Automated productivity evaluation of concreting works: The example of concrete pillar production

Fabian Pfitzner[1,2], Jonas Schlenger[1,2], and André Borrmann[1,2]

[1] Chair of Computational Modeling and Simulation, Technical University of Munich, Germany
[2] Leonhard Obermeyer Center, TUM Center of Digital Methods for the Built Environment

Abstract. Site schedules are usually developed by the rule of thumb based on the experience of on-site managers. While this approach can be suitable for smaller job sites, it is challenging to make good decisions for larger projects. Planning errors can result in massive delays and increasing costs. Significant improvements in other industries showed that data-driven productivity analysis of past processes advances the planning and execution of current and future projects. However, in the Architecture, Engineering & Construction (AEC) domain, automated productivity analysis of the construction phase has barely been investigated.
To overcome this deficiency, this paper presents a first approach for multi-level productivity analysis of shell constructions. We discuss several state-of-the-art vision-based technologies that serve as a foundation for large-scale evaluation of the progress on a construction site. A complete pipeline is introduced that uses different types of neural networks to extract productivity information from images at various levels of detail. The proposed workflow is demonstrated for the construction process of cast-in-place concrete pillars, implementing the first two layers. Finally, remaining challenges are discussed.

Keywords: construction monitoring, data mining, productivity

1 Introduction

For planning of construction projects, it is of utmost importance to make good estimations of how long individual construction processes will take. This highly influences the project end date and also the project cost. So-called expense values are reference values that describe the working hours required to complete a certain task, e.g. build one meter of concrete wall. They can be specific to the construction method used, differentiate between various types of building elements or other criteria and are based on the requirements of the ongoing planning phase [5]. In current practice, the expense values are often estimated by the project planner or construction manager using experience, tacit knowledge and gut feeling. Some companies perform time measurements of specific processes to obtain reference values. Yet another option is to rely on expense values provided by the literature. These are often calculated based on responses from expert interviews. All these options can be highly influenced by the subjective perception of the construction planner, interviewed expert, or the construction worker who notes down the time he/she spent on a construction task. Furthermore, the expense values are influenced by many different factors, like the construction method used, the specific circumstances of the construction site and project, as well as the company's internal processes [6, 21].

Within the last few years, a lot of progress has been made in the field of automated image analysis. Especially, AI-based approaches have found application in many fields, including the construction industry. Since acquiring regular images from a construction site is very affordable compared to other types of sensor data, they have become popular input data for many different types of automated analysis. AI-based approaches to process images can also provide progress-related information that is essential for automated productivity evaluation. They are characterized by fast execution times, which makes them suitable for application on large data sets [12].

Even though productivity on construction sites is a crucial topic, previous productivity analysis in construction is either based on surveys [6] or conducted in small-scale, well-controlled surroundings over a short period of time [10, 15]. The lack of comprehensive productivity analysis is related to the limited availability of large data sets and computational resources. Both result in uncertainties of the derived statements about productivity in construction and make their validity for a wide range of construction projects questionable. Therefore, a data-driven approach is required to objectively judge productivity. The aim of our work is to enhance the understanding of productivity in the context of the entire construction environment. For this, insights are given into how productivity is calculated according to the construction practice. Furthermore, state-of-the-art vision-based approaches are presented that can serve as a basis for automated productivity analysis. As the main contribution of this paper, we propose a methodology that allows to determine productivity-related values in varying levels of granularity. It is demonstrated with a prototypical implementation on the example of erecting cast-in-place concrete pillars, identifying the individual operational states *not started, rebar, formwork,* and *finished*. The pipeline is presented together with remaining challenges, especially in regard to fine-grained productivity evaluation.

2 Background

2.1 Construction Monitoring

A large amount of monitoring solutions are being developed to support construction environments. For this purpose, various sensor technologies such as laser scanners, cameras, and Bluetooth Low Energy (BLE) tracking systems have been considered [8,16]. A conventional approach to monitoring construction progress is through image-based methods, which rely on capturing images and processing them to obtain information about the construction site. The information extraction is increasingly performed using machine learning methods [20]. This information is essential to represent the as-performed construction state by a digital twin [14]. Studies that provide insights into the construction phase based on real-world construction sites are scarce. Further research is necessary to develop a comprehensive understanding of the benefits and limitations of using diverse technologies like UAVs and crane cameras for monitoring construction sites and creating an as-performed digital twin [9].

2.2 Vision-based AI methods in Construction

In recent years, there have been significant improvements in computer vision technology, particularly in image classification [2], object detection [7], and semantic segmentation [17]. These advancements have unveiled potential use cases for computer vision technology in the construction industry. Some potential applications of image-based methods in construction include identifying safety hazards, monitoring construction progress, and conducting quality control checks [20]. With the development of more sophisticated computer vision algorithms, these applications will become more accurate and reliable, leading to safer and more efficient construction environments.

In the context of progress monitoring and productivity analysis, existing vision-based approaches can be classified into two different groups. On the one hand side, some researchers focus on the construction workers or the equipment as main objects of interest. As an example, [10] detect construction workers on images to estimate their productivity. Based on their pose, they distinguish between effective, ineffective, and contributory work. Similarly, [15] track construction workers to identify their actions. Using YOWO (You Only Watch Once), a neural network that simultaneously executes object detection and classification, they detect workers in images and categorize their actions into standing, walking, transporting, drilling, and hammering. Focusing on the concrete bucket as the critical resource for the concreting process, [3] assesses the productivity of concreting works. Depending on the bucket's location, its current status is identified, while its change over time indicates different types of production scenarios. On the other hand, some researchers focus on the building elements that are the primary output of the construction processes to judge progress and productivity. In their literature review, [12] showed that many approaches are based on 3D reconstruction from a set of images, while others directly analyze the 2D images. As an example, [18] monitor the installation process of precast wall elements. They use a combination of several neural networks for objection detection, instance segmentation, and object tracking. With this methodology, they are able to identify when wall elements are moved by the tower crane or installed in their final location.

3 Productivity in Construction

To evaluate how well processes are executed, they need to be observed and measured. Only then it is possible to timely detect deviations from the plan and induce change. Many industries determine productivity values of processes to quantify their effectiveness. In the construction industry, different types of productivity-related values are significant as reference values for construction planning. They help to make more accurate estimations of a project's required time and cost. On a generic level, productivity is defined by the ratio of input to the output of an activity, as shown in eq. (1) [6]. The input can include labor, required construction materials, used equipment, and more. The output often refers to the building elements that are built as the result of the construction process.

$$Productivity = \frac{Output}{Input} \tag{1}$$

It is the objective of construction managers to constantly enhance productivity by increasing production quantities, reducing costs, and improving profit margins. Specifically, the income-to-expense ratio is of high interest to construction managers when it comes to calculating the project costs to be expected, since it determines the construction company's efficiency and profit [6]. Even though productivity values can reflect long-time average values, they can not be assumed as constant during a single construction project. Every project undergoes various changes of productivity over time. During the early phase, the construction workers need to set up and get familiar with the processes of this particular construction site. During this initial adoption phase, productivity will be lower than during the main phase. Towards the end of the project, there is another phase of lowered productivity because of a lower amount of workers on site and the characteristics of the finishing works. Therefore, productivity can be considered as constantly changing [6]. While productivity describes the actual number or volume of building elements per labor unit, the expense value defines the amount of labor required for a certain element. Expense values (Exp) have a significant impact on the overall productivity of the project and are therefore included in process productivity estimation [5].

$$Productivity_x = \frac{1}{Exp_x} \tag{2}$$

To estimate expense values, the total amount of working hours (H_t) is divided by the amount of produced components. The number of working hours can be further dissected by multiplying the number of workers (A_w), the working hours per day (H_d), and the duration of the process in days (d) [1,4,6].

$$Exp_x = \frac{H_t}{V_t} = \frac{A_w \cdot d \cdot H_d}{V_t} \tag{3}$$

Productivity and expense values can be calculated at various granularity levels. On the top-most level, they are determined on the basis of all types of processes on the construction site. Going more into detail, they are differentiated based on the type of construction work. As an example, productivity could be calculated separately for excavations, masonry, and reinforced concrete elements. These can be further dissected into productivity values for the individual operational steps. Reinforced concrete would, e.g., require work related to reinforcement, formwork, and concrete pouring. One level further down, one can group different types of building elements, like walls, pillars, and slabs. Finally, the productivity-related values can also be calculated for individual elements or element groups [5].

To assess productivity, it is crucial to compare the target expense values from the construction plan to the actual expense values achieved during a project. This supports determining if the project goals have been fulfilled. Such expense values support construction managers in identifying problems during the project and ensure that project goals will be met. The expense values can be set in comparison with the target expense values of a project to compute the productivity loss of several processes, as formalized in eq. (4).

$$\Delta Productivity_{loss} = \frac{\frac{1}{Exp_{target}} - \frac{1}{Exp_{actual}}}{\frac{1}{Exp_{target}}} \tag{4}$$

4 Proposed solution

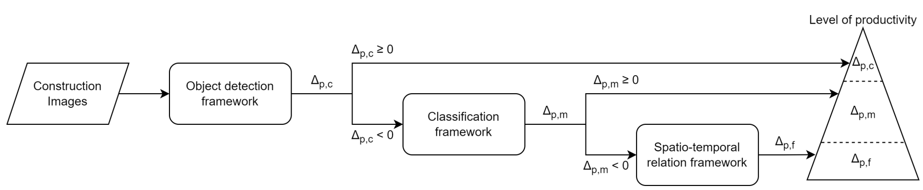


Fig. 1: Automated productivity evaluation pipeline

To compute the productivity on construction sites we use photographic images acquired from on-site environments as input source. In this research contribution, we focus on cast-in-place concrete pillars which require multiple individual steps to be built. The images get processed by diverse AI-based frameworks. Depending on the estimated productivity, a more in-depth investigation of a construction process is necessary. Therefore, our proposed pipeline, shown in fig. 1, is categorizing the productivity deltas into three levels: coarse $\Delta_{p,c}$, medium $\Delta_{p,m}$, and fine $\Delta_{p,f}$. One of the reasons for not computing the fine-grained productivity for all building elements on the total amount of images is to save computational costs and avoid hardware failures.

We apply multiple methods to extract the duration of concreting pillars. However, the equation to compute productivity Δ_p stays the same for all using eq. (4). The sum of working hours is computed by the number of daily working hours and the size of worker groups allocated to one building component. The volume of the pillars is derived from the BIM model. Having acquired all the site-related information stated above, the expense values and productivity are determined. As the starting point, the duration of the entire pillar construction process is derived by detecting two states of the classified pillars: *start* and *finish* (see fig. 2). When productivity deltas differ significantly from the expected outputs, more fine-grained methods are applied to dismantle the pillar concreting process in its individual parts: *preparation, rebar, formwork,* and *finished*.

In case further details are required, spatio-temporal activities (e.g., *standing, walking, placing*) are identified to quantify the time that was spent working on a particular pillar. The following subsections go more into detail about the productivity analysis on the three different granularity levels. It needs to be mentioned that the selected neural networks represent only one possibility to tackle the given task and might be replaced, e.g. when focusing on other building elements.

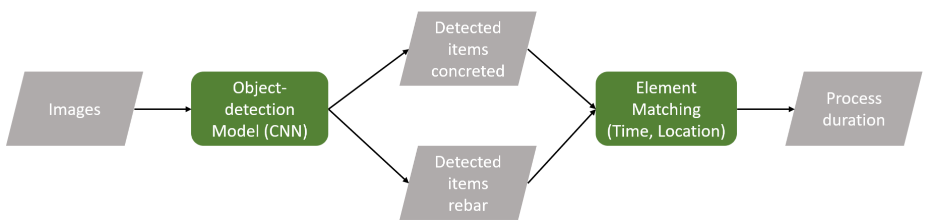
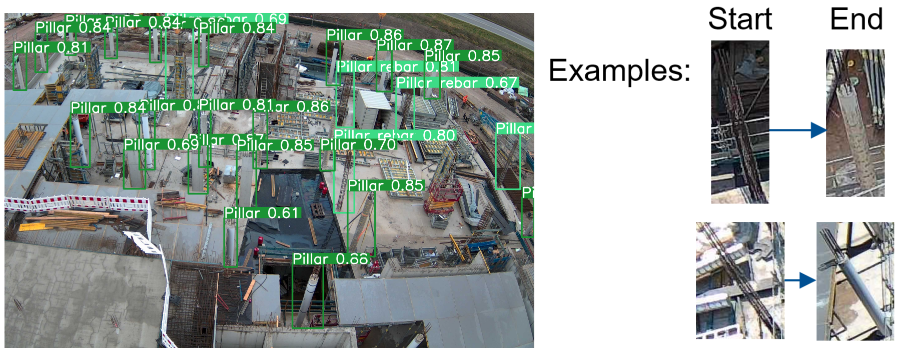


Fig. 2: Object detection pipeline

A framework to detect diverse objects on construction sites that forms the starting point of the pipeline was developed by [11]. The approach passes the captured images to a Convolutional Neural Network (CNN) using a one-stage anchor-based detector to identify the type and location of the pillars. Since the location of the detected pillars remains constant, objects can be monitored over multiple images within a construction project to estimate the diverse states of the building components. For example, the beginning of a pillar (rebar) and finish state (concreted column) is captured, as demonstrated in fig. 3. With the time and location-dependent information, the process duration is estimated allowing productivity computation.

The pillars with exceptionally long construction times and low productivity, highlighted in fig. 4 are usually of particular interest to better understand issues in the construction process. However, more detailed information is required to identify possible reasons for low productivity. The same applies if pillars are particularly relevant for overall the construction project. In that case, all pillars are examined in further detail.



(a) Pillars detected from crane camera (b) States of pillars

Fig. 3: Start- and finish time computed with object detection

The identified starting and end times of the pillar construction process are used to narrow down the time interval in which images are analyzed in closer detail. Based on the bounding boxes from the object detection, image sections between start and finish of a building component are cut out from the frequently captured images. These are classified into four different classes according to their current status to identify the operational steps. To do so, the approach developed by [13] is applied, which is described in further detail in the following paragraph.

As depicted in fig. 5, the input for the image classification are the image sections originating from the object detection. These are passed to a CNN that classifies them according to their status. It differentiates between the status *not started, rebar, formwork*, and *finished*. The process of pouring concrete is not considered in this approach since it cannot be detected by observing the pillar itself but would require to shift the attention to the detection of the construction equipment like the concrete bucket.

Since the accuracy of the CNN is limited and further decreased by clutter on the construction site and moving objects, some images are wrongly classified. However, having several images of the same pillar is used as an advantage to correct some of the erroneous image classifications. Based on the expected sequence of statuses, the transitions points from one status to the next are identified and then used to correct the status predictions of the CNN. As the final result, the start and end times of every construction phase of the pillar are provided [13]. Even though this gives more insights into the construction process, it still does not allow to directly measure the time that construction workers actively spent working on individual building elements. It also includes the time which a particular pillar remained in a certain status without any working being done.

Similar to the step described beforehand, the operational steps with exceptionally low productivity rates are of particular interest to identify reasons for delays. The level of productivity analysis requires measuring the actual time

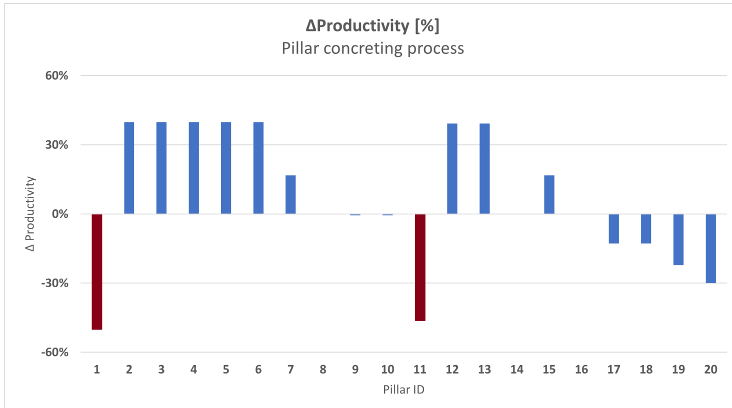


Fig. 4: Productivity of the concreting process of sample pillars: Pillars with significantly low productivity $\Delta_{p,c}$ compared to other values are highlighted in red

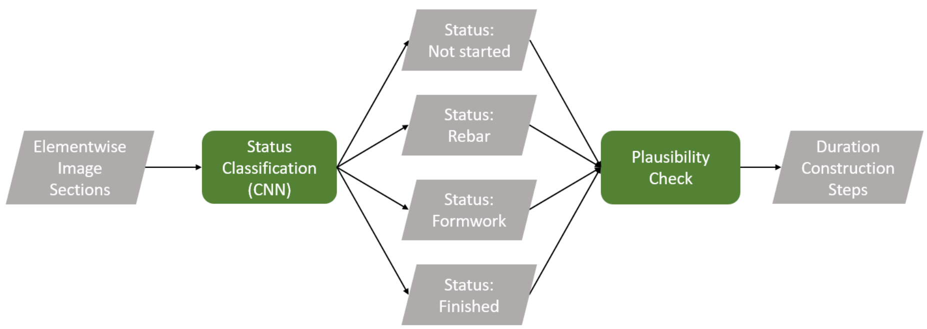


Fig. 5: Status classification pipeline for cast-in-place concrete pillars

that construction workers spent working on a specific pillar. This can help, e.g., to distinguish situations where workers were waiting for material deliveries from situations where the delay was caused by faulty execution and therefore required rework. The results from the classification step help to narrow down the time intervals of special importance. The points in time when status changes occur roughly indicated the time when construction personnel has worked on the building element. Observing the construction activities in a predefined time interval before and after the status change will suffice to analyze the construction process. To identify the time that construction workers were working on a specific pillar, an approach is required that takes into account the spatial location of the column but also the movement of the workers. A single image is often not sufficient to identify what a worker is currently doing.

Spatio-temporal action localization networks [19] are a natural fit for this task. They depend upon extracting information from images frame by frame and utilizing the relationships between them to create additional features within the network. The model is trained using the extracted features to create a comprehensive understanding of the activity being performed over time. This allows for accurate in-depth activity classification, and the identification of the proportion of time when workers are performing specific actions on-site. One of the shortcomings is that action classification networks require multiple frames per second since movements need to be tracked precisely. Based on current data limitations, the spatio-temporal action localization part of the framework has not been implemented yet and remains conceptual at this stage.

5 Conclusion

Traditional literature sources that rely on questionnaires may not provide a complete picture of productivity in construction sites, since their primary data source was based on surveys of experts [6]. The proposed objective measurement approach provides more accurate and reliable data.

In this paper, we presented a comprehensive method to measuring productivity for pillar production in construction sites. We gave a detailed introduction to past productivity measurements in on-site environments. In addition, we devised a methodology to compute on-site productivity on diverse levels of granularity based on real-world image data. Finally, we provided a framework using state-of-the-art machine learning methods for a new way to derive productivity on construction sites and support construction management.

While measuring the time taken to complete a specific building element such as a pillar is a useful metric, it is necessary to further extend the approach to detect the factor of time a worker spent efficiently creating a building component. In addition, it is essential to monitor the entire construction site, including all types of building elements, to provide accurate productivity values. By tracking workers' time spent on various building elements, we can obtain valuable insights into their performance. The scope of this objective measurement approach is to identify potential inefficiencies in the construction process, such as delays or bottlenecks, that may be hampering productivity. With this information, construction managers can improve decision-making to optimize their operations and enhance productivity. In conclusion, adopting a comprehensive approach to computing productivity that involves monitoring the entire construction site and utilizing objective measurements is essential. This approach provides valuable data that can help identify inefficiencies, optimize operations, and improve productivity in the construction phase.

Acknowledgements

We thankfully acknowledge Innovation Management Bau for their financial support and for providing us access to multiple construction sites to collect valuable data. In addition, we would like to thank Siemens Real Estate and Max Bögl for providing us access to the construction site in Garching. Furthermore, the research presented in this paper has received funding from the European Union's

Horizon 2020 research and innovation programme under grant agreement no. 958398, "BIM2TWIN: Optimal Construction Management & Production Control". The authors are very thankful for the support of the European Commission in funding this project.

References

1. H. Bauer. *Baubetrieb 2: Bauablauf, Kosten, Störungen*. Springer-Verlag, 2013.
2. X. Chen, C. Liang, D. Huang, E. Real, K. Wang, Y. Liu, H. Pham, X. Dong, T. Luong, C.-J. Hsieh, Y. Lu, and Q. V. Le. Symbolic discovery of optimization algorithms, 2023.
3. J. Gong and C. H. Caldas. Computer Vision-Based Video Interpretation Model for Automated Productivity Analysis of Construction Operations. *Journal of Computing in Civil Engineering*, 24(3):252–263, 2010.
4. P. N. Greitemann. *Bestimmung der Bauzeit von Bauprojekten zum Zeitpunkt der Realisierungsentscheidung*. 2018.
5. C. Hofstadler. *Bauablaufplanung und Logistik im Baubetrieb*. Springer, 2007.
6. C. Hofstadler. *Produktivität im Baubetrieb: Bauablaufstörungen und Produktivitätsverluste*. Springer-Verlag, 2013.
7. C. Li, L. Li, Y. Geng, H. Jiang, M. Cheng, B. Zhang, Z. Ke, X. Xu, and X. Chu. Yolov6 v3.0: A full-scale reloading, 2023.
8. H. Li, G. Chan, J. K. W. Wong, and M. Skitmore. Real-time locating systems applications in construction. *Automation in Construction*, 63:37–47, 2016.
9. D. G. J. Opoku, S. Perera, R. Osei-Kyei, and M. Rashidi. Digital twin application in the construction industry: A literature review. *Journal of Building Engineering*, 40, 8 2021.
10. A. Peddi, L. Huan, Y. Bai, and S. Kim. Development of human pose analyzing algorithms for the determination of construction productivity in real-time. *Construction Research Congress*, pages 886–895, 2009.
11. F. Pfitzner, A. Braun, and A. Borrmann. Object-Detection Based Knowledge Graph Creation: Enabling Insight into Construction Processes . In *ASCE International Conference on Computing in Civil Engineering 2023*, volume in press, Oregon, US, 2023.
12. M. Sami Ur Rehman, M. T. Shafiq, and F. Ullah. Automated Computer Vision-Based Construction Progress Monitoring: A Systematic Review. *Buildings*, 12(7):1–32, 2022.
13. J. Schlenger, S. Vilgertshofer, and A. Borrmann. Process-oriented progress monitoring of cast-in-place shell constructions based on computer vision. In *30th EG-ICE International Workshop on Intelligent Computing in Engineering 2023*, London, UK, 2023.
14. J. Schlenger, T. Yeung, S. Vilgertshofer, J. Martinez, R. Sacks, and A. Borrmann. A comprehensive data schema for digital twin construction. In *Proceedings of the 29th EG-ICE International Workshop on Intelligent Computing in Engineering*, Jun 2022.
15. G. Torabi, A. Hammad, and N. Bouguila. Two-Dimensional and Three-Dimensional CNN-Based Simultaneous Detection and Activity Classification of Construction Workers. *Journal of Computing in Civil Engineering*, 36(4), 2022.
16. P. Vähä, T. Heikkilä, P. Kilpeläinen, M. Järviluoma, and E. Gambao. Extending automation of building construction - survey on potential sensor technologies and robotic applications. *Automation in Construction*, 36:168–178, 2013.

17. W. Wang, J. Dai, Z. Chen, Z. Huang, Z. Li, X. Zhu, X. Hu, T. Lu, L. Lu, H. Li, et al. Internimage: Exploring large-scale vision foundation models with deformable convolutions. *arXiv preprint arXiv:2211.05778*, 2022.
18. Z. Wang, Q. Zhang, B. Yang, T. Wu, K. Lei, B. Zhang, and T. Fang. 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), 2021.
19. P. Weinzaepfel, Z. Harchaoui, and C. Schmid. Learning to track for spatio-temporal action localization, 2015.
20. S. Xu, J. Wang, W. Shou, T. Ngo, A.-M. Sadick, and X. Wang. Computer vision techniques in construction: A critical review. *Archives of Computational Methods in Engineering*, 28:3383–3397, 2021.
21. Z. You and C. Wu. A framework for data-driven informatization of the construction company. *Advanced Engineering Informatics*, 39(December 2018):269–277, 2019.