolute errors of less than 1 minute. Moreover, the estimation subsystem was found to be capable of predicting framing PT with a mean absolute error of 1.57 min per panel (compared to an actual PT range spanning from about 2 min to about 27 min). It was also shown to be capable of predicting framing CT with a mean absolute error of about 3 min, in comparison to an actual CT range spanning from about 2 min to about 58 min. This error decreased to 2.3 min when excluding from the analysis data for two outlier wall panels with respect to which random extended worker absences from the workstation resulted in lengthy delays. The subsystem realized a 36% reduction in the mean absolute error of CT predictions compared to the current estimation practice, which assumes an average linear fixed rate (metres per minute) of productivity. It is important to note, however, that the subsystem's scheduling performance, described below, is directly influenced by—and, hence, reflective of—the performance of these methods, as they collectively contribute to the generation of production schedules.

The real-time tracking and monitoring subsystem was also tested with reference to the framing workstation, as shown in Fig. 6-8. Based on 247 sensor readings collected at the workstation, instances of true positives (i.e., the sensor accurately detects an obstruction), true negatives (i.e., the sensor does not detect anything when there is no obstruction), false positives (i.e., the sensor erroneously detects an obstruction when there is none), and false negatives (i.e., the sensor fails to detect an obstruction when one is present) were 168, 79, 0, and 0, respectively, meaning that it achieved accuracy and precision rates of 100%. In other words, the sensor made no false detections, and it did not miss any detections.

As for the performance of DiTES in generating workstation-level production schedules, it was also validated and evaluated in reference to the case framing workstation. It should be noted that, since DiTES has not yet been implemented for the entire production line, its performance in generating operations-level production schedules and in estimating the total production time has not yet been tested. This is further discussed in the “Limitations and future work” section of this chapter.

In total 25 wall panels were considered for evaluation (this is roughly the output of a typical shift at the case framing workstation) in order to examine the performance of DiTES in scheduling production for a work shift. As noted above, in current practice the case company relies on an average linear fixed rate (measured in metres per minute) as the basis for estimating the full production capacity of the wall production line. Moreover, production schedules are generated based on the assumption that the wall production line operates at 85% of full capacity (Alsakka et al., n.d.). In this section, the scheduling performance of DiTES is compared to both the fixed-rate method used in current practice and actual production. Timestamps of framing STs and FTs, estimated using both DiTES and the fixed-rate method for each of the 25 panels, are provided in Table 6-2.

It should be noted that the first simulation replica completely halts production during scheduled breaks. This means that if a coffee or lunch break is scheduled to start before a framing cycle is completed, the cycle will stop when the break begins and resume afterward. For example, according to DiTES' schedule, the framing of Panel 23 is completed at 11:57:35 a.m., i.e., just before the scheduled lunch break at 12:00:00 p.m. Consequently, the framing of Panel 24 would start just before the lunch break. However, to facilitate a simpler comparison among the scheduling methods, a slight modification was made to DiTES' schedule. The lunch break was adjusted to start at 11:57:35 a.m. and end at 12:27:35 p.m. (The same logic applies to the coffee break.) On a different aspect, in the fixed-rate-method schedule, the framing ST of a panel is set equal to the framing FT of the previous panel, considering that any delays are factored into the assumption that the wall production line operates at 85% of its full capacity.

Table 6-2. Evaluation results.

Panel#	Actual		DiTES		Fixed-rate method	
	Framing ST	**Framing FT**	**Framing ST**	**Framing FT**	**Framing ST**	**Framing FT**
1	07:22:42 AM	07:44:43 AM	07:22:42 AM	07:37:55 AM	07:22:42 AM	07:31:20 AM
2	07:44:48 AM	07:56:18 AM	07:38:00 AM	07:50:03 AM	07:31:20 AM	07:41:27 AM
3	07:59:24 AM	08:08:18 AM	07:52:33 AM	08:03:13 AM	07:41:27 AM	07:49:29 AM
4	08:08:23 AM	08:18:23 AM	08:03:42 AM	08:14:22 AM	07:49:29 AM	07:58:55 AM
5	08:18:33 AM	08:36:11 AM	08:17:41 AM	08:32:18 AM	07:58:55 AM	08:05:10 AM
6	08:36:21 AM	08:43:55 AM	08:32:31 AM	08:40:09 AM	08:05:10 AM	08:10:19 AM
7	08:44:00 AM	08:48:47 AM	08:41:06 AM	08:43:59 AM	08:10:19 AM	08:11:48 AM
8	08:48:51 AM	09:00:17 AM	08:45:27 AM	08:56:09 AM	08:11:48 AM	08:21:55 AM
9	09:04:08 AM	09:15:19 AM	08:57:01 AM	09:08:16 AM	08:21:55 AM	08:30:15 AM
10	09:15:24 AM	09:21:21 AM	09:09:18 AM	09:13:46 AM	08:30:15 AM	08:32:41 AM
11	09:21:26 AM	09:32:11 AM	09:16:46 AM	09:28:10 AM	08:32:41 AM	08:42:54 AM
12	09:32:16 AM	09:42:41 AM	09:28:27 AM	09:37:41 AM	08:42:54 AM	08:52:57 AM
	Coffee break					
13	09:56:27 AM	10:10:18 AM	09:37:42 AM	09:49:49 AM	08:52:57 AM	09:02:48 AM
			Coffee break			
14	10:13:40 AM	10:41:43 AM	10:04:49 AM	10:15:51 AM	09:02:48 AM	09:12:21 AM
15	10:41:53 AM	10:47:41 AM	10:16:15 AM	10:21:40 AM	09:12:21 AM	09:15:34 AM
16	10:47:45 AM	11:01:22 AM	10:21:58 AM	10:33:23 AM	09:15:34 AM	09:24:49 AM
17	11:01:42 AM	11:10:01 AM	10:33:30 AM	10:42:09 AM	09:24:49 AM	09:34:52 AM
18	11:10:36 AM	11:23:17 AM	10:42:15 AM	10:56:04 AM	09:34:52 AM	09:45:06 AM
					Coffee break	
19	11:23:22 AM	11:33:47 AM	10:56:04 AM	11:05:58 AM	10:00:06 AM	10:10:09 AM
20	11:33:52 AM	11:44:12 AM	11:10:30 AM	11:22:46 AM	10:10:09 AM	10:20:16 AM
21	11:44:17 AM	11:56:03 AM	11:22:46 AM	11:35:13 AM	10:20:16 AM	10:30:33 AM
22	11:56:18 AM	12:04:12 PM	11:37:11 AM	11:45:51 AM	10:30:33 AM	10:40:36 AM
	Lunch break					
23	12:37:07 PM	12:46:16 PM	11:46:25 AM	11:57:35 AM	10:40:36 AM	10:50:04 AM
			Lunch break			
24	12:46:31 PM	12:53:04 PM	12:27:35 PM	12:36:33 PM	10:50:04 AM	11:00:18 AM
25	12:53:09 PM	01:01:28 PM	12:36:44 PM	12:46:47 PM	11:00:18 AM	11:10:30 AM

As the results show, both DiTES and the fixed-rate method were found to underestimate the total time needed to frame the 25 panels, but DiTES' underestimation was found to be significantly less pronounced than that of the fixed-rate method. DiTES predicted that Panels 1 to 23 could be framed before the lunch break, whereas, in reality, just Panels 1 to 22 were completed before the break, indicating a deviation of one panel in DiTES' schedule. Meanwhile, the fixed-rate method estimated that 18 panels could be framed before the coffee break, while DiTES estimated just 13 panels could be completed within that time frame. Meanwhile, only 12 panels were actually completed before the break. Furthermore, the fixed-rate method estimated that all 25 panels could be framed by 11:10:30 a.m., but this turned out to be an overly optimistic estimation.

6.5 Discussion

To visualize the deviations of the DiTES-based schedule and the fixed-rate-based schedule from actual production, the cumulative total time elapsing from Panel 1's ST to the FT of each panel is plotted in Fig. 6-9. For the purpose of comparison, the lunch break (30 min) was added to the fixed-rate-based schedule, even though the estimated schedule completion time was prior to the start of the break. The plot demonstrates that the DiTES-based schedule was found to underestimate the total elapsed time by 15 min, resulting in a percent error of 4.3%. Meanwhile, the fixed-rate-based schedule underestimated it by 81 min, resulting in a percent error of 23.9%. Although the DiTES-based schedule was found to be optimistic, it achieved a reduction of approximately 81% in the deviation from actual production compared to the current estimation and scheduling practice.

Moreover, the deviation of 15 min over a period of 5.5 hours (from 07:22:42 a.m. to 01:01:28 p.m.) can be considered acceptable. Unpredictable delays were the main source of error contributing to this deviation. The Pareto distribution that was used to model UCD was capped at 4 min, since sampled UCD values exceeding 4 min consistently and significantly overestimated the observed values (refer to Section 3.6 of the previous study (Alsakka et al., n.d.) for further details). Using a cap of 4 min means that actual UCD values exceeding 4 min will be underestimated by the estimation subsystem. Given this, and considering that CT is a function of UCD, as per Eq. (6.1), CT was occasionally significantly underestimated by DiTES.

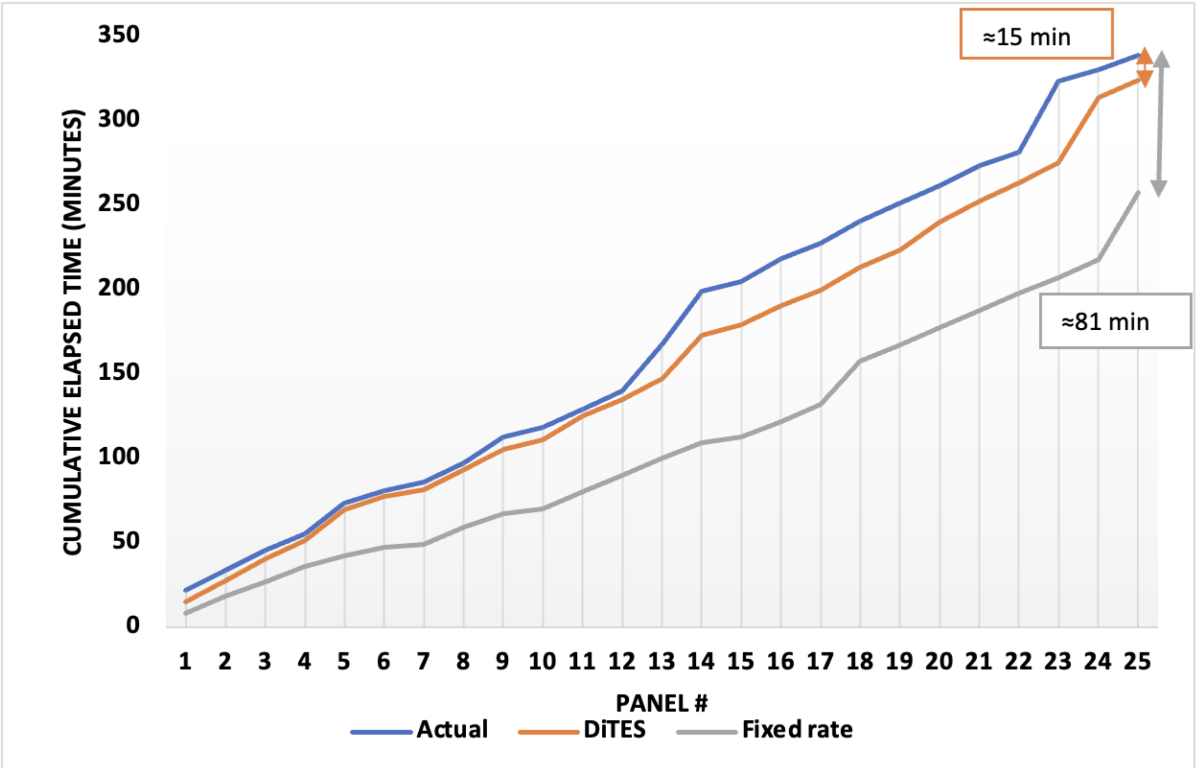


Fig. 6-9. Schedule deviations.

To visualize this, consider Fig. 6-10(a) which plots the individual prediction errors (i.e., difference between CT predicted by DiTES and actual CT) for each of the 25 panels. Despite the frequency of overestimations (i.e., 14 positive prediction errors) being higher than underestimations (i.e., 11 negative prediction errors), the arithmetic error was negative, amounting to -0.89 min, due to the larger magnitude of underestimations. Specifically, as the chart shows, the highest prediction error was an underestimation of -17 min for Panel #14. The actual value of UCD corresponding to this panel was approximately 16 min as the operator unpredictably left the workstation while framing it; meanwhile, the estimated UCD randomly sampled from the Pareto distribution was only 0.1 min. This implies that the estimated UCD alone contributed to about 16 min of the 17 min error. The individual prediction errors resulted in a total CT underestimation of about 22 min for the 25 panels as shown in Fig. 6-10(b). It should be noted that a portion of this underestimation was offset by a slight overestimation of ITD, calculated satisfying Eq. (6.2), which explains the smaller schedule deviation of 15 min.

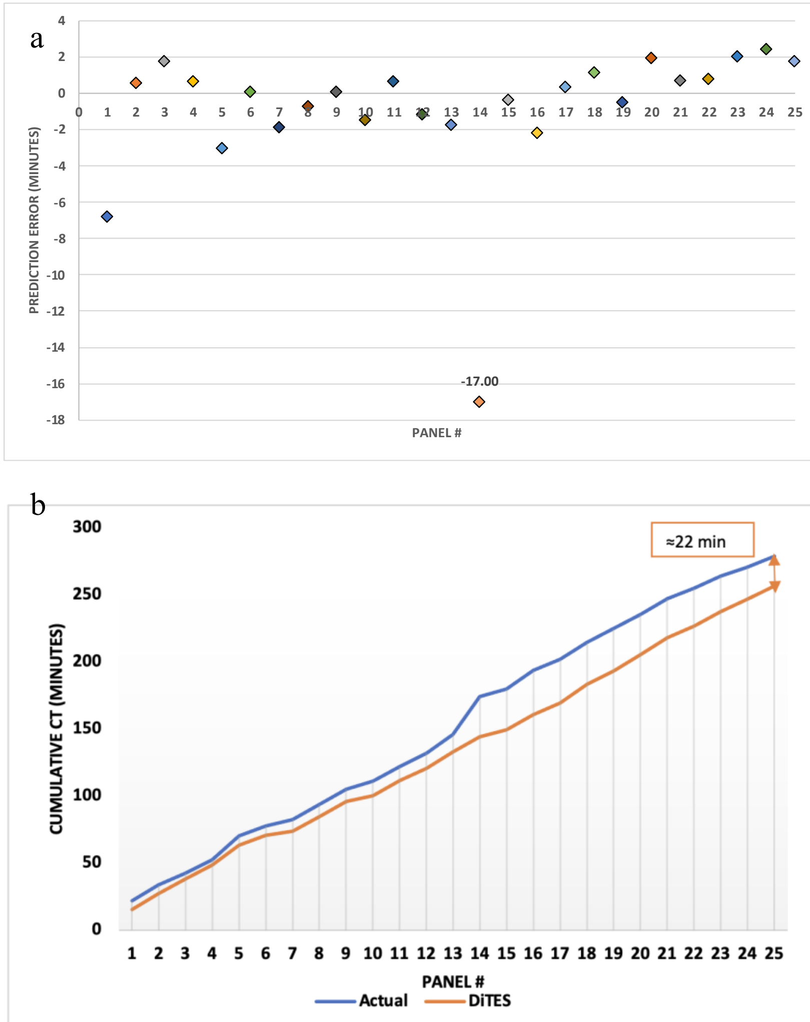


Fig. 6-10. (a) Individual CT prediction errors; (b) CT estimation.

Given the above, better control and management of operations may reduce this error by making UCD more controllable and, hence, more predictable. If achieving better control of operations is not feasible, another potential solution is to incorporate a buffer into the schedule in order to offset this underestimation. Another consideration is that the actual length of breaks does not typically conform exactly to the schedule. For instance, in the case of the studied work shift for the framing workstation, the coffee break lasted for 13.77 min and the lunch break lasted for 32.92 min. As such, the breaks totaled 46.69 min (compared to a total of 45 min of breaks accounted for in DiTES). These deviations are not significant and do not require any mitigation measures, but they do slightly contribute to the error.

While several constraints were encountered during the performance testing process, as explained in the "Limitations and Future Work" section of this chapter, it is noteworthy that DiTES exhibited a highly satisfactory level of performance, surpassing that of the commonly used fixed-rate method. The improvement in the schedule's performance before even employing the real-time tracking and monitoring subsystem was found to be significant. This subsystem enables real-time identification of deviations between the planned schedule and actual operations, helping management personnel to make timely decisions and proactively address deviations before they accumulate and become more challenging to manage. For example, upon identifying the 15-minute deviation observed earlier (towards the end of the first shift), and given its magnitude, management personnel may opt to communicate this deviation to the worker to expedite their work and prevent unnecessary delays or to adjust the schedule as necessary. In this way, management personnel will be better equipped to manage factory operations.

The results obtained also represent improvement compared to previous related work in which the total time needed to process wall panels at the same case framing workstation was predicted with a 10% error (Altaf et al., 2015). Finally, the implication of the results on the offsite construction industry are discussed in the following section.

6.6 Implications for industry practice and contributions to cultivating a lean offsite construction industry

The findings reveal that workstation-level production schedules generated using a single fixed production rate can be overly optimistic. Such optimistic schedules can negatively impact the well-being of the workforce, as employees may feel constant pressure to meet unrealistic deadlines. Additionally, these optimistic production schedules could result in committing to overly ambitious project delivery dates, leading to delays and associated costs. Although DiTES continues to underestimate the total time needed to process a batch of panels, the degree of underestimation is significantly smaller than that of the fixed-rate method. The degree of underestimation is also quantifiable, making it easier to account for by incorporating schedule buffers, for instance. Moreover, a digital twin enables a better understanding of various types of delays, thereby facilitating the identification and mitigation of their sources. Consequently, the proposed digital twin provides a promising production management solution for the offsite construction industry.

In addition to production estimation, scheduling, and real-time monitoring, DiTES can support the implementation of lean principles in offsite construction factories to improve production efficiency and productivity, as outlined below, to name a few examples:

- Production line balancing: The data collected by DiTES on CT, UW, and DW as well as the real-time monitoring data can be used to identify bottlenecks in a production line. Furthermore, the virtual factory replication feature of DiTES enables experimentation with transferring tasks between workstations or adjusting the resources allocated to workstations for the purpose of bringing CTs closer to one another and, thereby, balancing production lines. This supports the implementation of the Heijunka lean principle, which focuses on levelling out the work schedule and balancing production lines.
- Waste reduction: The data collected by DiTES on delays (i.e., PCD, UCD, and ID) and waiting times (i.e., UW and DW) provides insights on the amount of process waste present within operations. By leveraging the virtual factory replication feature, DiTES enables virtual experimentation with various strategies to mitigate sources of process waste. DiTES also allows offsite construction enterprises to estimate the potential benefits of implementing a given mitigation strategy and to evaluate its benefits in relation to the required capital investment. In this manner, DiTES can serve as a decision support tool for enhancing offsite production efficiency through optimization and waste reduction.
- Active involvement of management on the shop floor: Lean philosophy emphasizes the importance of management spending time on the shop floor to gain first-hand knowledge of work processes and to gain understanding of the challenges faced by workers. However, it can be challenging for management to grasp a global picture of the entire shop floor through observation alone. While consulting with workers may aid the manager in identifying local issues pertinent to the workstations, understanding the effect of these local issues at the factory level may be challenging. However, using DiTES, with its 3D virtual visualization of operations and measured metrics of

operations, management and office personnel can gain a comprehensive understanding of shop floor operations at both the workstation and factory levels, thereby enhancing communication and fostering productive discussions between shop floor workers and management.

In summary, aside from the functions of DiTES presented in this study, it also offers valuable support in implementing lean principles. Its capabilities contribute to increased production efficiency and improved decision-making, ultimately fostering an offsite construction environment conducive to the implementation of lean practices.

6.7 Conclusions

6.7.1 Summary

This study introduced a digital twin, DiTES, that replicates production operations in offsite construction factories while providing production scheduling and real-time monitoring capabilities. DiTES integrates a range of technologies, including computer vision, ultrasonic sensors, machine learning-based prediction models, and 3D simulation, to fulfill its intended functions. To evaluate its performance, DiTES was applied to a semi-automated wood-wall framing workstation in a panelized construction facility. The monitoring aspect of DiTES demonstrated 100% accuracy and precision as a result of the perfect performance of the integrated ultrasonic sensor. Furthermore, when assessing DiTES' performance in generating workstation-level production schedules, which reflects its estimation and data acquisition capabilities as well, a significant reduction of 81% in the deviation from actual production was observed when generating a schedule for framing 25 wall panels. This marks a substantial improvement compared to the fixed-rate method commonly employed in current practice, highlighting the system's potential and promising performance.

6.7.2 Limitations and future work

This study was subject to several limitations. First, regarding the evaluation of DiTES' performance, the workstations downstream of the framing workstation on the production line at the case company's production facility were modelled using rough estimates. Given the promising potential exhibited by DiTES during its preliminary evaluation as presented in this study, future research will focus on applying DiTES to entire production lines. This will allow for a more representative evaluation of its overall performance.

Second, since DiTES is yet to be applied to an entire production line, its performance in generating operations-level production schedules has not yet been tested. Given the significant reduction in deviations between the planned schedule and actual production at the workstation level realized by DiTES compared to the fixed-rate method, DiTES can be expected to similarly realize significant improvements for operations-level production schedules. However, quantifying the extent of improvement necessitates the implementation of DiTES across the entire shop floor.

Third, at present DiTES does not feature an optimization subsystem for finding optimal production sequences to minimize the production duration of building components. The fluctuation in CTs at workstations, influenced by various factors, as discussed earlier in this study, highlights the potential impact of the production sequence on the overall production time. For example, in a one-piece-flow production line, scheduling a series of components that can be quickly processed at the first workstation but that require longer processing times downstream may result in waiting times for resources at the first workstation. Adjusting the production sequence of components could help to mitigate these waiting times and reduce the total production time. Various optimization algorithms, including the genetic algorithm (GA), have been employed in the field of offsite construction to address this sequence optimization problem. For instance, GA has been applied in several studies to optimize the production sequence of concrete precast components with the objective of reducing production time (Du et al., 2021; Leu & Hwang, 2002; Z. Xu et al., 2020). In a similar fashion, an optimization algorithm could be integrated with DiTES to find the optimal production sequence of components that minimizes the total production time.

Finally, DiTES has only been evaluated through a case application to the framing workstation. Moreover, its performance in generating workstation-level production schedules has only been tested during a single shift. To comprehensively assess its performance, further evaluations should be conducted by applying DiTES to other workstations. In future applications of DiTES, several measures could be implemented to potentially enhance its performance, including: (1) replacing YOLOv4 in the training data acquisition subsystem with YOLOv7, the latest version in the YOLO object detection algorithm family (Wang et al., 2022), to potentially improve its measurement performance; (2) employing advanced feature engineering techniques to explore the potential benefits of engineered features in enhancing the predictive capabilities of the machine-learning models used in the estimation and scheduling subsystem; and (3) implementing robust control and management strategies to mitigate sources of UCD and ID at the workstations under study. By implementing these measures, the overall effectiveness of DiTES compared to current practice may be further improved.

Acknowledgements

This work was supported by the Alberta Innovates Graduate Student Scholarship, and by the Natural Sciences and Engineering Research Council of Canada (grant file no. ALLRP 577047-22). The authors gratefully acknowledge the technical support of Thomas Huynh and the technical review by Mr. Jonathan Tomalty.

References

References are provided in the Bibliographies chapter of this thesis.

Chapter 7

Summary of Results, Contributions, Limitations, and Future Work

7.1 Summary of key results and corresponding contributions

The offsite construction industry continues to rely on experience-based average production rates in the estimation and scheduling of production operations. Given that operations are subject to the continuous effects of a number of different sources of variability, the use of average production rates in these tasks results in high rates of error with respect to both production estimation and scheduling, including overly optimistic production schedules (as noted in preceding chapters). Hence, the research presented in this thesis included the development of a digital twin for dynamic and accurate production estimation, scheduling, and real-time monitoring of production operations in offsite construction. The proposed digital twin comprises three major subsystems, including (1) a computer-vision-based data acquisition subsystem, (2) an estimation and scheduling subsystem that considers the various of factors driving variability in production operations, and (3) a real-time tracking and monitoring subsystem. Several research needs have been identified in the literature in relation to the digital twin and its subsystems as discussed in the introductory chapter of this thesis. In addressing these research needs and developing the digital twin, including its subsystems, a number of key findings were obtained as summarized in the table below. This table also outlines the corresponding contributions to academia and industry practice that these findings represent. These results and contributions underscore the value and impact of this research in advancing knowledge in the academic sphere while also providing tangible benefits to the industry.

Table 7-1. Summary of main findings and contributions.

Main findings	Corresponding main contributions	
	Academia	**The offsite construction industry**
•A limited number of studies on computer-vision applications in offsite construction were identified. In recent years, the rate of publications has shown a general upward trend, albeit with a relatively modest growth rate ranging from one to eight per year between 2018 and 2021. •The proposed computer-vision-based data acquisition system can measure the start-time, productive or processing time, and cycle time (CT) of a process in offsite construction production with a mean absolute error of less than 1 minute.	•A review of computer vision applications in offsite construction, investigating the current practice, highlighting limitations, identifying research gaps, and proposing potential opportunities to apply computer vision in future research.	•An easy-to-use, time-effective, cost-effective, adaptable and sufficiently accurate system for the automated acquisition of process time data during production operations.
	•A novel method for deploying object detection algorithms pre-trained on common objects in order to study other custom objects without having to retrain them on the custom objects, thereby saving time and computational effort.	
	•Adding to the body of knowledge on computer vision applications in offsite construction, considering the dearth of related applications as of the time of the conducted research.	

Table 7 1. Summary of main findings and contributions (continued).

<table>
<tr><th rowspan="2">Main findings</th><th colspan="2">Corresponding main contributions</th></tr>
<tr><th>Academia</th><th>The offsite construction industry</th></tr>
<tr><td rowspan="3">•Process mapping and semi-structured interviews are effective tools for the identification and understanding of the effect of various factors on CTs, and they enable modellers to perform knowledge- and empirically-driven machine-learning-based modelling and, hence, to potentially improve the performance of CT estimation models.
•Having a comprehensive pool of influencing factors to consider in the development of CT estimation models helped to reduce the mean absolute error by 17%.</td><td colspan="2">•Understanding the significance of different sources of variability driving CT fluctuations at the workstation level in offsite construction factories.</td></tr>
<tr><td colspan="2">•A straightforward qualitative approach that helps researchers and practitioners to improve the performance and interpretability of machine-learning models developed to predict CTs or related process time variables.</td></tr>
<tr><td>•A preliminary list of the factors that could influence CTs at semi-automated wood framing workstations in offsite construction, providing a basis for future studies to build upon and further investigate.</td><td></td></tr>
<tr><td rowspan="2">•The developed CT estimation system was found to reduce the mean absolute error and sum of errors by 36% and 68%, respectively, compared to the current practice, which relies on average production rates.
•The performance of the estimation system confirmed the validity of using automatically collected data (and, specifically, computer vision data) for training machine-learning models and developing statistical models for the purpose of CT estimation.</td><td>•Enabling the development of CT estimation models, or, more generally, machine-learning-based prediction models using automatically collected data, thus saving researchers the time and effort that would otherwise be spent on manual data collection for training such models.</td><td rowspan="2">•A more accurate CT estimation method compared to the current practice of using average production rates.</td></tr>
<tr><td>•An evaluation and comparison of the performance of the feed-forward ANN, LR, and RF models in predicting CTs at a case semi-automated wall framing workstation.</td></tr>
</table>

Table 7 1. Summary of main findings and contributions (continued).

Main findings	Corresponding main contributions	
	Academia	**The offsite construction industry**
•Based on the cross-validation results, the feed-forward ANN model demonstrated slightly better performance than the LR model in predicting CTs at a case semi-automated wall framing workstation, while the RF model exhibited poorer performance. The LR model was found to be sensitive to multicollinearity, while the RF model was sensitive to features that had low frequencies in the training dataset. •Errors in the CT estimation system were found to be attributable primarily to random delays. It can thus be inferred that effective control of various sources of delays in production operations is critical for the success of any production estimation system.	•Better understanding of the effect of considering the various factors influencing CT, or, more generally, features, on the performance of machine-learning-based estimation models.	
•The proposed digital twin demonstrated an 81% reduction in schedule deviation from actual operations compared to the conventional fixed-rate method commonly employed in current practice. •The proposed digital twin was found to enable real-time tracking of production schedules and, hence, timely updating of schedules in response to unforeseen events. •The use of average production rates in production estimation and scheduling may result in overly optimistic production schedules.	•A novel digital twin design that integrates multiple technologies and methods, including computer vision, machine-learning-based prediction, statistical modelling, ultrasonic sensors, and 3D simulation, for production estimation, scheduling, and real-time progress monitoring, expanding upon the limited body of literature on digital twin applications in offsite construction.	•A more accurate production scheduling method compared to the current practice, which relies on average production rates. •Contribution to increased production efficiency and improved decision-making, ultimately fostering an offsite construction environment conducive to the implementation of lean practices.

7.2 Limitations and future work

The research presented in this thesis was subject to several limitations. The table below outlines these research limitations and suggests potential avenues of future research corresponding to these limitations.

Table 7-2. Limitations and avenues of future research.

Limitations	Avenues of future research
•Use of the presented computer-vision-based time process time data acquisition system is not advisable when the process of interest includes a subprocess that cannot be linked to object following the system's logic and yet shows high variability. •If workers at the workstation at which the data acquisition system is installed are frequently off task, the system will tend to overestimate productive times. •The object detection model used in the data acquisition system, YOLOv4, failed to detect the worker when certain parts of the worker's body were occluded.	•Integrate occlusion handling techniques into the data acquisition system. •Experiment with newer and more advanced object detection models, such as YOLOv7.
•The hypothesis underlying the inclusion of the ambient temperature feature was that, at the case factory, higher temperatures are associated with higher fatigue levels and slower work pace. Although relationships between the ambient temperature factor and processing time at the framing workstation were identified in the dataset used in this study, the ambient temperature was relatively low during the data acquisition period, so this hypothesis could not be adequately tested. •The data collected for the case study only covered a short period of operation, and this limited the opportunity to observe the effect of variations in scheduled workload on CT variability. •In the case study, CTs at the other workstations in the facility were modelled in the simulation using only rough estimates.	•Use data collected during the summer season to test the hypothesis regarding the effect of high temperatures on CT variability. •Use several months' worth of data to ascertain whether or not the scheduled workload contributes to CT variability.
•The proposed digital twin did not incorporate a schedule optimizer	•Integrate an optimization algorithm, such as genetic algorithms, into the digital twin design.

Table 7 2. Limitations and avenues of future research (continued).

Limitations	Avenues of future research
•The digital twin and its subsystems were only applied to a single workstation. As such, the performance of the digital twin in generating operations-level production schedules and in estimating production time were not tested.	•Apply the digital twin and its subsystems to an entire production line and evaluate their performance.

Bibliographies

Chapter 1 references

Adeli, H. (2001). Neural networks in civil engineering: 1989–2000. Computer-Aided Civil and Infrastructure Engineering, 16(2), 126–142. https://doi.org/https://doi.org/10.1111/0885-9507.00219

Ahmadian Fard Fini, A., Maghrebi, M., Forsythe, P. J., & Waller, T. S. (2021). Using existing site surveillance cameras to automatically measure the installation speed in prefabricated timber construction. Engineering, Construction and Architectural Management, 29(2), 573–600. https://doi.org/10.1108/ECAM-04-2020-0281

Ahn, S. J., Han, S. U., & Al-Hussein, M. (2020b). Improvement of transportation cost estimation for prefabricated construction using geo-fence-based large-scale GPS data feature extraction and support vector regression. Advanced Engineering Informatics, 43, Article no.101012. https://doi.org/10.1016/j.aei.2019.101012

Alsakka, F., Assaf, S., El-Chami, I., & Al-Hussein, M. (2023a). Computer vision applications in offsite construction. Automation in Construction, 154, Article no.104980. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104980

Alsakka, F., Yu, H., Hamzeh, F., & Al-Hussein, M. (2023c). Factors influencing cycle times in offsite construction. Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC), 723–734. https://doi.org/https://doi.org/10.24928/2023/0123

Alsakka, F., Yu, H., Hamzeh, F., & Al-Hussein, M. (2023d). Data- and knowledge-driven cycle time estimation in offsite construction. Under review in Journal of Computers and Industrial Engineering.

Altaf, M. S., Al-Hussein, M., & Yu, H. (2014). Wood-frame wall panel sequencing based on discrete-event simulation and particle swarm optimization. Proceedings of the 31st

International Symposium on Automation and Robotics in Construction and Mining, 254–261. https://doi.org/10.22260/isarc2014/0034

Altaf, M. S., Bouferguene, A., Liu, H., Al-Hussein, M., & Yu, H. (2018). Integrated production planning and control system for a panelised home prefabrication facility using simulation and RFID. Automation in Construction, 85, 369–383. https://doi.org/10.1016/j.autcon.2017.09.009

Attaran, M., & Celik, B. G. (2023). Digital Twin: Benefits, use cases, challenges, and opportunities. Decision Analytics Journal, 6, 100165. https://doi.org/https://doi.org/10.1016/j.dajour.2023.100165

Benjaoran, V., & Dawood, N. (2006). Intelligence approach to production planning system for bespoke precast concrete products. Automation in Construction, 15(6), 737–745. https://doi.org/10.1016/j.autcon.2005.09.007

Benjaoran, V., Dawood, N., & Scott, D. (2004). Bespoke precast productivity estimation with neural network model. Proceedings of the 20th Annual ARCOM Conference, 2, 63–74. https://doi.org/http://www.arcom.ac.uk/-docs/proceedings/ar2004-1063-1074_Benjoran_Dawood_and_Scott.pdf

Bhatia, A. P. S., Han, S. H., Moselhi, O., Lei, Z., & Raimondi, C. (2019). Simulation-based production rate forecast for a panelised construction production line. Proceedings of the 2019 CSCE Annual Conference, Article no. CON160. https://doi.org/https://www.csce.ca/elf/apps/CONFERENCEVIEWER/conferences/2019/pdfs/PaperPDFversion_160_0419122655.pdf

Caggiano, A. (2014). Manufacturing system BT - CIRP Encyclopedia of Production Engineering (L. Laperrière & G. Reinhart, Eds.; pp. 830–836). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20617-7_6562

Chryssolouris, G. (2006). *Manufacturing systems: Theory and practice* (2nd ed.). Springer. https://doi.org/10.1007/0-387-28431-1

Du, J., Dong, P., Sugumaran, V., & Castro-Lacouture, D. (2021). Dynamic decision support framework for production scheduling using a combined genetic algorithm and multiagent model. Expert Systems, 38(1). https://doi.org/10.1111/exsy.12533

Hu, X., Lu, M., & AbouRizk, S. (2015). BIM-based data mining approach to estimating job man-hour requirements in structural steel fabrication. Proceedings of the 2015 Winter Simulation Conference, 3399–3410. https://doi.org/10.1109/WSC.2014.7020173

Huang, T. S. (1996). Computer vision: Evolution and promise. In C. E. Vandoni (Ed.), 19th CERN School of Computing (pp. 21–25). https://doi.org/10.5170/CERN-1996-008

IBM. (2022). What is Computer Vision? | IBM. Retrieved April 19, 2022, from https://www.ibm.com/topics/computer-vision

Jiang, Y., Li, M., Li, M., Liu, X., Zhong, R. Y., Pan, W., & Huang, G. Q. (2022). Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly. Automation in Construction, 141, Article no.104397. https://doi.org/10.1016/j.autcon.2022.104397

Khandelwal, P. (2021). Offsite Construction Market by Material (Steel, Wood, Concrete, and Others) and Application (Residential, Commercial, and Industrial): Global Opportunity Analysis and Industry Forecast, 2021–2030.

Ko, C. H., & Wang, S. F. (2011). Precast production scheduling using multi-objective genetic algorithms. Expert Systems with Applications, 38(7), 8293–8302. https://doi.org/10.1016/j.eswa.2011.01.013

Kuhn, M., & Johnson, K. (2019). Feature engineering and selection: A practical approach for predictive models. Chapman & Hall/CRC Press. https://doi.org/https://doi.org/10.1080/00031305.2020.1790217

Lee, D., & Lee, S. (2021). Digital twin for supply chain coordination in modular construction. In Applied Sciences, 11(13). https://doi.org/10.3390/app11135909

Leu, S. Sen, & Hwang, S. T. (2002). GA-based resource-constrained flow-shop scheduling model for mixed precast production. Automation in Construction, 11(4), 439–452. https://doi.org/10.1016/S0926-5805(01)00083-8

Liu, H., Altaf, M. S., & Lu, M. (2015). Automated production planning in panelised construction enabled by integrating discrete-event simulation and BIM. Proceedings of the 5th International/11th Construction Specialty Conference, Article no.048. https://doi.org/https://dx.doi.org/10.14288/1.0076315

Liu, Z.-S., Meng, X.-T., Xing, Z.-Z., Cao, C.-F., Jiao, Y.-Y., & Li, A.-X. (2022). Digital Twin-Based Intelligent Safety Risks Prediction of Prefabricated Construction Hoisting. Sustainability, 14(9). https://doi.org/10.3390/su14095179

Lu, M. (2000). Productivity studies using advanced ANN models [PhD dissertation, University of Alberta]. https://doi.org/https://search.proquest.com/docview/304646918?accountid=188395

Lu, M., AbouRizk, S., & Hermann, U. (2000). Estimating Labour Productivity Using Probability Inference Neural Network. Journal of Computing in Civil Engineering, 14(4), 241–248. https://doi.org/https://doi.org/10.1061/(ASCE)0887-3801(2000)14:4(241)

Martinez, P., Barkokebas, B., Hamzeh, F., Al-Hussein, M., & Ahmad, R. (2021). A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. Automation in Construction, 125, Article no. 103620. https://doi.org/10.1016/j.autcon.2021.103620

Microsoft. (2022a). What is computer vision? https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/overview

Microsoft. (2022b). What is custom vision? https://docs.microsoft.com/en-us/azure/cognitive-services/custom-vision-service/overview

Mohsen, O., Mohamed, Y., & Al-Hussein, M. (2022). A machine learning approach to predict production time using real-time RFID data in industrialized building construction. Advanced Engineering Informatics, 52, Article no. 101631. https://doi.org/10.1016/J.AEI.2022.101631

Moselhi, O., Hegazy, T., & Fazio, P. (1991). Neural networks as tools in construction. Journal of Construction Engineering and Management, 117(4), 606–625. https://doi.org/https://doi-org.login.ezproxy.library.ualberta.ca/10.1061/(ASCE)0733-9364(1991)117:4(606)open_in_new

Parnaby, J. (1979). Concept of a manufacturing system. International Journal of Production Research, 17(2), 123–135. https://doi.org/https://doi.org/10.1080/00207547908919600

Rausch, C., Lu, R., Talebi, S., & Haas, C. (2021). Deploying 3D scanning based geometric digital twins during fabrication and assembly in offsite manufacturing. International Journal of Construction Management. https://doi.org/10.1080/15623599.2021.1896942

Shafai, L. (2012). Simulation based process flow improvement for wood framing home building production lines [MSc thesis, University of Alberta]. https://doi.org/doi.org/10.7939/R3SP65

Song, L., & AbouRizk, S. M. (2008). Measuring and Modelling Labour Productivity Using Historical Data. Journal of Construction Engineering and Management, 134(10), 786–794. https://doi.org/10.1061/(asce)0733-9364(2008)134:10(786)

Tran, H., Nguyen, T. N., Christopher, P., Bui, D. K., Khoshelham, K., & Ngo, T. D. (2021). A digital twin approach for geometric quality assessment of as-built prefabricated façades. Journal of Building Engineering, 41, 102377. https://doi.org/10.1016/j.jobe.2021.102377

Xu, Z., Wang, X., & Rao, Z. (2020). Automated optimization for the production scheduling of prefabricated elements based on the genetic algorithm and IFC object segmentation. Processes, 8(12), 1–25. https://doi.org/10.3390/pr8121593

Zhao, Y., Cao, C., & Liu, Z. (2022). A Framework for prefabricated component hoisting management systems based on digital twin technology. Buildings, 12(3), Article no. 276. https://doi.org/10.3390/buildings12030276

Chapter 2 references

Abbas, S. M., & Singh, S. N. (2018). Region-based object detection and classification using faster R-CNN. Proceedings of the 4th International Conference on Computational Intelligence & Communication Technology (CICT), 1–6. https://doi.org/https://doi.org/10.1109/CIACT.2018.8480413

Ahmadian Fard Fini, A., Maghrebi, M., Forsythe, P. J., & Waller, T. S. (2021). Using existing site surveillance cameras to automatically measure the installation speed in prefabricated timber construction. Engineering, Construction and Architectural Management, 29(2), 573–600. https://doi.org/10.1108/ECAM-04-2020-0281

Ahn, S., Crouch, L., Kim, T. W., & Rameezdeen, R. (2020a). Comparison of Worker Safety Risks between Onsite and Offsite Construction Methods: A Site Management Perspective. Journal of Construction Engineering and Management, 146(9), Article no. 05020010. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001890

Ahn, S., Han, S., & Al-Hussein, M. (2019). 2D drawing visualization framework for applying projection-based augmented reality in a panelised construction manufacturing facility: Proof of concept. Journal of Computing in Civil Engineering, 33(5), Article no. 4019032. https://doi.org/10.1061/(asce)cp.1943-5487.0000843

Albrecht, T., Lüthi, M., & Vetter, T. (2009). Deformable Models BT - Encyclopedia of Biometrics (S. Z. Li & A. Jain, Eds.; pp. 210–215). Springer US. https://doi.org/10.1007/978-0-387-73003-5_88

Alsakka, F., El-Chami, I., Yu, H., & Al-Hussein, M. (2023b). Computer vision-based process time data acquisition for offsite construction. Automation in Construction, 149, 104803. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104803

Alsakka, F., Khalife, S., Darwish, M., & Al-Hussein, M. (2020). Deploying discrete-event simulation and continuous improvement to increase production rate in a modular construction facility. Proceedings of the 2020 Winter Simulation Conference, 1676–1687. https://doi.org/10.1109/WSC48552.2020.9384105

Anand S. Rao, & Gerard Verweij. (2017). What's the real value of AI for your business and how can you capitalise? Retrieved April 1, 2023, from https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International Journal of Social Research Methodology, 8(1), 19–32. https://doi.org/https://doi.org/10.1080/1364557032000119616

Asadi, K., Ramshankar, H., Pullagurla, H., Bhandare, A., Shanbhag, S., Mehta, P., Kundu, S., Han, K., Lobaton, E., & Wu, T. (2018). Vision-based integrated mobile robotic system for real-time applications in construction. Automation in Construction, 96, 470–482. https://doi.org/https://doi.org/10.1016/j.autcon.2018.10.009

Azhar, M. I. H., Zaman, F. H. K., Tahir, N. M., & Hashim, H. (2020). People tracking system using DeepSORT. Proceedings of the 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 137–141. https://doi.org/10.1109/ICCSCE50387.2020.9204956

Barro-Torres, S., Fernández-Caramés, T. M., Pérez-Iglesias, H. J., & Escudero, C. J. (2012). Real-time personal protective equipment monitoring system. Computer Communications, 36(1), 42–50. https://doi.org/https://doi.org/10.1016/j.comcom.2012.01.005

Bernardin, K., Elbs, A., & Stiefelhagen, R. (2006). Multiple object tracking performance metrics and evaluation in a smart room environment. Proceedings of the 6th IEEE International Workshop on Visual Surveillance, in Conjunction with ECCV, 90(91), 1–8. https://www.semanticscholar.org/paper/Multiple-Object-Tracking-Performance-Metrics-and-in-Bernardin-Elbs/176d4c1bd97f011567d55e5fa9f875c671dedc60

Bertram, N., Fuchs, S., Mischke, J., Palter, R., Strube, G., & Woetzel, J. (2019a). Modular construction: From projects to products. Mckinsey & Company: Capital Projects & Infrastructure. Retrieved June 30, 2022, from https://www.mckinsey.com/business-functions/operations/our-insights/modular-construction-from-projects-to-products

Bertram, N., Mischke, J., & Sjödin, E. (2019b). Modular construction: Priorities for real-estate developers. Retrieved March 30, 2023, from https://www.mckinsey.com/capabilities/operations/our-insights/modular-construction-priorities-for-real-estate-developers#/

Bosche, F., Haas, C. T., & Akinci, B. (2009). Automated recognition of 3D CAD objects in site laser scans for project 3D status visualization and performance control. Journal of Computing in Civil Engineering, 23(6), 311–318. https://doi.org/https://doi.org/10.1061/(ASCE)0887-3801(2009)23:6(311)

Brown, M., & Lowe, D. G. (2007). Automatic Panoramic Image Stitching using Invariant Features. International Journal of Computer Vision, 74(1), 59–73. https://doi.org/10.1007/s11263-006-0002-3

Brunelli, R. (2009). Template Matching Techniques in Computer Vision: Theory and Practice. https://www.wiley.com/en-sg/Template+Matching+Techniques+in+Computer+Vision:+Theory+and+Practice-p-9780470517062

Caggiano, A. (2014). Manufacturing System BT - CIRP Encyclopedia of Production Engineering. 830–836. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20617-7_6562

Canny, J. (1986). A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Machine Learning with Applications, 6, Article no. 100134. https://doi.org/10.1016/j.mlwa.2021.100134

Chaudhuri, A., Mandaviya, K., Badelia, P., & K Ghosh, S. (2017). Optical Character Recognition Systems for Different Languages with Soft Computing. Springer Cham. https://doi.org/10.1007/978-3-319-50252-6

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

Chu, W., Han, S., Asce, A. M., Luo; Xiaowei, & Zhu, Z. (2020). Monocular Vision-Based Framework for Biomechanical Analysis or Ergonomic Posture Assessment in Modular Construction. Journal of Computing in Civil Engineering, 34(4), Article no. 04020018. https://doi.org/10.1061/(ASCE)

Chu, W., Han, S. H., Luo, X., & Zhu, Z. (2019). 3D human body reconstruction for worker ergonomic posture analysis with monocular video camera. Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019, 722–729. https://doi.org/10.22260/isarc2019/0097

COCO - Common Objects in Context. (2022). Retrieved June 2, 2022, from https://cocodataset.org/#detection-eval

Cox, R. F., Issa, R. R. A., & Ahrens, D. (2003). Management's perception of key performance indicators for construction. Journal of Construction Engineering and Management, 129(2), 142–151. https://doi.org/10.1061/(ASCE)0733-9364(2003)129:2(142)

Das, A., Gawde, S., Suratwala, K., & Kalbande, D. (2018). Sign language recognition using deep learning on custom processed static gesture images. Proceedings of the 2018 International Conference on Smart City and Emerging Technology (ICSCET), 1–6. https://doi.org/10.1109/ICSCET.2018.8537248

Data Bridge Market Research. (2022). Global Computer Vision Market – Industry Trends and Forecast to 2029. Retrieved April 20, 2022, from https://www.databridgemarketresearch.com/reports/global-computer-vision-market

Daudt, H. M. L., van Mossel, C., & Scott, S. J. (2013). Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Medical Research Methodology, 13(1), 1–9. https://doi.org/https://doi.org/10.1186/1471-2288-13-48

El-Gohary, K. M., Aziz, R. F., & Abdel-Khalek, H. A. (2017). Engineering approach using ANN to improve and predict construction labour productivity under different influences. Journal of Construction Engineering and Management, 143(8), Article no. 04017045. https://doi.org/https://doi.org/10.1061/(ASCE)CO.1943-7862.0001340

Fang, W., Ding, L., Zhong, B., Love, P. E. D., & Luo, H. (2018). Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach.

Advanced Engineering Informatics, 37, 139–149. https://doi.org/https://doi.org/10.1016/j.aei.2018.05.003

Fang, W., Zhong, B., Zhao, N., Love, P. E., Luo, H., Xue, J., & Xu, S. (2019). A deep learning-based approach for mitigating falls from height with computer vision: Convolutional neural network. Advanced Engineering Informatics, 39, 170–177. https://doi.org/https://doi.org/10.1016/j.aei.2018.12.005

Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc. https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/.

Gibb, A., Pavitt, T. C., & McKay, L. J. (2004). Designing for health and safety in cladding installation – implications from pre-assembly. Proceedings of International Conference on Building Envelope Systems and Technologies (ICBEST 2004), 1–7. https://hdl.handle.net/2134/8484.

Golparvar-Fard, M., Heydarian, A., & Niebles, J. C. (2013). Vision-based action recognition of earthmoving equipment using spatio-temporal features and support vector machine classifiers. Advanced Engineering Informatics, 27(4), 652–663. https://doi.org/https://doi.org/10.1016/j.aei.2013.09.001

Gong, J., & Caldas, C. H. (2011). An object recognition, tracking, and contextual reasoning-based video interpretation method for rapid productivity analysis of construction operations. Automation in Construction, 20(8), 1211–1226. https://doi.org/https://doi.org/10.1016/j.autcon.2011.05.005

Gonsalves, R., & Teizer, J. (2009). Human motion analysis using 3D range imaging technology. Proceedings of the 26th International Symposium on Automation and Robotics in Construction, 76–85. https://doi.org/https://doi.org/10.22260/ISARC2009/0044

Hafiz, A. M., & Bhat, G. M. (2020). A survey on instance segmentation: state of the art. International Journal of Multimedia Information Retrieval, 9(3), 171–189. https://doi.org/10.1007/s13735-020-00195-x

Hamledari, H., McCabe, B., & Davari, S. (2017). Automated computer vision-based detection of components of under-construction indoor partitions. Automation in Construction, 74, 78–94. https://doi.org/https://doi.org/10.1016/j.autcon.2016.11.009

Hercog, D., Bencak, P., Vincetič, U., & Lerher, T. (2022). Product Assembly Assistance System Based on Pick-To-Light and Computer Vision Technology. Sensors, 22(24), Article no. 9769. https://doi.org/https://doi.org/10.3390/s22249769

Hinze, J., & Appelgate, L. L. (1991). Costs of construction injuries. Journal of Construction Engineering and Management, 117(3), 537–550. https://doi.org/https://doi.org/10.1061/(ASCE)0733-9364(1991)117:3(537)

Hogarth, L. (2020). The Offsite Revolution in Construction. Calgary Construction Association. Retrieved June 30, 2022, https://cgyca.com/industry-updates/the-offsite-revolution-in-construction/

Hosang, J., Benenson, R., Dollar, P., & Schiele, B. (2016). What Makes for Effective Detection Proposals? IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(4), 814–830. https://doi.org/10.1109/TPAMI.2015.2465908

Host, K., Ivasic-Kos, M., & Pobar, M. (2020). Tracking Handball Players with the DeepSORT Algorithm. Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2020), 593–599. https://doi.org/10.5220/0009177605930599

Hsia, S.-C., Wang, S.-H., Wei, C.-M., & Chang, C.-Y. (2022). Intelligent Object Tracking with an Automatic Image Zoom Algorithm for a Camera Sensing Surveillance System. Sensors (Basel, Switzerland), 22(22), 1–18. https://doi.org/10.3390/s22228791

Huang, T. S. (1996). Computer Vision: Evolution and Promise. Proceedings of the 19th CERN School of Computing, 21–25. https://doi.org/10.5170/CERN-1996-008.21

Hymel, S. (2022). Object Detection Performance Metrics. Coursera. Retrieved June 1, 2022, from https://www.coursera.org/lecture/computer-vision-with-embedded-machine-learning/object-detection-performance-metrics-zDIgp

IBM. (2022). What is Computer Vision?. Retrieved April 19, 2022, from https://www.ibm.com/topics/computer-vision

Insafutdinov, E., Pishchulin, L., Andres, B., Andriluka, M., & Schiele, B. (2016). Deepercut: A deeper, stronger, and faster multi-person pose estimation model. Proceedings of the 2016 European Conference on Computer Vision, 34–50. https://doi.org/https://doi.org/10.1007/978-3-319-46466-4_3

Ippalapally, R., Mudumba, S. H., Adkay, M., & R., N. V. H. (2020). Object Detection Using Thermal Imaging. 2020 IEEE 17th India Council International Conference (INDICON), 1–6. https://doi.org/10.1109/INDICON49873.2020.9342179

Johnsson, H., & Meiling, J. H. (2009). Defects in offsite construction: timber module prefabrication. Construction Management and Economics, 27(7), 667–681. https://doi.org/10.1080/01446190903002797

Kaiming, H., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. 2961–2969. https://doi.org/10.48550/arxiv.1703.06870

KBV research. (2020 Global Computer Vision Market by Product Type, by Component, by Application, by Vertical, by Region, Industry Analysis and Forecast 2020-2026. Retrieved April 20, 2022, from https://www.kbvresearch.com/computer-vision-market/

Kolmogorov, V., & Zabih, R. (2002). Multi-camera Scene Reconstruction via Graph Cuts. In A. Heyden, G. Sparr, M. Nielsen, & P. Johansen (Eds.), Computer Vision — ECCV 2002 (pp. 82–96). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-47977-5_6

Krzanowski, W. J., & Hand, D. J. (2009). ROC curves for continuous data (1st Edition). Chapman and Hall/CRC. https://doi.org/https://doi.org/10.1201/9781439800225

Lee, D., & Han, K. (2020). Vision-Based Quality Assessment of Prefabricated Components Using Images and Camera Poses. Proceedings of the 2020 Construction Research Congress, 1021–1029. https://doi.org/https://doi.org/10.1061/9780784482889.108

Lee, S. J., Kwon, S. W., Jeong, M. K., Hasan, S. M., & Kim, A. (2020). Automated On-Site Quality Inspection and Reporting Technology for Off-Site Construction (OSC)-based Precast Concrete Members. Proceedings of the 2020 International Symposium on Automation and Robotics in Construction, 37, 1152–1159. https://doi.org/https://doi.org/10.22260/ISARC2020/0158

Lee, S., Jeong, M., Cho, C.-S., Park, J., & Kwon, S. (2022). Deep Learning-Based PC Member Crack Detection and Quality Inspection Support Technology for the Precise Construction of OSC Projects. Applied Sciences, 12(19), Article no. 9810. https://doi.org/10.3390/app12199810

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. http://www.cihr-irsc.ca

Liu, C., M.E. Sepasgozar, S., Shirowzhan, S., & Mohammadi, G. (2021). Applications of object detection in modular construction based on a comparative evaluation of deep learning algorithms. Construction Innovation, 22(1), 141–159. https://doi.org/10.1108/CI-02-2020-0017

Liu, S., Xu, Y., Guo, L., Shao, M., Yue, G., & An, D. (2021). Multi-scale personnel deep feature detection algorithm based on Extended-YOLOv3. Journal of Intelligent and Fuzzy Systems, 40(1), 773–786. https://doi.org/10.3233/JIFS-200778

Liu, Y., Li, S., Wang, J., Zeng, H., & Lu, J. (2015). A computer vision-based assistant system for the assembly of narrow cabin products. The International Journal of Advanced Manufacturing Technology, 76(1), 281–293. https://doi.org/https://doi.org/10.1007/s00170-014-6274-9

Luo, H., Xiong, C., Fang, W., Love, P. E. D., Zhang, B., & Ouyang, X. (2018). Convolutional neural networks: Computer vision-based workforce activity assessment in construction. Automation in Construction, 94, 282–289. https://doi.org/https://doi.org/10.1016/j.autcon.2018.06.007

Maalek, R., Lichti, D. D., & Ruwanpura, J. Y. (2019). Automatic recognition of common structural elements from point clouds for automated progress monitoring and dimensional quality control in reinforced concrete construction. Remote Sensing, 11(9), Article no. 1102. https://doi.org/https://doi.org/10.3390/rs11091102

Marasini, R., & Quinnell, P. (2010). Investigation of quality management practices in building construction sites in the UK. Association of Researchers in Construction Management, Leeds, UK, 1307–1316. https://www.semanticscholar.org/paper/Investigation-of-Quality-Management-practices-in-in-Marasini-Quinnell/5614d8954243e23e333119e894faee99cc3a15a0.

Martinez, P., Ahmad, R., & Al-Hussein, M. (2019a). A vision-based system for pre-inspection of steel frame manufacturing. Automation in Construction, 97, 151–163. https://doi.org/10.1016/J.AUTCON.2018.10.021

Martinez, P., Ahmad, R., & Al-Hussein, M. (2019b). Real-time visual detection and correction of automatic screw operations in dimpled light-gauge steel framing with pre-drilled pilot holes. Procedia Manufacturing, 34, 798–803. https://doi.org/10.1016/j.promfg.2019.06.204

Martinez, P., Al-Hussein, M., & Ahmad, R. (2019c). A scientometric analysis and critical review of computer vision applications for construction. Automation in Construction, 107, Article no. 102947. https://doi.org/https://doi.org/10.1016/j.autcon.2019.102947

Martinez, P., Al-Hussein, M., & Ahmad, R. (2020). Intelligent vision-based online inspection system of screw-fastening operations in light-gauge steel frame manufacturing. The International Journal of Advanced Manufacturing Technology, 645–657. https://doi.org/10.1007/s00170-020-05695-y/Published

Martinez, P., Barkokebas, B., Hamzeh, F., Al-Hussein, M., & Ahmad, R. (2021). A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. Automation in Construction, 125, Article no. 103620. https://doi.org/10.1016/j.autcon.2021.103620

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., McArthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Medical Research Methodology, 18(1). https://doi.org/10.1186/s12874-018-0611-x

Munn, Z., Pollock, D., Khalil, H., Alexander, L., Mclnerney, P., Godfrey, C. M., Peters, M., & Tricco, A. C. (2022). What are scoping reviews? Providing a formal definition of scoping reviews as a type of evidence synthesis. JBI Evidence Synthesis, 20(4). https://journals.lww.com/jbisrir/Fulltext/2022/04000/What_are_scoping_reviews__Providing_a_formal.2.aspx.

Nam, H., & Han, B. (2016). Learning multi-domain convolutional neural networks for visual tracking. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4293–4302. https://doi.org/10.1109/CVPR.2016.465

Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014). A survey of image classification methods and techniques. 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 554–557. https://doi.org/10.1109/ICCICCT.2014.6993023

NavLab. (2015). 3D Reconstruction with Multi-View Environment (MVE). Retrieved May 26, 2022, from http://www.cs.cmu.edu/~reconstruction/MVE.html

Orsic, M., Kreso, I., Bevandic, P., & Segvic, S. (2019). In defense of pre-trained imagenet architectures for real-time semantic segmentation of road-driving images. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 12607–12616. https://doi.org/10.1109/CVPR.2019.01289

Panahi, R., Louis, J., Aziere, N., Podder, A., & Swanson, C. (2022a). Identifying Modular Construction Worker Tasks Using Computer Vision. Proceedings of the 2021 ASCE International Conference on Computing in Civil Engineering, 959–966. https://doi.org/10.1061/9780784483893.118

Panahi, R., Louis, J., Podder, A., & Swanson, C. (2022b). Tracking Volumetric Units in Modular Factories for Automated Progress Monitoring Using Computer Vision. Proceedings of the 2022 Construction Research Congress, 822–829. https://doi.org/10.1061/9780784483961.086

Paneru, S., & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. Automation in Construction, 132, Article no. 103940. https://doi.org/https://doi.org/10.1016/j.autcon.2021.103940

Park, K., Ergan, S., & Feng, C. (2021). Towards Intelligent Agents to Assist in Modular Construction: Evaluation of Datasets Generated in Virtual Environments for AI training. Proceedings of the 38th International Symposium on Automation and Robotics in Construction, 38, 327–333. https://doi.org/https://doi.org/10.22260/ISARC2021/0046

Park, M.-W., Elsafty, N., & Zhu, Z. (2015). Hardhat-wearing detection for enhancing on-site safety of construction workers. Journal of Construction Engineering and Management, 141(9), Article no. 04015024. https://doi.org/https://doi.org/10.1061/(ASCE)CO.1943-7862.0000974

Ray, S. J., & Teizer, J. (2012). Real-time construction worker posture analysis for ergonomics training. Advanced Engineering Informatics, 26(2), 439–455. https://doi.org/https://doi.org/10.1016/j.aei.2012.02.011

Raynar, K. A. (1994). Integrating computer vision with computer aided design to produce construction as-built drawings. The Pennsylvania State University. https://www.proquest.com/openview/0e9c1c2182979958ad8381bb0ed9de98/1?pq-origsite=gscholar&cbl=18750&diss=y.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), Proceedings of the 28th International Conference on Neural Information Processing Systems (Vol. 1, pp. 91–99). https://doi.org/10.48550/arxiv.1506.01497

Resurchify. (2022). Top Journals in Building and Construction. Retrieved April 12, 2022, from https://www.resurchify.com/impact/category/Building-and-Construction

Roh, S., Aziz, Z., & Pena-Mora, F. (2011). An object-based 3D walk-through model for interior construction progress monitoring. Automation in Construction, 20(1), 66–75. https://doi.org/https://doi.org/10.1016/j.autcon.2010.07.003

Shapiro, S. C. (1992). Encyclopedia of artificial intelligence. John Wiley & Sons, Inc. https://dl.acm.org/doi/book/10.5555/531550.

Siam, M., Gamal, M., Abdel-Razek, M., Yogamani, S., Jagersand, M., & Zhang, H. (2018). A comparative study of real-time semantic segmentation for autonomous driving. Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops, 587–597. https://doi.org/10.1109/CVPRW.2018.00101

Sudharsan, B., Kumar, S. P., & Dhakshinamurthy, R. (2019). AI Vision: Smart speaker design and implementation with object detection custom skill and advanced voice interaction capability. Proceedings of the 11th International Conference on Advanced Computing (ICoAC), 97–102. https://doi.org/10.1109/ICoAC48765.2019.247125

Suzuki, S. (1985). Topological structural analysis of digitized binary images by border following. Computer Vision, Graphics, and Image Processing, 30(1), 32–46. https://doi.org/https://doi.org/10.1016/0734-189X(85)90016-7

Szeliski, R. (2010). Computer vision: algorithms and applications. Springer London. https://doi.org/https://doi.org/10.1007/978-1-84882-935-0

Tan, Y. S., Lim, K. M., & Lee, C. P. (2021). Hand gesture recognition via enhanced densely connected convolutional neural network. Expert Systems with Applications, 175, Article no. 114797. https://doi.org/https://doi.org/10.1016/j.eswa.2021.114797

Teizer, J. (2015). Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites. Advanced Engineering Informatics, 29(2), 225–238. https://doi.org/https://doi.org/10.1016/j.aei.2015.03.006

Tolba, A., El-Baz, A., & El-Harby, A. (2006). Face recognition: A literature review. International Journal of Signal Processing, 2(2), 88–103. https://doi.org/https://doi.org/10.5281/zenodo.1334652

Tseng, Y.-H., & Jan, S.-S. (2018). Combination of computer vision detection and segmentation for autonomous driving. Proceedings of the 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS), 1047–1052. https://doi.org/10.1109/PLANS.2018.8373485

Verified Market Research. (2021). Global Computer Vision Market Size by Application, By Geographic Scope and Forecast, 2021. Retrieved April 20, 2022, from https://www.verifiedmarketresearch.com/product/global-computer-vision-market-size-and-forecast-to-2025/

Walker, C. (2012). Glare and Polarized Light in Machine Vision Applications. Retrieved April 3, 2023, from https://chriswalkertechblog.blogspot.com/2012/08/glare-and-polarized-light-in-machine.html

Wang, J., Qiu, K., Peng, H., Fu, J., & Zhu, J. (2019). Ai coach: Deep human pose estimation and analysis for personalized athletic training assistance. Proceedings of the 27th ACM International Conference on Multimedia, 374–382. https://doi.org/https://doi.org/10.1145/3343031.3350910

Wang, Q., Cheng, J. C. P., & Sohn, H. (2017). Automated Estimation of Reinforced Precast Concrete Rebar Positions Using Colored Laser Scan Data. Computer-Aided Civil and Infrastructure Engineering, 32(9), 787–802. https://doi.org/https://doi.org/10.1111/mice.12293

Wang, Z. C., Yang, B., & Zhang, Q. L. (2021). Automatic detection and tracking of precast walls from surveillance construction site videos. In Life-Cycle Civil Engineering: Innovation, Theory and Practice (pp. 1439–1446). CRC Press/Balkema. https://doi.org/10.1201/9780429343292-191

Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B., & Fang, T. (2021). Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. Journal of Computing in Civil Engineering, 35(1), Article no. 04020056. https://doi.org/10.1061/(asce)cp.1943-5487.0000933

Wenkel, S., Alhazmi, K., Liiv, T., Alrshoud, S., & Simon, M. (2021). Confidence score: the forgotten dimension of object detection performance evaluation. Sensors, 21(13), Article no. 4350. https://doi.org/https://doi.org/10.3390/s21134350

Wu, B.-F., Tseng, W.-J., Chen, Y.-S., Yao, S.-J., & Chang, P.-J. (2016). An intelligent self-checkout system for smart retail. Proceedings of the 2016 International Conference on System Science and Engineering (ICSSE), 1–4. https://doi.org/10.1109/ICSSE.2016.7551621

Wu, C. (2022). VisualSFM: A Visual Structure from Motion System. Retrieved May 26, 2022, from http://ccwu.me/vsfm/index.html

Xiao, B., Xiao, H., Wang, J., & Chen, Y. (2022). Vision-based method for tracking workers by integrating deep learning instance segmentation in off-site construction. Automation in Construction, 136, Article no. 104148. https://doi.org/10.1016/j.autcon.2022.104148

Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A.-M., & Wang, X. (2021). Computer vision techniques in construction: a critical review. Archives of Computational Methods in Engineering, 28(5), 3383–3397. https://doi.org/https://doi.org/10.1007/s11831-020-09504-3

Yan, X., & Zhang, H. (2021). Computer Vision–Based Disruption Management for Prefabricated Building Construction Schedule. Journal of Computing in Civil Engineering, 35(6), Article no. 04021027. https://doi.org/10.1061/(asce)cp.1943-5487.0000990

Yan, X., Zhang, H., & Zhang, W. (2022). Intelligent monitoring and evaluation for the prefabricated construction schedule. Computer-Aided Civil and Infrastructure Engineering, 38(3), 391–407. https://doi.org/10.1111/mice.12838

Yu, G., & Qi, X. (2004). Disruption management: framework, models and applications. World Scientific Publishing Co. Pte. Ltd. https://econpapers.repec.org/RePEc:wsi:wsbook:5632.

Zhang, K., Tong, S., & Shi, H. (2019). Trajectory prediction of assembly alignment of columnar precast concrete members with deep learning. Symmetry, 11(5), 1541–1552. https://doi.org/10.3390/sym11050629

Zhang, Z., Pan, W., & Zheng, Z. (2020). Fighting Covid-19 through fast delivery of a modular quarantine camp with smart construction. Proceedings of the Institution of Civil Engineers: Civil Engineering, 174(2), 89–96. https://doi.org/10.1680/jcien.20.00025

Zhao, M., Li, T., Abu Alsheikh, M., Tian, Y., Zhao, H., Torralba, A., & Katabi, D. (2018). Through-wall human pose estimation using radio signals. Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition, 7356–7365. https://doi.org/10.1109/CVPR.2018.00768

Zheng, Z., Zhang, Z., & Pan, W. (2020). Virtual prototyping-and transfer learning-enabled module detection for modular integrated construction. Automation in Construction, 120, Article no. 103387. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103387

Chapter 3 references

Ahmadian Fard Fini, A., Maghrebi, M., Forsythe, P. J. & Waller, T. S. (2021). Using existing site surveillance cameras to automatically measure the installation speed in prefabricated timber construction. Engineering, Construction and Architectural Management, 29(2), 573–600. https://doi.org/10.1108/ECAM-04-2020-0281

Alexey. (2021). darknet: YOLOv4 / Scaled-YOLOv4 / YOLO - Neural Networks for Object Detection (Windows and Linux version of Darknet). Github. Retrieved June 8, 2022, from https://github.com/AlexeyAB/darknet

Altaf, M. S. (2016). Integrated Production Planning and Control System for Prefabrication of Panelised Construction for Residential Building. [PhD Thesis, University of Alberta]. https://doi.org/10.7939/R3QZ22N9G

Bertram, N., Fuchs, S., Mischke, J., Palter, R., Strube, G. & Woetzel, J. (2019a). Modular construction: From projects to products. Mckinsey & Company: Capital Projects & Infrastructure. Retrieved June 30, 2022, from https://www.mckinsey.com/business-functions/operations/our-insights/modular-construction-from-projects-to-products

Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. ArXiv Preprint ArXiv:2004.10934. https://doi.org/https://doi.org/10.48550/arXiv.2004.10934

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

COCO Consortium. (2022). COCO - Common Objects in Context. Retrieved June 9, 2022, from https://cocodataset.org/#home

Data Bridge Market Research. (2022). Global Computer Vision Market – Industry Trends and Forecast to 2029. Retrieved April 20, 2022, from https://www.databridgemarketresearch.com/reports/global-computer-vision-market

Engineering at Alberta. (2022). Simphony.NET 4.6. Retrieved July 4, 2022, from https://www.ualberta.ca/engineering/research/groups/construction-simulation/simphony.html

Hogarth, L. (2020). The Offsite Revolution in Construction. Calgary Construction Association. Retrieved June 30, 2022, https://cgyca.com/industry-updates/the-offsite-revolution-in-construction/

Huang, T. S. (1996). Computer Vision: Evolution and Promise. In C. E. Vandoni (Ed.), 19th CERN School of Computing (pp. 21–25). https://doi.org/10.5170/CERN-1996-008

IBM. (2022). What is Computer Vision? | IBM. Retrieved April 19, 2022, from https://www.ibm.com/topics/computer-vision

Ji, S., Xu, W., Yang, M. & Yu, K. (2012). 3D convolutional neural networks for human action recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(1), 221–231. https://doi.org/10.1109/TPAMI.2012.59

Karr, A. F., Sanil, A. P. & Banks, D. L. (2006). Data quality: A statistical perspective. Statistical Methodology, 3(2), 137–173. https://doi.org/https://doi.org/10.1016/j.stamet.2005.08.005

KBV research. (2020 Global Computer Vision Market by Product Type, by Component, by Application, by Vertical, by Region, Industry Analysis and Forecast 2020-2026. Retrieved April 20, 2022, from https://www.kbvresearch.com/computer-vision-market/

Lasi, H., Fettke, P., Kemper, H. G., Feld, T. & Hoffmann, M. (2014). Industry 4.0. Business and Information Systems Engineering, 6(4), 239–242. https://doi.org/10.1007/s12599-014-0334-4

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. European Conference on Computer Vision, 740–755. https://doi.org/https://doi.org/10.48550/arXiv.1405.0312

Liu, C., Shirowzhan, S., Sepasgozar, S. M. E. & Kaboli, A. (2019). Evaluation of classical operators and fuzzy logic algorithms for edge detection of panels at exterior cladding of buildings. Buildings, 9(2). https://doi.org/10.3390/buildings9020040

Long, J., Shelhamer, E. & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3431–3440. https://doi.org/https://doi.org/10.48550/arXiv.1411.4038

Lowe, D. G. (1992). Robust model-based motion tracking through the integration of search and estimation. International Journal of Computer Vision, 8(2), 113–122. https://doi.org/https://doi.org/10.1007/BF00127170

Martinez, P., Ahmad, R. & Al-Hussein, M. (2019a). A vision-based system for pre-inspection of steel frame manufacturing. Automation in Construction, 97, 151–163. https://doi.org/10.1016/J.AUTCON.2018.10.021

Martinez, P., Al-Hussein, M. & Ahmad, R. (2019c). A scientometric analysis and critical review of computer vision applications for construction. Automation in Construction, 107(September). https://doi.org/10.1016/j.autcon.2019.102947

Martinez, P., Barkokebas, B., Hamzeh, F., Al-Hussein, M. & Ahmad, R. (2021). A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. Automation in Construction, 125(February), 103620. https://doi.org/10.1016/j.autcon.2021.103620

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association, 46, 68–78. https://doi.org/10.1080/01621459.1951.10500769

Microsoft. (2022a). What is Computer Vision? Retrieved November 8, 2022, from https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/overview

Microsoft. (2022b). What is Custom Vision? Retrieved November 8, 2022, from https://docs.microsoft.com/en-us/azure/cognitive-services/custom-vision-service/overview

Mohsen, O. (2021). A Machine Learning Approach to Predict Production Time in Industrialized Building Construction [PhD Thesis, University of Alberta]. https://doi.org/10.13140/RG.2.2.23122.68804

Papageorgiou, C. P., Oren, M. & Poggio, T. (1998). A general framework for object detection. Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), 555–562. https://doi.org/10.1109/ICCV.1998.710772

Pearson, K. (1936). Method of moments and method of maximum likelihood. Biometrika, 28(1/2), 34–59. https://doi.org/https://doi.org/10.2307/2334123

Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 779–788. https://doi.org/https://doi.org/10.48550/arXiv.1506.02640

Sun, K., Xiao, B., Liu, D. & Wang, J. (2019). Deep high-resolution representation learning for human pose estimation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5693–5703. https://doi.org/10.1109/CVPR.2019.00584

Toshev, A. & Szegedy, C. (2014). Deeppose: Human pose estimation via deep neural networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1653–1660. https://doi.org/10.1109/CVPR.2014.214

Venture Beat. (2019). Why do 87% of data science projects never make it into production? Retrieved June 28, 2022, from https://venturebeat.com/2019/07/19/why-do-87-of-data-science-projects-never-make-it-into-production/

Verified Market Research. (2021). Global Computer Vision Market Size by Application, By Geographic Scope and Forecast, 2021. Retrieved April 20, 2022, from https://www.verifiedmarketresearch.com/product/global-computer-vision-market-size-and-forecast-to-2025/

Wang, C.-Y., Bochkovskiy, A. & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. https://doi.org/10.48550/arxiv.2207.02696

Wang, H. & Schmid, C. (2013). Action recognition with improved trajectories. Proceedings of the IEEE International Conference on Computer Vision, 3551–3558. https://doi.org/10.1109/ICCV.2013.441

Wang, M., Wang, C. C., Sepasgozar, S. & Zlatanova, S. (2020). A systematic review of digital technology adoption in off-site construction: Current status and future direction towards industry 4.0. Buildings, 10(11), 204. https://doi.org/10.3390/buildings10110204

Wang, R. Y. & Strong, D. M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems, 12(4), 5–33. http://www.jstor.org/stable/40398176?origin=JSTOR-pdf

Wang, Z. C., Yang, B. & Zhang, Q. L. (2021). Automatic detection and tracking of precast walls from surveillance construction site videos. In Life-Cycle Civil Engineering: Innovation, Theory and Practice (pp. 1439–1446). CRC Press/Balkema. https://doi.org/10.1201/9780429343292-191

Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B. & Fang, T. (2021a). Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. Journal of Computing in Civil Engineering, 35(1), 04020056. https://doi.org/10.1061/(asce)cp.1943-5487.0000933

Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B. & Fang, T. (2021b). Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. Journal of Computing in Civil Engineering, 35(1), Article no. 04020056. https://doi.org/10.1061/(asce)cp.1943-5487.0000933

Wolfram MathWorld. (2022). Least Squares Fitting. Retrieved July 4, 2022, from https://mathworld.wolfram.com/LeastSquaresFitting.html

Zhang, Z., Pan, W. & Zheng, Z. (2020). Fighting Covid-19 through fast delivery of a modular quarantine camp with smart construction. Proceedings of the Institution of Civil Engineers: Civil Engineering, 174(2), 89–96. https://doi.org/10.1680/jcien.20.00025

Zheng, Z., Zhang, Z. & Pan, W. (2020). Virtual prototyping-and transfer learning-enabled module detection for modular integrated construction. Automation in Construction, 120, Article no. 103387. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103387

Chapter 4 references

Adams, W. C. (2015). Conducting semi-structured interviews. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), Handbook of practical program evaluation (pp. 409–421). Wiley Online Library. https://doi.org/10.1002/9781119171386.ch19

Alsakka, F., Darwish, M., Yu, H., Hamzeh, F., & Al-Hussein, M. (2022a). The impacts of lean implementation revealed in the course of building a digital twin of a construction manufacturing facility. Proceedings of the 30th Annual Conference of the International Group for Lean Construction (IGLC), Edmonton, Canada, 890–901. https://doi.org/10.24928/2022/0200

Alsakka, F., El-Chami, I., Yu, H., & Al-Hussein, M. (2023b). Computer vision-based process time data acquisition for offsite construction, Automation in Construction. 149 Article no. 104803. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104803.

Alsakka, F., Hamzeh, F., Yu, H., & Al-Hussein, M. (2022b). A goal-oriented framework for implementing change in off-site construction in the industry 4.0 era. In Construction Research Congress 2022, pp. 511–520. https://doi.org/doi:10.1061/9780784483961.054

Ayinla, K., Cheung, F., & Tawil, A. R. (2019). Towards an ontology-based approach to measuring productivity for offsite manufacturing method. Proceedings of the 35th Annual ARCOM Conference, 587–596. https://www.arcom.ac.uk/-docs/proceedings/6319b483bc16f3ca561b2dd5069c24af.pdf

Benjaoran, V., Dawood, N., & Scott, D. (2004). Bespoke precast productivity estimation with neural network model. In F. Khosrowshahi (Ed.), Proceedings of the 20th Annual ARCOM Conference. https://www.arcom.ac.uk/-docs/proceedings/ar2004-1063-1074_Benjoran_Dawood_and_Scott.pdf

Damelio, R. (2011). The basics of process mapping (2nd ed.). CRC Press. https://books.google.ca/books?id=7jzOBQAAQBAJ&lpg=PP1&ots=RZEilL8-4B&lr&pg=PP1#v=onepage&q&f=false

El Sakka, F., Eid, K., Narciss, T., & Hamzeh, F. (2016). Integrating lean into modular construction: A detailed case study of company X. Proceedings of the 24th Annual Conference of the International Group for Lean Construction (IGLC), Boston, Massachusetts, 23–32. https://iglcstorage.blob.core.windows.net/papers/attachment-22f55cde-cb8f-4b0a-8462-23fbe49a8c98.pdf

Giachetti, R. E. (2011). Design of enterprise systems: Theory, architecture, and methods. CRC Press. https://doi.org/10.1201/9781439882894

Hu, X., Lu, M., & Abourizk, S. (2015). BIM-based data mining approach to estimating job man-hour requirements in structural steel fabrication. Proceedings of the 2015 Winter Simulation Conference, 3399–3410. https://doi.org/10.1109/WSC.2014.7020173

Kuhn, M., & Johnson, K. (2019). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. Chapman & Hall/CRC Press. https://doi.org/https://doi.org/10.1080/00031305.2020.1790217

Liker, J. (2004). The Toyota Way - 14 management principles the world's greatest manufacturer. In McGraw-Hill (Vol. 11).

Moghadam, M., & Al-Hussein, M. (2013). Resource optimization for modular construction through value stream map improvement. Proceedings of the 21st Annual Conference of the International Group for Lean Construction (IGLC), 503–512. https://iglc.net/Papers/Details/931

Ohno, T. (1988). Toyota Production System: Beyond Large-Scale Production (1st Edition). Productivity Press. https://doi.org/10.4324/9780429273018

Shafai, L. (2012). Simulation based process flow improvement for wood framing home building production lines. MSc thesis, University of Alberta, Edmonton, AB, Canada. https://doi.org/10.7939/R3SP65

Song, L., & AbouRizk, S. M. (2008). Measuring and Modelling Labour Productivity Using Historical Data. Journal of Construction Engineering and Management, 134(10), 786–794. https://doi.org/10.1061/(asce)0733-9364(2008)134:10(786)

Zhang, Y. (2017). A framework to improve modular construction manufacturing production line performance. MSc thesis, University of Alberta, Edmonton, AB, Canada. https://doi.org/10.7939/R31J97M21

Chapter 5 references

Adeli, H. (2001). Neural Networks in Civil Engineering: 1989–2000. Computer-Aided Civil and Infrastructure Engineering, 16(2), 126–142. https://doi.org/https://doi.org/10.1111/0885-9507.00219

Ahmadian Fard Fini, A., Maghrebi, M., Forsythe, P. J., & Waller, T. S. (2021). Using existing site surveillance cameras to automatically measure the installation speed in prefabricated timber construction. Engineering, Construction and Architectural Management, 29(2), 573–600. https://doi.org/10.1108/ECAM-04-2020-0281

Ahn, S. J., Han, S. U., & Al-Hussein, M. (2020b). Improvement of transportation cost estimation for prefabricated construction using geo-fence-based large-scale GPS data feature extraction and support vector regression. Advanced Engineering Informatics, 43(March 2019), Article no.101012. https://doi.org/10.1016/j.aei.2019.101012

Alsakka, F., Assaf, S., El-Chami, I., & Al-Hussein, M. (2023a). Computer vision applications in offsite construction. Automation in Construction, 154, 104980. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104980

Alsakka, F., El-Chami, I., Yu, H., & Al-Hussein, M. (2023b). Computer vision-based process time data acquisition for offsite construction. Automation in Construction, 149, 104803. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104803

Alsakka, F., Yu, H., Hamzeh, F., & Al-Hussein, M. (2023c). Factors influencing cycle times in offsite construction. Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC), 723–734. https://doi.org/https://doi.org/10.24928/2023/0123

Altaf, M. S., Al-Hussein, M., & Yu, H. (2014). Wood-frame wall panel sequencing based on discrete-event simulation and particle swarm optimization. Proceedings of the 31st

International Symposium on Automation and Robotics in Construction and Mining, 254–261. https://doi.org/10.22260/isarc2014/0034

Benjaoran, V., & Dawood, N. (2006). Intelligence approach to production planning system for bespoke precast concrete products. Automation in Construction, 15(6), 737–745. https://doi.org/10.1016/j.autcon.2005.09.007

Benjaoran, V., Dawood, N., & Scott, D. (2004). Bespoke precast productivity estimation with neural network model. Proceedings of the 20th Annual ARCOM Conference, 2, 63–74. https://doi.org/http://www.arcom.ac.uk/-docs/proceedings/ar2004-1063-1074_Benjoran_Dawood_and_Scott.pdf

Bhatia, A. P. S., Han, S. H., Moselhi, O., Lei, Z., & Raimondi, C. (2019). Simulation-based production rate forecast for a panelised construction production line. Proceedings of the 2019 CSCE Annual Conference, Article no. CON160. https://doi.org/https://www.csce.ca/elf/apps/CONFERENCEVIEWER/conferences/2019/pdfs/PaperPDFversion_160_0419122655.pdf

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. ArXiv Preprint ArXiv:2004.10934. https://doi.org/https://doi.org/10.48550/arXiv.2004.10934

Caggiano, A. (2014). Manufacturing System BT - CIRP Encyclopedia of Production Engineering (L. Laperrière & G. Reinhart, Eds.; pp. 830–836). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20617-7_6562

Casson, R. J., & Farmer, L. D. M. (2014). Understanding and checking the assumptions of linear regression: A primer for medical researchers. Clinical and Experimental Ophthalmology, 42(6), 590–596. https://doi.org/10.1111/ceo.12358

Chao, L. C. (2001). Assessing earth-moving operation capacity by neural network-based simulation with physical factors. Computer-Aided Civil and Infrastructure Engineering, 16(4), 287–294. https://doi.org/10.1111/0885-9507.00233

Chao, L. C., & Skibniewski, M. J. (1994). Estimating construction productivity: Neural-network-based approach. 8(2), 234–251. https://doi.org/https://doi.org/10.1061/(ASCE)0887-3801(1994)8:2(234)

Chen, T. (2013). An effective fuzzy collaborative forecasting approach for predicting the job cycle time in wafer fabrication. Computers & Industrial Engineering, 66(4), 834–848. https://doi.org/https://doi.org/10.1016/j.cie.2013.09.010

COCO Consortium. (2022). COCO - Common Objects in Context. Retrieved June 9, 2022, from https://cocodataset.org/#home

Data Bridge Market Research. (2022). Global Computer Vision Market – Industry Trends and Forecast to 2029. Retrieved April 20, 2022, from https://www.databridgemarketresearch.com/reports/global-computer-vision-market

Engineering at Alberta. (2022). Simphony.NET 4.6. Retrieved July 4, 2022, from https://www.ualberta.ca/engineering/research/groups/construction-simulation/simphony.html

Gedeon, T. D. (1997). Data Mining of Inputs: Analysing Magnitude and Functional Measures. International Journal of Neural Systems, 08(02), 209–218. https://doi.org/10.1142/S0129065797000227

Gurney, K. (1997). An introduction to neural networks (1st ed.). CRC Press. https://doi.org/https://doi.org/10.1201/9781315273570

H2O.ai. (2023a). Grid (Hyperparameter) Search - H2O 3.40.0.3 documentation. Retrieved April 26, 2022, from https://docs.h2o.ai/h2o/latest-stable/h2o-docs/grid-search.html

H2O.ai. (2023b). Overview - H2O 3.40.0.3 documentation. Retrieved April 26, 2022, from https://docs.h2o.ai/h2o/latest-stable/h2o-docs/index.html

H2O.ai. (2023c). Variable Importance - H2O 3.40.0.3 documentation. Retrieved April 26, 2022, from https://docs.h2o.ai/h2o/latest-stable/h2o-docs/variable-importance.html

Hu, X., Lu, M., & Abourizk, S. (2015). BIM-based data mining approach to estimating job man-hour requirements in structural steel fabrication. Proceedings of the 2015 Winter Simulation Conference, 3399–3410. https://doi.org/10.1109/WSC.2014.7020173

Huang, T. S. (1996). Computer Vision: Evolution and Promise. Proceedings of the 19th CERN School of Computing, 21–25. https://doi.org/10.5170/CERN-1996-008.21

IBM. (2022). What is Computer Vision? | IBM. Retrieved April 19, 2022, from https://www.ibm.com/topics/computer-vision

Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: a tutorial. Computer, 29(3), 31–44. https://doi.org/10.1109/2.485891

KBV research. (2020 Global Computer Vision Market by Product Type, by Component, by Application, by Vertical, by Region, Industry Analysis and Forecast 2020-2026. Retrieved April 20, 2022, from https://www.kbvresearch.com/computer-vision-market/

Kuhn, M., & Johnson, K. (2019). Feature Engineering and Selection: A Practical Approach for Predictive Models. Chapman & Hall/CRC Press. https://doi.org/https://doi.org/10.1080/00031305.2020.1790217

Liu, H., Altaf, M. S., & Lu, M. (2015). Automated production planning in panelised construction enabled by integrating discrete-event simulation and BIM. Proceedings of the 5th International/11th Construction Specialty Conference, Article no.048. https://doi.org/https://dx.doi.org/10.14288/1.0076315

Lu, M. (2000). Productivity studies using advanced ANN models [PhD dissertation, University of Alberta]. https://doi.org/https://search.proquest.com/docview/304646918?accountid=188395

Lu, M., AbouRizk, S., & Hermann, U. (2000). Estimating Labour Productivity Using Probability Inference Neural Network. Journal of Computing in Civil Engineering, 14(4), 241–248. https://doi.org/https://doi.org/10.1061/(ASCE)0887-3801(2000)14:4(241)

Martinez, P., Barkokebas, B., Hamzeh, F., Al-Hussein, M., & Ahmad, R. (2021). A vision-based approach for automatic progress tracking of floor paneling in offsite construction facilities. Automation in Construction, 125(February), 103620. https://doi.org/10.1016/j.autcon.2021.103620

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. Journal of the American Statistical Association, 46, 68–78. https://doi.org/10.1080/01621459.1951.10500769

Mohsen, O., Mohamed, Y., & Al-Hussein, M. (2022). A machine learning approach to predict production time using real-time RFID data in industrialized building construction. Advanced Engineering Informatics, 52, Article no. 101631. https://doi.org/10.1016/J.AEI.2022.101631

Moselhi, O., Hegazy, T., & Fazio, P. (1991). Neural networks as tools in construction. Journal of Construction Engineering and Management, 117(4), 606–625. https://doi.org/https://doi-org.login.ezproxy.library.ualberta.ca/10.1061/(ASCE)0733-9364(1991)117:4(606)open_in_new

O’brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. Quality & Quantity, 41(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

RapidMiner. (2023). RapidMiner | Amplify the Impact of Your People, Expertise & Data. Retrieved April 24, 2023, from https://rapidminer.com/.

Shafai, L. (2012). Simulation based process flow improvement for wood framing home building production lines [MSc thesis, University of Alberta]. https://doi.org/doi.org/10.7939/R3SP65

Simio. (2023). Simulation, Production Planning and Scheduling Software. Retrieved March 15, 2023, from https://www.simio.com/.

Song, L., & AbouRizk, S. M. (2008). Measuring and Modelling Labour Productivity Using Historical Data. Journal of Construction Engineering and Management, 134(10), 786–794. https://doi.org/10.1061/(asce)0733-9364(2008)134:10(786)

Sun, C., Dominguez-Caballero, J., Ward, R., Ayvar-Soberanis, S., & Curtis, D. (2022). Machining cycle time prediction: Data-driven modelling of machine tool feedrate behavior with neural networks. Robotics and Computer-Integrated Manufacturing, 75, 102293. https://doi.org/10.1016/J.RCIM.2021.102293

Tai, Y. T., Pearn, W. L., & Lee, J. H. (2012). Cycle time estimation for semiconductor final testing processes with Weibull-distributed waiting time. International Journal of Production Research, 50(2), 581 – 592. https://doi.org/10.1080/00207543.2010.543938

Tam, C. M., Leung, A. W. T., & Liu, D. K. (2002). Nonlinear Models for Predicting Hoisting Times of Tower Cranes. Journal of Computing in Civil Engineering, 16(1), 76–81. https://doi.org/10.1061/(asce)0887-3801(2002)16:1(76)

Tam, C. M., Tong, T. K. L., & Tse, S. L. (2010). Modelling hook times of mobile cranes using artificial neural networks. 6193(8), 839–849. https://doi.org/10.1080/0144619042000202771

Time and Date AS. (2023). Weather for Edmonton, Alberta, Canada. Retrieved March 16, 2023, from https://www.timeanddate.com/weather/canada/edmonton

Verified Market Research. (2021). Global Computer Vision Market Size by Application, By Geographic Scope and Forecast, 2021. Retrieved April 20, 2022, from https://www.verifiedmarketresearch.com/product/global-computer-vision-market-size-and-forecast-to-2025/

Wang, Z. C., Yang, B., & Zhang, Q. L. (2021). Automatic detection and tracking of precast walls from surveillance construction site videos. In Life-Cycle Civil Engineering: Innovation, Theory and Practice. 1439–1446. CRC Press/Balkema. https://doi.org/10.1201/9780429343292-191

Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B., & Fang, T. (2021). Vision-Based Framework for Automatic Progress Monitoring of Precast Walls by Using Surveillance Videos during the Construction Phase. Journal of Computing in Civil Engineering, 35(1), Article no. 04020056. https://doi.org/10.1061/(asce)cp.1943-5487.0000933

Wolfram MathWorld. (2022). Least Squares Fitting. Retrieved July 4, 2022, from https://mathworld.wolfram.com/LeastSquaresFitting.html.

Zayed, T. M., & Halpin, D. W. (2004). Simulation as a Tool for Pile Productivity Assessment. Journal of Construction Engineering and Management, 130(3), 394–404. https://doi.org/10.1061/(asce)0733-9364(2004)130:3(394)

Zayed, T. M., & Halpin, D. W. (2005). Pile Construction Productivity Assessment. Journal of Construction Engineering and Management, 131(6), 705–714. https://doi.org/10.1061/(asce)0733-9364(2005)131:6(705)

Zheng, Z., Zhang, Z., & Pan, W. (2020). Virtual prototyping-and transfer learning-enabled module detection for modular integrated construction. Automation in Construction, 120, Article no. 103387. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103387

Chapter 6 references

Alsakka, F., Assaf, S., El-Chami, I., & Al-Hussein, M. (2023a). Computer vision applications in offsite construction. Automation in Construction, 154, 104980. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104980

Alsakka, F., El-Chami, I., Yu, H., & Al-Hussein, M. (2023b). Computer vision-based process time data acquisition for offsite construction. Automation in Construction, 149, Article no. 104803. https://doi.org/https://doi.org/10.1016/j.autcon.2023.104803

Alsakka, F., Yu, H., Hamzeh, F., & Al-Hussein, M. (2023c). Factors influencing cycle times in offsite construction. Proceedings of the 31st Annual Conference of the International Group for Lean Construction (IGLC), 723–734. https://doi.org/https://doi.org/10.24928/2023/0123

Alsakka, F., Yu, H., Hamzeh, F., & Al-Hussein, M. (2023d). Data- and knowledge-driven cycle time estimation in offsite construction. Under review in Journal of Computers and Industrial Engineering.

Altaf, M. S., Al-Hussein, M., & Yu, H. (2014). Wood-frame wall panel sequencing based on discrete-event simulation and particle swarm optimization. Proceedings of the 31st International Symposium on Automation and Robotics in Construction and Mining, 254–261. https://doi.org/10.22260/isarc2014/0034

Altaf, M. S., Bouferguene, A., Liu, H., Al-Hussein, M., & Yu, H. (2018). Integrated production planning and control system for a panelised home prefabrication facility using simulation and RFID. Automation in Construction, 85(September 2017), 369–383. https://doi.org/10.1016/j.autcon.2017.09.009

Altaf, M. S., Liu, H., Al-Hussein, M., & Yu, H. (2015). Online simulation modelling of prefabricated wall panel production using RFID system. Winter Simulation Conference Proceedings, 3379–3390. https://doi.org/10.1109/WSC.2015.7408499

Amazon. (2023). What Is Digital Twin Technology? Retrieved June 16, 2023, from https://aws.amazon.com/what-is/digital-twin/

Attaran, M., & Celik, B. G. (2023). Digital twin: Benefits, use cases, challenges, and opportunities. Decision Analytics Journal, 6, Article no. 100165. https://doi.org/https://doi.org/10.1016/j.dajour.2023.100165

Autodesk. (2023). What is digital twin technology? and what are the benefits? Retrieved July 10, 2023, from https://www.autodesk.com/solutions/digital-twin

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. ArXiv Preprint ArXiv:2004.10934. https://doi.org/https://doi.org/10.48550/arXiv.2004.10934

Caggiano, A. (2014). Manufacturing system BT - CIRP encyclopedia of production engineering (L. Laperrière & G. Reinhart, Eds.; pp. 830–836). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-20617-7_6562

Chryssolouris, G. (2006). *Manufacturing systems: Theory and practice* (2nd ed.). Springer. https://doi.org/10.1007/0-387-28431-1

COCO Consortium. (2022). COCO - Common Objects in Context. Retrieved June 9, 2022, from https://cocodataset.org/#home

Cooper, J. P., Jackson, S., Kamojjala, S., Owens, G., Szana, K., & Tomić, S. (2022). Demystifying Digital Twins: Definitions, Applications, and Benefits. Journal AWWA, 114(5), 58–65. https://doi.org/https://doi.org/10.1002/awwa.1922

Du, J., Dong, P., Sugumaran, V., & Castro-Lacouture, D. (2021). Dynamic decision support framework for production scheduling using a combined genetic algorithm and multiagent model. Expert Systems, 38(1). https://doi.org/10.1111/exsy.12533

Grieves, M. (2014). Digital twin: Manufacturing excellence through virtual factory replication. In White Paper (Issue March). https://docplayer.net/37776975-Digital-twin-manufacturing-excellence-through-virtual-factory-replication.html

Gulewicz, M. (2022). Digital twin technology — awareness, implementation problems and benefits. Engineering Management in Production and Services, 14(1), 63–77. https://doi.org/doi:10.2478/emj-2022-0006

H2O.ai. (2023b). Overview - H2O 3.40.0.3 documentation. Retrieved April 26, 2022, from https://docs.h2o.ai/h2o/latest-stable/h2o-docs/index.html

IBM. (2023). What is a digital twin? | IBM. Retrieved June 6, 2023, from https://www.ibm.com/topics/what-is-a-digital-twin

Jiang, Y., Li, M., Li, M., Liu, X., Zhong, R. Y., Pan, W., & Huang, G. Q. (2022). Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly. Automation in Construction, 141(May), 104397. https://doi.org/10.1016/j.autcon.2022.104397

Khandelwal, P. (2021). Offsite Construction Market by Material (Steel, Wood, Concrete, and Others) and Application (Residential, Commercial, and Industrial): Global Opportunity Analysis and Industry Forecast, 2021–2030. Retrieved August 15, 2022, from https://www.alliedmarketresearch.com/offsite-construction-market-A12504

Ko, C. H., & Wang, S. F. (2011). Precast production scheduling using multi-objective genetic algorithms. Expert Systems with Applications, 38(7), 8293–8302. https://doi.org/10.1016/j.eswa.2011.01.013

Lawrence, S. R., & Sewell, E. C. (1997). Heuristic, optimal, static, and dynamic schedules when processing times are uncertain. Journal of Operations Management, 15(1), 71–82. https://doi.org/https://doi.org/10.1016/S0272-6963(96)00090-3

Lee, D., & Lee, S. (2021). Digital twin for supply chain coordination in modular construction. In Applied Sciences (Switzerland) (Vol. 11, Issue 13). https://doi.org/10.3390/app11135909

Leu, S. Sen, & Hwang, S. T. (2002). GA-based resource-constrained flow-shop scheduling model for mixed precast production. Automation in Construction, 11(4), 439–452. https://doi.org/10.1016/S0926-5805(01)00083-8

Liu, Z.-S., Meng, X.-T., Xing, Z.-Z., Cao, C.-F., Jiao, Y.-Y., & Li, A.-X. (2022). Digital Twin-Based Intelligent Safety Risks Prediction of Prefabricated Construction Hoisting. Sustainability, 14(9). https://doi.org/10.3390/su14095179

Ouelhadj, D., & Petrovic, S. (2009). A survey of dynamic scheduling in manufacturing systems. *Journal of Scheduling*, *12*(4), 417–431. https://doi.org/10.1007/s10951-008-0090-8

Parnaby, J. (1979). Concept of a manufacturing system. International Journal of Production Research, 17(2), 123–135. https://doi.org/https://doi.org/10.1080/00207547908919600

Pavel Orlov. (2023, April 19). What businesses should know about digital twins. Forbes. https://www.forbes.com/sites/forbestechcouncil/2023/04/19/what-businesses-should-know-about-digital-twins/?sh=4cde77771e4f

Rausch, C., Lu, R., Talebi, S., & Haas, C. (2021). Deploying 3D scanning based geometric digital twins during fabrication and assembly in offsite manufacturing. In International Journal of Construction Management. https://doi.org/10.1080/15623599.2021.1896942

Shafai, L. (2012). Simulation based process flow improvement for wood framing home building production lines [MSc thesis, University of Alberta]. https://doi.org/doi.org/10.7939/R3SP65

Simio. (2023). Simulation, production planning and scheduling software. Retrieved March 14, 2023, from https://www.simio.com/.

Tran, H., Nguyen, T. N., Christopher, P., Bui, D. K., Khoshelham, K., & Ngo, T. D. (2021). A digital twin approach for geometric quality assessment of as-built prefabricated façades. Journal of Building Engineering, 41(December 2020), 102377. https://doi.org/10.1016/j.jobe.2021.102377

Tronel, R. (2023). Deloitte, Siemens and the model-based enterprise: The future of the digital twin - thought leadership. Retrieved June 16, 2023 from https://blogs.sw.siemens.com/thought-leadership/2023/04/05/deloitte-siemens-and-the-model-based-enterprise-the-future-of-the-digital-twin/?utm_campaign=2023-5-global-deloitte_mbe_social_ads&utm_source=google&utm_medium=paid_search&utm_content=dmirsa_twin_

Wang, C.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. https://doi.org/10.48550/arxiv.2207.02696

Xu, Z., Wang, X., & Rao, Z. (2020). Automated optimization for the production scheduling of prefabricated elements based on the genetic algorithm and IFC object segmentation. Processes, 8(12), 1–25. https://doi.org/10.3390/pr8121593

Zhao, Y., Cao, C., & Liu, Z. (2022). A Framework for prefabricated component hoisting management systems based on digital twin technology. Buildings, 12(3), Article no. 276. https://doi.org/10.3390/buildings12030276