

# MONITORING CONCRETE POURING WITH KNOWLEDGE GRAPH-ENHANCED COMPUTER VISION: A CASE STUDY FROM MUNICH, GERMANY

Fabian Pfitzner<sup>1</sup>, Songbo Hu<sup>2</sup>, Alex Braun<sup>1</sup>, André Borrmann<sup>1</sup>, and Yihai Fang<sup>2</sup>

<sup>1</sup>Chair of Computational Modeling and Simulation, Technical University of Munich, Germany

<sup>2</sup>Department of Civil Engineering, Monash University, Australia

## Abstract

This paper introduces a novel approach for monitoring concrete pouring. Traditional manual tracking methods are tedious, while automated solutions, such as Computer Vision (CV)-enabled methods, are challenged with occulted data and limited adaptability to diverse crane behaviour patterns. We propose a knowledge graph-enhanced CV method that combines context knowledge with object recognition. This approach analyses tower crane behaviours and their interactions with workers, truck mixers, and building elements, providing a detailed and resilient interpretation of concrete pouring progress. Preliminary findings reveal the method's capacity to interpret incomplete data and comprehend complex site dynamics, demonstrating promising potential in a real-world scenario.

## Introduction

Concrete pouring is a common and critical construction activity, significantly impacting both the completion time and cost of construction projects (Wang et al., 2022). The crane plays a pivotal role in this activity, as the “crane-and-skip” method emerging as one of the most prevalent techniques for concrete pouring (Lu et al., 2003). In this process, concrete is mixed and then poured into the skip on the ground by workers; then the crane lifts the skip to one or more locations requesting concrete; once in position, the skip is tilted or manipulated to pour, before the empty skip is lowered back to the mixer for refill.

Traditionally, monitoring the progress of concrete pouring processes has been manual and approximate, where the total volume of concrete delivered to the site is used as an indirect indicator of pouring progress (Lu & Anson, 2004). This method merely provides a rough estimate of the progress, and cannot capture nuances related to the pouring process, such as the cycle time of crane lifts and the waiting time of truck mixers. Consequently, it provides limited insight into the bottlenecks of critical site resources (e.g., crane availability) and yields minimal contributions to crucial decisions that impact site productivity (e.g., hiring extra cranes) (Hu et al., 2021).

To gain insights into the pouring process, which requires a large volume of data, Computer Vision (CV) has emerged for automated data acquisition and analysis. For instance, Gong and Caldas (2010) developed a CV-based method to track crane hooks and concrete buckets (i.e., skips), enabling the analysis of concrete pouring states

(e.g., bucket readiness, pouring into specific columns) and cycle times. Nevertheless, these CV-based methodologies often confront adaptability challenges. They rely on the assumption of continuous and uninterrupted concrete pouring, following an overly rigid crane behaviour pattern in each pouring operation (e.g., the buckets have to return to the mixer after each pouring). Meanwhile, these methods often necessitate project-specific parameters, such as designated loading and unloading zones, to streamline data interpretation (Yang et al., 2014).

These rigid rules in data analyses often misalign with the dynamic environment of construction sites, leading to inaccurate data interpretations and limited generalisability in different construction projects. Compounding this issue are hardware limitations, such as cameras missing crucial frames or occlusions blocking the view. These limitations will also disrupt the crane behaviour pattern recognition, leading to inaccuracies or frequent relocating of cameras.

Therefore, an adaptable reasoning mechanism is required that ensures enhanced robustness and greater generalisability for monitoring concrete pouring. Knowledge Graphs (KGs) offer significant potential to elevate the reasoning capabilities of CV systems (Fang et al., 2020). KGs enable machines to comprehend complex relationships and enriched contexts within data, going beyond mere visual recognition (Pfitzner et al., 2023a).

By integrating KG and CV, the authors propose a universally adaptable interpretation of concrete pouring processes. In later sections, the paper includes a literature review on concrete pouring monitoring and KG-enhanced CV systems in the construction industry, followed by a detailed description of our methodology and a case study from an active construction site in Munich. It concludes with a discussion of the method's contributions, limitations and potential improvements.

## Literature review

### Methods of Monitoring Concrete Pouring

Over the past decades, continuous research has focused on monitoring concrete pouring, with relevant studies summarised in Table 1. Initially, this monitoring was mainly for quality control, aiming to predict concrete curing conditions and strength development (Moon & Yang, 2017). The process involved manual documentation of the concrete pouring time into formwork and environmental factors, such as humidity, to ensure compliance with curing standards.

Table 1: Literature on Monitoring Concrete Pouring

Reference	Year	Tracked item	Tracking Mechanism	Collected data	Outputs	Limitation
(Lu & Anson, 2004)	2004	Concrete Mixer	Retrieving information from manually prepared documents	Site arrival /departure time; Location and element type for pouring; Pouring method.	A productivity evaluation (i.e., volume/time) for different pouring methods	A coarse granularity of tracking at a dozen of truck trips level; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Moon & Yang, 2010)	2010	Concrete Mixer and concrete pump	Radio-frequency identification device (RFID)	Site arrival time, mixer entering/ leaving work zone time.	Total amount of poured/remaining concrete; Average pouring speed; Estimated time to complete.	A coarse granularity of tracking at a truck trip level; Manual installation of RFID tags; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Gong & Caldas, 2010)	2010	Bucket	Computer Vision (CV)	Buckets' presence in user-specified waiting zone and column zones; Buckets' absence (indicating their loading).	Durations of each user-defined pouring states for a particular pouring activity.	Assuming uninterrupted pouring; The case study merely focuses on the pouring of two columns; Relying on fixed states in the column pouring and site-specific spatial context; Requiring meticulous/dynamic planning of camera locations.
(Yang et al., 2014)	2014	Crane jib	Computer Vision (CV)	Crane jib trajectory and its overlap with site layout.	State durations of concreting pouring and non-concrete pouring activities.	Assuming concrete mixers' locations are fixed; Relying on site-specific spatial context and observed probability when differentiating concrete pouring states.
(Danel et al., 2022)	2022	Crane	Crane inbuilt motion trackers and load sensor	Slewing angle, trolley distance, hoist height, and load weight.	State durations of concreting pouring activities.	Assuming uninterrupted concrete pouring process; Relying on a standard crane behaviour pattern.
(Wang et al., 2022)	2022	Crane cable and bucket	Computer Vision (CV)	Concrete buckets' presence in a meticulously planned imaged area (e.g., on the top of the dam to be poured).	Cycle time of crane lifts.	Assuming uninterrupted concrete pouring processes; Relying on a standard crane behaviour pattern; Requiring meticulous locating of camera; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Kim et al., 2023)	2023	Sound	Classifying site sound with deep-learning algorithms to recognise concrete pouring	Acoustic signal from construction sites.	Start and finish time of concrete pouring; Abnormal concrete pouring such as the impact sound of the vibrator's formwork and the sound of concrete slowly leaking.	Focusing on the safety of concrete pouring; Not revealing productivity-related insights (e.g., cycle times).

Lu & Anson (2004) analysed hundreds of digital quality control records to explore the link between concreting speed and pouring methods. Their work provided significant insights for enhancing productivity through method selection. However, their reliance on manual data collection leads to a coarse tracking granularity (e.g., dozens of truck mixer trips) and restricts insights for site management.

The introduction of IoT devices creates a shift towards automated, sensor-based tracking methods. Moon & Yang (2010), for instance, employed RFID technology to monitor the movement of truck mixers on-site, enhancing tracking to a granularity of individual trips. Kim et al. (2023) used acoustic sensors combined with deep learning algorithms to pinpoint the state when concrete is poured into formwork. While it does not monitor every state of the concrete pouring process, such as loading the concrete, this method further refines the granularity for tracking to specific pouring states. Danel et al. (2022) further advanced the tracking by using in-built crane sensors to monitor every state of the concrete pouring process, thereby uncovering prolonged delays caused by resource miscoordination. However, IoT methods often struggle to gather context information efficiently, limiting their ability to correlate lift operations with overall construction progress.

Recent research has turned to CV for more nuanced monitoring. CV methods typically consist of two steps: object recognition and reasoning. Object recognition identifies concrete-related resources in images, while reasoning interprets the actions of these objects. Gong & Caldas (2010) were among the first to apply CV in this field. They detected concrete buckets and mapped them onto predefined work zones in the images to identify specific pouring states with an accuracy of 84.7%. However, the need for user input limited their method's applicability. In contrast, Wang et al. (2022) applied CV to dam construction, tracking concrete buckets and crane cable interactions to indicate pouring cycles, achieving an even higher accuracy (>99%). However, the method assumed an uninterrupted pouring process, which is unrealistic for building construction. Yang et al. (2014) focused on crane jib recognition to track crane movements in relation to site layout plans, allowing for the recognition of mixed crane operations (i.e., concrete pouring and other lifting activities). Yet, this required distinct separation of concrete and other materials' loading zones and relied on observed site-specific probabilities in the reasoning process, potentially limiting adaptability to other sites.

In addition to the generalisability issue, one of the main challenges with these CV methods is their struggle to correlate crane operations with specific building elements and to detect the particular causes of idling. For example, they were often unable to identify scenarios where concrete buckets were in position but left unattended. These limitations highlight the imperative need to develop a reasoning mechanism that can adapt to varied site-

specific parameters (e.g., site layouts) and effectively highlight the interactions between concrete pouring resources (e.g., cranes, buckets, workers, and mixers) and construction products (e.g., building elements).

### **Leveraging the Knowledge Graph to Enhance CV in Construction Applications**

Knowledge Graphs (KGs) are poised to significantly augment the reasoning capabilities of CV, providing a universally applicable, generalisable context. By focusing on universal heuristics, such as the spatial relationships among objects (e.g., workers), KGs actively process and interpret CV outcomes, transforming raw data into information and enabling the derivation of new insights. In the construction industry, KG-enhanced CV systems have shown particular promise in safety management. Fang et al. (2020) demonstrated this by developing a KG-based CV system that identifies potential hazards through the spatial relationships among workers, PPE, and heavy machinery. Lee and Yu (2023) employed a KG to standardise and identify safety hazards in mobile scaffolding use, effectively pinpointing common misuses, including absent outriggers or missing guardrails. Both KGs perform spatial analysis of relevant objects to identify unsafe conditions or behaviours.

Unlike unsafe behaviours that often occur in isolated instances, construction activities, such as concrete pouring, typically consist of multiple states following a specific sequence. Therefore, solely specifying spatial relationships may only allow for a rough productivity evaluation at a broader level (e.g., site level) rather than recognising specific construction activities (Pfitzner et al., 2023a). To link planned sequences to actual site conditions, Braun et al. (2020) proposed a graph-based method to derive construction sequences from BIM models, overlay the sequence with detected objects, and thereby compensate for data gaps caused by occlusions. This application underscores the capability of KGs to embed sequence information.

However, the sequence of concrete pouring states (i.e., lifting states) is more dynamic. Addressing this challenge, Hu et al. (2023) utilised a KG to organise heterogeneous data from motion trackers, weight sensors and images for recognising crane lifting states (e.g., pick-up, suspend). This approach embeds dependencies of states instead of a standard sequence to allow flexibility. Thus, it is able to understand complex operations, such as those with multiple unloads in one lift. Despite its effectiveness in crane operation recognition, Hu et al.'s approach falls short in differentiating concrete pouring from other lifting activities and in linking crane operations to specific building elements. Moreover, their method is not tailored for enriching CV data. This underscores the necessity for a specialised KG tailored for concrete pouring processes and CV data formats. Such a KG should be proficient in interpreting concrete pouring operations, distinguishing between various lifting activities and linking pouring operations to building elements seamlessly.

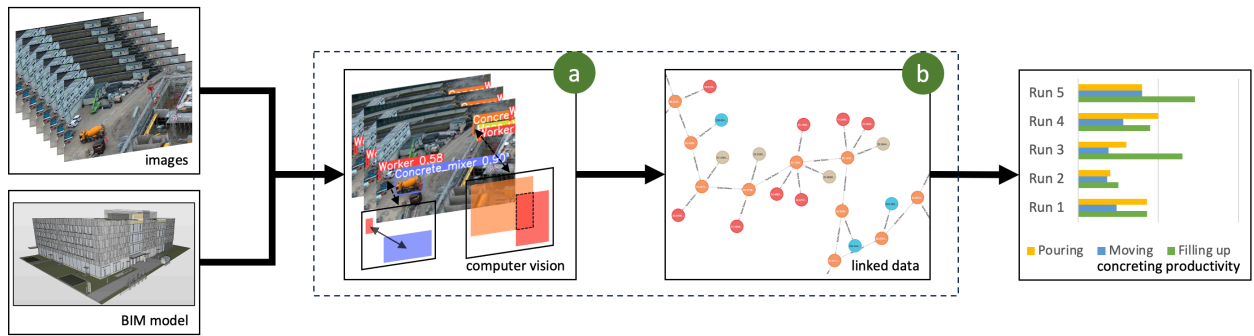


Figure 1: Overview of pipeline

## Methodology

### Overview of the approach

The scope of this study is to determine the utility of the crane in concrete pouring activities and associate crane operations with building progress. The underlying purpose of activity monitoring is to optimize the coordination of concrete delivery schedules, thereby mitigating overproduction and minimizing material waste during concrete pouring. Streamlining the concrete pouring process could condense construction timelines and elevate resource utilization.

We use semantic knowledge derived from the state dependencies in concrete pouring and BIM model to fuse CV object recognition results. As a result, we automatically monitor the pouring process in fine granularity (i.e., states), with the elapsed time of distinct process states recorded to provide precise insight into the required time for pouring diverse building zones. The results are anticipated to facilitate detailed control of construction progress.

The CV-based pipeline, shown in Figure 1, is developed to convert raw image and geometry data to precise process-level information. Instead of using data-heavy video streams, a low frequency of input images is chosen to avoid extensive computation times for processing long construction periods.

CV methods (a) are applied to extract information from the images. This step includes interpreting spatial dependencies in the context of the concrete pouring phase. In addition, the information is mapped to the BIM model using grid-based zones.

The knowledge graph (b) is used to determine unseen states and to link the as-planned geometry data. Unknown states are predicted based on their predecessors and successors. This predictive capability is particularly valuable in addressing data acquisition limitations, such as missing frames or occlusion, ensuring a robust analysis of individual pouring states while considering the as-planned quantities.

### CV-based state recognition

The computer vision part contains three steps: Object detection, spatial-temporal reasoning, and geometric projection.

A YOLOv8 model is used for object detection, covering the following classes: concrete\_bucket, hook, concrete\_mixer, hose, and worker. The model is deployed on image sequences representing the concrete pouring process. Figure 2 illustrates the distinct states of the concrete pouring process: Loading the concrete bucket at the concrete mixer, moving the bucket to the building element, and filling the building element's formwork with concrete. To derive these instant states from the images, spatial reasoning is necessary.

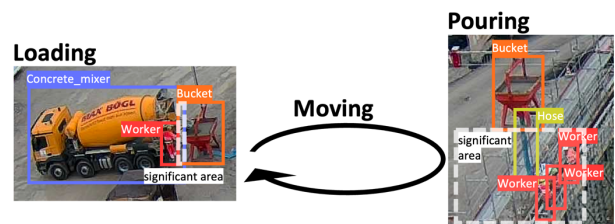


Figure 2: Spatial relationships of concreting cycle

In the case of crane-and-skip concrete pouring, the location of the concrete bucket and its relation to other elements defines the current action. Fig. 2 shows the essential scenarios of the concrete bucket to determine the process state: “Loading”, “Moving”, and “Pouring”. At the start of the process, the concrete bucket moves close to the concrete mixer to be filled by a worker. The bucket stays for a period of time (e.g., 30s) and then gets lifted by the crane. Intersection over Union (IoU) is used to determine the bounding boxes' spatial dependency.

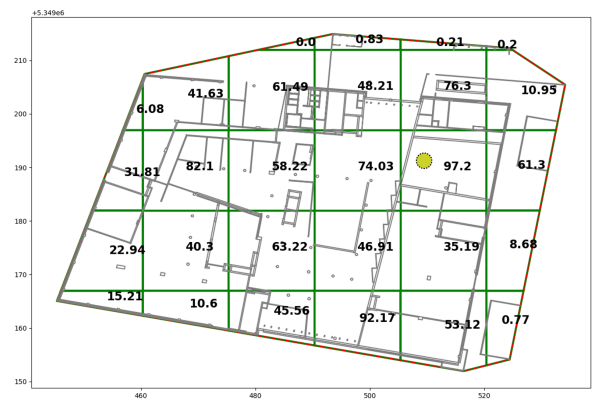


Figure 3: As-planned zones and volumes [m<sup>3</sup>] derived from the BIM model and projected base point of the hose (light green).

Subsequently, the bucket moves to the building element, and the pouring state starts. During the pouring state, a hose is rolled down, and a worker controls the nozzle of the hose to fill the formwork. Like in the “Loading” scenario, the activity is derived by interpreting the spatial relationships of the bounding boxes using IoU. During the “Pouring” state, workers and the hose are within the area underneath the bucket. The “Moving” state is computed based on the known order of process steps, discussed in the following section.

To compare the as-performed pouring volumes to the as-planned volumes, the image information is enriched. First, the as-performed quantities are computed using known volumes of the concrete mixer and bucket. Second, the as-planned volumes are generated using the BIM model, illustrated in Fig 3. The quantities of the BIM model are computed storey-wise to specific zones of a grid. The grid contains the quantity of all building elements within one zone. The location of the hose is projected to the BIM model using a perspective transformation approach presented by the authors in previous publications (Pfitzner et al., 2023b). As the hose traverses a specific zone, the as-performed data gets linked with the as-planned data. Based on the expected concrete amount of a specific zone, the amount of poured concrete is validated. Moreover, the amount of wasted concrete can be detected. A detailed investigation of geometry mapping is outside this paper’s scope and will be discussed by the authors in future publications.

### Ontological model for process reasoning

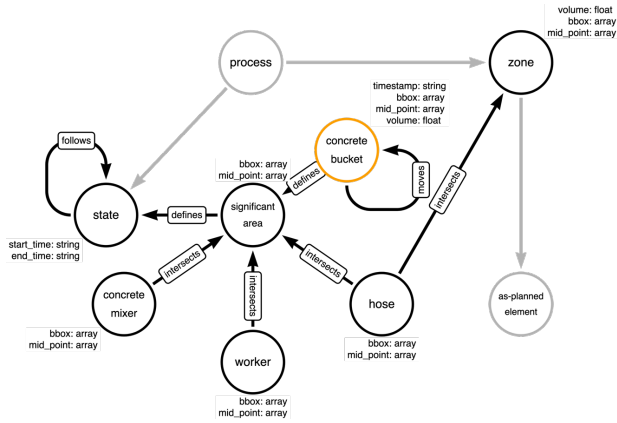


Figure 4: Ontological model

The ontological model, shown in Fig. 4, is designed according to the states representing the concrete pouring process. The process’s states are defined by the time-dependent actions of the concrete bucket. As such, the bucket nodes have a timestamp property and a relationship to their predecessors and successors. This ensures that the bucket’s current, past, and future actions can be determined based on the bucket’s significant area. The significant area is defined by typical construction-related dimensions; it encompasses the space beneath the bucket and incorporates a buffer, considering the objects utilising the bucket manifest beneath it.

Unseen states are parts of the process that cannot be detected on the frames. This is the case when the concrete bucket moves. The predecessor and successor relationships are utilised to compute the missing information. The unseen states of the concrete pouring cycle are determined based on the four different relationship patterns shown in Fig. 5. If the detected state changes, the bucket’s state is considered to be “Moving” in between, encompassing both scenarios “moving from mixer” and “moving from building elements”. In addition to that, the appearance of the concrete mixer is an indicator of pouring. If there is no mixer on-site, the crane is not involved in the pouring process.

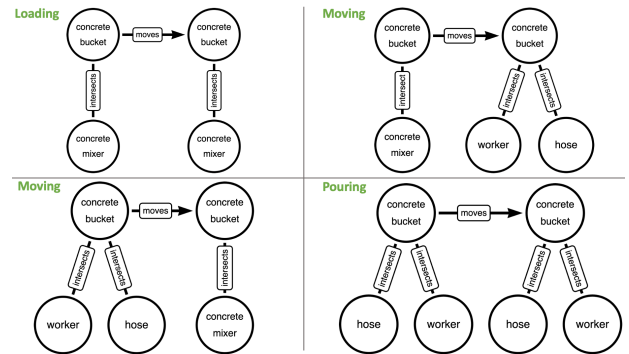


Figure 5: Relationship patterns for process reasoning

Once the relationship patterns are detected in the graph, the state information is stored in the equivalent state nodes. Subsequently, the process chain is created by sequential states containing start and end times. This has the advantage of querying the elapsed times straightforwardly. The start of the process chain is defined by the first time the concrete bucket gets loaded by the concrete mixer. The process chain ends once the bucket no longer returns to the concrete mixer and has moved away from the element. The link to the as-planned zones is created through the projected tip of the hose once it intersects with the corresponding zone. The process chain is used to investigate the elapsed time of individual states, the number of cycles, and the amount of concrete used during the process.

### Case Study

The case study, shown in Fig. 6, was conducted on a building construction project near Munich, Germany. The building was partially constructed using cast-in-place concrete pouring.

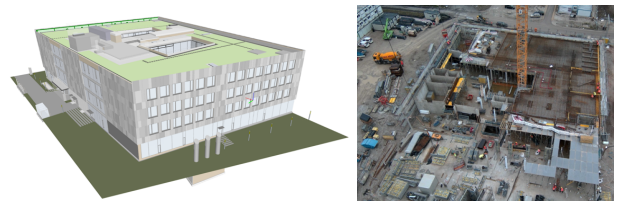


Figure 6: BIM model and crane camera perspective from the construction project

## Data and setup

The images were taken every 30 seconds from fixed crane cameras at diverse heights and perspectives. The dataset consists of a total number of 270k images. The annotated image dataset containing 325 images was split by 80/20. An additional model trained on the MOCS dataset (Xuehui et al., 2021) was included in the pipeline to detect workers. The model training on a Nvidia RTX 8000 GPU took 0.98 hours. To receive the as-planned geometry, we used an existing IFC model of the building. A Neo4j server was set up in a docker environment to host the labeled property graph. Four concrete pouring examples were investigated with a total number of 623 images.

## Results

The YOLOv8 model was trained on 79 epochs and reached an mAP score of 92.1 %, as shown in Table 2. The hose class performed weaker than the other classes due to fewer occurrences in the dataset.

Table 2: Object detection results

class	Precision	Recall	mAP <sup>0.5</sup>
bucket	92.6%	98.3%	97.7%
hook	97.6%	96.9%	99.2%
concrete mixer	98.1%	97.4%	97.8%
hose	87.3%	77.8%	73.5%
all	93.9%	92.6%	92.1%

The model was deployed batch-wise using a self-implemented PyTorch-based environment. The extracted image information was inserted into the Neo4j graph database using the Python library neomodel and the ontological model, illustrated in Figure 4. Using the distinct relationship patterns, the state nodes were generated and integrated into the graph. Based on the start- and end-time attributes of the state nodes, the elapsed time was queried. The buckets' bounding boxes, timestamps, and the particular concrete pouring states were annotated across two datasets to evaluate the ground truth of all states. The first dataset contained 119 samples; the second dataset contained 184 samples.

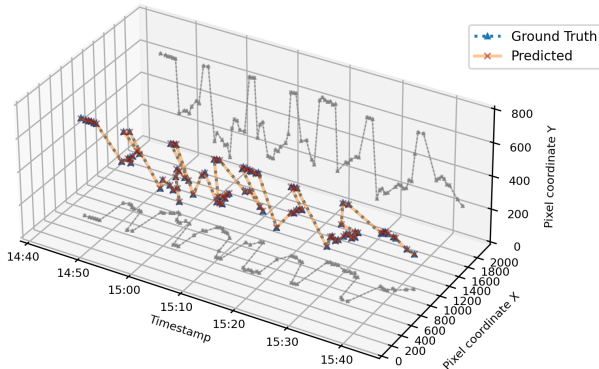


Figure 7: Comparison of the bucket's pixel coordinates between ground truth and detection result in dataset 1

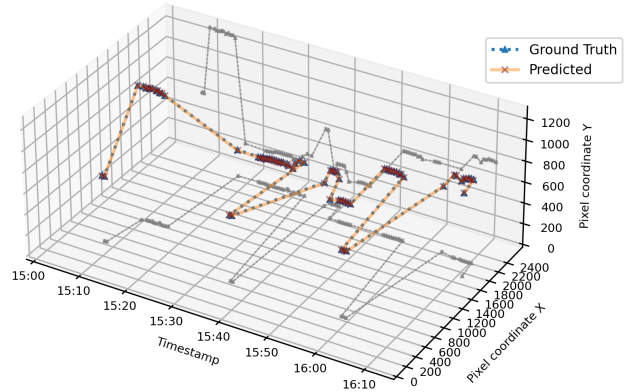


Figure 8: Comparison of the bucket's pixel coordinates between ground truth and detection result in dataset 2

Figures 7 and 8 illustrate the bucket detection accuracy by comparing the coordinates of the buckets' bounding boxes against ground truth data. The comparison is conducted over all timestamps across the two datasets, providing insights into the accuracy of the detection algorithm relative to the actual positions of the buckets. The predicted pouring states compared against the ground truth are shown in Figures 9 and 10. A state accuracy of 92.44% was achieved for the first dataset and 95.11% for the second dataset. The consistency and precision depicted in Fig. 7 and 8 underscore the robustness of the bucket tracking mechanism. However, Fig. 9 and 10 show some challenges in the state reasoning. Our method performed better on the second dataset due to the lower frequency of cycles.

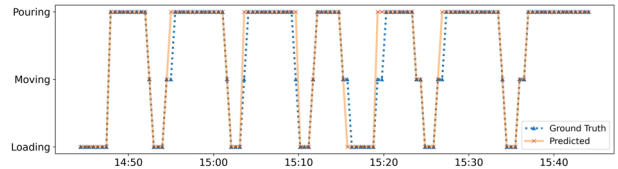


Figure 9: States of the pouring phase in dataset 1

In particular, states with shorter durations are more challenging to detect. It is important to note that exact moving times cannot be computed for the rare cases of moving time below the frames' interval (30 seconds).

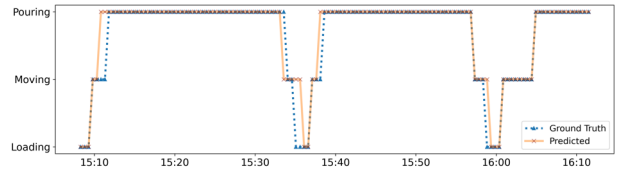


Figure 10: States of the pouring phase in dataset 2

We investigated the elapsed time of four different concrete pouring samples with the introduced pipeline. The times of the individual phases and the number of cycles are shown in Table 3. In general, pouring consumes most of the time. The results show that the process's time can significantly vary depending on the construction scenario.

In certain situations, precisely when waiting for the following concrete mixer's arrival, the moving time of the

bucket substantially differs. Moreover, factors like location and size of building elements play a substantial role. To allow further investigations, the amount of concrete volume was calculated based on the IFC model and included in the approach. For this step, the geometry from the IFC model was derived and processed using IfcOpenShell. Based on the building floor, the building elements' vertices were projected to the corresponding 2D floor, and polygons were generated to define the building elements' region, shown in Fig. 3. The grid was created using 15x15 metre cells. All building elements within a cell were considered and summed up to get the total volume amount of the cell. Finally, the correlating grid cell was determined based on the projected basepoint of the hose's bounding box and linked within the graph. Further studies of this topic will be included in future publications of the authors.

Table 3: Elapsed time of the specific cast-in-place states

Sample	No. 1	No. 2	No. 3	No. 4
No. cycles	7	2	7	3
Loading	0:13:06	0:05:04	0:20:44	0:03:31
Moving	0:04:12	0:05:05	0:23:24	0:11:11
Pouring	0:34:52	0:10:06	0:54:04	0:45:31

## Discussion and Future Work

The case study results validate the feasibility of our proposed KG-enhanced CV approach in monitoring and analysing the concrete pouring process. Capturing images at 30-second intervals ensures sufficient granularity, achieving 92.44% and 95.11% accuracy in identifying concrete pouring states. This accuracy surpasses previous efforts, such as the 84.7% reported by Gong & Caldas (2010).

With satisfactory accuracy, the case study also demonstrates our methods' superiority in flexibility and extensibility through our innovative adoption of KG in enriching CV systems. Traditional CV methods have a limited application scope, often requiring site-specific model training and manual annotations for critical areas, with a lack of correlation between crane operations and construction progress. Our approach uses CV to identify basic construction objects like workers, concrete mixers, and buckets, enhancing detection success.

The KG then streamlines their relationships with semantic context, enabling a generalisable data interpretation that adapts to changing construction contexts (e.g., concrete loading areas) without reprogramming. This facilitates efficient and resilient processing of diverse data across various construction stages, even different projects. Additionally, this method employs grid-based mapping to indirectly link crane operations with building elements. Although this connection is approximate, it can be further refined to recognise the specific elements being poured based on the positions of workers in a particular zone.

As a result of this method, the data on concrete pouring states offers a dynamic, real-time view of on-site resource

coordination, uncovering subtleties often missed in traditional monitoring. For instance, KG analysis reveals that prolonged "moving states" usually miss workers who operate the nozzle. This indicates time wasted mobilising workers, suggesting potential optimisations in pouring sequences. For instance, pouring activities should prioritise the occupied zones to minimise workers' travel. It also informs future development of our ontological model, where tracking worker movements could be valuable as it can highlight unnecessary movements.

Despite its innovative aspects, our approach has limitations. Currently, it focuses on crane-and-skip methods, with pump-based pouring scenarios outside our scope. Meanwhile, its success hinges on accurately detecting discrete states based on spatial relationships between construction elements, which can fail in cases of occlusion or distance. In particular, the hoses have a lower detection accuracy compared with other classes, suggesting the vulnerability of current reasoning rules. To address this, additional parameters and reasoning rules could be used to infer pouring states. For example, the distances between concrete buckets and workers, along with buckets' moving speed, can be used for reasoning. Future work will also explore graph-based machine learning, such as graph neural networks, for automated state classification.

## Conclusion

This paper introduces a KG-enhanced CV approach to monitoring concrete pouring. By integrating KG with CV, this method offers a dynamic and adaptable system capable of interpreting complex site dynamics and managing incomplete data. Achieving high accuracy rates of 92.44% and 95.11% in identifying concrete pouring states, our approach matches or surpasses earlier efforts in terms of data interpretation accuracy.

The strength of our method also lies in its ability to efficiently process big data with diverse quality and adapt to changing construction contexts, such as varying concrete loading areas, without needing reprogramming. This adaptability extends across different construction stages and projects, showcasing its potential for broad applicability. Additionally, the use of grid-based mapping to correlate crane operations with as-planned BIM models, although approximate at the moment, opens avenues for more precise recognition of specific pouring elements, even when cameras are positioned at a distance.

In general, our approach's real-time data analysis capability offers a nuanced view of concrete pouring-related resources, effectively highlighting miscoordinations. These insights pave the way for optimising resource allocation and improving site productivity. Anticipated further research aims to integrate this KG with semantic knowledge related to the pump-based concrete pouring method and enhance independence from the detection of site-specific or hard-to-recognise objects (e.g., hoses) by analysing the spatial relationships of easier-to-detect objects (e.g., buckets and

workers) or employing machine learning algorithms to discover patterns in KG's topology for classification.

In summary, our KG-enhanced CV method represents a significant stride in concrete pouring monitoring, offering improved accuracy, adaptability, and insight into construction processes. It promises to reshape traditional practices, leading to more efficient and effective construction project management.

### Acknowledgements

We thankfully acknowledge Innovation Management Bau GmbH 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 AG and Max Bögl.

### References

- Braun, A., Tuttas, S., Borrmann, A., & Stilla, U. (2020). Improving progress monitoring by fusing point clouds, semantic data and computer vision. *Automation in Construction*, 116. <https://doi.org/10.1016/j.autcon.2020.103210>
- Danel, T., Lafhaj, Z., Puppala, A., BuHamdan, S., Lienard, S., & Richard, P. (2022). Identifying tower crane activities with data: the case of the concrete pouring. *Engineering, Construction and Architectural Management*. <https://doi.org/10.1108/ECAM-10-2021-0936>
- Fang, W., Ma, L., Love, P. E. D., Luo, H., Ding, L., & Zhou, A. (2020). Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. *Automation in Construction*, 119. <https://doi.org/10.1016/j.autcon.2020.103310>
- Gong, J., & Caldas, C. H. (2010). Computer Vision-Based Video Interpretation Model for Automated Productivity Analysis of Construction Operations. *Journal of Computing in Civil Engineering*, 24(3). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000027](https://doi.org/10.1061/(asce)cp.1943-5487.0000027)
- Hu, S., Fang, Y., & Bai, Y. (2021). Automation and optimization in crane lift planning: A critical review. *Advanced Engineering Informatics*, 49. <https://doi.org/10.1016/j.aei.2021.101346>
- Hu, S., Wang, J., & Fang, Y. (2023). Semantic web-assisted progress monitoring of tower crane operations in construction projects. *The 30th EG-ICE: International Conference on Intelligent Computing in Engineering*.
- Kim, I., Kim, Y., & Chin, S. (2023). Deep-Learning-Based Sound Classification Model for Concrete Pouring Work Monitoring at a Construction Site. *Applied Sciences (Switzerland)*, 13(8). <https://doi.org/10.3390/app13084789>
- Lee, S. K., & Yu, J. H. (2023). Ontological inference process using AI-based object recognition for hazard awareness in construction sites. *Automation in Construction*, 153. <https://doi.org/10.1016/j.autcon.2023.104961>
- Lu, M., Anson, M., Tang, S. L., & Ying, Y. C. (2003). HKCONSIM: A Practical Simulation Solution to Planning Concrete Plant Operations in Hong Kong. *Journal of Construction Engineering and Management*, 129(5), 473–585. <https://doi.org/10.1061/ASCE0733-93642003129:5547>
- Lu, M., & Anson, M. (2004). Establish Concrete Placing Rates Using Quality Control Records from Hong Kong Building Construction Projects. *Journal of Construction Engineering and Management*. <https://doi.org/10.1061/ASCE0733-93642004130:2216>
- Moon, S., & Yang, B. (2010). Effective Monitoring of the Concrete Pouring Operation in an RFID-Based Environment. *Journal of Computing in Civil Engineering*, 24(1), 127. <https://doi.org/10.1061/ASCECP.1943-5487.0000004>
- Pfitzner, F., Braun, A., & Borrmann, A. (2023a). Object Detection Based Knowledge Graph Creation: Enabling Insight into Construction Processes. *ASCE International Conference on Computing in Civil Engineering* 2023. <https://doi.org/10.1061/9780784485224.023>
- Pfitzner, F., Braun, A., & Borrmann, A. (2023b). From Data to Knowledge: Construction Process Analysis Through Continuous Image Capturing, Object Detection, and Knowledge Graph Creation. *Automation in Construction*, under revision
- Wang, D., Wang, X., Ren, B., Wang, J., Zeng, T., Kang, D., & Wang, G. (2022). Vision-Based Productivity Analysis of Cable Crane Transportation Using Augmented Reality-Based Synthetic Image. *Journal of Computing in Civil Engineering*, 36(1). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000994](https://doi.org/10.1061/(asce)cp.1943-5487.0000994)
- 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>
- Yang, J., Vela, P., Teizer, J., & Shi, Z. (2014). Vision-Based Tower Crane Tracking for Understanding Construction Activity. *Journal of Computing in Civil Engineering*, 28(1). [https://doi.org/10.1061/\(asce\)cp.1943-5487.0000242](https://doi.org/10.1061/(asce)cp.1943-5487.0000242)