

Object Detection Based Knowledge Graph Creation: Enabling Insight into Construction Processes

Fabian Pfitzner;¹ Alexander Braun, Dr.-Ing.;²
and André Borrmann, Prof. Dr.-Ing.³

¹ Chair of Computational Modeling and Simulation, Technical University of Munich, Germany (corresponding author). ORCID: <https://orcid.org/0009-0006-9105-9455>.

Email: fabian.pfitzner@tum.de

² Chair of Computational Modeling and Simulation, Technical University of Munich, Germany.

ORCID: <https://orcid.org/0000-0003-1513-5111>. Email: alex.braun@tum.de

³ Full Professor, Chair of Computational Modeling and Simulation, Technical University of Munich, Germany. ORCID: <https://orcid.org/0000-0003-2088-7254>.

Email: andre.borrmann@tum.de

ABSTRACT

Compared to other industries, the construction sector shows low productivity worldwide. However, holistic, data-oriented methods for investigating potential bottlenecks within the as-performed construction stage are scarce. Our research presents an approach to acquiring raw data from job sites and its subsequent processing to high-level information. First, images were captured over a period of one year in high frequency using multiple crane cameras. Second, an end-to-end deep learning based approach was developed to derive and link information about construction activities, covering the classification and localization of specific on-site objects. This information was subsequently integrated into a knowledge graph. Finally, additional data sources like the weather were exploited to interpret different on-site scenarios. We demonstrate that construction-related activities like working times can be detected. The presented approach provides a significant step towards exposing correlations on construction sites by using multiple data processing steps and showcases the possibility of identifying process patterns.

INTRODUCTION

In contrast to significant process optimization in other industries, the construction domain is still behind, reflected in extremely low productivity (Farmer, 2016). This results in a waste of time, money, and resources. At the same time, reusing existing approaches from the production industry is not straightforward since the construction environment varies significantly from well-controlled surroundings. Yet, the potential for optimization using state-of-the-art digital twin (DT) technologies is vast (Opoku et al., 2021). Barbosa et al. (2017) imply that on-site execution must be improved and engineering processes be rethought to increase productivity. The practiced monitoring is based on manually written daily reports on many construction sites. These reports lack information and are outdated and unsuitable for covering complex on-site environments. Precise, well-documented real-world data is rare, slowing down ongoing developments (Guo et al., 2021). Solid documentation of construction activities is required to enhance the construction sector in the future. Our research provides a digital twin for construction sites that focuses on

process aspects, aiming to uncover reasons behind schedule deviations and potential sources for improvements. A data pipeline was developed by generating data from multiple building construction sites. Further, a graph meta-model was created to store and structure big data retrieved from job sites without losing the context of construction environments. We contribute to a real-world digital twin construction facing the demand of construction management and stakeholders.

BACKGROUND

The concept of *digital twin construction* developed by Sacks et al. (2020) provides an approach for creating a digitized representation of the on-site environment. Similar to digital twins in other industries, the ultimate objective of the framework is process improvement. Essential parameters for monitoring the subsequent on-site processes are time, location, resources, preconditions, and building elements.

Process Monitoring. Most construction monitoring methods have limitations (Li et al., 2016). Vision-based approaches, for example, can usually reach precision rates up to the meter scale only. Nevertheless, image-based monitoring methods retrieving construction site activity have shown a recent reputation (Guo et al., 2021). Torabi et al. (2022), for example, developed an approach that applies a unified Convolutional Neural Network (CNN) architecture for real-time spatiotemporal localization in a video stream to detect the different poses of construction workers and thereby evaluate their performance. At the same time, many open-source on-site datasets, as shown by Xuehui et al. (2021), have found popularity within the on-site monitoring domain. Braun & Borrmann (2019) conducted an approach to automatically label images to enhance the creation of training datasets. These datasets are highly demanded to train a network that enables processing diverse images or point clouds and retrieving information.

Object Detection. One of the major developments within the computer vision domain is object detection. Understanding image scenes based on neural networks has changed image processing fundamentally (Zou et al., 2019). Object detection describes detecting an object's location and classifying it in a 2D image. A bounding box determines the size and position of an object. The location and classification of the detected items are stored in annotation files. One of the object detection development milestones was CNN-based One-stage Detectors, e.g., You Only Look Once (YOLO), allowing extremely fast prediction speed (Zou et al., 2019). This was achieved by dividing an image into several regions and predicting the type and location of the objects at once using a single network. Such networks support processing big image datasets within a short amount of time.

Graph databases. In our research contribution, we are acquiring large amounts of data that must be structured. Graph technologies provide powerful tools for linked data, like modeling dynamic environments, finding predictive features, and uncovering patterns (Hodler et al., 2019). In the architecture, engineering, and construction (AEC) domain, linked data applications have significantly increased (Pauwels et al., 2017). A graph is a representation of connected data described with nodes and relationships. Ontologies define how and where data is stored in the graph by providing a schema. As-performed data can be stored similarly using a tailored graph meta-model, as Fang et al. (2020) demonstrated. They classify the entities into four categories: Human resources, equipment, materials, and environment. A data model supports structuring the

low-level information acquired from the construction site and simplifies obtaining higher-level information at a later stage. This can even be driven to the finest level of detail, as shown by Zheng et al. (2021), providing the framework *digital construction ontologies* for capturing individual construction processes.

Research Problem. A recent challenge is that many computer vision methods applied in construction target to balance deficiencies of algorithms by using prior knowledge, e.g., as-planned data, to support on-site understanding (Xu et al., 2021). However, this is not accomplished when neglecting the construction surroundings by obtaining data from controlled environments aimed at an academic audience rather than industry professionals (Mostafa & Hegazy, 2021). Therefore, an approach considering both a data-driven digital twin and integration in the real-world construction environment is needed to solve problems confronting the construction industry (Opoku et al., 2021). Consequently, data acquisition, storage, and processing must be applied in authentic construction scenarios. With this scientific contribution, we want to show an end-to-end approach to embedding AI-based tools in the construction environment while considering realistic on-site situations and expert knowledge. In conclusion, we are focusing on the following questions:

- In which way can raw data from on-site environments be acquired, processed, and stored?
- What is a suitable method to make as-planned vs. as-performed comparisons based on real-world scenarios considering the construction conditions?
- How can big data, retrieved from continuously monitored on-site environments, leverage its use case to support construction management?

METHODOLOGY

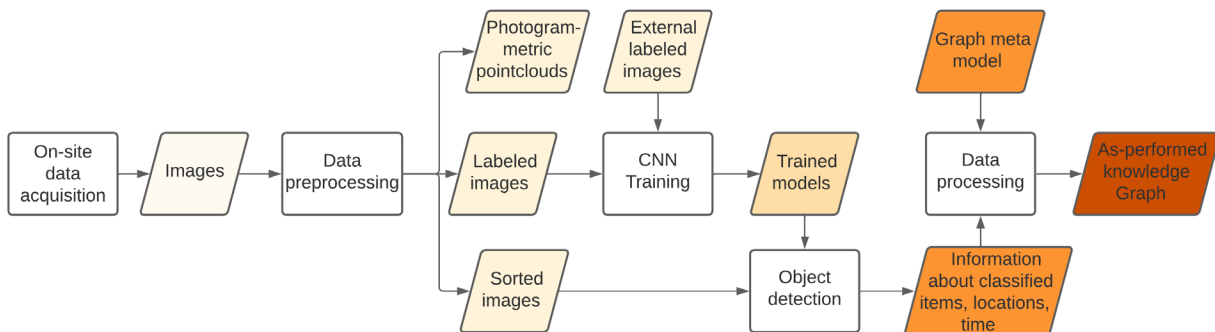


Figure 1. Flow of data to create a data-driven construction twin.

The data-driven digital twin construction framework is shown in Figure 1. The developed concept is designed in a general way, making it applicable to other data sources. Our data pipeline focuses exemplarily on processing one data source, the images. To keep the number of processes limited, we mainly concentrate on building construction sites and the shell construction phase. The as-planned data gets derived based on the available information, e.g., the 4D-BIM model, construction schedules, and construction reports. If parts of the as-planned data are incomplete, they are recreated synthetically based on acquired information and expert knowledge. Figure 1 distinguishes between process (rectangles) and data (rhombuses). The individual process parts are classified into *data acquisition*, *storage*, and *processing*. The data segments of the scheme

categorize into *raw data*, *processed data*, *information*, and *knowledge* and show the level of information highlighted with a color scale.

We use crane camera systems allocated to construction sites to collect raw data from real-world construction sites. In addition, further sensor systems are used for data capturing. However, in this contribution, we focus on the on-site images.

Examples of preprocessing tasks are undistorting images, sorting data types and time-sensitive data, and storing data. Even though the data is more structured after this step, it has no relationship to the construction actions. Thus, linking this data to the construction context is unavoidable.

We use deep neural networks to convert the image data into higher-level information. The object detection models are trained on labeled construction-based datasets to thoroughly handle the large amount of raw data. A combination of diverse datasets ensures higher prediction and reliability rates. After training, the weighted models are applied to the acquired datasets from the construction sites. The raw data gets converted into site-related information like classified construction items, locations, and temporal information as a by-product. We use graph databases to store and link the data.

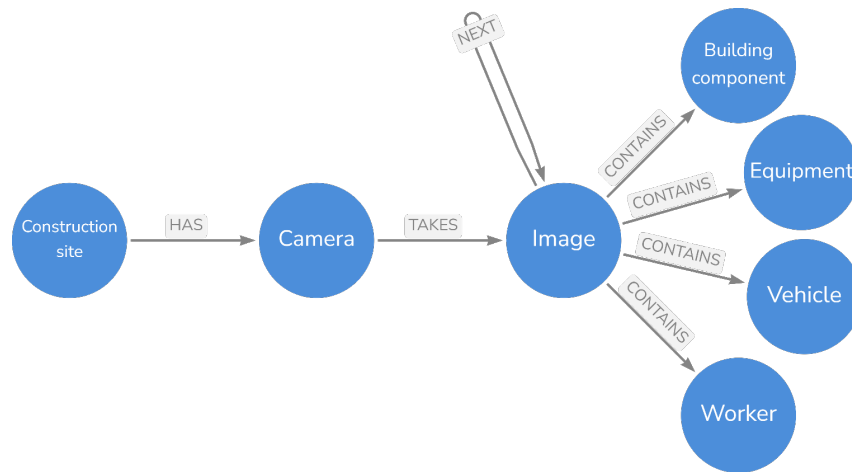


Figure 2. Graph meta-model for the as-performed construction graph.

Data Management. A well-defined data structure, or schema, is required to make meaningful, unambiguous use of any data. Figure 2 shows the graph meta-model structured into seven base categories: *Construction site*, *Camera*, *Image*, *Worker*, *Equipment*, *Vehicle*, and *Building Component*. The information from the image gets extracted, linked, and stored following the graph meta-model. The graph can be extended with off-site data, like weather or traffic data. Desired information, for example, the time and location of a specific item, is accessed with Cypher queries (Francis et al., 2018). This supports construction managers in checking whether processes follow the construction schedule and exposes potential schedule deviations. In addition, it is possible to reveal and investigate the site influence of weather or traffic. Transforming and fitting such data sources to the construction environment enables assembling information beyond traditional construction reports.

CASE STUDY

The case study was conducted on a construction site in Munich, Germany. We monitored the shell construction phase of the 1.500 square meters large building site with two camera systems holding

three cameras each. The cameras were set up to take an image every thirty seconds. A more detailed description of the used camera system can be found in Collins et al. (2022). We gained about 1.7M images from this construction site from April to August 2022. The images were stored at a local server and then regularly transferred to large cloud storage. Several preprocessing stages were conducted, including image sorting, image cleaning, and image undistorting, as well as anonymizing for privacy reasons.

To process the image data, we trained a Yolov5-7.0x model in PyTorch-Docker on the MOCS (Moving Objects on Construction) dataset developed by Xuehui et al. (2021) on an Nvidia RTX 8000 GPU. We finetuned the trained model with a self-labeled dataset of 2000 images based on the acquired data from on-site environments on 11 different classes: *Concrete silo*, *Slab panel*, *Worker*, *Concrete mixer*, *Other vehicle*, *Pump truck*, *Loader*, *Pillar rebar*, *Pillar*, *Slab rebar*, and *Formwork*.

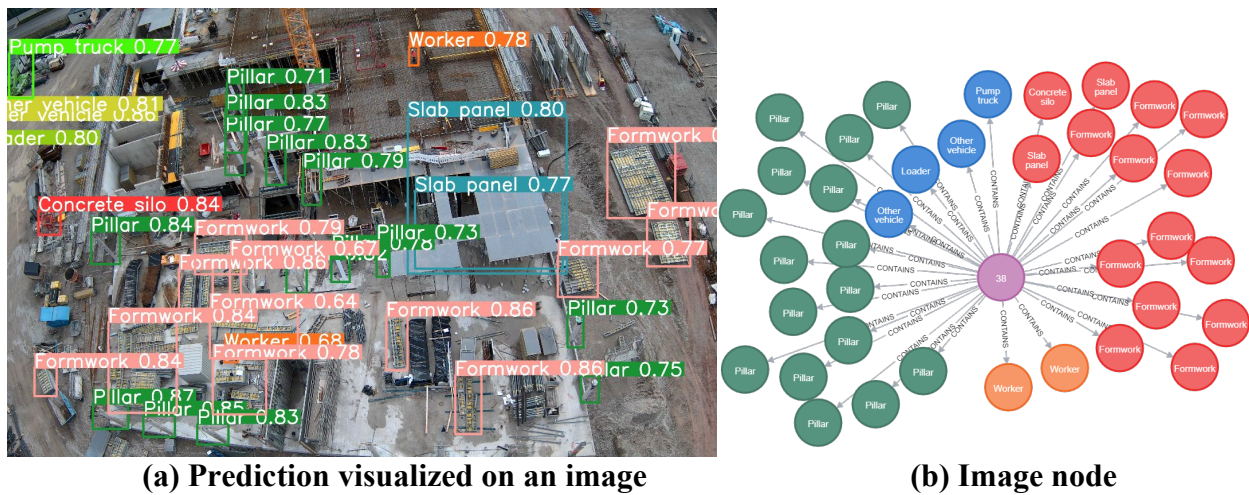


Figure 3. Processing construction-based image data.

We used the trained model to detect objects on the self-acquired on-site image dataset. A slicing approach developed by Akyon et al. (2022) was integrated into the pipeline to enhance the detection of smaller items on images, like workers or formwork elements. In addition, to transfer the position of the images to the site plans, we used a predefined transformation matrix based on perspective transformation, as shown by Pfitzner et al. (2022). To store the detected information, we set up a Neo4j graph database encapsulated in a docker container. The abovementioned meta-graph model enforced the linking and grouping of the eleven different predicted classes by creating individual nodes with the respective relationships based on the detected location, for example. A sample node of the graph is shown in Figure 3, highlighting the different on-site categories: In Figure 3 (a), the prediction of the image can be seen; Figure 3 (b) shows the equivalent node of the knowledge graph, adopting the on-site meta-graph model.

We use the query language Cypher (Francis et al., 2018) provided by Neo4j to answer specific questions, like when and how often certain items appeared on the construction site. For obtaining the information from the graph and plotting sample diagrams, we used the Python libraries Neomodel and Matplotlib. The Neo4j container processes the queries. Using the linked construction data, we determined process-related information, like the number of on-site personnel, for mirroring as-planned scenarios.

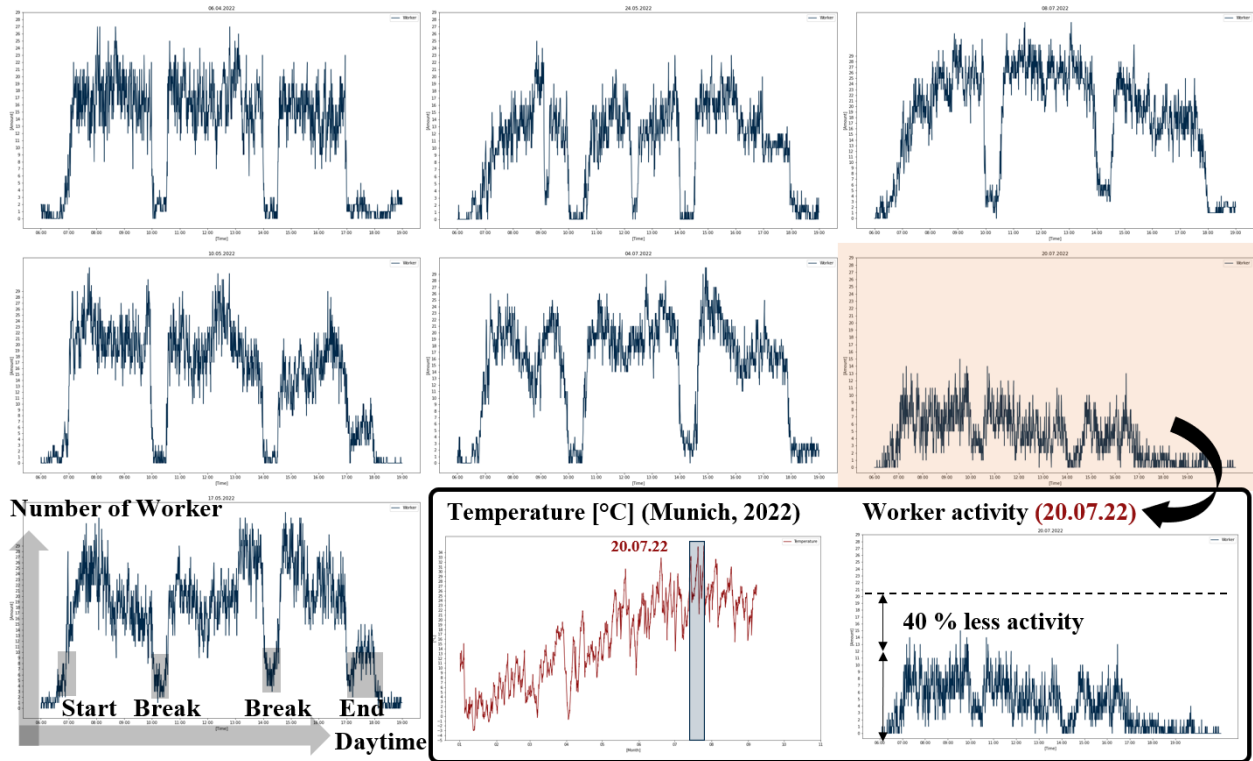


Figure 4. Number of on-site personnel on multiple days aligned with temperature data.

Construction site activity. Figure 4 demonstrates the plotted patterns based on the number of workers on-site. The daily course of the worker frequency shows that the start, end, and break times can be derived. Moreover, relationships between the construction site activity and the weather can be detected. During the hot summer days, the amount of on-site personnel decreased. The developed knowledge graph supports the recreation of specific construction processes in two ways: Top-down and bottom-up. Top-down enables recreating the construction processes using known dependencies, e.g., concreting a building component requires a concrete mixer, a concrete silo, and a group of workers. Bottom-up approaches build on dense clusters, e.g., many linked nodes, support a higher probability of on-site activity, found using graph data science algorithms.

DISCUSSION, LIMITATIONS, AND FUTURE RESEARCH

Table 1. Sample node statistics knowledge graph.

Labels	Number of nodes	Average rel. number
Worker	1.210.000	1,0
Equipment	920.000	1,7
Vehicle	240.000	1,2
Building Component	2.470.000	2,5

With the introduced method, we established a data-driven construction twin. In contrast to many other research contributions, the data was acquired from a real-world construction site scenario. The trained object detection model achieved a mAP_{0.5} of 88 %. Most of the classes reached mAP_{0.5}

scores above 85 %. Some classes, like *Formwork*, were difficult to detect due to their variety and thus reached lower scores (76%). With our knowledge graph, we demonstrated to link on-site data. Table 1 shows the enormous number of nodes generated in the graph based on the introduced categories. Overall, the graph contains 5 million nodes. On average, 38 items get detected on one image. The average number of relationships depends on whether an object is moving.

We must emphasize that we showed only a tiny fragment of potential data analysis of our approach. Possibilities go far beyond it, for example, a detailed investigation of the required resources associated with the daily concreting capacity across multiple job sites. An exemplary cross-domain analysis was performed by overlaying the weather data with the construction data. The construction managers in charge verified the temporary reduced on-site capacity. As highlighted in the methodology part, we aim to acquire data continuously over the construction phase. New possibilities like precise monitoring of individual items, combining external data, tracking project goals, and predicting future construction scenarios arose. Like other research contributions (Guo et al., 2021; Mostafa & Hegazy, 2021; Xuehui et al., 2021), we faced the problem of lacking open-source training data. Open datasets are required to push the development of on-site investigations further. Finally, more research needs to be done on on-site activity investigation supporting construction management decision-making (Zheng et al., 2021).

CONCLUSION

Our research contribution demonstrates that digitized construction monitoring yields many opportunities on the one hand. On the other hand, it still has many challenges and complications. Covering the construction site entirely, beyond types of building construction sites and shell construction phases, requires significantly more data. By now, the on-site data acquisition part is an immense challenge and holds back the returns from modern data science methods. Our method demonstrated the potential of an AI-based image processing approach enabling in-depth construction process analysis. We recommend researching additional data sources and possible acquisition techniques that may already exist in related industries. Further, real-world on-site collaborations with the industry must be tremendously pushed forward to learn more about dynamic process monitoring and develop big data understanding related to the construction environment.

ACKNOWLEDGMENTS

We thankfully acknowledge Innovation Management Bau GmbH for their financial support and for providing us access to multiple construction sites to acquire valuable data.

REFERENCES

- Akyon, F. C., Altinuc, S. O., & Temizel, A. (2022). *Slicing Aided Hyper Inference and Fine-tuning for Small Object Detection*. <https://doi.org/10.48550/arxiv.2202.06934>
- Barbosa, F., Woetzel, J., & Mischke, J. (2017). Reinventing construction: A route of higher productivity. *McKinsey Global Institute*.
- Braun, A., & Borrmann, A. (2019). Combining inverse photogrammetry and BIM for automated labeling of construction site images for machine learning. *Automation in Construction, 106*, 102879. <https://doi.org/10.1016/J.AUTCON.2019.102879>

- Collins, F., Pfitzner, F., & Schlenger, J. (2022). Scalable construction monitoring for an as-performed progress documentation across time. *Proceedings of 33. Forum Bauinformatik*. <https://mediatum.ub.tum.de/doc/1688376/document.pdf>
- Farmer, M. (2016). *The Farmer Review of the UK Construction Labour Model Modernise or Die Time to decide the industry's future*. www.cast-consultancy.com
- Francis, N., Green, A., Guagliardo, P., Libkin, L., Lindaaker, T., Marsault, V., Plantikow, S., Rydberg, M., Selmer, P., & Taylor, A. (2018). Cypher: An evolving query language for property graphs. *Proceedings of the ACM SIGMOD International Conference on Management of Data, 13*, 1433–1445. <https://doi.org/10.1145/3183713.3190657>
- Guo, B. H. W., Zou, Y., Fang, Y., Goh, Y. M., & Zou, P. X. W. (2021). Computer vision technologies for safety science and management in construction: A critical review and future research directions. *Safety Science, 135*. <https://doi.org/10.1016/j.ssci.2020.105130>
- Hodler, A. E., Farnham, B., Tokyo, S., Boston, B., Sebastopol, F., & Beijing, T. (2019). *Mark Needham Graph Algorithms Practical Examples in Apache Spark and Neo4j*.
- Li, H., Chan, G., Wong, J. K. W., & Skitmore, M. (2016). Real-time locating systems applications in construction. In *Automation in Construction* (Vol. 63, pp. 37–47). Elsevier. <https://doi.org/10.1016/j.autcon.2015.12.001>
- Mostafa, K., & Hegazy, T. (2021). Review of image-based analysis and applications in construction. In *Automation in Construction* (Vol. 122). Elsevier B.V. <https://doi.org/10.1016/j.autcon.2020.103516>
- Opoku, D. G. J., Perera, S., Osei-Kyei, R., & Rashidi, M. (2021). Digital twin application in the construction industry: A literature review. In *Journal of Building Engineering* (Vol. 40). Elsevier Ltd. <https://doi.org/10.1016/j.jobe.2021.102726>
- Pauwels, P., Zhang, S., & Lee, Y. C. (2017). Semantic web technologies in AEC industry: A literature overview. *Automation in Construction, 73*, 145–165. <https://doi.org/10.1016/J.AUTCON.2016.10.003>
- Pfitzner, F., Braun, A., & Borrmann, A. (2022). Towards data mining on construction sites: Heterogeneous data acquisition and fusion. *ECPPM: European Conferences on Product and Process Modeling in the Building Industry*.
- Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., & Girolami, M. (2020). Construction with digital twin information systems. *Data-Centric Engineering, 1*, e14. <https://doi.org/10.1017/dce.2020.16>
- Torabi, G., Hammad, A., & Bouguila, N. (2022). Two-Dimensional and Three-Dimensional CNN-Based Simultaneous Detection and Activity Classification of Construction Workers. *Journal of Computing in Civil Engineering, 36*(4). [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0001024](https://doi.org/10.1061/(ASCE)CP.1943-5487.0001024)
- Xu, S., Wang, J., Shou, W., Ngo, T., Sadick, A.-M., & Wang, Xiangyu. (2021). Computer Vision Techniques in Construction: A Critical Review. *Archives of Computational Methods in Engineering, 28*, 3383–3397. <https://doi.org/10.1007/s11831-020-09504-3>
- Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., & Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. *Automation in Construction, 122*. <https://doi.org/10.1016/j.autcon.2020.103482>
- Zheng, Y., Törmä, S., & Seppänen, O. (2021). A shared ontology suite for digital construction workflow. *Automation in Construction, 132*. <https://doi.org/10.1016/J.AUTCON.2021.103930>
- Zou, Z., Shi, Z., Guo, Y., Ye, J., & Member, S. (2019). Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*.