Automated Construction Site Analysis Using Images: A Review

Dr. Manjusha Tatiya, Vaishnavi Katikar, Samruddhi Bagal, Gajendra Thakur, Maharudra Ganjure

Artificial Intelligence and Data Science/ Indira College of Engineering and Management / Savitribai Phule Pune University, Pune, India

Corresponding Author Email: hodai_ds@indiraicem.ac.in vaishnavi.katikar@indiraicem.ac.in | ORCID: https://orcid.org/0009-0006-8784-3810

bagalsamruddhi22@gmail.com| ORCID: https://orcid.org/0009-0005-5994-9151

gajendrathakur1031@gmail.com | ORCID: https://orcid.org/0009-0006-5156-6196

maharudraganjure@gmail.com | ORCID: https://orcid.org/0009-0009-6055-9918

Abstract

Keeping construction projects on track is a major challenge for project managers. The outdated method of sending authorities to a site for manual inspection is inefficient, expensive, and difficult to scale across multiple projects. To address these problems, recent research has progressively turned to computer vision and machine learning to automate progress monitoring. This paper reviews the current state of these automated techniques, creating findings from key recent studies. Current research establishes significant success in using AI to analyze site images. Deep learning models like Convolutional Neural Networks (CNNs) are now extensively used for detecting construction stages and classifying materials. For more detailed tasks, such as identifying specific building components, object detection algorithms like YOLO and Mask R-CNN have also proven effective. Additionally, recent studies are actively addressing the unique challenges of monitoring progress in complex indoor environments. Despite these advancements, a notable gap remains between the capabilities of these AI models and the general needs of project management. The literature consistently highlights persistent challenges such as image obstructions, poor lighting conditions, and the need for more granular, activity-level tracking. This review consolidates the progress made in the field and highlights the critical next steps needed to bridge the gap from specialized AI tools to fully integrated, reliable construction monitoring platforms.

Keywords— Convolutional Neural Networks, Deep Learning, Object Detection, YOLO, Mask R-CNN, 3D Scene Reconstruction, Image Segmentation.

I. INTRODUCTION

Effective progress monitoring is fundamental to successful project management, particularly within the construction industry, where it is crucial for adhering to deadlines, managing resources, mitigating risks, and making informed decisions [1]. The construction sector, however, consistently faces significant challenges, including project delays, low productivity, and substantial cost

overruns [2]. Inefficient or inaccurate monitoring of work-in-progress is a major contributor to these issues, with studies indicating that more than half of construction projects experience delays and over two-thirds exceed their budgets [3]. The push for full-scale digitalization is a direct response to these persistent problems, with estimates suggesting that it could generate trillions of dollars in savings and dramatically increase industry profitability [4]. For decades, the standard approach to progress monitoring has relied on traditional methods that are laborious, time-consuming, costly, and susceptible to errors [5]. These practices depend heavily on manual site inspections, paper-based reports, and the subjective visual assessments of project managers, making the process slow, inaccurate, and often visually unfriendly [5][6]. This dependency on human judgment introduces inconsistencies and limits the ability of project teams to make timely and effective control decisions [6]. To overcome these limitations, the construction industry has seen a significant shift

ISSN: 2583-6129
DOI: 10.55041/ISJEM05298

toward digital transformation, marked by the accelerated adoption of automated technologies [7]. Central to this evolution is the use of computer vision (CV), a field of artificial intelligence that enables computers to interpret and understand visual data from sources like site images and videos. By leveraging CV in conjunction with deep learning, a subset of machine learning, it is possible to automate the analysis of visual data, offering far greater accuracy and efficiency than traditional methods. This automated approach facilitates the core tasks of progress monitoring: collecting as-built data, analyzing it, estimating progress against as-planned models, and visualizing the results [8]. Despite the promise of automation, significant challenges persist, particularly in achieving accurate progress tracking at the schedule activity level (e.g., formwork, reinforcement, concrete placement) [5]. Automated systems often face difficulties in complex and cluttered indoor construction environments, where obstructions, variable lighting conditions, and poor line-of-sight can compromise data quality and completeness [9]. Furthermore, many existing vision-based methods require Building Information Models (BIM) with an elevated Level of Development (LoD), which are not always available. This often restricts progress reports to a simple binary status— such as ’built’ or ’not built’— which fails to capture the nuanced status of partially completed tasks. An activity that is halfway complete on a monitoring date would be incorrectly reported as having zero progress, obscuring potential issues until the task is fully finished [10]. A critical gap, therefore, exists for a system that can accurately quantify the work-in progress for partially completed construction activities and is accessible to non-technical users. Many advanced deep learning models are developed without considering the practical deployment platforms needed by project managers who lack a programming background [11]. This study addresses these challenges by developing a deep learning-based computer vision model, built upon a Mask Recurrent Convolutional Neural Network (Mask R-CNN), specifically designed to automatically quantify the work-in-progress of indoor construction elements. Furthermore, this paper details the process of deploying the trained model on a user-friendly, cloud-based platform called Streamlit, making sophisticated progress monitoring tools accessible for practical application [12]. By developing a model tailored for quantifying partial work completion and demonstrating its deployment, this study contributes significantly to the digitalization of construction project management [11]. The advancements presented aim to provide project managers with a more accurate, automated, and accessible tool for progress monitoring, ultimately enabling better and more timely decision-making to keep projects on schedule and within budget. The following sections will detail the literature review, research methodology, model development and deployment, and a discussion of the results and their implications [12].

II. LITERATURE REVIEW

Computer-assisted construction progress monitoring has gone from traditional manual on-site inspections to modern image-based and AI-driven technologies. Early research concentrated on tracking progress through comparison with visual records while more recent work combines deep learning, semantic segmentation, and Building Information Modeling (BIM) for detailed, ongoing insights. This paper consolidates contributions up to date under four specific themes: vision-based monitoring methods, machine learning and image processing methods, semantic segmentation solutions, and computer vision- based reporting.

2.1. Vision-Based Monitoring Methods

Initial methods used static images, time-lapse photography, and video to track project progress. Yang et al. (2015) [13] categorized these methods into: a. Project based monitoring, a comparison between as-built conditions and planned 4D models to detect deviations. b. Monitoring of operations, which watches the movements of equipment and workers to analyze productivity and safety. While these techniques automated the tracking of progress, they were limited in extent by manual data analysis, by gaps in human–machine interaction, and by issues of scalability. For instance, it became common to show deviations based on BIM models with colorful overlays (e.g., traffic light’s metaphor), but the real-time and automatic reasoning about progress was kind of limited [3,10].

2.2. Machine Learning and Image Processing Strategies

To replace the manual inspection with algorithms, machine learning (ML) and image processing were utilized. Greeshma and Edayadiyil (2022) designed CNN-based classifier to detect a masonry activity from the dataset of 356 images. Their model achieved recall

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 12 | Dec – 2025 DOI: 10.55041/ISJEM05298
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

with more than 80produced automatic performance reports through Python and Excel [14]. This study showed that task specific monitoring can be achieved reasonably accurate using low-cost ML tools like OpenCV and CNNs. But it was limited to one activity scenario (masonry), and it did not seem scalable to more construction phases. Additionally, the absence of a broad set of normalized datasets limits the robustness of these ML-based methods [15-19].

2.3. Semantic Segmentation and Activity-Level Frameworks

Pal et al. made a breakthrough to monitor the progress who suggested the Activity-Level Progress Monitoring System (ALPMS) [13]. Unlike the binary models that preceded it (built' / not built'), ALPMS: Produce orthogonalized surface views by projective transformation and NeRF [20]. Utilized deep learning-based semantic segmentation (Mask- RCNN) to detect partial completion. The dataset was divided into 1,723 training images, 490 validation images, and 245 testing images [13]. Offers percentage completion tracking with less than 6 This feature is noteworthy because construction times more than once also deal with multiple activities combined related to one BIM object. Partial progress was not overseen in existing methods, and they led to misinterpretations of on-site status. With the help of as-built point clouds and semantic segmentation, ALPMS increased the level of detail and accuracy of reporting [13].

2.4. Computer Vision for Reporting and BIM Integration

Shamsollahi et al. (2022) performed a systematic review on computer vision applications for the monitoring of construction projects and had discovered that the applications [21] can be largely divided into three categories: 3D Scene Reconstruction–estimating point cloud /mesh models from single/multiple images (e.g., Structure-from-Motion, RGB-D cameras) [22, 23]. Object Detection and Tracking – identifying and positioning construction resources, such as construction materials, plant, and technicians. Image Segmentation: Defining progress at an activity level from images taken [24]. They emphasized the growing importance of BIM integration, and the comparison between going as-built models versus planned schedule. Nevertheless, there are still challenges regarding computation time, application to different site conditions, and a lack of annotated datasets.

2.5. Synthesis and Research Gaps

The sequenced literature follows a distinct evolutionary path:

- From vision-based documentation (Yang et al., 2015) [10]
- To ML based recognition of specific tasks (Greeshma Edayadiyil, 2022) [14]
- To activity-levels segmentation (Pal et al., 2024) [13]
- Towards comprehensive review of computer vision focusing on BIM integration (Shamsollahi et al., 2022) [21].

However, despite these advances there are still several large gaps:

- There is limited generalization over more than one of the construction phases (sub structure, super-structure, facades, interior works). Most of the models are trained for specific tasks.
- Hazards and safety checks are seldom included in the monitoring of progress.
- Computation speed is a bottleneck for both scalability and for deploying real-time.
- Challenges stay in terms of dataset, and there are not many large-scale, standardized benchmarks for multi-phase progress monitoring

Figure1: Evolution of Automated Construction Progress Monitoring Approaches

2.6. Relevance to the Proposed Study

The proposed study would fill these gaps by: Design and implementation of a multi-model machine learning approach for real-time monitoring of various construction stages.

- Including risk detection (such as earthquake- prone area risk) and progress tracking.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 12 | Dec – 2025 DOI: 10.55041/ISJEM05298
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

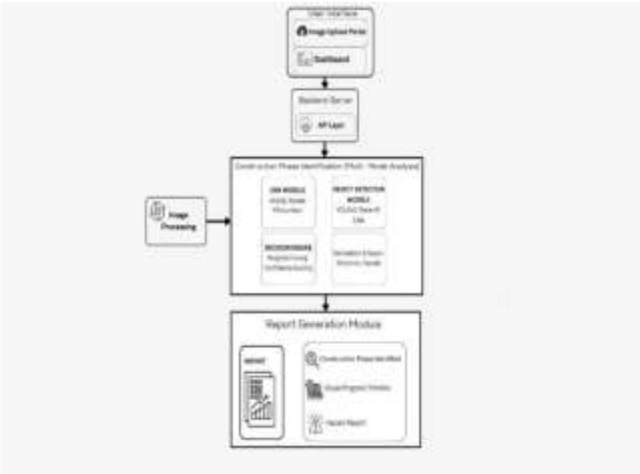

Figure 2: Proposed System Architecture

- Using advanced image analytics and BIM integration to provide scalable, automated reporting.

- Contributions Actively fosters the creation and use of benchmark datasets that allow fair, generalizable, and repeatable benchmarks of progress tracking research.

This exercise synthesizes potential of multi- model ML frameworks to grow beyond activity specific solutions into a seamless, intelligent monitoring system that addresses industry needs in terms of efficiency, accuracy, and safety.

III. LIMITATIONS

3.1. Dataset Generalizability

- Limitation: The models, while trained on extensive and augmented datasets, remain not entirely representative of all indoor and outdoor construction situations and thus are restricted in generalizing to unseen conditions or highly variable site configurations.[2][3][1]

- Challenge: Unconventional lighting, heavy occlusions, uncommon materials, or uncommon geometric configurations may reduce prediction performance if such situations weren't properly addressed in the training data.[3][2]

IV. DATA COLLECTION CONSTRAINTS

- **Limitation**: Reliance on camera-based data (pictures, videos, and time-lapse) limits the ability to evaluate progress to areas with visual access.[1][3]

- **Challenge**: Important operations may go unnoticed due to occlusions (static from stored materials, dynamic from personnel or equipment) and restrictions in camera placement or coverage, which could result in insufficient status reporting.[2][3][1]

4.1. Partial Activity Recognition

Limitation: Although the framework can approximate partial accomplishment, it has trouble measuring advancement Display sequences that are when visually similar tasks: or overlap. Take place on the same component simultaneously (for example, insulation and framing installed almost simultaneously).[3]

Challenge: Accuracy at the activity level may be impacted by current segmentation and classification models' inability to distinguish visually blended activities or their confusion of related stages.[4][3]

4.2. Model Deployment and Scalability

Limitation: Demonstrations tend to be in controlled or semi-controlled settings. Large- scale deployment to operational live sites introduces new challenges: 1. Integrating into current workflows and user habits. 2.Returning training and acceptance by site personnel. 3.Responsible use of privacy, security, and data ownership.[1][2][3]

Challenge: Real-time or high-frequency monitoring requires a lot of computational resources for 3D reconstruction, inference, and synchronization.[2][3]

4.3. Manual Inputs and Preprocessing

Limitation: Various workflow steps involve expert human effort, including: 1. Ground truth annotation for training models. 2.Manual identification of reference points for alignment of BIM/point cloud. 3.Quality assurance periodically in the pipeline of capture.[3][1]

Challenge: Human inputs can cause delays, subjectivity, and inconsistency, curtailing the path to full automation.[1][2][3].

4.4. Lighting and Environmental Effects

Limitation: Indoor locations with poor lighting, dust, reflections, or moving machinery may make segmentation and detection less dependable.[2][3][1]

Challenge: Although basic preprocessing can be beneficial, more domain-invariant techniques and additional sensors (such LIDAR and infrared) would be

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 12 | Dec – 2025 DOI: 10.55041/ISJEM05298
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

needed for wider robustness.[3]

V. CONCLUSION

This study develops a strong, scalable system to automate and improve building progress monitoring by utilizing deep learning and computer vision, incorporating Mask R-CNN and sophisticated 3D data fusion. It offers real- time integration with 4D BIM for visualization, helps project managers with earned value and schedule management, and achieves mean absolute errors in activity-level quantification on representative case studies below 6.

Notwithstanding these successes, there is still much to be done to address issues with workflow automation, generalization, occlusion management, incomplete progress measurement, and user adoption and scaling. The following areas should be the focus of future research:

- Growing and broadening annotated datasets (multi-modal imaging, cross-project, and cross-region data).
- Creating more resilient fusion and multi- sensor techniques to deal with changing illumination and occlusion.
- Investigating weakly-supervised or unsupervised methods to lessen the requirement for manual labeling.
- Developing smooth, intuitive user interfaces and APIs for industry integration and on-site deployment

REFERENCES

[1]Navon, R., 2007. Research in automated measurement of project performance indicators. Autom. ConStruct. 16 (2), 176–188. https://doi.org/10.1016/j.autcon.2006.03.00 3.

[2]P.R. de Almeida, M.Z. Solas, A. Renz, M.M. Buhler, P. Gerbert, S. Castagnino, C. Roth- ¨ baller, Shaping the Future of Construction: A Breakthrough in Mindset and Technology, World Economic Forum, Switzerland, 2016. https://www.weforum.org/reports/shapingt he-future-of-construction a-breakthrough-in- mindset-and-technology (Feb. 9, 2022).

[3]Han, K., Cline, D., Golparvar-Fard, M., 2015. Formalized knowledge of construction sequencing for visual monitoring of work-in- progress via incomplete point clouds and lowLoD 4D BIMs. Adv. Eng. Inf. 29 (4), 889–901.https://doi.org/10.1016/j.aei.2015.10.00 6.

[4]M. Purdy, P. Daugherty, How AI Boosts Industry Profits and Innovation, Accenture, 2017. https://www.accenture.com/no-en/insight-ai-industry-growth (Feb. 9, 2022).

[5]Golparvar-Fard, M., Pe˜ na-Mora, F., Savarese, S., 2015. Automated progress monitoring using unordered daily construction photographs and IFC based Building In formation Models. J. Comput. Civ. Eng.29(1) https://doi.org/10.1061/(asce)cp.1943- 5487.0000205

[6]H.K. Kevin, G. Fard, Multi-Sample Image- based Material Recognition and Formalized Sequencing Knowledge for Operation-Level Construction Progress Monitoring, Computing in Civil and Building Engineering 364–372 (2014), https://doi.org/10.1061/9780784413616.04 6.

[7]Bednar, P.M., Welch, C., 2020. Socio- technical perspectives on smart working: creating meaningful and sustainable systems. Inf. Syst. Front 22 (2), 281–298. https://doi. org/10.1007/s10796-019-09921-1.

[8]Moragane, H., Perera, B.A.K.S., Palihakkara, A.D., Ekanayake, B., 2024. Application of computer vision for construction progress monitoring: a qualitative investigation. Construct. Innovat. 24 (2), 446–469. https://doi.org/10.1108/CI-05-2022-0130.

[9]Deng, H., Hong, H., Luo, D., Deng, Y., Su, C., 2020. Automatic indoor construction process monitoring for tiles based on BIM and

[10] computer vision. J. Construct. Eng. Manag. 146 (1) https://doi.org/10.1061/(asce)co.1943-7862.0001744.

[11] Yang, J., Park, M.-W., Vela, P. A., Golparvar-Fard, M. (2015). Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future. Advanced Engineering Informatics, 29(2),211–224. https://doi.org/10.1016/j.aei.2015.01.011

[12] Khorasani, M., Abdou, M., Hern´andez Fern´andez, J., 2022. Streamlit basics. In: Web Application Development With Streamlit. Apress, Berkeley, CA. https://doi.org/ 10.1007/978-1 4842-8111-6-2.

[13] Biyanka Ekanayake, Johnny Kwok Wai Wong, Alireza Ahmadian Fard Fini, Peter Smith, Vishal Thengane, Deep learning-based computer vision in project management: Automating indoor

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 12 | Dec – 2025 DOI: 10.55041/ISJEM05298
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

construction progress monitoring https://doi.org/10.1016/j.plas.2024.100149

[14] Pal, A., Lin, J. J., Hsieh, S.-H., Golparvar-Fard, M. (2024). Activity-level construction progress monitoring through semantic segmentation of 3D-informed orthographic images. Automation in Construction, 157, 105157. https://doi.org/10.1016/j.autcon.2023.1051 57

[15] Greeshma, A. S., Edayadiyil, J. B. (2022). Automated progress monitoring of construction projects using machine learning and image processing approach. Materials Today: Proceedings, 65, 554–563. https://doi.org/10.1016/j.matpr.2022.03.13 7

[16] C. Kim, H. Son, C. Kim, n Automated construction progress measurement using a 4D building information model and 3D data, Automation in Construction, Elsevier B.V. 31 (2013)75–82, https://doi.org/10.1016/j.autcon.2012.11.04 1.

[17] N.R. Howes, Managing Software Development Projects for Maximum Productivity, IEEE Trans. Software Eng. 10 (1) (1984) 27–35, https://doi.org/ 10.1109/TSE.1984.5010195.

[18] A.M.A. Aziz, Minimum performance bounds for evaluating contractor's performance during construction of highway pavement projects, Construction Management and Economics 26 (5) (2008) 37–41, https://doi.org/10.1080/014461908019187 48.

[19] Breunig, M. M., Kriegel, H., Ng, R. T., and Sander, J. (2000). "LOF : Identifying Density Based Local Outliers." 65(1), 93–104. https://doi.org/10.1145/ 335191.335388

[20] H. Mahami, F. Nasirzadeh, A. Hosseininaveh Ahmadabadian, F. Esmaeili, S. Nahavandi, Imaging network design to improve the automated construction progress monitoring process, Construction Innovation 19 (3) (2019) 386–404, https://doi.org/10.1108/CI-07- 2018-0059.

[21] T. Muller, A. Evans, C. Schied, A. Keller, Instant neural graphics primitives with ¨ a multiresolution hash encoding, ACM Trans. Graph. 41 (4) (2022), https://doi.org/ 10.1145/3528223.3530127. Association for Computing Machinery. 20 A. Pal et al.

[22] Shamsollahi, D., Moselhi, O., Khorasani, K. (2022). Construction progress monitoring and reporting using computer vision techniques – A review. In Proceedings of the 39th International Symposium on Automation andRobotics in Construction (ISARC 2022). Concordia University.

[23] Kang Z., Yang J., Yang Z. and Cheng S. A review 472 39th International Symposium on Automation and Robotics in Construction (ISARC 2022) of techniques for 3D reconstruction of indoor environments. ISPRS International Journal of Geo Information, 9(5):330, 2020.

[24] Eder M., Moulon P., and Guan L. Pano Popups: Indoor 3D Reconstruction with a PlaneAware Network. In Proceedings of the 2019 International Conference on 3D Vision (3DV), Pages 76-84, Quebec City, Canada, 2019.

[25] Paneru S. and Jeelani I. Computer vision applications in construction: Current state, opportunities challenges. Automation in Construction, 132:103940, 2021