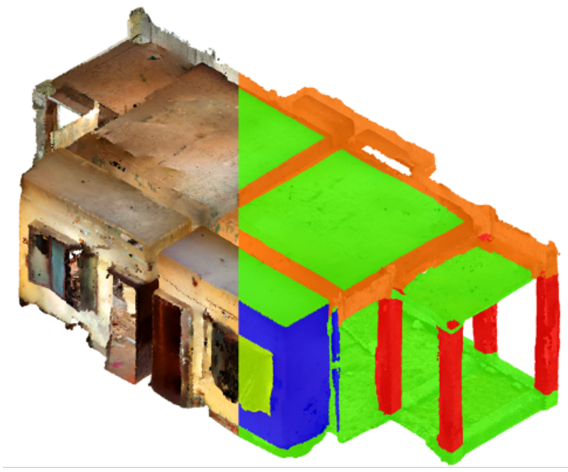
DEPARTMENT OF CIVIL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY
MADRAS
CHENNAI – 600036

A Hybrid Self-Supervised Approach Towards Computer Vision Based Construction Progress Monitoring



A Thesis

Submitted by

Varun Kumar Reja

For the award of the degree

Of

DOCTOR OF PHILOSOPHY

July 2023

"I'm a scientist; because I invent, transform, create, and destroy for a living, and when I don't like something about the world, I change it."

- ***Rick Sanchez***

I dedicate this thesis to my mother, father, and my wife for always believing in me!

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Varun Kumar Reja, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and IT at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with the Indian Institute of Technology, Madras, India.

This research is supported by the Australian Government Research Training Program.

Signature: Varun Kumar Reja

Date: 07-June-2023

THESIS CERTIFICATE

This is to certify that the thesis titled **A Hybrid Self-Supervised Approach Towards Computer Vision Based Construction Progress Monitoring** submitted by me to the Indian Institute of Technology, Madras for the award of the degree of **DOCTOR OF PHILOSOPHY,** is a bona fide record of research work carried out by me under the supervision of Prof. Koshy Varghese and Prof. Quang Phuc Ha. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

*The research work has been carried out at IIT Madras and the University of Technology Sydney.

Production Note:
Signature removed prior to publication.

Chennai 600036

Varun Kumar Reja

Research Scholar

Date: Wednesday, 07 June 2023

Dr. Koshy Varghese
Professor
Department of Civil Engineering
Indian Institute of Technology Madras, India
Research Guide

Dr. Quang Phuc Ha
Associate Professor
School of Electrical and Data Engineering
University of Technology Sydney, Australia

Research Guide

LIST OF PAPERS BASED ON THIS THESIS

REFEREED JOURNALS BASED ON THIS THESIS

1. **V.K. Reja**, K. Varghese, Q.P. Ha, Computer Vision-Based Construction Progress Monitoring, in: Automation in Construction 138, 2022 pp. 104245. https://doi.org/10.1016/j.autcon.2022.104245

REFEREED JOURNALS (OTHER)

1. **V.K. Reja**, S. Goyal, K. Varghese, B Ravindran, Q.P. Ha, ConPro-NET: A Hybrid Self Supervised Architecture for Progress Quantification of Construction Projects, in: Automation in Construction (Submitted)
2. **V.K. Reja**, M.S. Pradeep, K. Varghese, Digital Twin for Construction Project Management: Applications and Future Research Directions, IEEJ Transactions on Advanced Information & Communication, Sensing, and Medical Technologies Supporting Smart Cities (Submitted)

PRESENTATIONS IN CONFERENCES

1. **V.K. Reja**, M.S. Pradeep, K. Varghese, A Systematic Classification & Evaluation of Automated Progress Monitoring Technologies in Construction, in 39th Int. Symp. Autom. Robot. Constr. (ISARC 2021), Bogota, Columbia, 13th July 2022.

2. **V.K. Reja**, K. Varghese, Q.P. Ha, As-Built Data Acquisition for Vision-Based Construction Progress Monitoring: A Qualitative Evaluation of Factors, 10th World Construction Symposium, Sri Lanka, 24th June 2022.

3. **V.K. Reja**, P. Bhadaniya, K. Varghese, Q.P. Ha, Vision-based Progress Monitoring of Building Structures using Point-Intensity Approach, in 38th Int. Symp. Autom. Robot. Constr. (ISARC 2021), Dubai, UAE, 2nd November 2021.
4. **V.K. Reja**, K. Varghese, Digital Twin Applications for Construction Project Management, in: Joint Indo-Japanese Smart City Conference (IJSCS-2022) 15th March 2022.
5. V. Thelakulla, **V.K. Reja**, K. Varghese, A Web-Based GIS Tool for Progress Monitoring of Linear Construction Projects, in 38th Int. Symp. Autom. Robot. Constr. (ISARC 2021), Dubai, UAE, 2nd November 2021.
6. D. Nath, **V.K. Reja**, K. Varghese, A Critical Review of Literature on Collaboration in Construction, in: Indian Lean Construction Conference (ILCC 2021) 6th December 2021.
7. D. Nath, **V.K. Reja**, K. Varghese, A framework to measure collaboration in a construction project, in: 9th World Constr. Symp. 2021 Reshaping Constr. Strateg. Struct. Cult. Transform. Towar. 'Next Norm., The Ceylon Institute of Builders - Sri Lanka, 25th July 2021.
8. A.K. Saxena, **V.K. Reja**, K. Varghese, IoT Enabled Framework for Real-time Management of Power-Tools at Construction Projects, in: 37th Int. Symp. Autom. Robot. Constr., 16th October 2020.
9. **V.K. Reja**, K. Varghese, Impact of 5G Technology on IoT Applications in Construction Project Management, in: 36th Int. Symp. Autom. Robot. Constr. (ISARC 2019), Banff, Canada, 20th May 2019.

PUBLICATIONS IN CONFERENCE PROCEEDINGS

1. A.K. Rai, **V.K. Reja**, and K. Varghese, Discrete Event Simulation Based Approach for Tracking Performance of Segmental Production at Precast Yard, in 40th Int. Symp. Autom. Robot. Constr. (ISARC 2023), Chennai, India (Accepted)
2. A.G. Mallya, **V.K. Reja**, and K. Varghese, Impact of Reinforcement Design on Rebar Productivity, in 40th Int. Symp. Autom. Robot. Constr. (ISARC 2023), Chennai, India (Accepted)
3. M. Jain, **V.K. Reja**, and K. Varghese, Assessment of Factors Affecting Productivity of Pilot Tube Micro-tunneling Operation through Case Study, in: Indian Lean Construction Conference (ILCC 2022), Hyderabad, India
4. A.K. Rai, **V.K. Reja**, and K. Varghese, Application of Operational Management Tools at Precast Yard, in: Indian Lean Construction Conference (ILCC 2022), Hyderabad, India
5. A.G. Mallya, **V.K. Reja**, and K. Varghese, Application of Lean Principles to Improve Rebar Productivity In Heavily Reinforced Structures, in: Indian Lean Construction Conference (ILCC 2022), Hyderabad, India
6. M. Jain, **V.K. Reja**, and K. Varghese, Exploring The Critical Factors Affecting The Productivity of Microtunneling Pipe Installation, in: 38th International No-Dig, Helsinki, Finland.
7. **V.K. Reja**, M.S. Pradeep, K. Varghese, A Systematic Classification & Evaluation of Automated Progress Monitoring Technologies in Construction, in 39th Int.

Symp. Autom. Robot. Constr. (ISARC 2021), Bogota, Columbia pp. 120-127. DOI: https://doi.org/10.22260/ISARC2022/0019

8. **V.K. Reja**, K. Varghese, Q.P. Ha, As-Built Data Acquisition for Vision-Based Construction Progress Monitoring: A Qualitative Evaluation of Factors, Proceedings of the 10th World Construction Symposium, 2022, pp. 138-149. https://doi.org/10.31705/WCS.2022.12

9. **V.K. Reja**, P. Bhadaniya, K. Varghese, Q.P. Ha, Vision-based Progress Monitoring of Building Structures using Point-Intensity Approach, in 38th Int. Symp. Autom. Robot. Constr. (ISARC 2021), Dubai, UAE, 2021, pp. 33-40. DOI: https://doi.org/10.22260/ISARC2021/0049

10. P. Bhadaniya, **V.K. Reja**, K. Varghese, Mixed Reality-Based Dataset Generation for Learning-Based Scan-to-BIM, Pattern Recognition. - Lect. Notes Comput. Sci. 12667 (2021) 389–403. https://doi.org/10.1007/978-3-030-68787-8**.**

11. **V.K. Reja**, K. Varghese, Digital Twin Applications for Construction Project Management, in: Joint Indo-Japanese Smart City Conference (IJSCS-2022)

12. V. Thelakulla, **V.K. Reja**, K. Varghese, A Web-Based GIS Tool for Progress Monitoring of Linear Construction Projects, in 38th Int. Symp. Autom. Robot. Constr. (ISARC 2021), Dubai, UAE, 2021, pp. 33-40. https://doi.org/10.22260/ISARC2021/0007

13. D. Nath, **V.K. Reja**, K. Varghese, A Critical Review of Literature on Collaboration in Construction, in: Indian Lean Construction Conference (ILCC 2021)

14. D. Nath, **V.K. Reja**, K. Varghese, A framework to measure collaboration in a construction project, in: Proc. 9th World Constr. Symp. 2021 Reshaping Constr. Strateg. Struct. Cult. Transform. Towar. 'Next Norm., The Ceylon Institute of Builders - Sri Lanka, 2021: pp. 2–13. https://doi.org/10.31705/WCS.2021.1.

15. A.K. Saxena, **V.K. Reja**, K. Varghese, IoT Enabled Framework for Real-time Management of Power-Tools at Construction Projects, in: Proc. 37th Int. Symp. Autom. Robot. Constr., 2020: pp. 992–999. https://doi.org/10.22260/ISARC2020/0137.

16. **V.K. Reja**, K. Varghese, Impact of 5G Technology on IoT Applications in Construction Project Management, in: 36th Int. Symp. Autom. Robot. Constr. (ISARC 2019), Banff, Canada, 2019: pp. 209–217. https://doi.org/10.22260/ISARC2019/0029.

ACKNOWLEDGEMENTS

I express my gratitude to all those who have supported me throughout my PhD journey.

I would like to thank Prof. Koshy Varghese, my supervisor at IIT Madras, for his valuable guidance and support throughout my PhD. His way of supervision has provided me with flexibility in the way of working, and his trust has always motivated me to come up with outcomes to match his expectations. He encouraged me and provided various opportunities for travel, collaboration, and interdisciplinary learning. He facilitated several academic and administrative tasks smoothly, making my research the core focus throughout my PhD. He constantly pushed me to strive for excellence in my area of study. His strong ethical and highly professional way of working has inspired me to follow his footsteps even after graduating from this institute and wherever I go. He gave me ample opportunities to explore other research areas to develop a wide breadth of knowledge before diving into the depth of a particular area in focus. This will help me shape my career and research direction in the long term.

Apart from technical supervision, he has always encouraged me to grow my network by discussing and allowing me to meet highly distinguished researchers in this field. Due to him, I have been given the opportunity to attend multiple international and national conferences and develop a strong academic network of highly distinguished research in this field. Words can only inadequately express my gratitude to him.

I would like to thank Prof. Quang Ha, my supervisor at UTS. It was a memorable and enjoyable experience in my Joint PhD program to have worked with him. He advised me with his strong technical expertise, particularly in machine learning. His weekly review meetings have always kept me on my toes to progress towards my objectives. He has always encouraged me to publish in high-quality journals and conferences. He was very kind to introduce me to several distinguished academics at ISARC 2022 in Bogota, Colombia, which we attended together. He allowed me to work with Larry with UAVs and learn and experience a different field of research. It was a real privilege to work with him and the research group at UTS. My time with him has helped me to 'grow' in many ways. I am grateful to him.

I respectfully acknowledge Prof. Benny Raphael, Prof. Ashwin Mahalingam, Prof. Raghunathan Rengasamy, my doctoral committee members at IIT Madras and Prof. Ricardo P. Aguilera, my co-supervisor at UTS for enriching me with constructive comments at various stages of my candidature. These comments helped me to improve my research and to think about new ideas.

I would also like to thank Prof. Nikhil Bugalia for his invaluable suggestions on my research topic. We had many friendly discussions, which helped me to plan for my PhD and beyond. He supported me as a mentor in many things, including academic and non-academic.

It has been an honour to be the inaugural recipient of the Prime Minister's Research Fellowship, and therefore I am indebted to the Ministry of Education, Government of India. With its perks and thorough review process on the institute and national levels, this prestigious scheme has motivated me to do high-quality research from India's best institute.

I would like to acknowledge the IIT Madras administration for providing me with their financial support in the form of the PMRF Scholarship and the MHRD Scholarship. I would also like to thank them for providing me with all the necessary facilities to do my research and on-campus accommodation in this lush, green, and beautiful campus.

Next, I would like to thank, Prof. Balaraman Ravindran, Head, Robert Bosch Centre for Data Science and Artificial Intelligence (RBCDSAI) for providing valuable support and guidance for the final objective of this work. I would also like to acknowledge Shreya Goyal, Post Doc. at RBCDSAI, for providing her domain expertise.

I would like to thank the Global Engagement Office at IIT Madras and the Graduate Research School at UTS for their kind cooperation and facilitation of the Joint Doctoral Degree Program. I would like to acknowledge the support of the collaboration between IIT Madras and the University of Technology Sydney, which helped me pursue a Joint Degree and made my admission possible even during the pandemic. It has been a fantastic experience working between two research groups across different departments and institutions, thanks to everyone from the BTCM, IIT Madras and SEDE UTS groups I've worked with and from whom I've learned so much.

I would like to thank the University of Technology Sydney for providing me with UTS Presidents Scholarship and Collaborative Research Degree Scholarship, which helped financially assist my research and my stay in Sydney.

I thank my fellow research collaborators and fellow groupmates at IIT Madras and UTS for the stimulating discussions and collaborative teamwork. I am grateful to everyone

I've collaborated with throughout my research, particularly Ashish, Parth, Deepjyoti, Vineela, Mohit, Ashutosh, Amith, Megha, Larry, Shreya, Saish, and Madhav.

Next, I would like to thank the participant of my survey for their feedback on the questions of importance. Their expertise and responses helped me gather data about the industry's current state.

I would like to thank Engineering Unit at IIT Madras for their permission and support in the data collection for my case study. They and the project's various stakeholders helped me with data collection by following proper safety protocols and regulations. Prof. Koshy facilitated this connection, Thanks to him.

Next, I would like to thank the CMRF group, my PhD colleagues, the IAARC@IITM team and other friends from IIT Madras for the valuable discussions and time I had with them on this journey.

I would like to thank all my family of friends (some in particular) from NIT Calicut (Shaunish, Adarsh), L&T (Sagar, Saurabh), IIT Madras (Niranjan, Anuj, Parikshit, Velmurugan, Pavan, Kaushik, Rohit, Anusha, Shahid, Srinath, Sreelakshmi, Vaishnav, Binay, Nandheesh, Rohit, Gokul), and UTS (Sid, Soumya, Biyanka, Prerna, Larry, Ignacio, and Hubert).

I thank my extended family Rohit Bhaiya, Divas, Rakhi Bua, Babli Maasi, Bobby Mama, Father-in-Law, Mother-In-Law, Sarthak, Abhay, Riya, Agrim for their comprehensive support and discussions at various stages.

Finally, my special thanks to my mother, father, wife (Jagrati), son (Vivaan), Jolly, and Anand for their support throughout my PhD candidature, especially during a hard time of the global epidemic. They have constantly encouraged me and taken over some of my responsibilities so that I focus on my research. Jagrati deserves special applause for being with me throughout the highs and lows of the whole journey and having confidence and faith in my abilities. She has been a better parent than me during this period.

Finally, a special thanks to my pets Mojo and Jojo for being their always and giving me tranquillity of mind while I was lost in the thoughts, watching them sitting or swimming quietly.

Varun Kumar Reja

Chennai, India, 2022

ABSTRACT

KEYWORDS: Computer Vision, Construction Progress Monitoring, Unsupervised Segmentation, Self-supervised Classification, Point Clouds, Deep Learning, Data acquisition, 3D reconstruction, As-built modelling, Feature Engineering.

Progress monitoring is one of the essential tasks while executing a construction project. Effective monitoring leads to accurate and timely analysis of the project's progress which is required to make vital decisions for project control. Conventional progress monitoring techniques are error-prone, time-consuming, and require human effort. Therefore, this research aims at automation of the monitoring process of construction progress through computer vision to enable effective control of projects.

The significance of research on computer vision-based construction progress monitoring is high as it provides advantages over conventional technologies in terms of reduced labour costs, improved quality and increased productivity. However, there are direct benefits of using vision-based progress monitoring which can help in better and informed project management. Firstly, computer vision can be used to automate the process of construction progress monitoring, which can lead to improved accuracy and efficiency. This is because computer vision can be used to extract data from images and videos without the need for manual intervention. This can save time and money, and it can also reduce the risk of human error. Secondly computer vision can be used to provide increased visibility into the construction process. This can be helpful for project managers and stakeholders, as it can help them to track progress and identify potential

problems early on. This can help to prevent delays and cost overruns. Finally, computer vision can be used to improve decision-making in construction projects. This is because it can provide project managers with real-time data about the progress of the project. This data can be used to make informed decisions about scheduling, resource allocation, and quality control. With this motivation, this research is focused on achieving three key objectives.

The first objective is to explore the state-of-the-art of progress monitoring in construction in the literature and in practice. The key takeaway from the first objective is establishing the need to work towards a robust, autonomous, and implementable progress monitoring technology for progress monitoring of construction projects. Computer vision is identified as an appropriate technology that fulfils all the essential requirements for monitoring projects autonomously. Therefore, the second objective aims to develop an integrated framework for Computer Vision-Based Construction Progress Monitoring (CV-CPM).

The developed three-stage framework discusses in detail the various tools, technologies, and algorithms involved in the process of CV-CPM. The element identification stage is found to be one of the key stages of the framework. Previously, researchers have experimented with automating this task using heuristics-based approaches, which require manually applying constraints and, therefore, an ample amount of complex coding and domain knowledge. Recently, learning-based approaches are being explored, which take the input as the as-built data from the projects, captured in the form of 3D point clouds. However, the existing supervised approaches require much effort in manually labelling the training data. Also, the training data cannot be reused in other projects because each construction project involves a unique set of elements, and as such, preventing them from being generalised. Therefore, there is the need for a hybrid approach for as-built modelling to overcome the individual shortcomings of the heuristics and learning-based approaches for element identification in point clouds.

Finally, the third objective aims to develop a novel hybrid self-supervised approach for element identification for CV-CPM. In this context, the proposed hybrid network using deep learning based on a contrasting approach concatenated with a set of handcrafted features can extract specific features to differentiate between various elements on a construction site. This hybrid feature vector enables the network to segment various building elements from the construction point cloud data and classify them into six object classes, i.e., wall, beam, column, door, window, and slab. The model is trained and evaluated on the S3DIS dataset with the classes relevant to construction stages. The results are evaluated using the standard metrics for precision, recall, F1-score, and overall accuracy. Finally, the developed pipeline titled 'ConPro-NET' is tested on a mid-construction dataset. The results showed that ConPro-NET achieved an overall accuracy of 80.86% on the S3DIS test dataset and 80.95% on the case study dataset. The hybrid feature model gave a significant improvement of 24.49% over deep learning features and 14.54% over handcrafted features on the S3DIS dataset.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AEC		Architecture, Engineering and Construction
AI		Artificial Intelligence
AR		Augmented Reality
BIM		Building Information Modelling
BoQ		Bill of Quantities
CAD		Computer-Aided Design
CCTV		Closed-Circuit Television
CMVS		Clustering Multi-View Stereo
CNN		Convolutional Neural Network
CV-CPM		Computer Vision-based Construction Progress Monitoring
DT		Digital Twin
DL		Deep Learning
GPS		Global Positioning System
GPU		Graphical Processing Unit
GUI		Graphical User Interface
HSI		Hue-Saturation-Intensity
ICP		Iterative Closest Point
IFC		Industry Foundation Class
IMU		Inertial Measurement Unit
IoT		Internet of Things
LiDAR		Light Detection and Ranging
LoD		Level of Detail
LPM		Level of Progress Monitoring
MEP		Mechanical, Electrical, and Plumbing
ML		Machine Learning
MR		Mixed Reality
MVS		Multi-View Stereo

PCA		Principal Component Analysis
PMVS		Patch-based Multi-View Stereo
PSR		Poisson’s Surface Reconstruction
QR		Quick Response
RANSAC		Random Sample Consensus
RFID		Radio Frequency Identification
RGBD		Red, Green Blue Depth
SfM		Structure-from-Motion
SGM		Semi-Global Matching
SLAM		Simultaneous Localisation and Mapping
SMS		Short Message Service
TLS		Terrestrial Laser Scanning
UAV		Unmanned Aerial Vehicle
UGV		Unmanned Ground Vehicle
VR		Virtual Reality
XR		Extended Reality

CHAPTER 1.

INTRODUCTION

Construction Progress monitoring is one of the key tasks in construction processes. It is one of the toughest challenges a construction manager must encounter. It is considered as a critical success factor for projects to be delivered on time and within budget and as one of the most difficult tasks due to the complexity and interdependency of activities. Accurate construction progress measurement has been shown to be critical to the success of a building project (C. Kim, Son, & Kim, 2013a). Despite project control being very important, the construction industry does not have efficient monitoring systems compared to other industries (Navon & Sacks, 2007)(K. Han, Degol, & Golparvar-Fard, 2018).

The progress of a project is represented in terms of percentage which is calculated using earned value of individual activities. The basic concept of progress monitoring involves computing actual progress and comparing it with expected progress at a particular instance and make timely decisions for corrective actions.

The chapter has been divided into five sub sections. The first sub-section is on motivation. The second sub-section is on the broad research gaps identified. The third sub-section states the objectives of the thesis. The fourth sub-section presents the overall methodology of research, and the fifth sub-section presents the thesis organization.

1.1 MOTIVATION

Construction industry interests for timely and accurate information on the progress of the construction project are increasing. Construction managers need to have a check on progress for schedule completion of the projects (Hamledari, McCabe, Davari, & Shahi, 2017). They also want information about the activities which are ahead of schedule and behind schedule to reallocate the resources. Progess monitoring also helps to reduce the risk of project failure by identifying potential problems early on, project managers can take steps to mitigate those risks. This can help to ensure that projects are completed on time and within budget.

Progress monitoring can help in enhancing data-driven decision making. Manual visual observations and traditional progress monitoring help to obtain feedback about progress measurement but are time-consuming, error-prone, and infrequent (Marianna Kopsida & Brilakis, 2020). By conducting research in this field, we aim to develop data-driven monitoring approaches that provide accurate and reliable information for decision-making processes. This involves utilizing advanced technologies such as data analytics, machine learning, and artificial intelligence to process and analyze monitoring data.

A recent article highlighted the benefits of progress monitoring in quantitative terms with the data collected from the industry (Unearth, 2023):

- Cost Savings: Effective progress monitoring can help identify delays, inefficiencies, and potential issues early in the construction process. By taking prompt corrective actions, construction companies can avoid unnecessary expenses associated with rework, schedule overruns, and material wastage. Cost savings through progress monitoring have been reported to range from 5% to 20% of the total project budget.
- Time Reduction: Timely monitoring allows project manag(Unearth, 2023)ers to identify bottlenecks, track progress, and take necessary steps to avoid delays. By optimizing resource allocation and scheduling, construction projects can minimize time overruns and potentially accelerate the overall project timeline. This can lead to earlier

project completion, reduced financing costs, improved cash flow, and earlier revenue generation.

- Improved Productivity: Progress monitoring provides valuable data on labor productivity, equipment utilization, and material flow. By analyzing this information, construction companies can identify areas for improvement and implement measures to enhance overall productivity. Improved productivity translates to better project performance and increased profitability.

However, construction progress monitoring currently is a cumbersome task and reporting and analysing of results can be very costly and resource consuming (T. Omar & Nehdi, 2016a). These numbers of savings can improve if progress monitoring is done in an efficient and automated manner. Therefore, there is a need for an automated platform that can recognize the progress in real-time, analyse and report the results (Pučko, Šuman, & Rebolj, 2018).

Many technologies like RFIDs, Barcodes, GIS, LiDAR, RGB-D Cameras, Photogrammetry, videogrammetry, laser scanners etc., have been used for extracting the data and evaluating the progress from construction sites (T. Omar & Nehdi, 2016a). Many of these technologies though researched upon have not delivered an implementable level of automation and accuracy and therefore are not being used in practice .

There is no study on level of adoption of the existing state of the art practices for progress monitoring in construction, particularly in the context of Indian construction industry. According to the MoSPI (Ministry of statistics and Program Implementation) report, in general there is a delay and cost overrun observed in the completion of many construction projects. Ineffective progress monitoring has been attributed as one of the primary reasons for these.

There is also no study on the challenges faced in the industry in context of the applied approaches of progress monitoring. Even the list of the available progress monitoring methods was last evaluated and classified in 2016 (T. Omar & Nehdi, 2016a). There have been a lot of technical advancement in these methods in the last five years, and a new classification and evaluation of technologies is required (Alizadehsalehi & Yitmen, 2016a, 2021). (Extended literature review for this is presented in Section 2.3)

Even the selection of progress monitoring technology is subjective and is no scientific approach has been developed towards this, the factors which needs to be considered while selecting a progress monitoring technology for a project are not available directly in literature.

Over the past decade or so, the nature of construction and infrastructure management is shifting towards a paradigm in which there is an increasing demand for understanding various aspects of what is happening at the field from a 3D perspective. There are many challenges in this domain which have been identified and worked upon by the researchers but still the accuracy of point cloud models, noise in the data captured, incorrect identification of structural elements remains to be the key problems faced (Z. Ma & Liu, 2018).

Images and videos for documentation and visualization are currently being used at construction projects (Dimitrov & Golparvar-Fard, 2014). It is convenient to capture images and they can be transferred in real time. The computational power changed in the last decade and the availability of mobile cameras have significantly driven the monitoring process to use a vision-based approach. The decrease in camera cost and the improvement in communication technologies have motivated this research to be based on the inputs from vision-based technologies. The decreasing cost of laser scanners is also a driver to adopt vision-based technologies for monitoring projects in future.

Therefore, the computer vision-based progress monitoring methods have been explored very recently. However, these approaches have contributed to different parts of the whole pipeline for example in data acquisition (Q. Wang, Tan, & Mei, 2020), 3D reconstruction (Z. Ma & Liu, 2018) or as-built modelling (Hyunsoo Kim & Kim, 2021). There are numerous methods, tools, techniques, and algorithms at every stage of the process, however there is no study that assembles these components in a systematic and structured manner. There is a critical need for a framework which encompasses these components and the comparison withing various options at each stage. Also, there is a requirement that these should be positioned and compared for better and informed selection. (Extended literature review for this is presented in Section 3.1 and summaries in section 3.3)

The areas of priority for research for onsite implementation of computer vision based progress monitoring have not been identified till now and there is a need to explore these with future research strategies.

Next in these computer vision-based methods, researchers have used heuristics (Kang, Patil, Kang, Koo, & Kim, 2020) as well as supervised learning-based models (Park & Cho, 2022) to estimate the progress. However, these approaches have their individual shortcomings and cannot be applied for practical implementation. Not only this, but there are also other challenges involved for practical implementation and feasibility of these approaches on site.

Some of these challenges include the lack of smooth and automated workflow for scan to BIM, use of laborious and inefficient heuristic-based approaches and use of supervised learning approaches which requires a lot of labelled data. Also the data being used for training the models is of fully constructed buildings rather than under construction buldings. (Extended literature review for this is presented in Section 5.2)

Therefore, with this motivation and background the next section summarises the specific research gaps in the three key areas that this research will focus upon.

1.2 RESEARCH GAPS

The key area wise research gaps this thesis focuses on are listed below:

A. Construction Progress Monitoring - State of the Art: Literature and Practice

1. There is currently no structured information on the type of progress monitoring technologies being used by construction companies (India and UAE).
2. There is no information on the challenges faced by the construction companies using the conventional progress monitoring technologies.
3. A systematic classification and evaluation of existing progress monitoring technologies does not exist.
4. There is no structured information about the factors and their relative importance that influence the selection of a progress monitoring technology for a project.

B. Computer Vision Based Construction Progress Monitoring (CV-CPM)

1. There are several approaches in literature focusing on CV-CPM, however they are siloed and therefore there is a need for an integrated framework for CV-CPM.
2. There are numerous tools, techniques, algorithms these methods focus on and therefore there is a requirement that these should be positioned and compared for informed selection.
3. The areas of priority for research for onsite implementation of CV-CPM have not been identified till now and there is a need to explore these with future research strategies.

C. Element Identification and Recognition for CV-CPM

1. The solutions involving Scan-to-BIM are still not matured, various studies have been conducted, still there is lack of smooth and automated workflow.

2. Currently proposed heuristics-based techniques require a lot of effort in hard coding and domain knowledge about the facility being constructed.
3. Learning-based techniques have potential to be used in the three key steps of vision-based progress monitoring but have been only sparsely explored.
4. The data being used currently is of fully constructed buildings & spaces, the constructions stage data is something different and should be generated and utilized.
5. Existing methods are based on supervised learning of features and requires a lot of labelled data that is not available for construction domain and hence cannot be implemented easily.

1.3 THESIS OBJECTIVES

Within the scope of this study, the following objectives are expected to be achieved for mentioned challenges:

1. To evaluate the state of the art of progress monitoring in construction.
    a. To evaluate existing solutions for progress monitoring using literature, actual site-data and challenges that are faced.
    b. To provide a systematic classification and evaluation of the available progress monitoring technologies.
    c. To list out the factors that affect selection of progress monitoring technology for a project.
2. To develop an integrated framework for vision-based progress monitoring in construction.
    a. Develop an integrated framework that captures the process requirements of construction progress monitoring and enables the characterization and categorization of current and future work in the area.

b. Utilize the framework to position and compare various concepts and tools adopted by published research studies.

c. Identify areas and strategies for future work in the area.

3. To develop a pipeline using hybrid self-supervised approach for automatic capture of constructed elements from point clouds and using it for progress monitoring in construction.

   a. To study the combined used of heuristic and learning based approaches (hybrid approach).

   b. To develop a customised approach for object segmentation specifically for construction

   c. To utilize concepts from feature engineering to and improve the performance of the method.

1.4 THESIS SCOPE

The thesis begins evaluation and classification of various available technologies for achieving objective 1. However, the scope for objective 2 and objective 3 is limited to 3D computer vision technologies for construction progress monitoring.

1.5 RESEARCH METHODOLOGY

The following Figure 1.1 shows the overall methodology adopted for this research. The identification of this research area (Progress Monitoring in Construction) was done through the practical experience from industry, current on-going trend in the literature (Patel, Guo, & Zou, 2021) as well as the estimated impact on the industry (Unearth, 2023). The state of the art of the technologies used in construction industries were documented and were found to be in-effective.

In the next step, challenges pertaining to progress monitoring in the industry were identified and then a preliminary review of literature was conducted to check the potential solutions available. It was observed that there is a considerable gap in research and implementation due to various technological and economic reasons. With this as motivation, gaps were identified and objectives aiming towards these gaps were set. The scope of the study was also constrained to solve it in the given timeframe. In final step, the work towards each objective was conducted and the findings were used as an input for solving the next research gap. The first two objectives adopted a qualitative methodology, the third objective adopted an iterative design methodology followed by a quantitative method for testing.

It should be noted that, this research contains both qualitative and quantitative methodologies, based on the specific gap to be solved at each objective. Therefore, a mixed methodology is used to integrate qualitative and quantitative approach together.

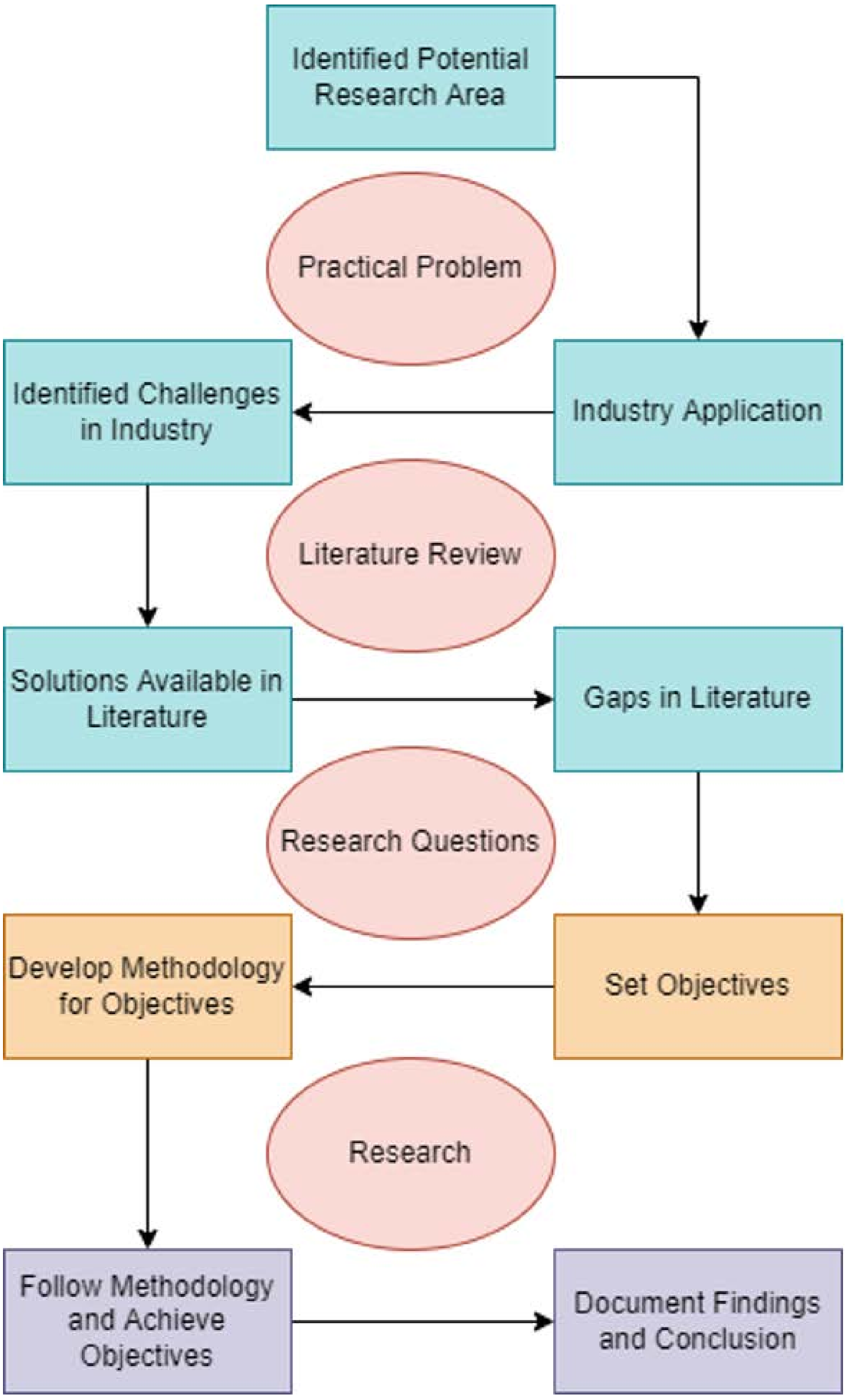


Figure 1.1 Overall methodology for research.

An overall methodology for each objective set is shown below in Figure 1.2. These will be discussed in the separate chapters for each objective. The red outline boxes show the contribution towards fulfilling each objective.

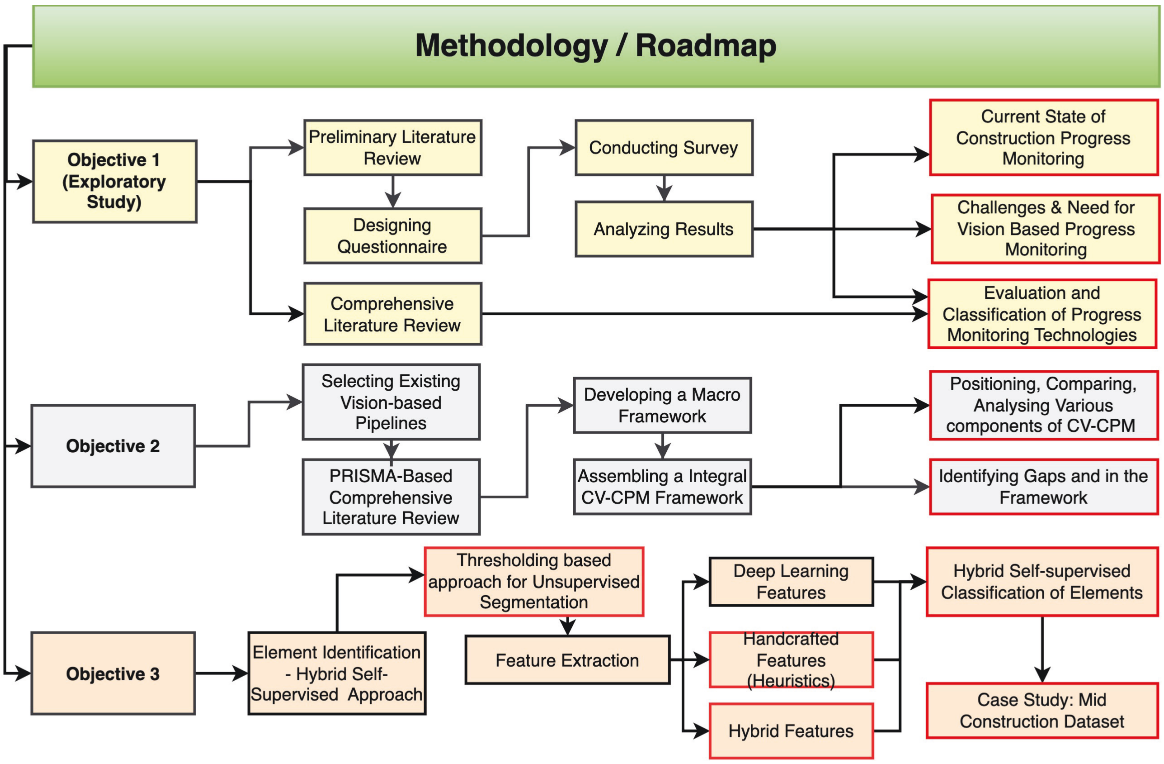


Figure 1.2 Consolidated methodology for each research objective.

1.6 THESIS ORGANISATION

This thesis is structured into seven comprehensive chapters. The first chapter serves as the 'Introduction,' laying the foundation for the research by presenting the background information, discussing the motivation behind formulating the problem, identifying research gaps, and outlining the objectives to be achieved.

The second chapter is dedicated to the evaluation and classification of progress monitoring technologies in the construction industry. This chapter primarily focuses on accomplishing objective 1, providing an in-depth analysis and assessment of existing technologies used for monitoring construction progress.

Moving forward, the third chapter undertakes a detailed literature review, specifically exploring computer vision-based approaches for construction progress monitoring. This chapter synthesizes and examines the existing body of knowledge in the field, shedding light on the various methodologies and techniques employed.

In the fourth chapter, a comprehensive and elaborate 'Integrated Framework for Computer Vision-Based Construction Progress Monitoring (CV-CPM)' is introduced. This framework is meticulously defined, outlining its components and functionalities, and its development centers around accomplishing objective 2.

The fifth chapter delves into the 'Hybrid Self-Supervised Approach for CV-CPM,' a novel methodology specifically designed to enhance the computer vision-based progress monitoring system. Objective 3 is the main focus of this chapter, as it explores the development and implementation of the hybrid self-supervised approach to further improve the accuracy and efficiency of the CV-CPM system.

Chapter six presents a captivating case study, illustrating the practical application and real-world results of the developed CV-CPM system on an under-construction project. This chapter provides valuable insights into the system's performance and effectiveness in a real-life scenario.

Finally, the seventh chapter concludes the thesis, summarizing the key findings, highlighting the contributions made by the research, discussing any limitations or challenges encountered, and offering recommendations for future studies and advancements in the field of computer vision-based construction progress monitoring.

By organizing the thesis into these seven chapters, the research aims to provide a comprehensive and systematic exploration of the topic, presenting a clear progression from problem formulation and literature review to the development and evaluation of innovative approaches for construction progress monitoring.

CHAPTER 2.

IDENTIFICATION, CLASSIFICATION AND EVALUATION OF PROGRESS MONITORING TECHNOLOGIES IN CONSTRUCTION

In the previous chapter, the motivation, various research gaps and objectives were presented. This chapter is focused on the objective 1 of this thesis. This chapter have eight sub-sections. The first sub-section is on Introduction. The second sub-section is on the questionnaire survey-based field study. The third sub-section states the literature review method. The fourth sub-section presents the classification of automated progress monitoring technologies. The fifth sub-section presents the factors affecting progress monitoring technology selection. The sixth sub-section presents the result of a questionnaire survey for identifying the relative importance index (RII) of the various factors. The seventh sub-section is discussion and followed by the final sub-section on conclusions.

2.1 INTRODUCTION: PROGRESS MONITORING IN CONSTRUCTION

Progress monitoring is one of the essential tasks while executing a construction project. Effective monitoring will lead to an accurate and timely analysis of the project's progress which is required to make vital decisions for project control (Ekanayake, Wong, Fini, & Smith, 2021b). On the other hand, inefficient and delayed updates regarding the project's progress, estimated by comparing the as-built status with the as-planned status, will lead to time and cost overruns (Pushkar, 2018). Automated progress monitoring techniques are preferred over the conventional manual data entry method as the latter is time-consuming and complex, especially if the project scope is vast. Numerous tools and technologies are being used for progress monitoring of construction projects (Alizadehsalehi & Yitmen, 2019; T. Omar & Nehdi, 2016a). Therefore, it is necessary to systematically classify and evaluate them based on their advantages and limitations for successful and appropriate implementation. Hence, this chapter identifies several progress monitoring techniques and evaluates them by highlighting their advantages and limitations. Several factors affecting the selection of these technologies for implementation have also been identified and ranked.

2.1.1 Research Gaps and Sub-objectives

A project life cycle in a construction industry involves several stages, like designing, planning, scheduling, execution, monitoring, controlling, and demolition. Monitoring and control to minimize time and cost overruns are crucial for a construction project. Accurate and real-time progress monitoring is essential for achieving project objectives with expected KPIs.

Progress monitoring also plays a vital role in avoiding unexpected circumstances and eliminating disputes and legal challenges among the stakeholders. Automating various tasks in monitoring and controlling will reduce the complexity involved in manual documentation and calculations in a project to a considerable extent. Hence, a prompt and feasible automated progress monitoring technology is essential in the present-day construction sector (M. Kopsida, Brilakis, & Vela, 2015) (T. Omar & Nehdi, 2016b).

Automation in progress monitoring has evolved over the past two decades, with several technologies with varying levels of automation used in projects. With several technologies available, there is not enough clarity on the type of technology appropriate for a specific case or project.

Existing literature has focused on the specific technology of progress monitoring, for example, specifically, vision-based (Ekanayake et al., 2021b) or tag-based (P. R., Raphael, & Vaidyanathan, 2019). For a robust implementation, firstly, there is a critical need to identify and classify these technologies and, secondly, to evaluate them based on their advantages and limitations. Therefore, two sub-objectives towards objective 1 are to:

1. Identify, classify, and evaluate technologies available for progress monitoring of construction projects.

2. To list factors that enable appropriate technology selection for the project-specific use case.

This chapter contributes to objective 1 of this thesis. The overall methodology followed for objective 1 has been shown in Figure 1.2

2.2 RESEARCH METHODOLOGY – OBJECTIVE 1

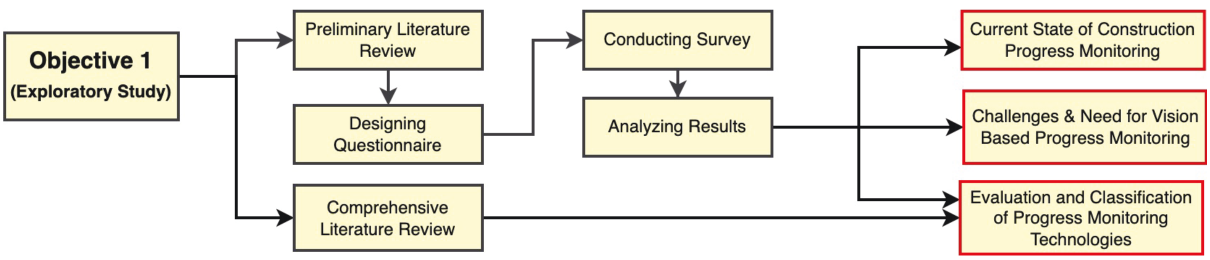


Figure 2.1 Research Methodology for Objective 1

The research method for achieving objective 1 is shown in Figure 2.1. First, a preliminary review of literature was conducted by searching various papers on progress monitoring technologies. These papers were collected from Scopus and Web of Science databases using a keyword search-based method, followed by snowballing technique. A total of 61 papers with 49 journal articles and 12 conference papers were identified from the databases, and an exhaustive review with analysis was performed.

These papers were analysed to gent preliminary information about the state of the art of progress monitoring from a literature standpoint. With this information and authors practical on-site experience, a preliminary questionnaire survey was designed to evaluate the state of the art of progress monitoring from industry standpoint. Three key information were to be derived from the survey; these were:

1. Which technologies are being practically used for progress monitoring and what is the frequency of their use at various projects?
2. What are the challenge the task force is facing by using these technologies?
3. What are the other mobile or desktop applications being used for storing the progress data?

Once the questionnaire was designed for getting these answers, a survey was conducted with participants who were working or having experience in executing at least one construction project in the last 2-3 years. Next, they obtained data was analysed and the findings were presented in form of visual and statistical graphs and charts.

The information helped the author to have the holistic view of progress monitoring from both the literature standpoint and implementation standpoint. Hence the evaluation of technologies was performed. Also, the classification of technologies was done by segregating the technologies based on the type of science (fundamental technology) they use for progress estimation.

2.3 QUESTIONNAIRE SURVEY-BASED FIELD STUDY

The data was collected through a questionnaire survey to study the progress monitoring technologies currently used in projects. This section describes the participant demographics, the survey's results, and conclusions. The questionnaire is attached in appendix A for reference.

2.3.1 Demographics of participants

The survey was sent to about 200 participants working or having experience in executing at least one construction project in the last 2-3 years. The participants from the survey were part of one of the following organizations.

1. L&T Construction
2. Navnirman Builders
3. Bara architects
4. URC Construction (P) Ltd
5. APCRDA
6. NCBS-TIFR
7. L&W construction
8. Total Environment Building Systems Pvt Ltd, Bengaluru
9. L&T Oman LLC
10. CDM Smith
11. NBCC INDIA LTD

2.3.2 Survey Results and Discussion

Sixty-eight responses were received for the survey. That account for 34% response rate. Distributions of cost of the projects in INR Crores where the survey participants have worked in the last 2 to 3 years is shown in the following Figure 2.2and Figure 2.3. It can be seen from the distribution in Figure 2.3, that the mix of projects by varying project cost i.e., helped to capture a decent diversity of the Indian and Arabian markets.

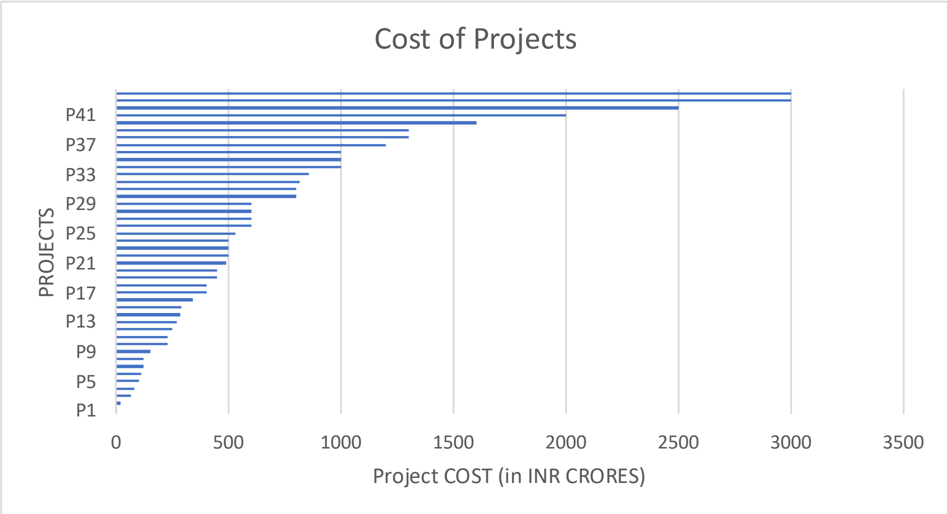


Figure 2.2 Distribution of individual project cost

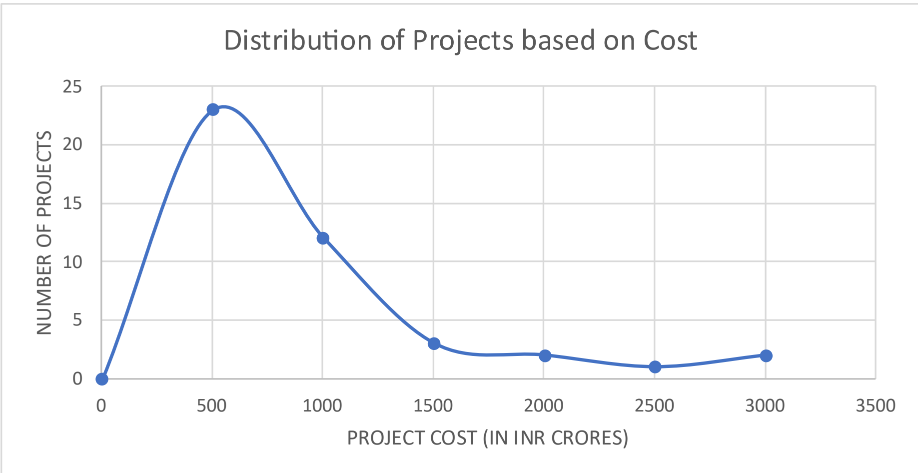


Figure 2.3 Distribution of the number of projects in the cost bracket

The average cost of the 50 projects reported was found as INR 715 Crores. The project's cost ranged from INR 6 Crores to INR 3000 crores which helped to capture the technologies used from very small projects to mega projects.

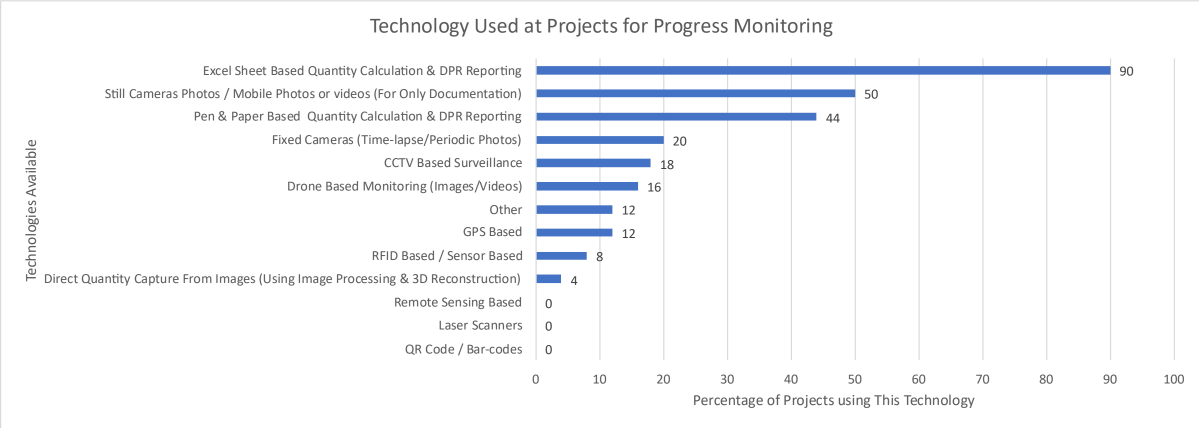


Figure 2.4 Progress monitoring technology being used as reported by survey participants on different projects.

Figure 2.4 shows the various progress monitoring technology being used as reported by survey participants on different projects. Manual entry on Excel Sheets and calculation of quantities by pen and paper-based methods are the major methods of progress monitoring used in Indian construction sites. It is also observed that 50% of sites use still cameras for documentation of site photographs and videos. This documented information is presented in site meetings for visual understanding of the project progress. A very small number of sites (<4%) reported usage of image processing or 3D reconstruction techniques for quantity calculation from images.

The use of laser scanners and remote sensing methods is not observed in the 50 projects (including the high-cost projects also), there can be two explanations for this as per this survey, first is because of the high expertise required to operate a laser scanner, second is that many scans from different locations are required to get an accurate point cloud which makes it time consuming.

Because this was a preliminary study, an option of mentioning a technology was also provided if it was not listed in the survey. The other technologies reported for capturing progress were:

1. Android Mobile applications
2. Total Station (For Stock and Earthwork Qty. estimation)
3. Smart Glasses (visualization)
4. BIM (reporting and tracking progress)

The following Figure 2.5 shows the tools that have been reported being used for progress monitoring at various projects.

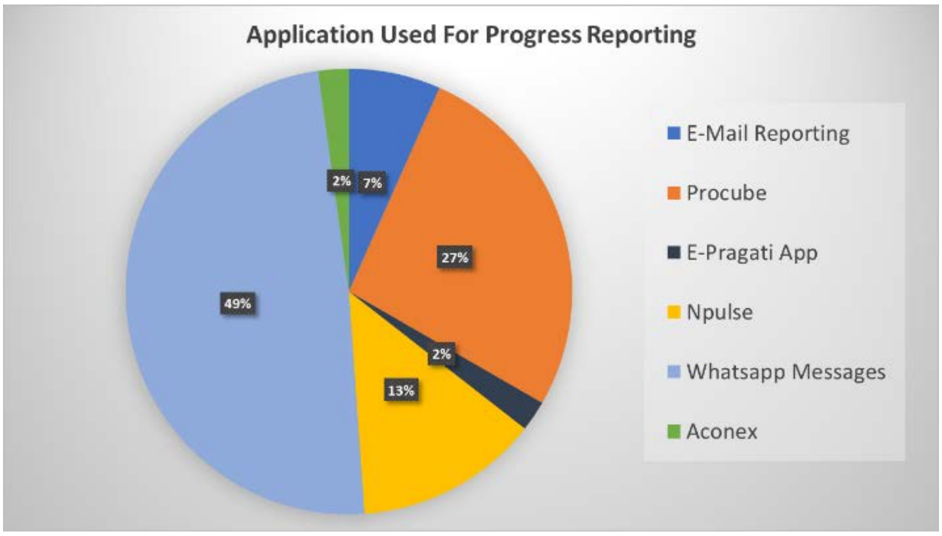


Figure 2.5 Media Application used for Progress Reporting

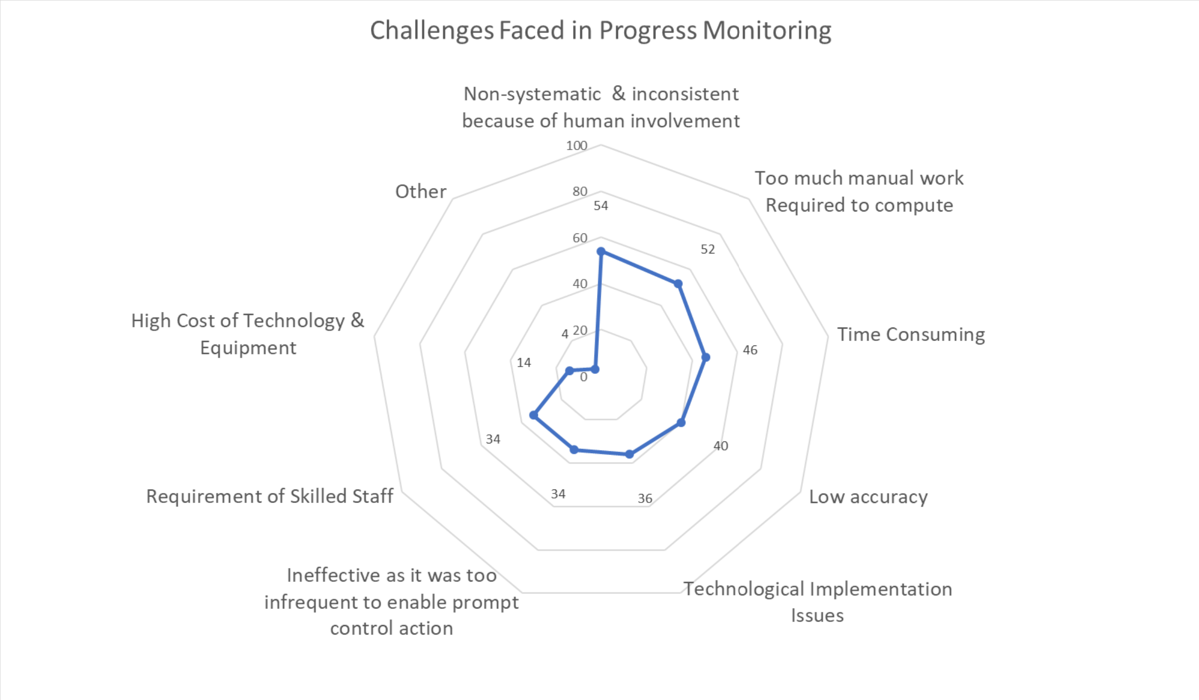


Figure 2.6 Challenges of the used technologies in progress monitoring on construction sites

The prominent challenges as shown in Figure 2.6 are:

1. Non-systematic & inconsistent because of human involvement (reported at 54% projects)
2. Too much manual work required to compute (reported at 52% projects)
3. The used method is time consuming (reported at 46% projects)
4. Low accuracy of data (reported at 40% projects)

The other challenges faced in progress monitoring reported were:-

1. unwillingness of the engineers to use applications developed (reluctance to change)
2. network connectivity issues.
3. alignment of corresponding clusters and segments towards digital solutions developed in project management.

The technology adoption at individual level was highlighted as one of the points which impacts the adoption of any pilot technology being introduced in the projects, interestingly these adoption practices should be looked into so as to implement the technologies being developed.

2.3.3 An overview of reporting of progress on major projects across India

From Table 2.1 it is evident that even on large value projects the pen and paper based and excel sheet-based methods supported by manual inspection are the most common ways of evaluating project progress. CCTV based monitoring; RFID have been used in some projects.

The projects which are deploying video cameras and drones e.g., Project 2 and 3, are using this data for only documentation purpose. This also indicates that, the high initial cost of technology at selection stage is not a concern and constrain for these projects and organisations.

Table 2.1 Progress monitoring technology reported on high value projects.

Sl. No.	Project Name (Anonymized)	Project Cost (INR Crores)	Organization (Anonymized)	Technology Used for Progress Monitoring
1	Project 1	3000	Organization 1	Excel Sheet Based Quantity Calculation & DPR Reporting
2	Project 2	3000	Organization 2	Excel Sheet Based Quantity Calculation & DPR Reporting
				RFID Based / Sensor Based
				CCTV Based Surveillance
				Drone Based Monitoring (Images/Videos)
				GPS Based
				Still Cameras Photos / Mobile Photos or videos (For Only Documentation)
3	Project 3	2500	Organization 3	Pen & Paper Based Quantity Calculation & DPR Reporting
				Excel Sheet Based Quantity Calculation & DPR Reporting
				CCTV Based Surveillance
				Fixed Cameras (Time-lapse/Periodic Photos)
				Still Cameras Photos / Mobile Photos or videos (For Only Documentation)
				Smart glasses
4	Project 4	1600	Organization 4	Pen & Paper Based Quantity Calculation & DPR Reporting
				Excel Sheet Based Quantity Calculation & DPR Reporting

				Still Cameras Photos / Mobile Photos or videos (For Only Documentation)
5	Project 5	1300	Organization 5	Pen & Paper Based Quantity Calculation & DPR Reporting
				Excel Sheet Based Quantity Calculation & DPR Reporting

2.4 LITERATURE REVIEW – PROGRESS MONITORING TECHNOLOGIES

The reference literature for the review was collected from Scopus and Web of Science databases using a keyword search-based method, followed by snowballing technique. Initially a gross total of 120 articles were obtained. Then the articles were shortlisted based on their abstracts which were focusing on progress monitoring technologies, with a total database of 53 articles. Next eight additional articles were added for completeness for covering the newer technologies (Extended Reality) which rarely covered by the selected papers. Finally a total of 61 papers with 49 journal articles and 12 conference papers were identified from the databases, and an exhaustive review with analysis was performed. The chronological distribution of the selected papers is shown in Figure 2.7.

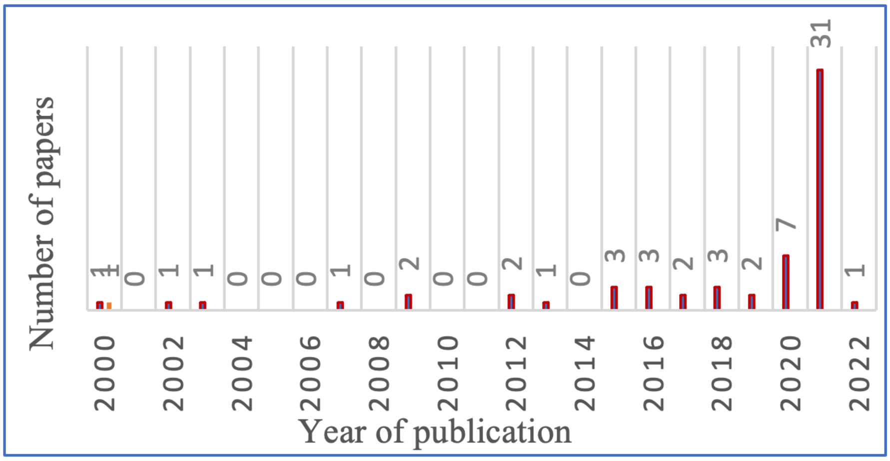


Figure 2.7 Chronological distribution of selected papers

The search attributes used in the review with the keywords used and search scope is as shown in

Table 2.2. The relevant articles for the construction domain were filtered after reading the abstracts. The filtered articles were considered for meta-analysis.

Table 2.2 Search attributes

Search attributes	Values used in the search
Databases	Web of Science, Scopus
Language	English
Year of Publication	2000-2022
Type of the document	Journal articles, Conference papers
Keywords	Progress monitoring technologies, Automated progress monitoring.

2.5 CLASSIFICATION OF AUTOMATED PROGRESS MONITORING TECHNOLOGIES

The selected papers contained various automated techniques' case studies, challenges, and benefits. Based on the meta-analysis, seventeen state-of-the-art progress monitoring techniques were identified.

Each of these technologies has unique advantages and limitations in construction site monitoring. Hence, there is a need for a detailed analysis to identify these so that they can be used effectively. A detailed and systematic review of the above-mentioned technologies is presented in Table 2.3, along with their advantages and limitations with the relevant references.

As shown in Table 2.3, these techniques were classified into six major categories as enhanced Information Technologies (IT), tag-based methods, geospatial technologies, building information modelling and associated commercial software, computer vision-based approaches, and extended reality and are discussed below:

1. Enhanced IT: These include handheld computing devices (Personal Digital Assistants or PDAs, handheld personal computers), Interactive Voice Response or IVR, multimedia tools, and e-mails. These are the most basic techniques, which are information technology-based communication tools. These technologies are primarily of lower cost with a limited level of automation but can increase the chances of communication between stakeholders, thereby helping site information tracking (T. Omar & Nehdi, 2016b).

2. Tag-based techniques: These involve tags and codes that can be attached to various resources on-site and are primarily used for material tracking and inventory, employee badge scanning, and equipment tracking. These include barcodes, quick-response or QR codes, radio-frequency identification, or RFID tags, and ultra-wideband or UWB tags. Each tag's working principle is based on Automatic Identification and Data Capture (AIDC). It must be noted that a tag-based technology cannot directly extract spatial element information, visually represent the site changes, and collaborate with other vision-based techniques (T. Omar & Nehdi, 2016b)(Guven & Ergen, 2021)(Alaloul, Qureshi, Musarat, & Saad, 2021).

3. Geospatial techniques: These include fundamental technologies based on location-based sensors like Geographic Information System (GIS) and Global Positioning System (GPS). These techniques are used for geo-referenced data capture, analysis, and modelling. GIS can be used in large infrastructure projects where there is a need to store and handle vast amounts of data. It can be a useful geospatial tool, which uses location as the primary focus in database management, whereas GPS aids in the spatial analysis and navigation of different activities on the site.

4. Building Information Modelling Based: BIM is a process involving different tools, technologies, and contracts, which aids in better visualization of construction sites for accurate project management. BIM also aids in stakeholder management practices of the construction industry for different aspects of communication, collaboration, engagement, and satisfaction.

This can be used along with commercial scheduling software like MS Project so that progress monitoring can be done efficiently (Deng, Gan, Das, Cheng, & Anumba, 2019).

5. Computer vision Based: It is an emerging field focusing on information retrieval through visual inputs. These inputs include digital images, videos, thermal images, as-built point clouds, panoramas, and photospheres. These techniques involve fixed surveillance, photogrammetry, videogrammetry, range imaging, and 3-D laser scanning. Computer vision sub-domains include learning, 3D scene modelling, video tracking, 3D pose estimation, object recognition, scene reconstruction, object detection, and event detection, which can be used for progress monitoring (Ekanayake, Wong, Fini, & Smith, 2021a) (Paneru & Jeelani, 2021b).

Table 2.3 Detailed review of the progress monitoring technologies

No		Technology	Advantages	Limitations	Ref
1.	**Enhanced IT**	**Hand-held computing**	• Small-sized, portable, handy, flexible devices with several features, that can be integrated with other technologies.	• Some devices are costly, and suitable applications need to be developed so that integration can be made more efficient.	(T. Omar & Nehdi, 2016b)
		Interactive Voice Response or IVR	• Efficient and quick means of sending information from sites.	• Manual errors might occur while responding to multiple choices. • Difficulty in retracing the already answered correct messages.	
		Multimedia	• Flexible tool to aid remote progress monitoring by safe documentation and visualization of project information.	• Manual site data capture results in errors.	
		Electronic Mails or E-Mails	• E-procurement tool for quality supply chain management. • Attached with images, documents, videos, forms, etc, where the site personnel respond with the easy retracing of answered questions.	• Improper internet connection, and difficulty in responding from small devices might cause a delay in information exchange.	
2.	**Tag based**	**Bar-codes**	• Cost-effective, accurate, easy to use, portable, and flexible. • No need for an external device to read the codes.	• Direct line of sight required for data capture, time taking in item tracking. • Labels can get destroyed or lost due to adverse weather conditions.	(T. Omar & Nehdi, 2016b)(Keyvanfar, Shafaghat, & Awanghamat, 2021)
		Quick Response or QR codes	• Portable and flexible technology with a better storage capacity. • QR code reading applications can be installed on devices easily. • Lightweight wireless QR code pocket printers can be used in sites.	• Might get affected by harsh environmental conditions. • More effective in indoor tracking compared to outdoors. • Monitoring of reinforcement, concreting, and those which are not easily accessible are difficult to be monitored using QR codes.	(Z. Wang et al., 2021a)

No	Technology		Advantages	Limitations	Ref
		Radio Frequency Identification or RFID tags	• Use radio waves that can be read accurately outside the line of sight as well, without direct contact compared to light waves. • Reliable, portable, flexible, reusable, technique that can withstand harsh environmental site conditions. • Supports indoor tracking of materials, facility and building component management, and information flow in large projects. • Capable of identifying individual items and can read multiple items in an instant simultaneously. • Storage capacity is higher than that of a barcode and is unaffected by differences in illumination.	• Costlier than barcodes. • If there are metals or liquids or moisture in the nearby area, the results can be erroneous. • Using RFID can be time-consuming, and costly if a single tag is used to track each one among several materials and equipment. • Limitation of the battery operation time, and there is insufficient accuracy in location identification if it is not depending on a fixed network. • There are insufficient international standards, multi-protocol tags and readers, and also a concern on the return of investment.	(M. Kopsida et al., 2015)(Guven & Ergen, 2021)(Alaloul et al., 2021)(Alizadehsalehi & Yitmen, 2016b)
		Ultra-wide band or UWB tags	• More accurate than RFID with strong signals even in obstructions. • Provide real-time resource tracking, 3-D coordinates for position sensing and consume low energy.	• UWB tags are not cost-effective compared to RFID. • No daily necessity embedded tool or mini device.	(Alaloul et al., 2021)
3	**Geospatial**	**Geographic Information System or GIS**	• Optimal location for construction equipment can be found, with efficient capture, storage, and analysis of geo-referenced information with minimum redundancy. • Creation of geographical maps of high quality, by visually representing the construction schedule to monitor the plant and equipment, that can be provided to the clients. • Can be used as a forecasting tool for early identification of time and space conflicts, and better safety regarding worksite considerations.	• Difficult to use in indoors.	(Cheng & Chen, 2002)
		Global Positioning System or GPS	• Accurate location of positions while material tracking in construction supply chain management can be facilitated by using the Global Positioning System or GPS.	• Difficult to use in indoors. • Tagging several construction elements using GPS tags is very expensive.	(Sbiti, Beddiar, Beladjine, Perrault, & Mazari, 2021)(Xue, Hou, & Zeng, 2021)

No	Technology		Advantages	Limitations	Ref
4.	BIM and commercial software		• Visualization- clash detection & information management. • Schedule updates- automatic quantity take off & cost-estimation. • Enhanced collaboration & information exchange among stakeholders • Integrated with supply chain for product design & material delivery. • Integration with LPS or Last Planner System for lean construction. • Knowledge based systems that are active, along with various simulations in BIM contribute to data analysis. • Choosing inputs using a hybrid video and laser scans, as-built BIM can be generated automatically.	• Limited in monitoring, scheduling, and decommissioning phases. • Errors resulting from the manual navigation of BIM model and need for constant automated updates, especially for fast-tracked projects. • Limited interoperability, even with other data acquisition techniques. • Commercial software like MS Project, Primavera, etc cannot provide digital drawings and visualisation for construction.	(M. Kopsida et al., 2015)(Mutis, Joshi, & Singh, 2021)
5.	Computer Vision	**Fixed surveillance (crane cameras, closed circuit television cameras, etc)**	• Safe, cost-effective, fully automated techniques, with low labour requirements that can be deployed in multi-building as-built point cloud extraction in conjunction with BIM. • Inefficient site coverage should be resolved by the deployment of more crane cameras. • Can be used for documenting daily or weekly progress in the construction site. • The data acquired from the CCTV images can be directly transferred to the head office from the site through the internet.	• Crane camera images may result in noisy as-built point clouds and may get affected by heavy winds. Mounting the camera may require extra effort, and there is limited flexibility due to the motion range of the cranes. • 3-D point clouds may be fragmented if there is incomplete site coverage. • As the position of cameras is fixed, there is a limitation in the application of CCTV cameras in huge projects. There is a need to arrange several cameras, and the data clashes between different cameras have to be fixed.	(M. Kopsida et al., 2015)(Guven & Ergen, 2021)(Alaloul et al., 2021)(Alizadehsalehi & Yitmen, 2016b)(Xue et al., 2021)(Masood, Aikala, Seppänen, & Singh, 2020)
	Comp	**Photogrammet ry**	• Automatic identification of objects using cameras and image processing algorithms by integration with n-D BIM.	• There is a limitation due to difference in lighting conditions, which may affect the resolution. Thermal	

No	Technology	Advantages	Limitations	Ref
		• Lesser equipment cost and technical requirement, along with portability for the image capturing devices and improved flexibility. • High resolution compared to satellite imaging, for representing the geometric attributes and high texture representation. • The recorded images can be analysed using software packages with computer vision techniques and machine learning algorithms for automatic updates with as-built -3D models by reconstruction.	images along with wireless sensors and BIM can be used to overcome this problem. • Object edge detection might not be proper, occlusions, noisy images and presence of shadows will affect the accuracy of progress estimation. • The location from which the photos are taken has to be matched with the checkpoints in the drawings, which can be a difficult problem.	(Xue et al., 2021)
	Videogrammetry	• Can be used for both indoor and outdoor progress tracking. • Moving equipment can be tracked.	• Less accurate than laser scanning and photogrammetry, may get affected due to occlusions.	(Alaloul et al., 2021)(Mutis et al., 2021)
	Range or depth cameras	• Easier as-built point cloud generation directly, as it contains depth information. • Higher resolution compared to normal digital cameras, and higher portability with lower technical pre-requisite.	• Cost is lower than that for laser scanners but will be generally higher than that of normal digital cameras. • Range of shooting is limited and is mostly used in automated indoor construction progress monitoring.	(Xue et al., 2021)
	3-D Laser scanners	• High-resolution, precise, and accurate progress monitoring technology unaffected by illumination, and is used for quality control, structural health monitoring, condition assessment of structures, and tracking of components, along with active collaboration between the stakeholder teams. • Automatic comparison being done between as-built and as-planned point clouds so that progress deviation detection becomes easier, and the schedule is updated accordingly.	• Highly expensive equipment with low portability, limited texture information, time-consuming data acquisition, mixed pixel restoration, need for sensor calibrations regularly, greater warm-up time. • Operation requires larger technical knowledge, and might not be suitable for progress monitoring continuously. • Accuracy of the data acquisition using laser scanning might be affected due to occlusions and in the site.	(Guven & Ergen, 2021)(Alaloul et al., 2021)(Alizadehsalehi & Yitmen, 2016b)(Xue et al., 2021)(Masood et al., 2020)(Arif & Khan, 2021a)

No	Technology	Advantages	Limitations	Ref
6.	**Extended reality**	• Enables accurate visualization of the construction site from various angles. • Worksite planning in construction, visualization of equipment operation for inspection, comparison between as-planned and as-built images can be done. • They can be used easily in both interior and exterior locations, under different construction phases, and is cheap with the requirement of minimal training and set-up time.	• Automation quality depends on the technology in the device used. • Stationary methods are limited in portability, have less cost-effectiveness, and need additional time for setting up the equipment when compared to mobile methods. • Installation of fiducial markers requires additional investment in time and cost.	(M. Kopsida et al., 2015)(Sbiti et al., 2021)(Xue et al., 2021)(Ali, Lee, Lee, & Park, 2021)

6. Extended Reality Based: This relatively newer technology allows a combined real and virtual environment, supporting human-machine interactions. These techniques can be further classified into augmented reality (AR), virtual reality (VR), and mixed reality (MR), based on the difference in visualization. These techniques can be employed for the collection of digital data and can handle computing and network technologies in progress monitoring.

2.6 FACTORS AFFECTING PROGRESS MONITORING TECHNOLOGY SELECTION

As shown in Table 2.3, each technology is characterized by its advantages and limitations. Apart from these, key factors should be considered before choosing the appropriate technology for progress monitoring in a construction project. These parameters have been identified by several authors through their research (Alizadehsalehi & Yitmen, 2019; Ekanayake et al., 2021b; Hannan Qureshi et al., 2022; Ibrahimkhil, Shen, & Barati, 2021; M. Kopsida et al., 2015). However, three factors i.e. project type, statutory requirements and operating range are added based on the authors industry experience and the learnings form the preliminary questionnaire survey.

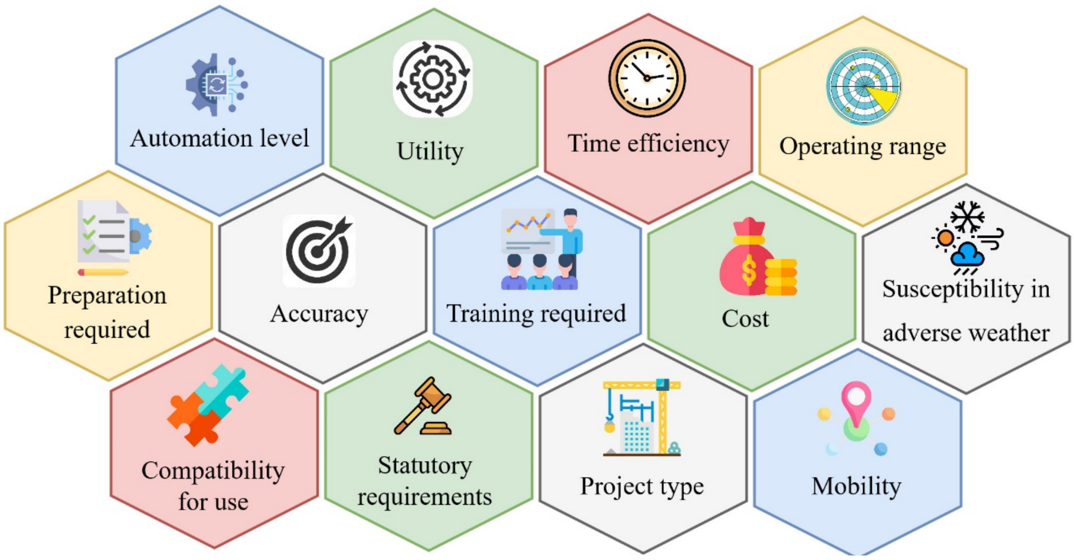


Figure 2.8 Factors affecting selection of technology.

Figure 2.8 shows the key factors to be considered while selecting the appropriate technology for progress monitoring. The context of these factors for selection is described as follows:

1. **Level of Automation:** The extent of manual/computer control while using the technique..
2. **Time efficiency:** The speed of data acquisitions as well as data processing.
3. **Operating range:** The distance up to which the employed technology works.
4. **Utility:** Adaptability of the technology used in interior and exterior construction progress monitoring. In other words, whether the technology is a general case solution.
5. **Preparation required:** The level of preparation required while setting up the equipment or process at the deployment stage.
6. **Accuracy:** The reliability of the collected data along with precision.
7. **Training required:** The amount of training or knowledge a user requires before using a particular technique.
8. **Cost:** The amount of financial and computational costs incurred to adopt and implement the technology.
9. **Susceptibility in adverse weather:** The extent of use of the technology in harsh environmental conditions like low visibility
10. **Compatibility for use :** The level by which a particular technology can be integrated with other technologies or the existing Enterprise Resource Management (ERP) system.
11. **Statutory requirements:** The legal codes and procedures to be followed while using a technique enforced by the authorities.
12. **Mobility:** The ease, flexibility, and portability of the related equipment.
13. **Project type & characteristics**: The type of the project and characteristics of a particular construction project where the technology can be used.

Note: The difference between utility and Project Type and Characteristics is highlited below for better clarity:

Utility - Adaptability of the Technology: Utility refers to the value or benefit derived from using a particular technology for a specific purpose. In this case, the utility is focused on the adaptability of the technology used in interior and exterior construction progress monitoring. It assesses whether the technology can effectively and efficiently fulfill the monitoring requirements in both interior and exterior construction settings.

The utility of the technology would depend on factors such as its compatibility with different construction environments, its ability to capture and analyze relevant data, its ease of integration with existing systems or processes, and its overall effectiveness in facilitating progress monitoring activities. A technology with high adaptability would be versatile and flexible enough to be applied in various construction scenarios.

Project Type & Characteristics: Project type and characteristics refer to the specific nature and attributes of a construction project. Different construction projects can vary significantly in terms of their scale, complexity, scope, timeline, and specific requirements. The project type and characteristics provide context to understand the unique aspects of a particular construction endeavor.

When considering the application of technology in construction projects, understanding the project type and characteristics is essential. This includes factors such as the type of construction (residential, commercial, industrial, etc.), the size of the project, the specific construction methods or materials used, the presence of any regulatory or environmental considerations, and other project-specific factors.

2.7 RANKING OF FACTORS BASED ON A QUESTIONNAIRE SURVEY

Currently, there is no scientific method followed on projects site to determine the progress monitoring technology to be used. The current method of deciding that is based on the experience of the technical staff employed for the project. Therefore, an attempt in the direction to make this decision objective and scientific has been made in this research.

An indicative survey from the experts from various companies and project sites has been done to give these factors a relative importance score (Appendix B). However, this is just an initial step towards this study and future in-depth research is require validating these factors and detailed framework development.

To evaluate the relative importance of these factors, an indicative questionnaire survey with factors description and factors definition was designed and sent to 40 experienced professionals from the construction industry. The participants were postgraduate engineers working on different construction projects across India in several organisations with at least two years of field experience. The survey participants were selected so that each participant had relevant experience and knowledge about the factors that are considered while selection of progress monitoring technologies on construction projects. The definitions of these factors were also included in the survey for uniform understanding. Each factor was given a rating of high (3), medium (2) or low (1) for the importance they hold in the selection process. Finally, 33 responses were received from the participants of the survey.

Table 2.4 RII Scores for various factors computed from the questionnaire survey.

Sl. No.	Factors	RII
1	Project Type and Characteristics	0.93
2	Cost	0.89
3	Time efficiency	0.88
4	Level of Automation	0.80
5	Statutory requirements	0.75
6	Operating Range	0.73
7	Training Required	0.67
8	Preparation Required	0.63
9	Accuracy	0.60
10	Susceptibility in adverse weather	0.55
11	Mobility	0.54
12	Utility	0.48
13	Compatibility for use	0.42

Below are the results after analysing the Relative Importance Index (RII) as calculated from Equation 2.1. RII is used to analyse the survey results for factors.

$$RII = \frac{\Sigma(W_n)}{A*N} \qquad (2.1)$$

Where,

W = Constant expressing the weighting given to each response

A = Highest rating (In this case, A=3)

n = Frequency of responses

N = Total number of responses

The RII of the factors that are identified for progress monitoring technology selection can be utilized in order of their relative importance if to avoid conflict of ambiguity in decision making for selecting appropriate technology. The “Project type and characteristics” is the most important factor to be considered. Next cost and time efficiency of the technology being implemented are the next factors which are considered important.

2.8 HYPOTHETICAL CASE FOR DEMONSTRATING TECHNOLOGY SELECTION

In the context of a large-scale bridge construction project, let's compare the factors in order of their importance to select a progress monitoring technology between laser scanning and photogrammetry.

1. Project Type and Characteristics: The project involves constructing a long-span suspension bridge over a river with complex geometry and unique design elements.
2. Cost: Laser Scanning: The initial cost of laser scanning equipment and software is \$100,000. Photogrammetry: The initial cost of photogrammetry equipment and software is \$50,000.
3. Time Efficiency: Laser Scanning: It takes approximately 3 hours to complete a comprehensive scan of the entire bridge construction site. Photogrammetry: It takes approximately 6 hours to capture aerial photographs of the bridge construction site.
4. Level of Automation: Laser Scanning: Laser scanning technology offers a high level of automation, capturing detailed point cloud data automatically. Photogrammetry: Photogrammetry requires manual processing of aerial photographs to generate 3D models, which is less automated.
5. Statutory Requirements: Both laser scanning and photogrammetry meet the statutory requirements set by local building codes and regulations.

6. Operating Range: Laser Scanning: Laser scanners have a range of up to 300 meters, allowing for comprehensive coverage of the bridge construction site. Photogrammetry: Aerial photographs can cover a wide area, providing sufficient coverage of the bridge construction site.
7. Training Required: Laser Scanning: Operators require specialized training to operate laser scanning equipment and process point cloud data effectively. Photogrammetry: Operators need training to capture high-quality aerial photographs and process them into accurate 3D models.
8. Preparation Required: Laser Scanning: Some site preparation is needed to set up targets for registration and ensure line-of-sight coverage of the bridge structure. Photogrammetry: Aerial photography requires flight planning and coordination with drone operators, but minimal site preparation is needed.
9. Accuracy: Laser Scanning: Laser scanners provide highly accurate point cloud data with millimeter-level precision. Photogrammetry: Photogrammetry can achieve sub-centimeter accuracy in generating 3D models from aerial photographs.
10. Susceptibility in Adverse Weather: Both laser scanning and photogrammetry can be affected by adverse weather conditions such as heavy rain or strong winds. However, laser scanning is generally more robust against adverse weather conditions.
11. Mobility: Laser Scanning: Laser scanners are less mobile due to their size and the need for setup on tripods or mounts. Photogrammetry: Photogrammetry using drones offers greater mobility and flexibility to capture images from various angles and locations.
12. Utility: Laser Scanning: Laser scanning is well-suited for capturing detailed as-built information, precise measurements, and monitoring structural deformations.

Photogrammetry: Photogrammetry excels in providing visual representations and 3D models of the bridge construction site.

13. Compatibility for Use: Both laser scanning and photogrammetry can be integrated with Building Information Modeling (BIM) systems and other project management tools for seamless data exchange and collaboration.

Considering these factors and their relative importance as provided, the project team determines that laser scanning is the preferred progress monitoring technology for the bridge construction project. While it has a higher initial cost and requires specialized training, laser scanning offers superior accuracy, automation, and is well-suited for capturing detailed as-built information and monitoring structural deformations, which are critical for a large-scale bridge construction project.

2.9 DISCUSSIONS

Progress monitoring is crucial for accurate project control. Choosing the appropriate automated technology based on required parameters is vital in the monitoring stage of a project. Automated technologies can be integrated based on the requirements and can be highly efficient in reducing project overruns compared to manual methods. The technology must be chosen without overselling, such that the investment returns from the project can be made higher.

An idealized situation would enable a higher efficiency in all these parameters, which is not practically possible in a single technology. Therefore, selecting a suitable technology that produces the maximum output based on these parameters would be the goal in the monitoring phase of a construction project.

Another important consideration is that newer technologies might face challenges about their widespread acceptance, as the construction companies might tend to reject the pilot integrated automation technology proposals (Fathi, Dai, & Lourakis, 2015). This happens mainly due to a lack of technical knowledge on automated technologies and the tendency to continue adopting conventional techniques. So integrated proposals through research should be added with proper inspection and maintenance guidelines, followed by proper incentives to the enterprises, so that widespread adoption can be facilitated. Moreover, data collection by a single resource tracking is never sufficient for accurate progress monitoring. Hence, applying data fusion techniques is vital to track multiple resources in a construction site.

2.10 CONCLUSIONS

This chapter provides a systematic review, evaluation and then classification of various automated progress monitoring technologies from field survey and literature.

Using field survey, the data was collected for the technologies and tools being used in construction projects. The challenges that construction sites face with the existing monitoring technologies were also acquired and documented. It was observed that majorities of the construction sites relied on excel sheet based as well as pen and paper-based method of reporting progress, and hence were struggling to complete the projects on time. The survey results also reported that the technologies being used were time consuming, requires several manual tasks and are not accurate.

In literature sixty-one relevant publications to understand the state-of-the-art to guide future research. It also identifies the benefits and limitations associated with each technology (in Table 2.3), along with the factors affecting their selection.

It is to be noted that each construction project is unique and has its specific characteristics. As all the technologies have their advantages and shortcomings, selecting a technology that suits a particular project is extremely important. The technologies can be combined and integrated to minimize cost overruns. This review provides a basis for this selection, as it systematically identifies the scope for each automated technology. In addition, more review efforts are recommended to identify suitable mounting methods that can be used in combination with the techniques. The RII scores of various factors have been computed through a questionnaire survey-based method can be used in the technology selection process. However, these importance values are the preliminary research findings, more in-depth work is required for their validation in future.

It can be noted that, after the development of smart phone, there have been increase in the use of mobile cameras world-wide. The images and videos are now being commonly used for documentation of information. Even the cost of these devices has decreased, and they are now affordable. The processing capabilities have also increased in near future and advanced with the developments in cognitive computing. Therefore, the field of computer vision which involves extracting and processing information from inputs such as images, videos, point clouds is believed to be one of the disruptive technologies of current and near future. Therefore, this research scopes down to explore the use of computer vision for construction progress monitoring in the following chapters. [1]

---

[1] Parts of this chapter have been published in the following articles:

1. Reja, V. K., Pradeep, M. S., & Varghese, K. (2022). A Systematic Classification and Evaluation of Automated Progress Monitoring Technologies in Construction. *Proceedings of the 39th ISARC*, 120–127.
2. Reja, V. K., Varghese, K., & Ha, Q. P. (2022a). As-built data acquisition for vision-based construction progress monitoring: A qualitative evaluation of factors. *Proceedings of the 10th WCS, 24-26 June 2022, Sri Lanka.*, (June), 138–149.

CHAPTER 3.

COMPUTER VISION-BASED TECHNOLOGIES FOR CONSTRUCTION PROGRESS MONITORING

In the previous chapter, various progress monitoring technologies were introduced, analysed, and classified and finally the scope of this research was fixed to the computer vision technologies. This chapter reviews the existing computer vision-based progress monitoring methods in the literature systematically. This chapter have four sub-sections. The first sub-section is on Introduction. The second sub-section is on the methodology followed. The third sub-section states the overview of vision-based monitoring of construction progress. The fourth sub-section presents conclusions. This chapter contributes to the objective 2 of this thesis. The overall methodology followed for objective 2 has been shown in Figure 1.2

3.1 INTRODUCTION: COMPUTER VISION-BASED TECHNOLOGIES

Monitoring the progress of construction projects is crucial, as it provides vital inputs for managers to make timely and informed decisions. Improper progress monitoring leads to losing control of the project, resulting in time and cost overruns. Traditional progress-monitoring methods require manual data entry, which proves to be tedious, time-consuming, and prone to human error (Teizer, 2015). Therefore, its effectiveness can be improved through automation.

In the previous chapter, it was observed that technologies investigated for their applicability in automated progress monitoring in construction include RFID, barcodes and QR codes, laser scanning, photogrammetry, videogrammetry, range imaging (RGB-D), web-based CCTV, and structural sensing (T. Omar & Nehdi, 2016b). Among these, the impact of vision-based technologies was initially limited due to the sophistication of the data-acquisition devices and the high computing power required for processing. However, with the increasing ubiquity of both devices and high-performance computing, it has become feasible to implement vision-based technologies in applications that automate construction processes. Computer vision enables computers to derive numeric information from digital images, videos, depth images and 3D point clouds, process the information, and take action. Sub-domains of computer vision relevant to this research include scene reconstruction, 3D pose estimation, motion estimation, object detection, object recognition and labelling, learning, and 3D scene modelling.

Vision-based technologies have been studied on a wide range of construction applications, such as workforce tracking, resource tracking, condition assessment, quality inspections, safety management, and automated layout generation (Paneru & Jeelani, 2021a). Several studies on vision-based construction progress monitoring have also been reported. However, the processes for effective progress monitoring are diverse, intricate, and complex. In addition, many concepts and technologies can be applied to automate each stage of the process. An integrated framework to characterise and categorise the current and future studies in this area will enable systematic investigation and documentation in this domain.

Review papers in this domain have focused on specific aspects of the progress monitoring process. One study has presented a review of data acquisition technologies (T. Omar & Nehdi, 2016b), while another (Z. Ma & Liu, 2018) have comprehensively studied 3D reconstruction techniques. A couple of studies (Ekanayake et al., 2021b) (Patel et al., 2021) have recently presented a bibliometric analysis of the literature and highlighted specific challenges to be addressed. However, there are several stages in the process, and a broad study of both literature and practice has resulted in identifying the following key stages: (i) data acquisition and 3D reconstruction, (ii) generation of as-built models, and (iii) monitoring progress. Currently, there is no integrated framework that addresses the details of the processes of Computer Vision-

Based Construction Progress Monitoring (CV-CPM), from data acquisition to progress estimation. Therefore, to address this gap, the following sub-objectives are identified:

1. Develop an integrated framework that captures the process requirements of vision-based construction progress monitoring and enables the characterisation and categorisation of current and future work in the area.
2. Utilise the framework to position and compare the various concepts and tools adopted by published research studies.
3. Identify areas and strategies for future work in the area.

3.2 RESEARCH METHODOLOGY – OBJECTIVE 2

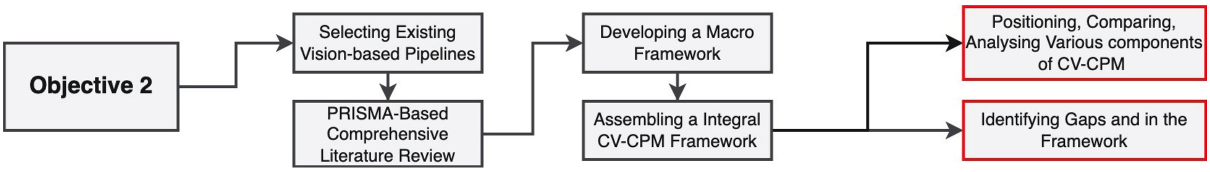


Figure 3.1 Research methodology for Objective 2

As shown in Figure 3.1, as a first step towards these objectives, relevant articles were selected through the broader steps defined in the PRISMA framework. These include identification, screening, eligibility, and inclusion. An extensive keyword search in the Scopus and Web of Science (WoS) reputable scientific databases was conducted. The keywords used were "progress monitoring in construction," "construction progress monitoring," "progress monitoring," "automated progress monitoring," "computer vision in construction progress monitoring," and "vision-based progress monitoring in construction." One hundred eighty-two articles were retrieved from the keyword search. Primary screening of the articles was carried out by reading their abstracts. In order to capture relevant developments in this rapidly evolving area, only key articles appearing after 2014 were defined as eligible for meta-analysis. Other articles which address specific parts of the framework were also referred to and positioned appropriately to support the framework. For an article to qualify as a key article for review, the following two criteria were set:

- “The article should present an applied computer vision-based progress monitoring pipeline by demonstrating it with a case study or experimental results.”
- “Pipelines addressing both indoor and outdoor progress monitoring, as well as pipelines ending at different stages of a type of progress estimation, were considered (i.e., only visualisation as well as visualisation with quantification).”

Twenty-four key articles were selected and included for meta-analysis to formulate the framework. Among other articles referred for specific tools and techniques, fifty-four articles are from the civil engineering domain, and eight are from the computer science domain.

The key articles were systematically reviewed, analysed, and categorised based on the process they addressed and the concepts they used, following the methodology defined by (Omair & Alturki, 2020). The framework evolved as more studies were characterised, categorised, and positioned within it, and the application contexts of concepts and technologies used were studied and compared. After the reviewed works were positioned in the framework, the processes, concepts, and technologies that need to be further developed/explored were identified.

The chapter characterises and classifies the key studies in this area, based on the three macro-stages of the CV-CPM process.

3.3 OVERVIEW OF VISION-BASED MONITORING OF CONSTRUCTION PROGRESS

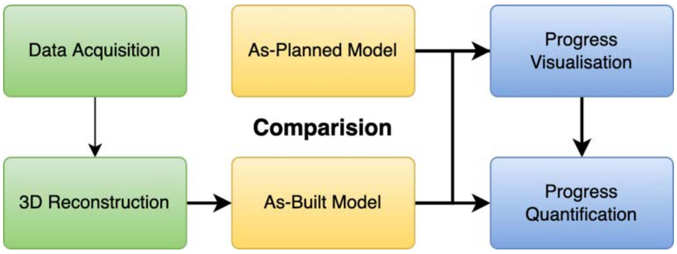


Figure 3.2 Macro-level conceptualization of vision based progress monitoring

This section presents the macro-level stages proposed for vision-based progress monitoring in construction and references the key progress monitoring processes within these stages. Each of these processes is also referred to as a pipeline. Figure 3.2 shows the six macro-level processes that constitute vision-based progress monitoring

Table 3.1 Review of pipelines for vision-based progress monitoring.

Papers	Data Acquisition			3D Reconstruction	As-Built Output	Progress Monitoring				Identified Elements	Operational State	Key Focus	LPM
	Technology	Method	Indoor/ Outdoor			Visualization	Comparison	Quantification	Schedule Updating				
(Khairadeen Ali, Lee, Lee, & Park, 2021)	Laser Scanning and Photogrammetry	Manual	Indoor	SLAM	3D Mesh	Extended Reality XR (VR+AR)	Visualization	N	N	Target Cuboidal Objects	N	Immersive VR (iVR)–based visualization of progress	L-2 Visualization and comparison
(Pour Rahimian, Seyedzadeh, Oliver, Rodriguez, & Dawood, 2020)	RGBD Camera	Manual	Both	Virtual photogrammetry	Segmented Images	Interactive VRE	Object Detection	N	N	Building Elements (Walls, floor, beams etc.,)	Y	CNN-based object detection and use of gaming engine	
(Alex Braun, Tuttas, Borrmann, & Stilla, 2020)	Digital Images	UAV	Outdoor	Commercial Software	Segmented Image with Point Cloud	3D Viewer	Image and Point cloud-based Object Detection	N	N	Columns	Y	Computer vision-based element detection using images, camera poses, as-planned-BIM, precedence relationships	
(Vincke, de Lima Hernandez, Bassier, & Vergauwen, 2019)	Laser Scanning or Photogrammetry	Manual	Both	Commercial Software	Point Cloud	VRE	Commercial Software	N	N	Not specified	N	Gaming engine utility for visualization of as-built and as-planned data	
(K. Han et al., 2018)	Photogrammetry / LS	UAV	Both	SfM & MVS	Point Cloud	3D Viewer	Thresholding	N	N	Construction Materials	Y	Uses Geometry and Appearance based reasoning	
(A. Braun, Borrmann, Tuttas, & Stilla, 2014)	Photogrammetry	Manual	Outdoor	VSfM / SGM	Point Cloud	3D Viewer	Projection Thresholding	N	N	Columns	N	Uses Graph Theory and precedence	

Papers	Data Acquisition			3D Reconstruction	As-Built Output	Progress Monitoring				Identified Elements	Operational State	Key Focus	LPM
	Technology	Method	Indoor/ Outdoor			Visualization	Comparison	Quantification	Schedule Updating				
												relationships for Progress	
(K. K. Han, Cline, & Golparvar-Fard, 2015)	Photogrammetry / LS	Manual	Both	SfM & MVS	BIM	3D Viewer	Ontology based Deviations	N	N	Not specified	Y	Uses Construction Sequence Ontology	
(Zollmann et al., 2014)	Photogrammetry	UAV	Outdoor	SfM	3D Mesh	AR viewer	Interactive AR	N	N	Building Façade	N	AR-based Progress monitoring using mobile aerial 3D reconstruction and visualization	
(Karsch, Golparvar-Fard, & Forsyth, 2014)	Photogrammetry (Unordered Images)	Manual	Both	Model-assisted SfM	Point Cloud	ConstructAide GUI	Visual Colour Coded	N	N	Not specified	N	Uses user assisted SfM method	
(Dimitrov & Golparvar-Fard, 2014)	Photogrammetry (Unordered Images)	Manual	Both	SfM & MVS	BIM	web-based viewer	Material Recognition	N	N	Construction Materials	Y	Appearance-based material classification using LM & HSI features	
(Arif & Khan, 2021b)	Videography	Fixed	Both	NA	Object in Image	Image viewer	MATLAB Image Processing	Y	N	Columns, beams, block masonry	N	MATLAB-based element cropping and quantity estimation	L-3 Quantification
(S. Kim, Kim, & Lee, 2020)	Laser Scanning and Photogrammetry	UAV	Both	Manual Modelling	Mesh Model/ 3D BIM	3D Viewer	Overlapping & BoQ	Y	N	Structural Components	N	Hybrid data collection using drone-based photogrammetry and Laser Scanning	
(Z. Wang et al., 2021b)	Video Camera	Fixed	Outdoor	NA	Object Detected & Segmented Images	3D Viewer	Object Detection	Y	N	Precast Walls	N	Object detection, segmentation and multiple objects tracking from	L-3 Quantification

Papers	Data Acquisition			3D Reconstruction	As-Built Output	Progress Monitoring				Identified Elements	Operational State	Key Focus	LPM
	Technology	Method	Indoor/ Outdoor			Visualization	Comparison	Quantification	Schedule Updating				
												images to detect precast walls.	
(Marianna Kopsida & Brilakis, 2020)	RGBD Camera	Manual	Indoor	Commercial Software	3D Mesh	Mixed Reality	Rays Thresholding	Y	N	big regular-shaped objects	N	Mixed Reality–based progress monitoring	
(Mahami, Nasirzadeh, Hosseininaveh Ahmadabadian, & Nahavandi, 2019)	Photogrammetry	Manual	Both	SfM & MVS with Coded Targets	BIM	3D Viewer	BoQ	N	N	Walls	N	Enhanced method for 3D reconstruction using coded targets	
(Maalek, Lichti, & Ruwanpura, 2019)	Laser Scanning	Manual	Both	Commercial Software	Point Cloud	3D Viewer	Overlapping Distance Threshold	Y	N	Rectangular Columns	N	Extraction of columns using geometric and relationship-based reasoning	
(M. Bassier et al., 2019)	NA	NA	Both	NA	Point Cloud	3D Viewer	Thresholding	Y	N	Walls	N	Percentage of completion of in-situ cast concrete walls	
(Bognot, Candido, Blanco, & Montelibano, 2018)	Videogrammetry	UAS	Outdoor	VSfM / CMPMVS	3D BIM	3D Viewer	Thresholding	Y	N	Building Façade	N	Uses UAS images, low-cost photogrammetry, and GIS	
(Pushkar, Senthilvel, & Varghese, 2018)	Photogrammetry	Manual	Both	Commercial Software	Point Cloud	3D Viewer	Object Detection	Y	N	Masonry Walls	Y	Use of feature engineering for object detection and data classification	
(Bosché, Ahmed, Turkan, Haas, & Haas, 2015)	Laser Scanning	Manual	Both	Commercial Software	BIM	3D Viewer	Object Recognition	Y	N	MEP Components	N	Use of Hough Transform to detect parametric features with Scan Vs BIM	

Papers	Data Acquisition			3D Reconstructio	As-Built Output	Progress Monitoring				Identified Elements	Operational State	Key Focus	LPM
	Technology	**Method**	**Indoor/ Outdoor**			**Visualization**	**Comparison**	**Quantification**	**Schedule Updating**				
(Golparvar-Fard, Peña-Mora, & Savarese, 2015)	Photogrammetry (Unordered Images)	Manual	Both	SfM & MVS	Point Cloud	D4AR Viewer	Voxel Occupancy	Y	N	Columns, Foundation Walls	N	Use of machine-learning scheme built upon a Bayesian probabilistic model	
(K. K. Han & Golparvar-Fard, 2014)	Photogrammetry (Unordered Images)	Manual	Both	SfM & MVS	Point Cloud	web-based viewer	Material Recognition	Y	N	Construction Materials	Y	Appearance-based material classification	
(H. Omar, Mahdjoubi, & Kheder, 2018)	Photogrammetry	Manual	Outdoor	Commercial Software	Point Cloud	NA	Enclosed Volume Calculation	Y	Y	Columns	N	Photogrammetry-based monitoring with automated notification system and schedule update	L-4 Schedule Update with notifications
(Pučko et al., 2018)	RGBD Camera	Manual	Both	Commercial Software	4D BIM	3D Viewer	Commercial Project Management Software	N	Y	Not specified	N	Continuous monitoring using helmet-mounted scanners	

These processes are grouped into three stages, as follows:

i. Acquisition of as-built data and conversion into a point cloud using 3D reconstruction techniques.
ii. Generation of the as-built model from the point cloud (with reference to the as-planned model, if available)
iii. Comparison of the as-built and as-planned models to assess progress by visualisation and/or by quantifying the work done.

All studies on vision-based progress monitoring focus on one or more of these stages. Table 3.1, a key output of this study, was developed through a detailed review of the published works on vision-based construction progress monitoring. The articles were identified through the systematic process described in the introduction. The categories shown in Table 3.1 were refined iteratively based on objectives, solutions and concepts addressed by the reviewed papers.

The studies referred to in Table 3.1 are arranged based on the category of progress monitoring, as indicated in the last column. The first column contains the reference to the study, and the subsequent columns contain the categories and sub-categories based on the macro-level model. The cells in Table 3.1 identify the technology, concept, and algorithm explored in each category and sub-category. It can be observed in Table 3.1 that for each category, there is a wide range of alternatives for all process components, construction elements, and the key foci that the studies have explored, as discussed briefly in the following paragraphs.

Data Acquisition and 3D Reconstruction: This section of the table documents the data acquisition technologies, type of mounting, and usage environment (indoor/outdoor) that the studies use. Depending on the input type, the data undergoes 3D reconstruction using various commercial software, as shown in Table 3.1. The workflow from data acquisition to 3D reconstruction is discussed in detail in Chapter 4, subsection 4.1.

As-built Modelling: As shown in Table 3.1, there are varied output data types that are generated by the pipelines. The generation of these as-built models requires specific as-built modelling techniques, which are classified and discussed in detail in subsection 4.2.

Progress Monitoring: Table 3.1 also indicates the level of progress monitoring (LPM) carried out, judging by the types of platforms used to implement progress estimation. It was seen that various progress monitoring pipelines have ended at different levels of monitoring. Therefore, this study classifies these into four levels. These are: (L-1) only visualising the as-built model, (L-2) visualisation incorporating a comparison with the as-planned model, (L-3) quantification of progress, and (L-4) quantification with schedule update and warning notifications.

For progress visualisation, different immersive and non-immersive environments have been used. For quantification, numerous approaches have been used based on the availability of an as-planned model. The details are discussed in Chapter 4, subsection 4.3.

In addition to the three process components—data acquisition and 3D reconstruction, as-built modelling, and progress monitoring—Table 3.1 also characterises the pipelines on two additional parameters. First, the types of elements being worked on, and second, whether the structure's current operational state can be recognised (e.g., a wall's stage can be brickwork- completed, plastered, or painted).

3.4 CONCLUSION

As inferred from Table 3.1, numerous combinations of tools, concepts, and algorithms need to be evaluated to arrive at the appropriate approach for a specific progress monitoring need. Though all the combinations may not be technically viable, a significantly large set of options needs to be experimentally explored and documented to ensure systematic progress in this domain. Therefore, an integrated framework encompassing all the processes will support the systematic study.

In the next chapter, the three macro-level steps are expanded to formulate the detailed framework wherein the concepts/processes/technologies mentioned in these studies are characterised, positioned, and discussed.[2]

[2] Parts of this chapter have been published in the following articles:

1. Reja, V. K., Varghese, K., & Ha, Q. P. (2022b). Computer vision-based construction progress monitoring. *Automation in Construction*, *138*, 104245.

CHAPTER 4.

INTEGRATED FRAMEWORK FOR COMPUTER VISION-BASED CONSTRUCTION PROGRESS MONITORING (CV-CPM)

4.1 INTRODCUTION: PROPOSED FRAMEWORK FOR CV-CPM

Automating the process of construction progress monitoring through computer vision can enable effective control of projects. Systematic classification of available methods and technologies is necessary to structure this complex, multi-stage process. Using the PRISMA framework, relevant studies in the area were identified. The various concepts, tools, technologies, and algorithms reported by these studies were iteratively categorised, developing an integrated process framework for Computer-Vision-Based Construction Progress Monitoring (CV-CPM). This framework comprises data acquisition and 3D reconstruction, as-built modelling, and progress assessment. Each stage is discussed in detail, positioning key studies, and concurrently comparing the methods used therein. The four levels of progress monitoring are defined and found to strongly influence all stages of the framework. The need for benchmarking CV-CPM pipelines and components are discussed, and potential research questions within each stage are identified. The relevance of CV-CPM to support emerging areas such as Digital Twin is also discussed.

The integrated framework for Computer Vision-based Construction Progress Monitoring is shown in Figure 4.1. This framework was derived after evaluating the existing literature studies in detail using the macro-level framework lens described in the previous chapter (Section 3.3, Figure 3.2) and the methodology defined in (Omair & Alturki, 2020). This section explores the micro-level details of each of the three main components.

The following three sub-sections are on the various processes, algorithms, methods, and technologies that could be applied to the relevant subsections of the framework.

This chapter is divided into seven subsections. The first sub-section introduces this chapter. The second sub-section is on data acquisition and 3D reconstruction. The third sub-section is on as-built modelling. The fourth sub-section is on progress monitoring. The fifth sub-section presents discussion. The sixth sub-section aligns the recently published works to the CV-CPM framework. This is followed by conclusions in the seventh sub-section.

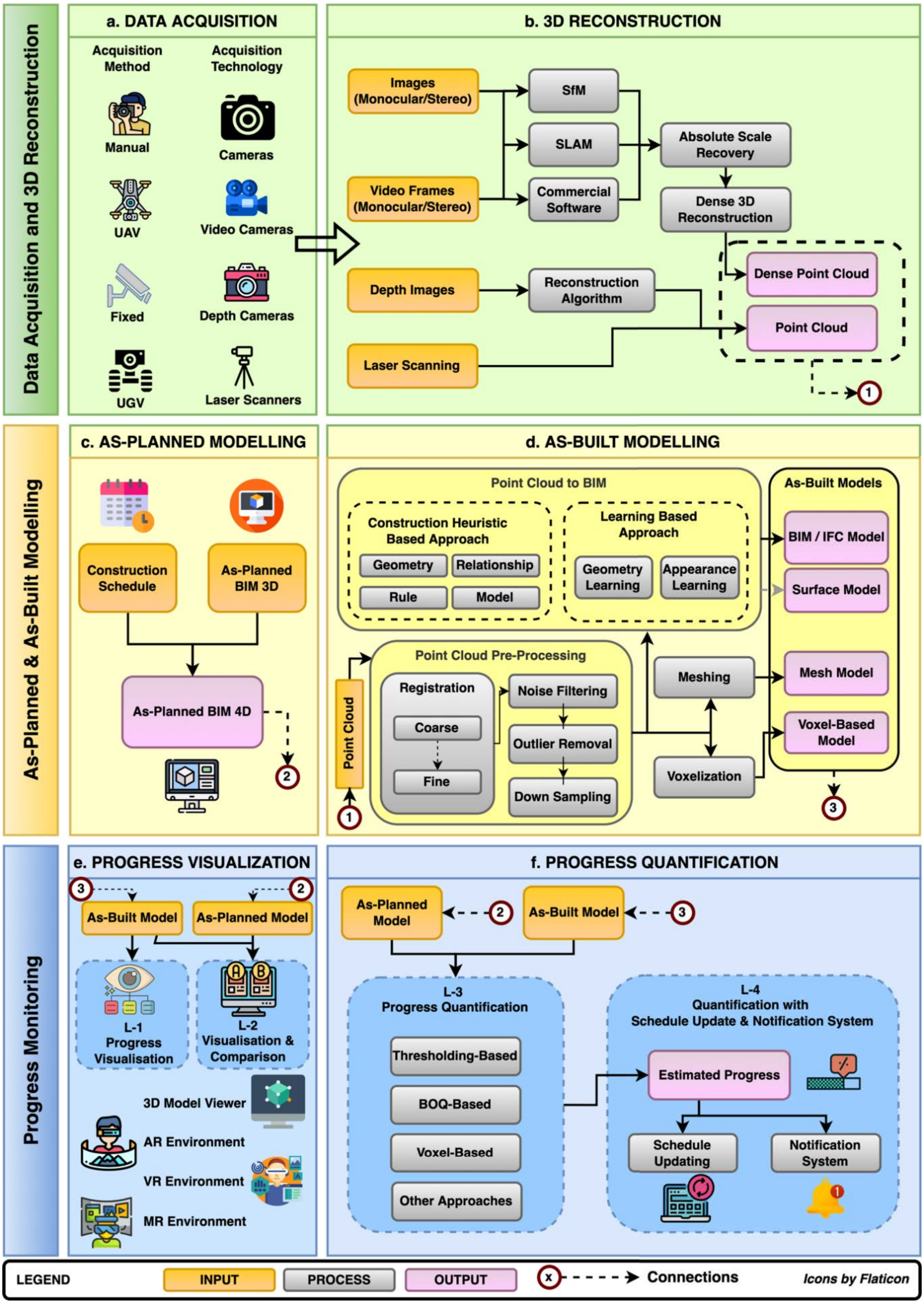


Figure 4.1 Integrated framework for computer vision-based construction progress monitoring (CV-CPM)

4.2 DATA ACQUISITION AND 3D RECONSTRUCTION

This section addresses the data acquisition (Figure 4.1 (a)) and 3D reconstruction (Figure 4.1 (b)) components of the framework. The data acquisition depends on the acquisition technology as well as the sensor mounting method selected. There are various steps involved in 3D reconstruction algorithms. The decision as to which technologies to use is a multi-faceted issue. Defining these based on the characteristics of a construction project and other factors is the primary contribution of this section.

4.2.1 Data acquisition: technologies and methods

Data acquisition depends on selecting the technology and the sensor mounting method (Figure 4.1 (a)). As seen in previous chapter (Table 3.1), several combinations of sensing technologies and sensor mounting methods have been explored. Digital cameras (Golparvar-Fard et al., 2015)(Dimitrov & Golparvar-Fard, 2014), video cameras (Bognot et al., 2018), laser scanners (Bosché et al., 2015), and range imaging (or RGB-D cameras) (Marianna Kopsida & Brilakis, 2020)(Pour Rahimian et al., 2020)(Pučko et al., 2018) are the vision-based technologies used for collecting data for progress monitoring. They generate inputs to the framework in the form of image frames or point clouds. The sensor mounting method for data acquisition can be in the form of fixed devices, handheld devices, robotic systems mounted on unmanned ground vehicles (UGV) (A. Adán, Quintana, Prieto, & Bosché, 2020), unmanned aerial vehicles (UAV) (Bognot et al., 2018), or a combination of these systems (Asadi et al., 2020).

A systematic review can be found in (T. Omar & Nehdi, 2016b) on technologies used for progress tracking in construction. The study finds technology selection a "multi-faceted issue" which depends on "required degree of accuracy, project size, level of automation, and the ultimate purpose of progress tracking".

Figure 4.2 shows a dual matrix illustrating the data acquisition technology and method selection criteria. The technologies and methods are listed on the concentric rings in their respective matrix, where colours and icons are added for easier comprehension.

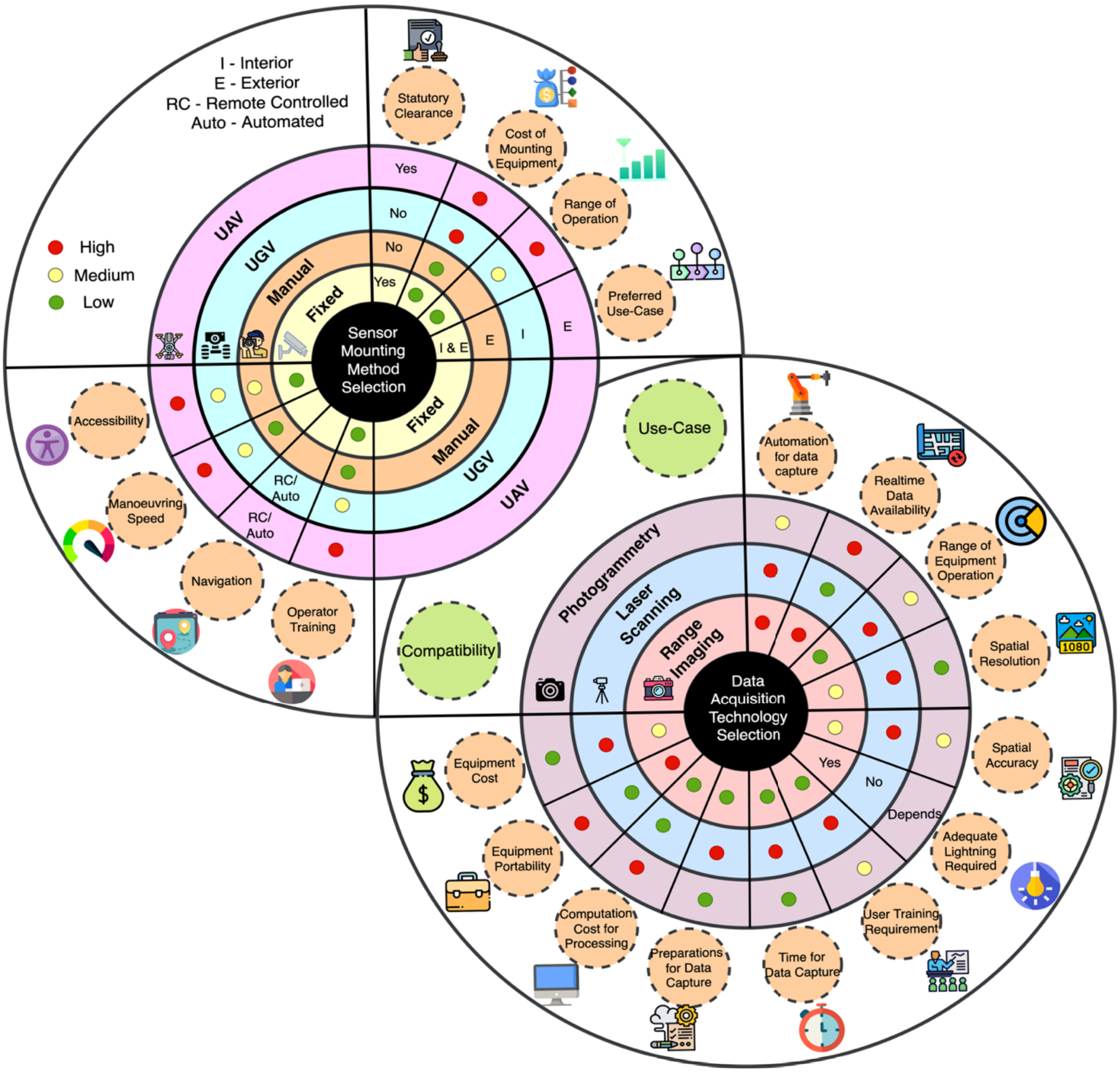


Figure 4.2 Dual matrix for sensor mounting method and data acquisition technology selection.

The key factors that govern their selection have been identified and are positioned along the outer ring. These factors have been adopted from literature as well as authors on experience.

Eight key factors for the sensor mounting method include statutory clearance, cost of mounting equipment, range of operation, preferred use case, operator training, navigation, manoeuvring speed, and accessibility. In addition, twelve key factors for data acquisition technology have been identified. They include the level of automation for data capture*, real-time data availability, range of equipment operation, spatial resolution, spatial accuracy*, adequate lightning requirement, user training requirement*, time for data capture*, preparations for data capture*, computation cost for processing, equipment portability*, and equipment cost*. Kindly note that the factors that are marked with * are adopted from (Marianna Kopsida, Brilakis, & Vela, 2015). The other factors included here are from the author's own experience with these technologies.

This matrix is constructed based on the review and classification of literature. The relative comparison indicator for technology versus sensor mounting method is indicated along the radial direction. The low, medium, and high ratings are shown using green, yellow, and red colour codes, respectively. The proposed rating is a preliminary step towards assisting in the selection of a combination. As the proposed ratings are based on the authors' perspective, experimental studies are required to make them more objective.

4.2.2 3D reconstruction

Following the image-based input from the previous step, the next step is to generate a point cloud model from a series of algorithms for 3D reconstruction (Figure 4.1 (b)). As observed from the literature presented in the previous chapter, most photogrammetry-based progress monitoring pipelines use commercially available software to generate a 3D point cloud from optical camera images or depth images.

Optical camera images do not contain depth information. SfM (Structure-from-Motion) and SLAM (Simultaneous Localisation and Mapping) are the two conceptual approaches to add depth information for sparse 3D reconstruction.

Depth images are captured using RGB-D cameras, for example, Microsoft Kinect. They have depth information (XYZ coordinates) along with colour information (RGB values) on a per-pixel basis. This information is used for mapping and performing dense 3D reconstruction using intrinsic and extrinsic camera parameters (Q. Wang et al., 2020). The output from a laser scanner is a 3D point cloud, and hence does not require 3D reconstruction.

The work in (Z. Ma & Liu, 2018) focuses on generating point clouds from monocular and stereo images, and video frames using SfM for photogrammetry and its corresponding algorithms for 3D reconstruction. In this chapter, only the basic concept of SfM is presented, and a detailed comparison with SLAM is made to facilitate decision making for 3D reconstruction in construction.

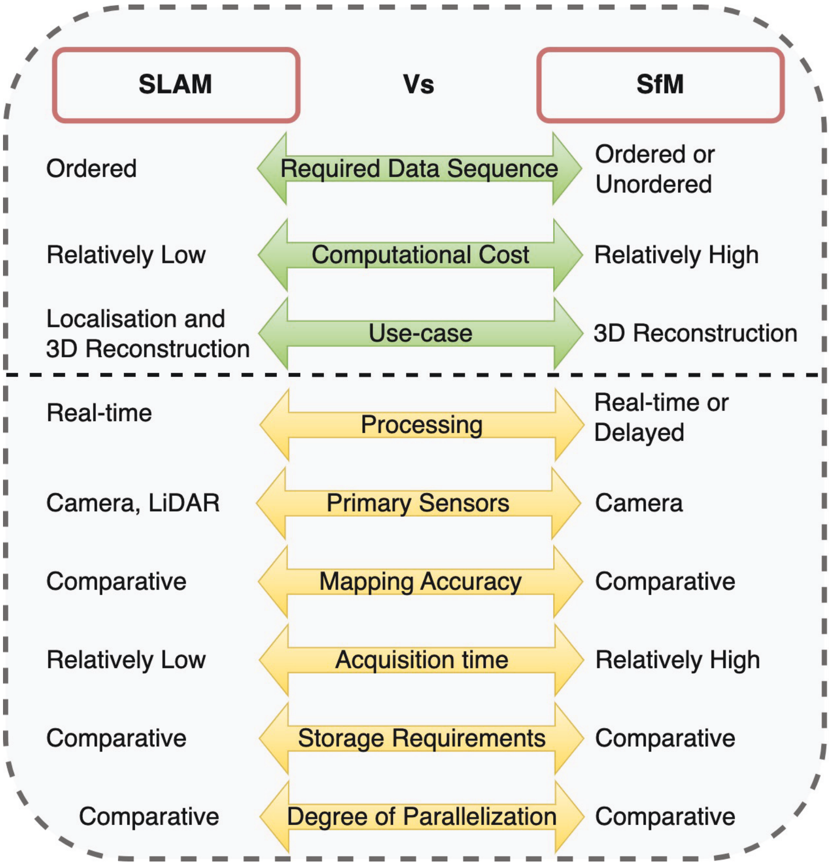


Figure 4.3 Comparison between SLAM and SfM for 3D reconstruction in construction

SfM (Visual SFM (Wu, 2011) or Open SfM (Mapillary, 2018)) is a technique consisting of a combination of algorithms for photogrammetric 3D reconstruction from numerous image frames (Figure 4.1 (b). It is an offline approach for estimating the scene's camera motion information. The process is to match all the images to each other, find the correspondences, and then delete mismatched images to obtain relative camera positions and the structure without any prior geometric or semantic information. The concept of SfM is based on stereoscopic photogrammetry, as stated in (Furukawa, Curless, Seitz, & Szeliski, 2009). A detailed review of the 3D reconstruction by SfM and its algorithm can be obtained in (Z. Ma & Liu, 2018) and (Golparvar-Fard, Peña-Mora, & Savarese, 2011).

SLAM (Bailey & Durrant-Whyte, 2006) is a more general framework, in real-time, where a mapping system starts motion from an unknown location in an unknown environment, and during movement, simultaneously keeps track of its position with respect to the surrounding environment, building an incremental map. SLAM may depend on sensors like cameras, LiDAR, GPS, IMU, etc. Modern visual SLAM has gradually developed into a multi-feature, multi-sensor, and deep learning-based method that can be used for dynamic and haphazard construction environments.

As shown in the literature presented in the previous chapter, the use of Visual SLAM in progress estimation has been sparsely explored in construction. With the rapid advances in autonomous technologies for data acquisition (UAV/UGV based), SLAM-based reconstruction will become more relevant in construction. Figure 4.3 shows the comparison between SfM and SLAM to guide the selection process for 3D reconstruction applications for progress monitoring. The three key factors are discussed below in detail, namely:

1. **Data sequence:** SfM is based on feature matching of image pairs is highly dependent on the quality and the sequence of the image frames obtained. It need not necessarily require ordered image frames. On the other hand, SLAM builds an incremental map in real-time and requires sequential frames to estimate the previous and next poses. Hence, construction sites collecting images from multiple location-aware sensors should be able to work on SfM. SLAM can be selected for sites requiring autonomous navigation of robots in volatile unmanned environments.
2. **Computational cost:** The process of SfM is computationally costly because of unordered data compared to SLAM, which works on ordered data. Bundle

adjustment (non-linear optimisation) (Triggs, McLauchlan, Hartley, & Fitzgibbon, 2000) in SLAM is applied only on the last “N keyframes”, as opposed to the entire graph in SfM, to give a real-time performance in budget. In scenarios where the availability of parallel GPUs is not a constraint for computation, SfM will give better results than SLAM.

3. **Path-planning or Autonomy:** Pre-planning the robot surveillance paths for construction sites is tedious and complex. Planned paths are subjected to uncertainty on construction sites due to their dynamic nature. Both SfM and SLAM can be used if the sensor is mounted on a UAV or a UGV. SfM is predominantly used in known environments where surveillance paths can be planned. In contrast, SLAM is advantageous for autonomous applications in random locations and unknown environments.

Apart from these three key features, other sub-factors which can be considered are shown in Figure 4.3. In general, SLAM generates a sparse map compared to SfM and has limitations in loop closure, scaling, and drifting issues. However, SLAM can overcome SfM's drawback when dealing with a featureless or repetitive scene with potential false matches. Also, SfM has been tried on a larger scale, whereas SLAM can be used for small-scale tasks (Mitsugami, 2011; Sun, Zhang, Wang, & Zhang, 2021).

It is evident from the review that both technologies have their advantages and disadvantages, and selection should be based on project requirements and compatibility. There are multiple algorithms for SLAM (Sun et al., 2021) and SfM ((Wu, 2011),(Mapillary, 2018)), and when various types of embedded sensors are included, the options for selection increase. For example, the embedded camera sensor can be monocular, binocular or RGB-D. Based on the type of sensor, the core of the reconstruction algorithm changes, as well as the accuracy and the computation time. Therefore, numerous combinations of these variations increase the complexity of the framework and the resulting selection process.

4.2.3 Absolute scale recovery and dense 3D reconstruction

The output obtained by the image-based pipeline is primarily sparse and not to scale. Therefore, as shown in Figure 4.1 (b), it requires absolute scale recovery and dense 3D reconstruction. Laser scans do not require this. Table 4.1 shows the techniques and algorithms that can be used in the associated sub-processes.

The scale recovery sub-process determines the absolute scale of the generated sparse point cloud by comparing it with the local coordinates of the point in the sparse point cloud. Scale recovery has been implemented using, manual techniques, pre-measured objects, and geo-registration, as shown in Table 4.1. The first two techniques require manual measurements and feeding ground truth, whereas geo-registration requires automated sensing of camera parameters to apply the transformation.

Table 4.1 Sub-processes and associated techniques

Sub-Process	Techniques/ Algorithms	Existing Use in Progress Monitoring Pipelines	Other Works
Absolute Scale Recovery	Manual	-	(Rashidi, Brilakis, & Vela, 2015)
	Pre-measured Object	(Mahami et al., 2019)(K. K. Han & Golparvar-Fard, 2014)	
	Geo Registration	(Zollmann et al., 2014)(Bognot et al., 2018)	
Dense 3D Reconstruction	Multi-view Stereo MVS	(K. K. Han & Golparvar-Fard, 2014)(K. Han et al., 2018)	(Wu, 2011)(Mapillary, 2018) (Mitsugami, 2011)
	Clustering Multi-View Stereo (CMVS)	(Mahami et al., 2019)	
	Patch-based Multi-View Stereo (PMVS)	(Mahami et al., 2019)	

Sub-Process	Techniques/ Algorithms	Existing Use in Progress Monitoring Pipelines	Other Works
	MVS with voxel colouring	(Golparvar-Fard et al., 2015)(Dimitrov & Golparvar-Fard, 2014)	
	SGM	(A. Braun et al., 2014)	
	Meshing using PSR	(Bognot et al., 2018)(Zollmann et al., 2014)	
Registration	**Coarse Registration**		(M. Bueno, González-Jorge, Martínez-Sánchez, & Lorenzo, 2017)(Martín Bueno, Bosché, González-Jorge, Martínez-Sánchez, & Arias, 2018) (Mitsugami, 2011)
	Manual	(Golparvar-Fard et al., 2015)(S. Kim et al., 2020)(K. Han et al., 2018)	
	Marker-Based	(Zollmann et al., 2014)(Maalek et al., 2019)	
	Sensor-Based	(Zollmann et al., 2014)	
	Feature-Based	(Mahami et al., 2019)(Lei, Zhou, Luo, & Love, 2019)	
	Fine Registration		
	ICP	(Zollmann et al., 2014)(C. Kim, Son, & Kim, 2013b)(Khairadeen Ali et al., 2021)(A. Braun et al., 2014) (Pučko et al., 2018)(C. C. Kim, Kim, Son, & Kim, 2013)(Alex Braun et al., 2020)	
	Image Registration		
	Geo-referencing	(Karsch et al., 2014)(K. K. Han et al., 2015)(Roh, Aziz, & Peña-Mora, 2011)(K. K. Han & Golparvar-Fard, 2014) (Golparvar-Fard et al., 2011)	
Noise and Outlier Removal	Manual	(S. Kim et al., 2020)	(Lee, Son, Kim, & Kim, 2013)(Y.-F. Liu, Cho, Spencer, & Fan, 2016)
	RANSAC+PCA-based	(Maalek et al., 2019)	
	Tensor voting algorithm	(C. Kim et al., 2013b)	
Down Sampling	Point-space strategy Algorithm	-	(Son, Kim, & Kim, 2015)

The next sub-process, Dense 3D reconstruction, recovers the scene details. The algorithms used are multi-view stereo (MVS), clustering views for multi-view stereo (CMVS), patch-based multi-view stereo (PMVS), and semi-global matching (SGM). The open-source algorithms which, perform dense 3D reconstruction include VisualSfM (Wu, 2011), OpenSfM (Mapillary, 2018), and Bundler SfM (Mitsugami, 2011).

A dense 3D reconstruction step is required if the reconstructed point cloud is sparse, typically in the case of reconstruction from monocular or stereoscopic images or video frames. The reconstruction output from depth images is already dense enough to be used directly and does not require dense 3D reconstruction.

Only a handful of CV-CPM pipelines have stated the method used for absolute scale recovery and dense 3D reconstruction; others have not explicitly mentioned it in the paper. The choices available for absolute scale recovery and dense 3D reconstruction add additional variables to the CV-CPM framework. Once a dense point cloud is generated, the next step is to pre-process it through registration, noise filtering, outlier removal and down sampling. These sub-processes and model generation are discussed in the next section.

4.3 MODEL GENERATION

In this section, first, the characteristics of the as-planned model are discussed. All the steps in pre-processing and the process of generating the as-built model are then discussed in detail (Figure 4.1 (d)). The existing techniques for as-built model generation are categorised into heuristic-based and learning-based methods. The definition of these categories and the mapping of existing pipelines to these categories/sub-categories is the primary contribution of this section. This section also presents the trade-off between the levels of information the various as-built outputs provide for progress estimation and the computational complexity required to generate them.

4.3.1 As-Planned model

As-planned models may or may not be available for a particular facility or project. Generally, they are prepared in the design phase of the project. These can be 2D/3D CAD or 3D/4D BIM models. The approach to progress monitoring depends on the availability and type of the as-planned model, as it facilitates comparison with the as-built model.

4.3.2 As-Built model

The following sub-sections include the details for the various sub-processes and options for as-built modelling (Figure 4.1 (c)).

4.3.2.1 Point cloud pre-processing

Point cloud pre-processing is a crucial step towards as-built modelling to improve the point cloud quality. The level of automation in pre-processing has not been adequate, and manual intervention is still used to get the required quality. Table 4.1 outlines the critical sub-processes of point cloud pre-processing, which broadly consist of registration, noise filtering, outlier removal, and down-sampling, along with various techniques and algorithms used in the corresponding literature. The table also shows studies that have addressed fundamental aspects of these sub-processes.

Registration involves merging multiple partial point clouds obtained into a single file or registering an as-built point cloud over an as-planned BIM model for progress monitoring. There are two stages of registration, coarse and fine.

Coarse registration aligns a set of point clouds using correspondences between them. In contrast, the fine registration algorithm further matches multiple point clouds by estimating a rigid transform and minimising the distance between the corresponding matched points. The selection of these approaches depends on the type of data acquisition technology used and location capture, as registration depends on localising the data. While the manual and marker-based methods require iterative human efforts, the sensor and feature-based approaches are reported to be more automated. The Iterative Closest Point (ICP) algorithm is the most used in CV-CPM pipelines for fine registration. It searches for an optimum solution, and the convergence speed is directly proportional to the accuracy of coarse registration. For image-based methods, image registration over the BIM model is made using georeferencing.

Next, noise and outliers are unwanted points and are removed to improve the point cloud quality prior to further processing. Finally, due to the merging of point clouds, the overlapped regions become much denser, affecting processing efficiency. Hence down-sampling is performed to make the point cloud uniform. Relevant works which include these steps are shown in Table 4.1.

Although pre-processing of data has been presented in some progress monitoring pipelines, there is no detailed exploration of the available techniques. There is a need for a comprehensive study and further research to customise pre-processing methods based on the input pipeline's requirement. There is also a need to explore the other techniques mentioned in Table 4.1 for better results using the CV-CPM framework.

4.3.2.2 Types of as-built models:

Table 4.2 Relative computational requirements and output accuracy for the level of progress monitoring, based on the type of as built models.

Type of as-built models	Level of progress monitoring			
	L-1 and L-2		L-3 and L-4	
	Progress Visualization		Progress Quantification	
	Computation for model generation	Quality of visual output	Computation for progress quantification	Accuracy of Qty. take-offs
IFC/BIM	Very High	Very High	Very Low	Very High
Surface Model	High	High	Low	High
Mesh Model	Low	Medium	High	Medium
Voxel Model	Low	Medium	High	Low
Point Cloud	Very Low	Low	Very High	Very Low

Existing pipelines have used several combinations of as-built and as-planned model comparisons for progress estimation. Based on the algorithm and method used, the as-built output can be a pre-processed point cloud (Maalek et al., 2019), voxel model (Hübner, Weinmann, & Wursthorn, 2020), mesh model (Marianna Kopsida & Brilakis, 2020)(Zollmann et al., 2014), surface model (Macher, Landes, & Grussenmeyer, 2017)(Xiong, Adan, Akinci, & Huber, 2013), or IFC/BIM models (Mahami et al., 2019)(Pučko et al., 2018)(Bosché et al., 2015) (Figure 4.1 (d)).

The level of progress monitoring proposed for a project plays a vital role in deciding the type of as-built model to be generated. Table 4.2 compares computational requirements and quality of output for the different levels of progress monitoring with the corresponding types of as-built models. Here, green represents a performance improvement, and red indicates a decrement in performance.

It can be seen from Table 4.2 that for L-1 or L-2 based monitoring, voxel-based, mesh-based, or point cloud models are recommended to be used because of low computational requirements. Even if the point cloud quality is low, it is adequate for visualisation. While visual output quality is better for IFC-based or surface models, it is not recommended for progress visualisation due to very high computing requirements.

For L-3 or L-4 based monitoring, while it will take higher computation to generate an IFC or surface-based model initially, the computation required for progress quantification will be substantially lower with higher accuracy; hence it is recommended. On the other hand, the mesh model, voxel model, and point clouds are easy to generate but will take higher computation to perform the comparison and result in low accuracy. Mesh-based and voxel-based modelling are explored domains that are easier to develop using established algorithms and software and therefore were briefly introduced.

Hence, this trade-off of required computation, accuracy, speed, and application case should be considered while designing the progress monitoring pipeline for a specific usage requirement. The relative comparison provided here is a preliminary step based on the authors' perspective. To arrive at more objective ratings, the parameters must be assessed quantitatively (for computation and accuracy) and qualitatively (for visual output quality). Further experimentation is required to make these assessments.

Progress visualisation (L-1 & L-2) is typically being implemented at construction sites. However, to take critical decisions, quantifying progress is essential. IFC-based as-built models can provide direct and accurate progress quantification (L-3) and facilitate schedule updating and notifications generation (L-4). Therefore, existing studies have explored various techniques to achieve point-cloud to IFC/BIM model conversion. The following sub-section discusses and classifies these approaches in detail.

4.3.2.3 Point Cloud to BIM

This section focuses on as-built modelling methods resulting in IFC classes (Figure 4.1 (d)). Only two existing studies have classified point cloud-to-BIM approaches (Pătrăucean et al., 2015)(Zeng, Chen, & Cho, 2020). One has classified them as local and auxiliary heuristics based on as-planned BIM involvement (Pătrăucean et al., 2015), whereas the other extends the classification based on geometry, rule, model, and feature-based techniques (Zeng et al., 2020).

Table 4.3 summarises the approaches taken by the referenced studies to explore the point cloud-to-IFC/BIM conversion. This study broadly classifies these approaches as heuristics-based and learning-based and into additional subclasses, extending the classification by (Pătrăucean et al., 2015) and (Zeng et al., 2020).

The use of heuristics has been based on geometry, rule-based, relationship-based, and model-based constraints. Learning-based techniques can be generally classified as geometric and appearance learning. The following section discusses and analyses the two approaches of element recognition in detail

Table 4.3 Parametric as-built modelling approaches in the literature

Approaches		(Li et al., 2020)	(Pour Rahimian et al., 2020)	(Xiong et al., 2013)	(Hamledari, McCabe, & Davari, 2017)	(Zeng et al., 2020)	(Chen, Kira, & Cho, 2019)	(Iwaszczuk et al., 2018)	(Alex Braun et al., 2020)	(Marianna Kopsida & Brilakis, 2020)	(Pučko et al., 2018)	(Golparvar-Fard et al., 2015)	(Antonio Adán, Quintana, Prieto, & Bosché, 2018)	(Bosché et al., 2015)	(Nikoohemat, Diakité, Zlatanova, & Vosselman, 2020)	(Maalek et al., 2019)	(Ochmann, Vock, & Klein, 2019)	(C. Wang, Cho, & Kim, 2015)	(Pu & Vosselman, 2009)	(Ha Tran & Khoshelham, 2020)	(Hübner et al., 2020)	(Franz, Irmler, & Rüppel, 2018)	(Macher et al., 2017)	(Barazzetti, 2016)	(Díaz-Vilariño, Khoshelham, Martínez-Sánchez, & Arias, 2015)	(Arnaud, Christophe, Gouiffes, & Ammi, 2016)	(K. K. Han et al., 2015)	(X. Liu et al., 2019)	(Chen, Fang, & Cho, 2017)
Heuristic Based	**Geometry**	x		x									x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x
	Rule	x													x	x	x	x	x	x	x	x	x	x	x	x			
	Relationship	x													x	x	x	x	x								x		
	Model		x						x	x	x	x	x	x															
Learning-Based	**Learning – Geometry**	x	x	x	x	x	x																						
	Learning – Appearance	x	x		x			x																					

A. Heuristics-based approach

Heuristics-based approaches (Figure 4.1 (d)) use pre-coded domain knowledge to identify elements from as-built point clouds. Table 4.4 shows the comprehensive sub-classification of identified heuristics with their definitions as formulated by the authors and the context of their applications in the literature. Following are a few insights from Table 4.4 for the four types of heuristics:

1. **Geometrical constraints** have been applied based on shape, dimensions, point density, and associative geometry. They require only basic geometric information as inputs; therefore, they are widely utilised. The challenge lies in retrieving and feeding this information for all the elements, as construction sites may have unconventional elements.
2. **Rule-based constraints** have been applied based on the direction of the surface normal, orientation, principal axis direction and other rules. Apart from a few complex elements, most elements can be identified by applying these rules.
3. **Relationship-based constraints:** The construction sequence is of utmost importance, as it can accurately define the objects' dependencies from schedules, enabling better element detection for progress monitoring. These time-based relationships can be helpful while there is limited visibility of the scene or occlusions. Space-based constraints have been used in the literature using ontological relations (K. K. Han et al., 2015) and shape grammar (H. Tran, Khoshelham, Kealy, & Díaz-Vilariño, 2019), but they are challenging to formulate and apply in practice.

4. **Model-based constraints:** Model-based element retrieval uses reverse engineering to recognise the elements. The point cloud is registered over the as-planned BIM model, and the elements are identified based on overlap occupancy. Although the method results in high accuracy, its use is limited because it requires a detailed as-planned model. It has specific point cloud quality and reference BIM model requirements (Rebolj, Pučko, Babič, Bizjak, & Mongus, 2017). The approach also faces computational challenges in automatic registration as the acquired point cloud data is voluminous.

It can be summarised that heuristics-based techniques usually require a significant amount of prior information. Generation of this information requires domain knowledge about construction and the type of elements being recognised. A recommended solution is to define a set of metaheuristics that generally applies to any construction dataset as a subset of all the heuristics; using them will make element identification viable with less detailed information requirements.

Though using some heuristics makes element identification easier, a pure heuristic-based approach for CV-CPM can be unreliable and challenging to be implemented.

B. Learning-based approach

Learning-based techniques (Figure 4.1 (d), Table 4.3) are used on point clouds for 3D shape classification, object detection, and segmentation. They are based on training features using a neural network and computed weights to predict object or material classes. Two types of learning approaches have been used:

1. **Geometric Learning** extracts geometrical features using descriptors and feeds neural networks on statistics like distance, area, and angles. The geometric features are used to identify the element's physical presence.
2. **Appearance Learning** extracts features like HSI (Hue-Saturation-Intensity) colour values, material reflectance, and surface roughness as features. The appearance-based features describe its operational state.

As observed from Table 4.3, only a few studies have detected elements for as-built modelling using learning-based techniques in the past. Most of these methods detect geometric features and do not recognise the elements' operational state. A few studies recognised the operational state by using image processing techniques (K. K. Han & Golparvar-Fard, 2015)(K. Han et al., 2018) and learning appearance-based features (Hamledari, McCabe, & Davari, 2017).

For progress monitoring applications, learning-based approaches have been predominantly used on 2D image data, and there are very few studies involving 3D point cloud learning. Recent advances in machine learning and deep neural networks have motivated researchers to explore this area. However, their application in point cloud-based element recognition for construction has been limited because of challenges such as:

1. 3D point cloud data are unstructured, large, and with varying point densities, making 3D feature detection computing-intensive and time-consuming.

2. Scan data are generally noisy and have occluded components and cluttered scenes. Also, the data are often susceptible to changes in environmental factors such as rain, fog, dust, or mist.
3. Learning models do not perform well when there is a significant difference between the training and test data.
4. Supervised learning requires extensive annotated training and testing datasets for various categories of incomplete building elements: currently, the data is limited.

ScanNet (Dai et al., 2017), S3DIS (Armeni et al., 2016), and ModelNET-40 (Zhirong Wu et al., 2015) are 3D point cloud datasets consisting of different classes of building elements that are currently used in studies. However, these data sets represent completed buildings; for construction progress monitoring studies, data sets are required to be acquired from buildings under construction. For material recognition, one such example of an image-based dataset is the Construction Material Library (CML) created in (Dimitrov & Golparvar-Fard, 2014) and (K. K. Han & Golparvar-Fard, 2015), which consist of more than 3000 images categorised into 20 construction material classes.

Table 4.4 Classification of applied heuristics to obtain specific elements.

Heuristics Classification		Elements and Applied Heuristics
Geometry-based Dimensional and associative restrictions that can be applied on point clouds, segmented planes, objects, or volumetric representation.	Shape and dimension	Wall: dimensional constraints(Pu & Vosselman, 2009)(Macher et al., 2017) Wall: floor to roof distance (Arnaud et al., 2016) Column: dimensional constraints (Chen et al., 2017) Floor: altitude of plane (Arnaud et al., 2016) (Hübner et al., 2020) Floor: location of centroid (Díaz-Vilariño et al., 2015) Roof: altitude of plane (Arnaud et al., 2016) Roof:location of centroid (Díaz-Vilariño et al., 2015) Pipe: dimensional constraints (Bosché et al., 2015) Pipe: circular cross-section (Bosché et al., 2015) Window: rectangular (Arnaud et al., 2016) (Previtali, Díaz-Vilariño, & Scaioni, 2018) Door: narrow regions along the trajectory of the device (Franz et al., 2018) Door: rectangle shape constraints (Arnaud et al., 2016) (Quintana, Prieto, Adán, & Bosché, 2018) (Previtali et al., 2018) Opening: shape (Arnaud et al., 2016) (Antonio Adán et al., 2018) Secondary components : shape (Antonio Adán et al., 2018)
	Point density	Floor: point density (Chen et al., 2017) Roof: point density (Chen et al., 2017) Window: low point density (Arnaud et al., 2016) (Pu & Vosselman, 2009) Wall: 2D countour generation (Franz et al., 2018)
	Associative geometry	Opening: the location on the wall (Arnaud et al., 2016) Beam: plane projection and line fitting (Chen et al., 2017) Window & door: distance from the wall, ceiling and adjacent wall (Arnaud et al., 2016) (Previtali et al., 2018)
Rule-based These are the general heuristics most of the building elements follow.	Surface normal	Floor: surface normal opposite to direction of gravity (Arnaud et al., 2016) (Hübner et al., 2020) (Ochmann et al., 2019) (C. Wang et al., 2015) Roof: surface normal in direction of gravity (Arnaud et al., 2016) (Ochmann et al., 2019) (C. Wang et al., 2015)
	Orientation	Wall: orthogonal to ground (Arnaud et al., 2016) (Hübner et al., 2020) Wall: verticallity constraint (Pu & Vosselman, 2009)(C. Wang et al., 2015) (Hübner et al., 2020)(Ochmann et al., 2019) Window & door: alignment (Previtali et al., 2018)
	Principal axis	Door: axis constraints (Franz et al., 2018)
	Other	Window: never stand-alone (Pu & Vosselman, 2009) Door: never stand-alone (Pu & Vosselman, 2009)
Relationship-based Relationship-based heuristics can be time-based or space-based.	Time-based (e.g., the sequence of construction objects)	Sequential relationships (K. K. Han et al., 2015)

Heuristics Classification		Elements and Applied Heuristics
	Space-based (e.g., connections between objects or spaces).	Wall: intersects with ground (Pu & Vosselman, 2009) (K. K. Han et al., 2015) Roof: on top of walls and intersect on top of walls (Pu & Vosselman, 2009) Window: positioned in walls (Pu & Vosselman, 2009) (K. K. Han et al., 2015) Door: positioned in walls (Pu & Vosselman, 2009) (K. K. Han et al., 2015) Opening: the location on the wall (Arnaud et al., 2016)
Model-based Uses as-planned BIM model to detect elements using overlap with as-built data.	Derived from existing model	Pipes: alignment and comparision (Bosché et al., 2015) Elements: voxel occupancy (Golparvar-Fard et al., 2015) Elements: ray thresholding (Marianna Kopsida & Brilakis, 2020) Elements: commercial software (Pučko et al., 2018) Elements: thresholding (Alex Braun et al., 2020) Elements: machine learning (Pour Rahimian et al., 2020) Secondary components: RGB comparison (Antonio Adán et al., 2018) Door: RGB comparison (Quintana et al., 2018)

There are two approaches to generate the data. The first approach is by generating realistic, rendered elements-based graphic models to create synthetic data. The second approach is to capture and process real-world data with the high computing power of state-of-the-art data acquisition devices.

If the challenges mentioned earlier are solved, the learning-based techniques open doors to autonomous element detection and identification for robust progress monitoring.

C. Hybrid Approaches

As discussed, both the heuristics-based approach and the learning-based approach face a set of challenges. While the former requires extensive hard coding of the domain knowledge, which is complex and challenging to define for all situations faced in construction, the latter approach requires extensive training datasets, higher computational resources, and knowledge of advanced computing techniques. Though heuristics-based approaches may work for general geometric components, they fail to perform when the construction scene becomes complex and do not follow conventional rules, relationships, and geometries. Learning-based approaches are more adaptive, as they can generalise on-site data and are not based on structured domain knowledge. This comparison will also be discussed in detail in section 5.3.2.

Therefore, to utilise the advantages of both the approaches and overcome their individual shortcomings, there is a need to explore the hybrid approach, i.e., using heuristics in conjunction with learning algorithms.

4.4 PROGRESS MONITORING

Progress monitoring requires a visual or quantitative comparison between the as-planned and as-built models. Four progress monitoring levels identified as a part of this study are L-1, L-2, L-3, and L-4, as discussed in section 2. The following sub-section discusses the key technologies and methods used for implementing these levels.

4.4.1 Progress visualisation and comparison (L-1 and L-2)

L-1 based monitoring is currently the predominant method used at construction sites and only requires visualising the 3D as-built model built from spatial data obtained from the site. As an extension to L-1, L-2 based monitoring requires a visual comparison between as-planned and as-built status at a given time. This is also used on several projects and explored by numerous research studies.

These visualisations or comparisons can be made using selected environments, as shown in Figure 4.1 (e). The literature review in previous chapter, shows that most pipelines have used non-immersive environments such as a 3D viewer or web-based viewer. A few pipelines have used immersive Extended Reality (XR) environments, which include Augmented Reality (AR) (Golparvar-Fard et al., 2011)(Zollmann et al., 2014), Virtual Reality (VR) (Pour Rahimian et al., 2020)(Vincke et al., 2019), and Mixed Reality (MR) (Khairadeen Ali et al., 2021)(Marianna Kopsida & Brilakis, 2020).

While a few studies used traffic light-based coding to show progress (Golparvar-Fard et al., 2011), the use of comparative and interactive UI has also been investigated (Zollmann et al., 2014). Interactive UIs are gaining popularity, as they can pass information among stakeholders and assign tasks or comments for the workforce by adding features like annotating PDF notes into the environments. Advancements in gaming engines can be used to represent point clouds and BIM models by importing and overlaying in a single VR environment utilising Unity3D (Pour Rahimian et al., 2020)(Vincke et al., 2019).

Based on the review of literature and practice, this study finds that the capability of visualisation environments is underutilised for progress monitoring applications. With advancements in computing skills, rapid development in immersive technologies, and required bandwidths supported by the 5G standard, these platforms can provide the required information in real-time. Focused research is required to generate interactive UI and integrate AR and VR environments to give the user the best of both worlds and utilise gaming engines to take progress visualisation to the next level.

4.4.2 Progress quantification, schedule updating, and notifications (L-3 and L-4)

The next level of progress monitoring, L-3, involves quantifying the completed elements. Therefore, in continuation to Table 4.2, Figure 4.4 shows a trade-off for selecting the quantity estimation method corresponding to the type of model generated. Here, on the left-hand scale, the green-to-red colour variation indicates the increase in complexity from low to high. Similarly, the green-to-red colour variation indicates the decrease in the ease of progress quantification from high to low on the right-hand scale.

While the IFC/BIM models can produce direct BoQs from IFC-based element geometries (Mahami et al., 2019), other models require comparison with the as-planned model. For surface models and mesh models, object detection (based on models overlapping) (Pour Rahimian et al., 2020), detection by ray thresholding (Marianna Kopsida & Brilakis, 2020), or enclosed volumetric comparison (H. Omar et al., 2018) can be made for estimating quantities of completed elements.

For comparison using voxel occupancy, the as-planned model is voxelised and overlapped with as-built point clouds, after which thresholding is done to detect the occupied voxels (Golparvar-Fard et al., 2015). Although it works well for smaller models, it may be computationally costly for larger models; hence, it is not a preferred option. For direct usage of the point cloud, the most straightforward and widely used method is Scan-vs-BIM, which uses thresholding-based element identification. This method requires an as-planned 3D BIM and has its shortcomings (M. Bassier et al., 2019), but the as-built model generation step can be skipped, saving computation. Various Commercial project management software are also used for detecting deviations between as-planned and as-built models that can result in the quantification of contractual scheduling metrics (Pučko et al., 2018)(Vincke et al., 2019).

The selection of a model for progress monitoring depends on the project's characteristics (type, budget, duration), the required level of progress monitoring, and other factors. A detailed study of these factors is required to converge onto an appropriate pipeline for progress monitoring based on the project-specific use case.

For L-4 progress monitoring, updating schedules and generating reports/notifications are included with progress quantification. This is challenging; therefore, as shown in literature in previous chapter, only a few pipelines have explored this area. Though

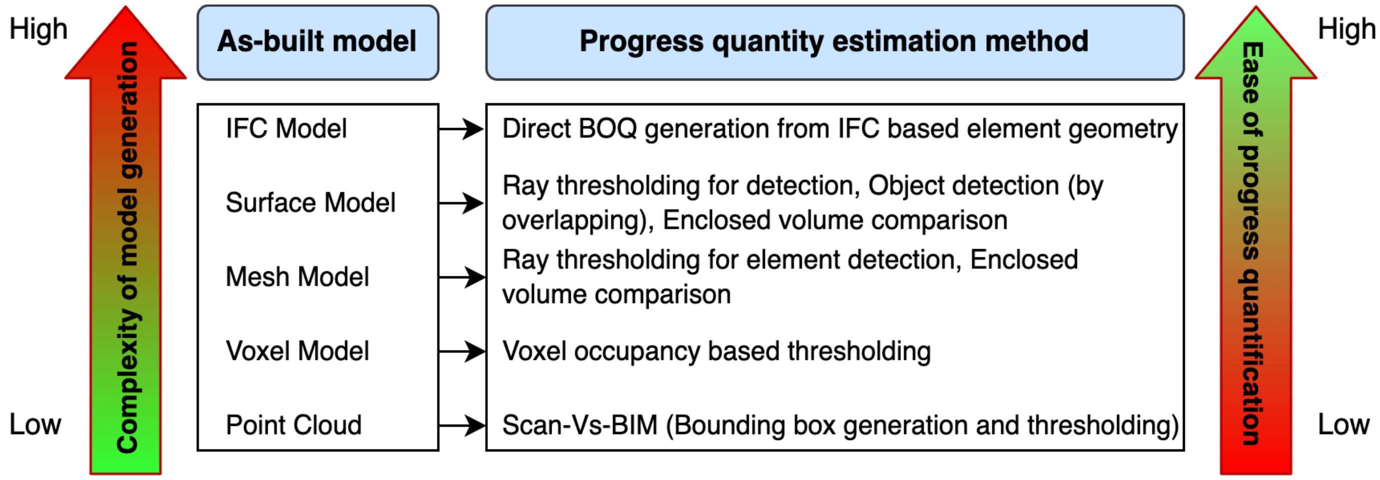


Figure 4.4 Trade-off for selecting Quantity Estimation method corresponding to the type of as-built model for progress quantification.

schedule updating has not been fully automated yet in practice, conceptual schemes have been proposed (Son, Kim, & Kwon Cho, 2017).

Some schemes have utilized the concept of precedence relationship graphs and articulation points, where all objects must be finished before the element linked to the articulation point can be built (A. Braun et al., 2014). Few studies have compared as-planned and as-built BIM for schedule updating (Mahami et al., 2019)(Hamledari, McCabe, Davari, et al., 2017). As the as-built IFC/BIM model is updated, parametric information has also been utilised to generate the updated schedules automatically (Hyunjoo Kim, Anderson, Lee, & Hildreth, 2013).

The primary challenge in schedule updating is that schedule LoD is different from BIM LoD, resulting in mismatch and incompatibility. Secondly, it requires an as planned 4D model for smooth implementation, that is not always available for a construction project. Other challenges include deciding the frequency of updating 4D-BIM. Although a few research papers on CV-CPM have mentioned the concept of schedule updating in their pipelines, they have not specified the details. Therefore, there is a necessity to conduct and integrate focused studies for tackling these specific challenges.

Likewise, only a few studies have explored an automatic notification system that can trigger notifications while monitoring and controlling a construction site (H. Omar et al., 2018). These can be obtained via SMS or e-mail, containing warnings, reports, or graphical outputs based on the settings and LoD required. The frequency and content of notifications should be selected to give construction managers broad insights about areas to target for meeting schedules. As discussed earlier, providing managers with real-time information/alerts about any deviation in the schedules or discrepancies in the as-built model from the as-planned model is essential for in-time decision making and can improve the project's performance.

In addition to providing updated information on progress status, features to suggest interventions and control strategies and forecast resulting outcomes will automate project controls even further.

Conventional control decisions taken by the project managers are based on the know-how of issues specific to the project. Automated control suggestions can be generated using cognitive computing on the progress status and the data from sources such as the organisation's Enterprise Resource Portal (ERP). Such an approach can recognise patterns in activities that suffer delays and suggest interventions to bring the project back on track. There have been selected deployments using AI-based simulation engines for addressing complications inherent in scheduling and project coordination (ALICE Technologies, 2021). Such deployments can also evaluate alternatives and provide real-time suggestions. This is an emerging area with significant potential, and further exploration is required.

This section presented a micro-analysis of the stages of the CV-CPM framework. Various gaps were identified while discussing each section in detail. The following section discusses the application of the CV-CPM framework to enable broader evaluation of various pipelines/components to develop benchmarks. It also summaries specific challenges identified within each stage.

4.5 DISCUSSION

As presented and discussed in the previous section, numerous combinations of input-process-output options are possible at each stage of the CV-CPM framework. For an integrated pipeline across all stages, the combinations of options are immense. Given the pace at which technology is developing, the number of options will continue to increase. However, only a limited combination of options will be feasible, and this research has proposed guidelines to structure these combinations.

A systematic approach is essential to explore and document the performance of specific approaches/algorithms and integrated CV-CPM pipelines. This section discusses requirements for benchmarking to enable a systematic study. The section also summarises potential research questions within each stage of CV-CPM and discusses its relevance to support the emerging area of Digital Twin.

4.5.1 Benchmarking

Benchmarking evaluates pipelines, algorithms, devices, and tools using standard datasets and test methods. Figure 4.5 outlines the broad requirements to benchmark pipelines and components of CV-CPM, and these are discussed below:

1. **Datasets:** Though existing datasets contain relevant information for building elements (Dai et al., 2017)(Armeni et al., 2016), these represent completed elements and not the schedule-based construction progress of the elements as required for progress monitoring. Hence, there is a requirement for creating an open-source dataset that contains the data on elements as construction progresses. As shown in Figure 4.5, these can be synthetically created or acquired from the real world. The dataset type will change based on the particular stage under investigation. For example, for 3D reconstruction, the type of dataset required for benchmarking reconstruction algorithms will be in the form of raw images, videos, and depth images.
2. **Testbeds:** While existing studies in the area have contributed significantly to knowledge about the applicability of specific algorithms and tools used, the experimental set-ups and procedures used for these investigations vary. As a result, repeatability, and comparison of results across research groups are not feasible.

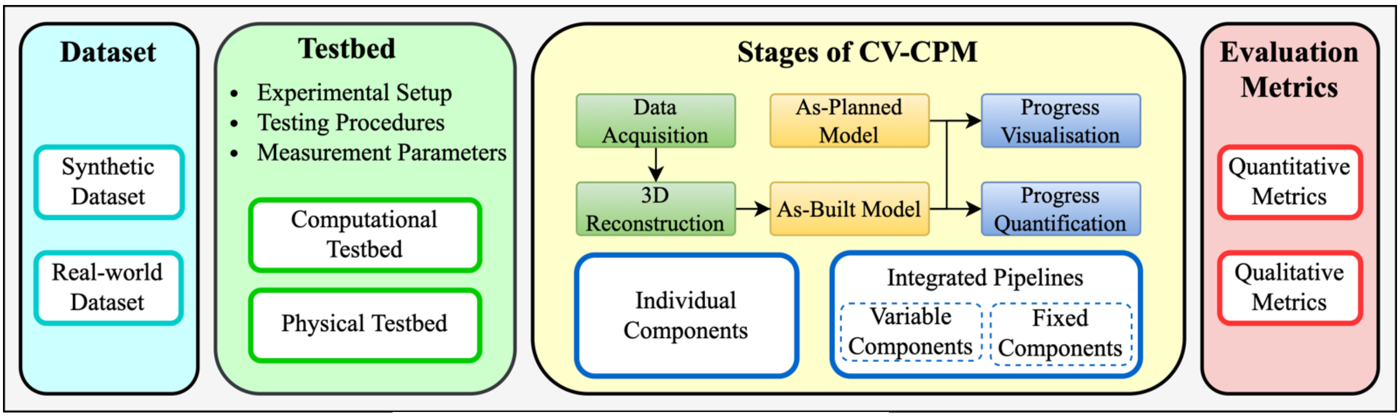


Figure 4.5 Benchmarking of CV-CPM

Hence, for each CV-CPM stage, there is a need to design standard testbeds that define the experimental setup, procedures, and measurement parameters. As shown in Figure 4.5, the testbed can be either computational for benchmarking software (algorithms) or physical for benchmarking the hardware (devices/sensors). For example, a testbed for data acquisition can have sensors capturing data with various physical constraints, like the speed of capture, distance from an object, lightning condition, resolution settings etc., that can be varied to perform multiple experiments. Further, the testbeds should be flexible to accommodate the requirements for collaborative explorations in the area.

3. **Component of CV-CPM:** As shown in Figure 4.5, for CV-CPM, the components can be benchmarked independently or as pipelines integrating multiple components. Each component can be evaluated for its independent performance. Additionally, the pipelines can also be benchmarked for their effectiveness for overall progress monitoring by varying specific components within the pipeline. This will standardise the approach to sequence and investigate various input-process-output combinations and enable a systematic comparison of results with other studies.

4. **Evaluation Metrics:** There are several factors to be considered for evaluation. While most benchmarking studies within this framework require the definition of quantitative metrics, qualitative measures will also be required to evaluate subjective outcomes. For example, for benchmarking visualisation environments, qualitative factors like ease of use, skills required for navigation, immersive features for comparison, and training requirements can be compared from a user feedback-based evaluation. Correspondingly, for benchmarking the algorithms, the

evaluation metrics can be in the form of accuracy, processing time, computational cost etc., which are quantitative.

The integrated CV-CPM framework and the guidelines proposed for each stage is expected to assist in developing a strategy and a roadmap for benchmarking. As a wide range of studies is required, the technology roadmap needs to be developed collaboratively by the research community. A starting point for these studies can be to investigate the subjective ratings proposed in this study and quantify these ratings through controlled experimental testbeds. The CV-CPM framework can help in structuring and prioritising the areas to be explored in developing a roadmap so that they form a standard reference for benchmarking studies.

4.6 ALIGNING THE RECENTLY PUBLISHED WORKS ON THE DEVELOPED CV-CPM FRAMEWORK

Table 4.5 Aligning the recently published work to the CV-CPM framework.

Paper Reference	Data Acquisition	3D Reconstruction	As-built Modelling	Progress Monitoring	Level of Progress Monitoring	Elements Identified	Key Contribution- Remarks
(Kavaliauskas, Fernandez, McGuinness, & Jurelionis, 2022)	Manual, Laser Scanner	Direct Point Cloud	Point Cloud	Thresholding based approach using overlapping	L-3	Columns, Wall	alignment of point cloud data with the IFC and automatic object detection
(Halder et al., 2022)	UGV, 360 Camera	NA	360 Projected Image spheres	BIM based AR Overlap	L-2	Structural Elements	remote AR solution to provide a real time visual stream of the construction work registered/aligned with the 3D geometric model of the building that is extracted from the BIM model
(Puri & Turkan, 2020)	Manual, Laser Scanner	Direct Point Cloud	Point Cloud	Comparison of real and virtually generated point cloud	L-3	Bridge Components	construction progress monitoring using lidar and 4D design models

It can be seen from Table 4.5 that the recently published works can directly be aligned with the CV-CPM framework. Using this, these works can be reviewed and classified easily, as it becomes well-defined that which technology or method have been used at the particular stage in the pipeline. This also shows that the developed CV-CPM framework is comprehensive.

Table 4.6 Aligning the recently published works for benchmarking requirements

Paper Reference	**Dataset**	**Testbed**	**Evaluation Matrix**
(Kavaliauskas et al., 2022)	Real world	Computational	Quantitative
(Halder et al., 2022)	Real world	Computational and Physical	Quantitative and Qualitative
(Puri & Turkan, 2020)	Real world	Computational	Quantitative

It can be seen from Table 4.6 that the recently published works can also be alighen dfor the benchmarking requirements that are mentioned in the previous section.

4.7 CONCLUSION

CV-CPM has the potential of creating an immense impact by providing real-time, accurate, reliable information to construction managers. Though a significant amount of work has been done in the last decade, specific challenges remain due to the construction industry's dynamic nature and complexities at sites. Some of these challenges have been addressed; however, significant gaps need to be filled to make pipelines accurate and automated to meet rising user expectations of real-time feedback.

This study has comprehensively reviewed individual pipelines and formulated an integrated CV-CPM framework, shown in Figure 4.1. The proposed framework positions all the reviewed papers, including recent papers in this area. Hence, it can be inferred that it is holistic and robust.

In addition to the basic concepts and references to the work done in this area, Table 4.7 summarises the key contributions of this chapter for the micro-level stages of the framework with reference to other related works.

It was found that the four levels of progress monitoring identified in this study strongly influence the technology selection for the pipeline at each stage. For implementing progress monitoring on a construction project, it is recommended that the pipeline and components required are selected based on the chosen level.

Table 4.7 Summary of stage-wise specific contributions

Stage	Published review papers		Detailed contribution of this research
	Reference	Their contribution	
Data Acquisition	(T. Omar & Nehdi, 2016b)	Categorised, listed, and compared various data acquisition technologies and studies in detail.	• Identified and proposed factors and ratings that can be used to guide selection of sensors-mount combinations for data acquisition.
3D Reconstruction	(Z. Ma & Liu, 2018)	Presented and overall pipeline for 3D reconstruction using SfM. Categorised, listed, and compared techniques for SfM technologies and studied in detail.	• Comparison SLAM and SfM technologies for construction progress monitoring.
As-Built Modelling	(Pătrăucean et al., 2015)	Presented an overview of the heuristic-based modelling process and classification of categories.	• Defined types of as-built models based on existing work and proposed guidelines for selecting as-built model type for progress monitoring based on project requirements. • Detailed review of as-built modelling approaches and rationale for using a hybrid approach.
Progress Visualization	NA	NA	• Categorised progress monitoring into four levels and associates these levels as the requirements driver for deciding on options in the upstream stages of process.
Progress Quantification	NA	NA	

Among several potential research areas, advancements in the as-built modelling stage are required to facilitate the quantification of progress. To enable this, exploring a hybrid approach that combines learning with heuristics is recommended.

The need and requirements for benchmarking and future research directions derived from the CV-CPM framework have been presented in the Discussion section. Addressing these requirements and following a well-developed roadmap in this area is essential to move CV-CPM research from laboratory studies to field applications. [3]

[3] Parts of this chapter have been published in the following articles:

1. Reja, V. K., Varghese, K., & Ha, Q. P. (2022). Computer vision-based construction progress monitoring. *Automation in Construction*, *138*, 104245.

CHAPTER 5.

CONPRO-NET – A HYBRID SELF-SUPERVISED LEARNING ARCHITECTURE FOR PROGRESS ESTIMATION OF CONSTRUCTION PROJECTS

In the previous chapter, the CV-CPM framework have been discussed in detail. It was also discussed that a hybrid approach which involves both the heuristic and learning based approach should be explored for better and improved performance. Therefore, in this chapter a hybrid approach is utilized for progress estimation. This chapter contributes to objective 3 of this thesis. The overall methodology followed for objective 3 has been shown in Figure 1.2

This chapter is divided into six sub-sections. The first subsection introduces this chapter. The second subsection is a specific literature review about the works which have explored the quantification of construction progress, the segmentation and classification methods that have been already used and finally the specific research gaps for this chapter. The third sub-section presents the Hybrid self-supervised approach for progress quantification. The fourth section is on experimentation and results over the S3DIS dataset. The fifth sub-section is on the interpretation of results and discussion. Finally, the conclusions of this chapter are presented in sub-section six.

5.1 INTRODUCTION: AUTOMATED PROGRESS ESTIMATION

From the last chapter, it is clearly evident that, among several automated technologies, computer vision-based construction progress monitoring (CV-CPM) is one of the leading technologies being explored. Using 3D as-built point clouds as inputs, researchers have experimented with heuristics-based approaches, which involve applying geometrical constraints and therefore require a significant amount of hard coding and domain knowledge. Recently, learning-based approaches have been explored; nevertheless, the existing supervised approaches also require substantial effort for manual labelling of the training data and cannot be generalised for different construction projects.

Therefore, this research investigates combining heuristics and learning based methods and develops ConPro-NET, a novel hybrid self-supervised learning-based method for element identification from construction point clouds. Figure 5.1 shows the Framework of the proposed hybrid self-supervised approach for construction progress quantification. After pre-processing the unlabelled data, first, a customized distance thresholding-based approach is adopted on pre-processed point clouds for the unsupervised segmentation of elements. Next, these segmented objects are given as input to the feature extraction module, which uses a contrastive learning approach. Contrastive learning uses positive and negative pairs of the same object to learn their match to each other and hence learn the features of each object. The learnt features are refined by performing clustering and further they are augmented with a set of handcrafted features defined based on local geometric and visual properties to form the hybrid feature vector. Handcrafted features include the surface area, average z-coordinate, average R, G and B values and local covariance features (linearity, planarity, verticality, and scattering) of an element.

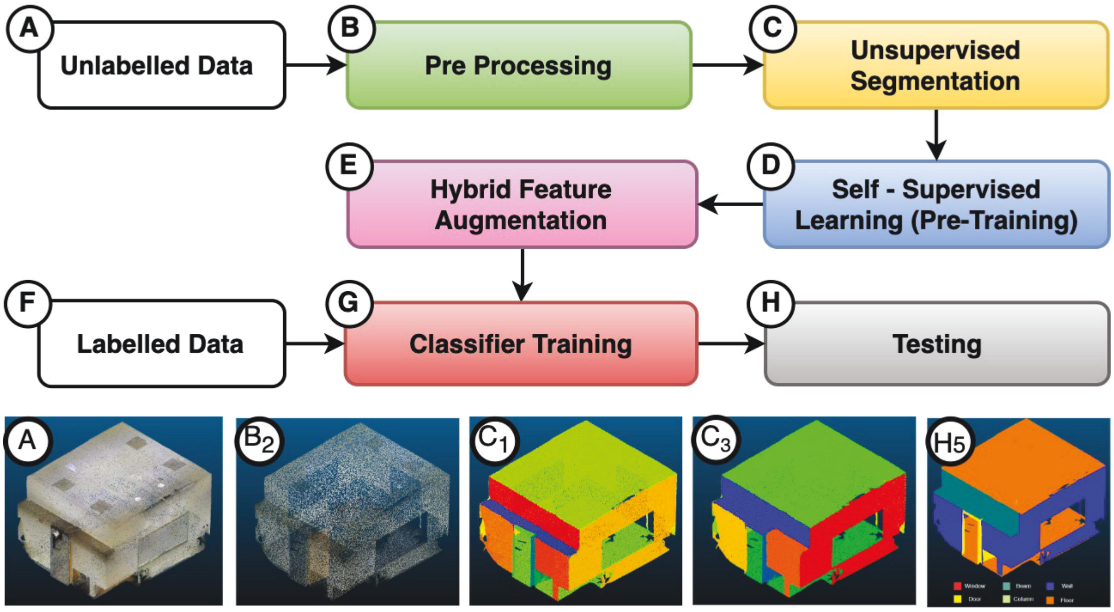


Figure 5.1 Framework of the proposed hybrid self-supervised approach for element identification – ConPro-NET

The pre-training and feature engineering step is followed by the downstream task, classification in this case with a test set which has labelled data for six object classes, i.e., wall, beam, column, door, window, and slab. The classification model is trained and evaluated on the remaining S3DIS dataset and a stage-wise collected data for progress monitoring of an under-construction building. The classification model is evaluated on matrices such as precision, recall, and F1-score. The results show that the proposed hybrid-self-supervised approach has achieved an overall classification accuracy of 80.86% on the S3DIS dataset and 80.95% on the case study dataset.

5.2 RESEARCH METHODOLOGY – OBJECTIVE 3

The research method commonly employed while designing a computer vision pipeline iteratively is known as the "iterative design" or "iterative development" method. This approach involves a cyclical process of designing, implementing, testing, and refining the computer vision pipeline in multiple iterations to progressively improve its performance and address any limitations or challenges that arise during the development process.

In the context of computer vision, an iterative design approach allows to gradually enhance the pipeline by incorporating feedback from testing and evaluation stages. Each iteration involves refining the pipeline's algorithms, adjusting parameters, fine-tuning models, and incorporating new insights gained from the analysis of results. This iterative process continues until the desired performance or objectives are achieved.

The iterative design method enables researchers to incrementally improve the accuracy, efficiency, and robustness of the computer vision pipeline. It also allows for flexibility and adaptability, as adjustments and modifications can be made based on the evolving requirements and challenges encountered during development.

Hence an iterative design method was deployed by conducting multiple trails before finally freezing the ConPro-NET pipeline.

5.3 QUANTIFICATION OF CONSTRUCTION PROGRESS: RELATED WORKS

In the past, several technologies have been explored for construction progress monitoring. In literature these technologies have classified under six broad categories, these are: conventional IT based, tag-based, geospatial-based, BIM-based, computer vision based, and extended reality based. The scope of this research and the reviewed work is limited to computer vision-based technologies (CV-CPM).

Most of the existing methods for progress monitoring have used an image based 2D comparison using computer vision to monitor progress. However, these approaches detect the presence of the elements based on comparison with existing models or drawings (Arif & Khan, 2021b). The have limited knowledge representation and are dependent on estimating pose of the object in the image which is a complex and requires additional information to solve in a featureless scene. Therefore, the recent methods have shifted towards 3D datasets and corresponding approaches.

For 3D data, the computer vision-based construction progress monitoring (CV-CPM) framework holistically defines the entire process. It consists of three key steps, i.e., data acquisition with 3D reconstruction, as-built modelling, and progress estimation. Recent studies on progress monitoring have focused on various tools, technologies, and methods in these pipelines (Ekanayake et al., 2021b)(Alizadehsalehi & Yitmen, 2021).

The first step towards element identification is to segment the point clouds. These segments of points are then clustered together to form a meaningful object. These objects are then classified by applying a heuristics-based approach or a learning-based approach. These steps are discussed in detail in the following subsections.

5.3.1 Segmentation of point clouds

After 3D reconstruction which was discussed in section 4.2.2 and Table 4.1, Segmentation of the 3D point cloud is a first step towards element detection. Data-driven methods such as Random Sample Consensus (RANSAC) (Kang et al., 2020)(Yang, Cheng, & Wang, 2020) and Hough Transform (HT) (Rausch & Haas, 2021) or region-growing (Maarten Bassier, Bonduel, Van Genechten, & Vergauwen, 2017)(Khaloo & Lattanzi, 2017) have been predominantly adopted for segmenting points generated in indoor building environments. These methods use geometrical properties of the data for performing segmentation, in some cases knowledge driven methods are also used. Another approach used three defined heuristics for segmentation of large point clouds, these were the distance threshold for the region growing, the threshold for the minimum number of points needed to form a valid planar region, and the decision criterion for adding points to a region (Poux, Mattes, Selman, & Kobbelt, 2022).

On the other hand, unsupervised segmentation approaches which have been used successfully include Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Makantasis, Doulamis, Doulamis, & Ioannides, 2016)(B. Wang, Yin, Luo, Cheng, & Wang, 2021). Extending this, Hierarchical DBSACAN (HDBSCAN) with KMeans for clustering was also used to boost performance for segmentation of indoor scenes (J. W. Ma & Leite, 2022). RandLA-Net, a deep learning-based encoder-decoder network, was used to segment the scaffolds from the acquired point clouds (J. Kim, Chung, Kim, & Kim, 2022).

Most of the existing studies work on point-level based segmentation approaches. These approaches require an additional step of extracting geometric elements from the segmented data using neighbour network approach which depends on the accuracy of the previous step (Hyunsoo Kim & Kim, 2021). Therefore, going forward applying deep learning networks on object-level features are expected to be much adaptable than point-level features from machine's perspective, this has also been validated (J. W. Ma & Leite, 2022). Therefore object-level classification approaches are discussed in the next sub section.

5.3.2 Classification of segmented elements

For CV-CPM once the as-built point cloud is segmented, and clusters are formed, these clusters of points are classified and measured to detect the progress. As discussed and shown in Section 4.3.2.3, Table 4.3 and Table 4.4, there are two key approaches found in the literature to facilitate this classification:

(a) Heuristic-based Approach: These are in form of geometrical constrains, conventional rules, spatial and temporal relationships or constraints derived from a model (Maalek et al., 2019)(K. K. Han et al., 2015).

(b) Learning based Approach: They use data to learn features using neural networks and then predict the object class (Perez-Perez, Golparvar-Fard, & El-Rayes, 2021).

Both the above approaches have their respective advantages and disadvantages. The heuristic-based approaches require to hardcode constraints, rules and relationships which is complex and laborious. The learning-based approach requires high amount of labelled data for each element class and are computationally complex. This study focusses on the usage of this hybrid approach as it has not explored earlier. The next subsections discuss the usage of point cloud-based heuristics and learning approaches used in literature.

5.3.2.1 Heuristic Based Approaches

Primarily three types of heuristics have been used in previous literature. Geometry based approaches have used shape and dimension (Bosché et al., 2015)(Macher et al., 2017)(Antonio Adán et al., 2018)(Díaz-Vilariño et al., 2015)(Quintana et al., 2018), point density of the point clouds (Pu & Vosselman, 2009)(Franz et al., 2018), associative geometry (Arnaud et al., 2016)(Chen et al., 2017)(Previtali et al., 2018) to recognize elements. Rule based approaches uses rules such as surface normal (Hübner et al., 2020)(C. Wang et al., 2015)(Arnaud et al., 2016), orientation (Pu & Vosselman, 2009)(Ochmann et al., 2019)(Previtali et al., 2018) and direction of principal axis (Franz et al., 2018). Relationships based approaches have been utlized by applying spatial (K. K. Han et al., 2015)(Pu & Vosselman, 2009)(Arnaud et al., 2016) as well as temporal relationships (K. K. Han et al., 2015).

Though this is a widely utilized approach for element identification, it requires a large amount of preliminary information about building elements to be fed prior. In case the information is fed manually, some elements which cannot be defined as simple shapes can go unrecognized (Zeng et al., 2020). It works well if the as-planned 3D BIM model is available so that the geometric properties can be extracted easily. This approach may lead to errors and ambiguity if multiple objects have a similar geometry.

5.3.2.2 Learning based Approaches

Learning-based approaches uses machine learning to train 3D feature descriptors to classify the point cloud data into different categories. Local descriptors are used to store features at the point level, while global descriptors are used to store features at the object level. Hence, there are two techniques in feature-based learning methods. First is point-level classification, which extracts each point's local features, then classifies them individually into an object category. The other technique is a segment-level classification, which first requires subdividing the point cloud into meaningful segments corresponding to different building components. Then, these segments are classified into different object classes.

The object level 3D detection of class-specific building primitives from point cloud scans is essential for Scan-to-BIM (Y. Xu, Shen, & Lim, 2021). Most of the Scan to BIM approaches depends on detecting detect geometric and semantically rich features (Z. Ma & Liu, 2018).

Recently deep learning techniques have endured considerable progress in applications involving the understanding of 3D scenes, which include instance segmentation (Jiang et al., 2020) and object detection (He, Zeng, Huang, Hua, & Zhang, 2020). A likely solution to 3D modelling problems is to leverage the advanced object-level feature encoding capability of deep convolutional neural networks (DCNNs). The majority of learning-based approaches are based on point cloud segmentation (Chen et al., 2019)(Perez-Perez et al., 2021)(Zeng et al., 2020).

PCIM by Park et al. can automatically recognize construction objects and their properties with supervised deep learning approaches. Furthermore, it can store information in the original point cloud data with a hierarchical structure, rather than converting it to a solid or rigid model. The PCIM framework generates XML files to represent detected object and their properties while preserving the original point cloud data (Park & Cho, 2022). Chen et al. (Chen et al., 2019) converted the point cloud into a graph representation, where vertices represented points and edges represented connections between points within a fixed distance. Then an edge-based classifier and a point-based classifier were successively used to determine the type of building element. Finally, the detected object was matched with the corresponding BIM entity based on the nearest neighbour. Iwaszczuk et al. (Iwaszczuk et al., 2018) studied RGBD information's influence to label structural elements in an indoor scene using an encoder-decoder CNN framework. This architecture also worked in the fusion of RGB and RGBD by taking advantage of redundancy in the information. Xu et al. (Y. Xu, Shen, Lim, & Li, 2021) presented a two-stage 3D object-detection method using region-based convolutional neural networks (R-CNN). Scan2BIM-NET (Perez-Perez et al., 2021) used: two convolutional neural network (CNN) and one recurrent neural network (RNN) for semantically segmenting the structural, architectural, and mechanical components present in point cloud data. It classifies beam, ceiling, column, floor, pipe, and wall elements.

5.3.3 Semi, self, and un-supervised learning

In the modern times, where data driven approaches are taking an edge over traditional approaches, the bottleneck situation is to find the correctly labelled data. Hence, in the case of learning-based methods, unsupervised, self-supervised and semi-supervised learning methods carries equal importance as supervised learning schemes, particularly when the labelled data is unavailable.

Unsupervised learning is a classical machine learning approach where labels for input data is not required for a model to learn (Hastie, Tibshirani, & Friedman, 2009). In this approach no corresponding output labels for input data are available. Instead of generating a mapping between input and output labels in supervised methods, the unsupervised schemes try to model the underlying distribution or structure of the data. These methods broadly revolve around two kinds of algorithms: clustering based unsupervised schemes; association rule based unsupervised schemes. Clustering based schemes try to take advantage of neighbourhood properties of the input data points and group them on the basis of their similarity or differences. Association rules try to find the relationship between several variables in the dataset, for example, occurrence of one event 'a' led to occurrence of another event 'b'.

Semi-supervised is a machine learning approach which utilises small amount of labelled data with a large amount of unlabelled data (Jaiswal, Babu, Zadeh, Banerjee, & Makedon, 2020; Zhu, 2005). It can be considered as an intermediate case of unsupervised learning and supervised learning where the target is to improve the performance of unsupervised models with small amount of labelled data while skipping the expensive process of getting the entire dataset labelled.

Self-supervised learning methods is another machine learning approach within contrast with semi supervised method which do not require any labels and utilise the underlying structure of the data to predict the outcomes (Jaiswal et al., 2020). It automatically learns the representation within the data and attempts to solve the tasks designed for supervised learning methods.

In construction, finding labelled data in a practical scenario is difficult and working on approaches which analyse building point cloud in an unsupervised, self-supervised or semi-supervised manner is the need of the hour. Hence this study proposes to use combination of unsupervised and self-supervised learning methods for segmentation and classification of components respectively.

5.3.4 Research Gaps, Objectives & Contributions

Regarding progress monitoring and in particularly identification of construction elements, the current research has identified following directions for the recognition and classification of point cloud objects:

- The aforementioned methods rely on supervised learning and necessitate a large, labelled dataset to train the neural network for feature extraction. However, due to the unique nature of construction elements and the potential for significant variations, obtaining labelled data for all element classes is impractical.
- These methods rely solely on deep learning features and have yet to explore handcrafted features, which could potentially aid in distinguishing elements from one another with greater ease.

- Traditional methods of object segmentation that involve edge-based or region growing techniques cannot be directly applied to under-construction data, as visible surfaces are often grey in colour, and element shapes are three-dimensional with multiple planar sides, making them non-planar.

Therefore, to address these potential research gaps, this research contributes by proposing a hybrid self-supervised approach. The work is towards fulfilling the third objective of this thesis with its sub-objective.

Thesis Objective 3: To develop a pipeline using hybrid self-supervised approach for automatic capture of constructed elements from point clouds and using it for progress monitoring in construction.

- To study the combined used of heuristic and learning based approaches (hybrid approach).
- To develop a customised approach for object segmentation specifically for construction
- To utilize concepts from feature engineering and to improve the performance of the method.

The key contributions of this study are mentioned below:

- An unsupervised segmentation approach has been developed that utilizes plane segmentation and distance-based thresholding to merge them to form meaningful objects.
- A self-supervised feature learning and classification algorithm has been developed based on contrastive learning to predict the object classes.

- Handcrafted features have been identified and incorporated into the deep learning-based features, resulting in a hybrid approach that improves performance.
- The application of ConPro-NET has been demonstrated on an under-construction building dataset.

5.4 CONPRO-NET – A HYBRID SELF-SUPERVISED APPROACH FOR PROGRESS QUANTIFICATION

The schematic diagram of the ConPro-NET for assessing the progress of construction projects is illustrated in Figure 5.2. The proposed framework processes unlabelled data after pre-processing and performs unsupervised segmentation of various 3D components of the construction point clouds. The resulting segmented components are randomly partitioned and then used as input to a self-supervised learning pipeline. A classification framework, ContraSim model, is trained using the parts of different objects to learn the similarity and dissimilarity between pairs of parts, labelling them as positive or negative pairs. The learned features are augmented with handcrafted features extracted using the geometrical properties of the point cloud objects. Unsupervised clustering is then applied to the features, and cluster IDs are assigned as pseudo-labels to the data.

Subsequently, the original data and their pseudo-labels are used to train another classification model, Clusterify model, which performs a 6-class classification task to refine the quality of the features generated by the previous model. This step improves the learned features for the point cloud data and augments them with handcrafted features to capture geometric details more precisely. The pre-training step allows the learned features of the point cloud data to be used for downstream tasks with small amounts of labelled data, such as classification, segmentation, and object recognition. In this study, the downstream task of point cloud object classification is performed on new and unseen construction data. Each component of the proposed framework is discussed in detail in the following subsections.

5.4.1 Dataset

In deep learning, large datasets are required for pre-training the initial weights for unsupervised feature extraction. Pre-training a model with large datasets can allow the model to learn intrinsic features of a specific type of data, which can later be used for downstream tasks such as classification and segmentation. However, a large dataset for construction data is currently unavailable. The building dataset, which contains finished or completed elements, is the closest representation of construction data. The ScanNet (Dai et al., 2017) and Stanford 3D Indoor Scene (S3DIS) datasets (Armeni et al., 2016) are the two available large building datasets.

The ScanNet dataset is an indoor RGB-D video dataset that only contains walls, floors, windows, and doors as relevant classes to construction. On the other hand, the S3DIS dataset consists of six large-scale indoor areas with 271 rooms, where the components are categorized into 13 categories, with seven classes being relevant to the problem: floor, wall, beam, column, door, window, and ceiling. Therefore, the S3DIS dataset was used for pre-training without labels (as shown in Figure 5.2 A). The distribution of these classes across the six areas is presented in Table 5.1. Additionally, to avoid confusion between roof and floor elements, they were merged into a single class named "slab".

Among the six areas, areas 1, 3, 5, and 6 were used for training, while areas 2 and 4 were used for testing purposes.

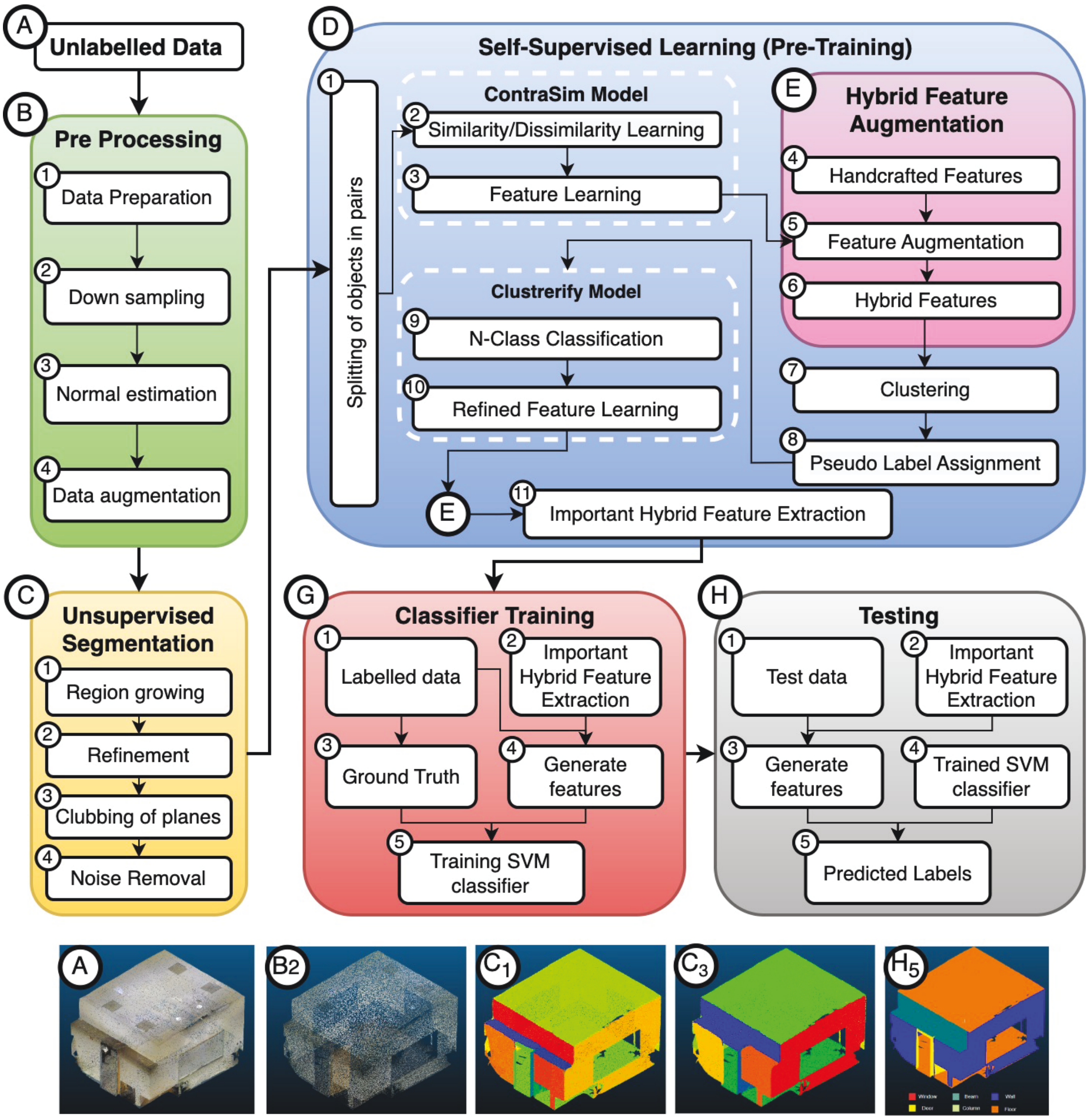


Figure 5.2 Framework diagram of ConPro-NET: the proposed hybrid self-supervised approach for construction progress quantification

Table 5.1 Area-wise and class-wise distribution of elements in the S3DIS dataset

Areas	Total Rooms	Beam	Ceiling	Column	Door	Floor	Wall	Window
Area 1	44	61	55	57	86	44	234	29
Area 2	40	11	81	19	93	50	583	8
Area 3	23	13	37	12	37	23	159	8

Areas	Total Rooms	Beam	Ceiling	Column	Door	Floor	Wall	Window
Area 4	49	3	73	38	107	50	282	40
Area 5	68	3	76	74	127	68	343	52
Area 6	48	68	63	54	93	49	247	31
Total	**272**	**159**	**385**	**254**	**543**	**284**	**1848**	**168**

5.4.2 Pre-processing

The data samples taken from S3DIS dataset are pre-processed before giving them to feature extraction models. As shown in Figure 5.2 (B). The steps taken for pre-processing the data samples, such as data preparation, down sampling, normal estimation, and data augmentation are described below.

5.4.2.1 Data Preparation

To prepare the S3DIS dataset for feature extraction, all buildings from the six distinct areas undergo a point removal process to exclude irrelevant classes. The S3DIS dataset comprises a total of 13 classes, namely clutter, ceiling, floor, wall, beam, column, window, door, chair, table, bookcase, sofa, and board. However, for the current study, only classes relevant to construction environments, including wall, door, roof, floor, window, beam, and column, are utilized. Therefore, all points associated with classes other than these seven are removed in this step.

5.4.2.2 Down Sampling

Construction point cloud datasets are usually very large and projects face, data transfer, storage, and computational challenges (Boje, Guerriero, Kubicki, & Rezgui, 2020. Therefore, to avoid high computation and processing time, all the building point clouds were down- sampled with a distance threshold of 0.015m. The Sub- sampling tool in cloud compare was used to down sample the point clouds. It works by sampling points from the point cloud such that the distance between two points in the point cloud is not less than the threshold specified. The value of the distance threshold was selected such that no element after down sampling had lesser than 2048 points, which is one of the hyper-parameters for contrasting based approach and will be discussed later. Although this is an optional step, it is utilized it in this study to so as to attain a practically implementable pipeline.

5.4.2.3 Normal Estimation

Estimating normal for unorganised point clouds and point clouds with noisy edge could be a challenging task. In order to estimate normal for real world data, the method should be able to robust to noise occurring and other outliers, and it should be sensitive to sharp feature while being computationally efficient.

For the proposed pipeline, the normal estimation algorithm is adopted from the work proposed by (Boulch & Marlet, 2012), which uses Randomized Hough Transform for estimating normal from the point clouds. Randomized Hough Transform (RHT) proposed in (L. Xu, Oja, & Kultanen, 1990) is a variation of conventional Hough Transform, widely used for detection of lines, curves, and other shapes. RHT is proposed to overcome the issues in the Hough Transform such as high computational complexity and low detection accuracy. The normal estimation algorithm is described as:

Algorithm: 5.1 – Algorithm for Normal Estimation

procedure $Normal\ Estimation(P)$
  if $dist(P, Edge) \geq Threshold$ **then**
    $N_p \leftarrow p_1, p_2, p_3$
    $Normal \leftarrow N_p$
  else if $dist(P, Edge) \leq Threshold$ **then**
    $Dom_{N_p} = dominant(N_{p1}, N_{p2})$
    $Dom_{N_p} \leftarrow p_1, p_2, p_3$
    $Normal \leftarrow Dom_{N_p}$
  end if
  return $Normal$

In the Algorithm 5.1, if the selected point P on a piece-wise planar surface is far away from any sharp feature or edge on the surface, then 3 points $p1$, $p2$, $p3$ are picked in the neighbourhood N_P, which defines a planar patch and hence normal can be selected accordingly. In another case, if the selected point P, lies near to a sharp feature or edge, then the neighbourhood Np is partitioned into two planes, $Np1$, $Np2$ with respect to the position of 3 points, and the dominant one is chosen for selecting the normal with point P. A plane is considered as dominant if all 3 points are contained by that plane.

5.4.2.4 Data Augmentation

To avoid the class imbalance, data augmentation was performed by creating desired number of building elements for each of the six classes. Data Augmentation was carried out on the building elements from the original S3DIS dataset and not on the building elements obtained after the segmentation pipeline to make sure that the newly added elements are free from any errors. The data was augmented by subjecting each building element to rotation, jittering, scaling, and shifting. The rotation was done only about the z-axis to maintain the orientation of the elements intact. Jittering is done on every point, and it adds a very small variation to the location of each point. Each building element is then scaled by a factor between 0.8 to 1.25 and is further shifted along all 3 axes by the same amount which lies between -0.1 meters to 0.1 meters. The details of elements used at each stage for the S3DIS dataset after data augmentation are shown in Table 5.2.

Table 5.2 Statistics of the dataset used at each stage for the S3DIS dataset after data augmentation.

Stage	**Pretraining**	**SVM Training**	**Testing**
Data	**S3DIS**	**S3DIS**	**S3DIS**
Wall	1500	300	76
Door	1199	240	60
Slab	1500	301	76
Window	1202	240	60
Beam	1201	115	61
Column	1201	240	61

5.4.3 Unsupervised Segmentation

Automatic shape segmentation is an essential step while analysing the building point clouds. The large point cloud scans of buildings have multiple components that are required to be identified and segmented for further processing. The automatic shape segmentation could be performed in a supervised or unsupervised fashion depending upon the availability of data and resources. An unsupervised or self-supervised approach of learning with no or less labels is always beneficial in case of less or no labelled data.

Getting labelled data is a very expensive process and creates bottleneck situation in any learning-based project. Hence as shown in Figure 5.2 (C), an attempt is made towards approaching the problem of segmentation with no labelled data in an unsupervised manner.

The unsupervised segmentation algorithm used in this work is a region growing based approach, which involves steps such as normal estimation, region growing, merging of planes by estimating correct set of parameters. Apart from the conventional voxel-based methods in 3D point clouds, the region growing method is another popular method to group different segments of a point cloud in 3D or an image in 2D. Region growing is a widely used method in unsupervised segmentation in 2D images, and it hence a promising segmentation approach for 3D point clouds as well. However, normal estimation is a critical step in the region growing algorithm where angles of several normal are used in growing the regions of a 3D point cloud. In the coming sections, several steps employed in the unsupervised segmentation will be described.

5.4.3.1 Region Growing

Region growing is used for estimating the planes from the normal estimated in the previous step. Region growing is a general method used for segmenting various regions in an image. A seed point is taken, and it is compared with neighbourhood points and merged in it if found similar. In this way pixels belonging to the same regions are grown and regions are identified. In case of 3D point clouds, taking the region growing method to 3D, regions are segmented with the help of normal corresponding to the point in consideration. Angles of the normal are compared to grow the region with respect to the seed point. Algorithm 5.2 describes the algorithm corresponding to region growing in point clouds with the help of normal extracted in the previous step.

The algorithm makes use of three parameters, number of nearest neighbours, angle threshold and curvature. In this work, for every point, the algorithm iterates through its nearest neighbours' points and keep on adding points to the current region if its normal is within a threshold angle with respect to the normal of the seed point. In the Alg. 5.2, line 1 and line 2, refers to initialization of regions and list of all points $ListP\ oints$, within a point cloud P. There is a point picked up with minimum curvature ($PminCurv$), which is added to the set of seed points (see line 7). Now, for every seed point, its neighbourhood point is found (see line 10). In this set Fc, every point is checked for its angle of the normal (θ) if it is lesser than a certain angle threshold ($\theta threshold$). If it is lesser than the threshold angle, then the current point is added to the current region, which is again tested for its curvature value against a pre-specified threshold value (see line 16).

If the curvature is found to be lesser than the threshold value then this point is added to the seeds point list (see line 17), followed by adding current region to global segment list (see line 22).

Algorithm: 5.2 – Algorithm for region growing

```
Regions ← ϕ
List_Points ← 1, 2, .., |P|
while List_Points ← ϕ do
    R_c ← ϕ
    S_c ← ϕ
    List_Points → P_minCurv
    S_c ← S_c ∪ P_minCurv
    List_Points ← List_Points − P_minCurv
    for i = 0 : size(S_c) do
        F_c ← NearestNeighbor(S_c(i))
        for j = 0 : size(F_c) do
            Current neighbour point P_j ← F_c(j)
            if List_Points ∈ P_j & θ(Normal(F_c(i), F_c(j))) < θ_threshold then
                R_c ← R_c ∪ P_j
                List_Points ← List_Points − P_j
                if Curv(P_j) < Curv_threshold then
                    S_c ← S_c ∪ P_j
                end if
            end if
        end for
    end for
    Regions ← Regions ∪ R_c
end while
return Regions
```

In these experiments, the number of nearest neighbours were chosen as 30 and the curvature threshold as 0.98. The nearest neighbour value is decided more from the computation perspective. The curvature threshold and angle thresh- old was decided based on the experiments carried out for various values of these parameters. A curvature threshold of 0.98 and angle threshold of 5^o worked best in the visual results of the experiments carried out. For higher values of thresholds, planar elements of wall and roof were getting combined into one element. Here 5^o threshold angle have been used based on the visual analysis of various angle threshold. Within a cluster, the seed point gets updated to a new point if its curvature is less than 98^{th} percentile of all the curvatures in the point cloud.

5.4.3.2 Cluster Refinement

In this step, the points which are not assigned to any of the regions in the region-growing step, are assigned to their nearest clusters. These points essentially belong to the sharp edges and hence do not satisfy the angle threshold condition with any of the seed points. The cluster refinement process works by first calculating the centres x_{center}, y_{center}, z_{center} for all the clusters, which is computed as the mean of the x, y, z coordinates of the clusters, respectively. It then iterates through all the unassigned points and assigns them to the cluster nearest to them. The angle threshold condition used in region-growing is not imposed during the cluster refinement step. Some of the unassigned points in the region- growing step could be outliers as well, which serves as noise in the dataset. In order to avoid assigning these outliers to any of the clusters, a check is performed to ensure that the distance between the new point being added, and the cluster centre is less than the maximum distance between the cluster centre and any point already present in the cluster i.e., the newly added point should not be at a distance greater than any of the already present cluster points from the cluster centre.

5.4.3.3 Merging the planes

The plane segmentation step outputs planes from the input point cloud. These set of planes alone do not have any significance in terms of building elements as they are plane segments from various elements in the building. For using them to form a meaningful information, these planes need to be separated based on the element they belong. For example, a laser scan of a column may result in four separate planes, the idea is to club these planes together, to make it a column element. To implement this, a centre-to- centre distance threshold can be applied to the set of planes acquired from the previous step. The value of this threshold should be selected such that:

- All the parts of the same elements should get clubbed together
- No parts of two different elements should get clubbed together

However, in a realistic situation, fulfilling both these criteria on the entire data is a tedious task. Hence the objective is to find a distance threshold which can fulfil these criteria to a maximum extent. Selecting a random threshold value can give following results as depicted in Figure 5.3:

- Case (a) No planes get clubbed together
- Case (b) Fewer planes from same element gets clubbed together
- Case (c) All planes from same element gets clubbed together
- Case (d) All planes from same element and some planes from other elements get clubbed together

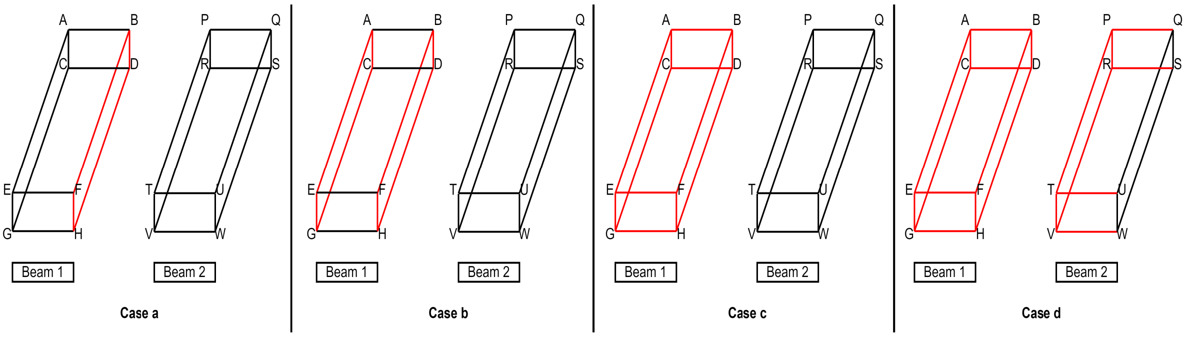


Figure 5.3 Illustration of clubbing of planes into elements with different range of thresholds. Case (a) No planes get clubbed together, Case (b) Fewer planes from same element gets clubbed together, Case (c) All planes from same element gets clubbed together, Case (c) All planes from same element gets clubbed together, Case (d) All planes from same element and some planes from other elements get clubbed together.

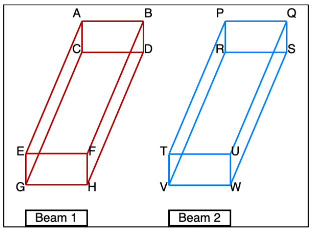


Figure 5.4 Illustration of ideal threshold value selection, where all planes of the same elements get clubbed together and two meaningful elements are generated.

In the following illustration, the distance threshold has been applied with reference to Plane BDFH of Beam 1 for explaining the different scenarios created by varying the value of threshold t.

In the Figure 5.3, in case a, where the threshold t is less than distance between any two planes. It can be seen in Figure 5.3 that no planes get clubbed together. In case b, as t increases, the planes with the nearest centre-to-centre distance gets clubbed together, but not all of them from 'Beam 1' gets clubbed together. In case c, as t is increased further, all the planes of 'Beam 1' gets clubbed together and this is the ideal scenario as it gives us a meaningful object as 'Beam 1'. In case d, as t is increased further, the planes from the nearby objects 'Beam 2' gets merged with 'Beam 1' and create a meaningless object. Hence the ideal value of t should be as in case c. Kindly note that when the threshold will be applied it will be applied to all the planes, and hence multiple objects will get clubbed together based on the threshold set. In this case of two beams, one should get two separate objects for 'Beam 1' and 'Beam 2' as shown in Figure 5.4 with a correct threshold value. Hence, one of the novel contributions in this work is the threshold finding method, which gives us the value of a best possible threshold, helping us to club not all but most of the planes in the building point clouds accurately. This is based on creating a matrix of distances of all detected plane and manually getting the threshold range out of it.

5.4.3.4 Noise Removal

After merging the planes into meaningful objects, there are few left out clusters of points which consist of less than 2048 points. These tiny clusters are usually points which gets left out and adds up to noise. Even visual assessment of these clusters does not give any sense of identification. Therefore, a noise removal step is applied, by automatically removing the clusters which have less than 2048 points.

5.4.4 Self-Supervised Classification

The building point clouds have been pre-processed and segmented to extract complete objects like wall, beam, columns by merging the planes obtained by using unsupervised segmentation technique as discussed in previous sections. These objects are given as the input to the model pre-training and feature learning framework which aims towards extracting high-quality features from these objects for classification. The pre-training step is followed by the object classification downstream task. The entire dataset is divided into two main sets: 80% of the data is unlabelled and is used for training the feature learning pipeline while the remaining 20% of the data is labelled and used for object classification downstream task.

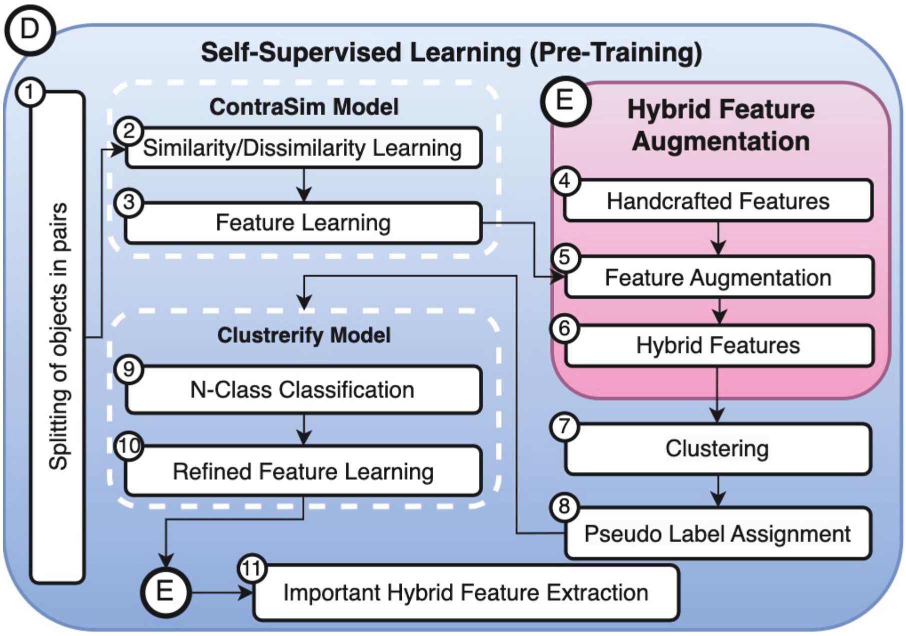


Figure 5.5 Detailed Methodology for Self-Supervised Classification

Therefore, as shown in Figure 5.5, the feature extraction is conducted in a self-supervised manner with contrastive learning, clustering the features and assignment of cluster IDs as pseudo labels. Further, to refine the learned features, a 6-class classification training is performed using the point cloud data and pseudo labels (see Figure 5.6). In this work, a ContraSim model for contrastive feature learning is proposed, which uses cross entropy loss to learn the features from part objects. The feature learned from this stage are clustered using K-means ++ algorithm and pseudo labels are assigned to each cluster.Further Clusterify model is proposed to obtain refined features by performing a 6-class classification task with the labels obtained from the previous step for the original point cloud data. Specific Hand-crafted features are extracted to further enhance feature representation and capture local geometric properties. These features are used with contrastive features in clustering to obtain pseudo labels and further with learned features using Clusterify model.

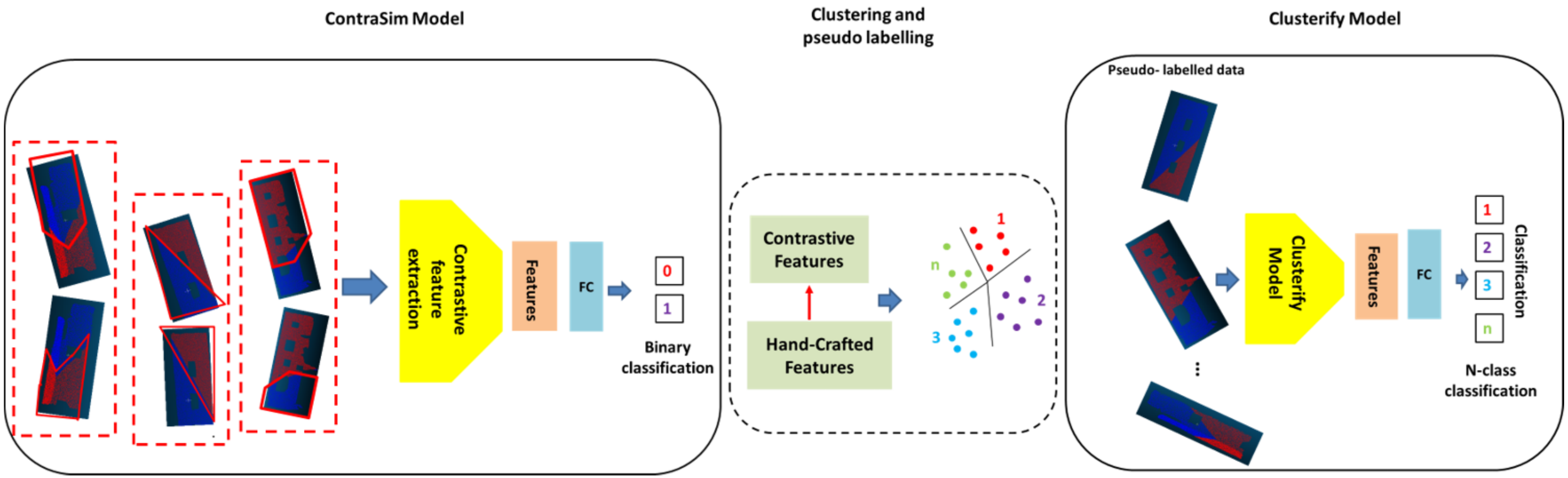


Figure 5.6 The self-supervised feature learning framework - ContraSim and Clusterify

Figure 5.6 depicts the contrastive feature extraction using ContraSim and refinement of those features using Clusterify model. The proposed feature extraction models and classification framework is described in the below sections.

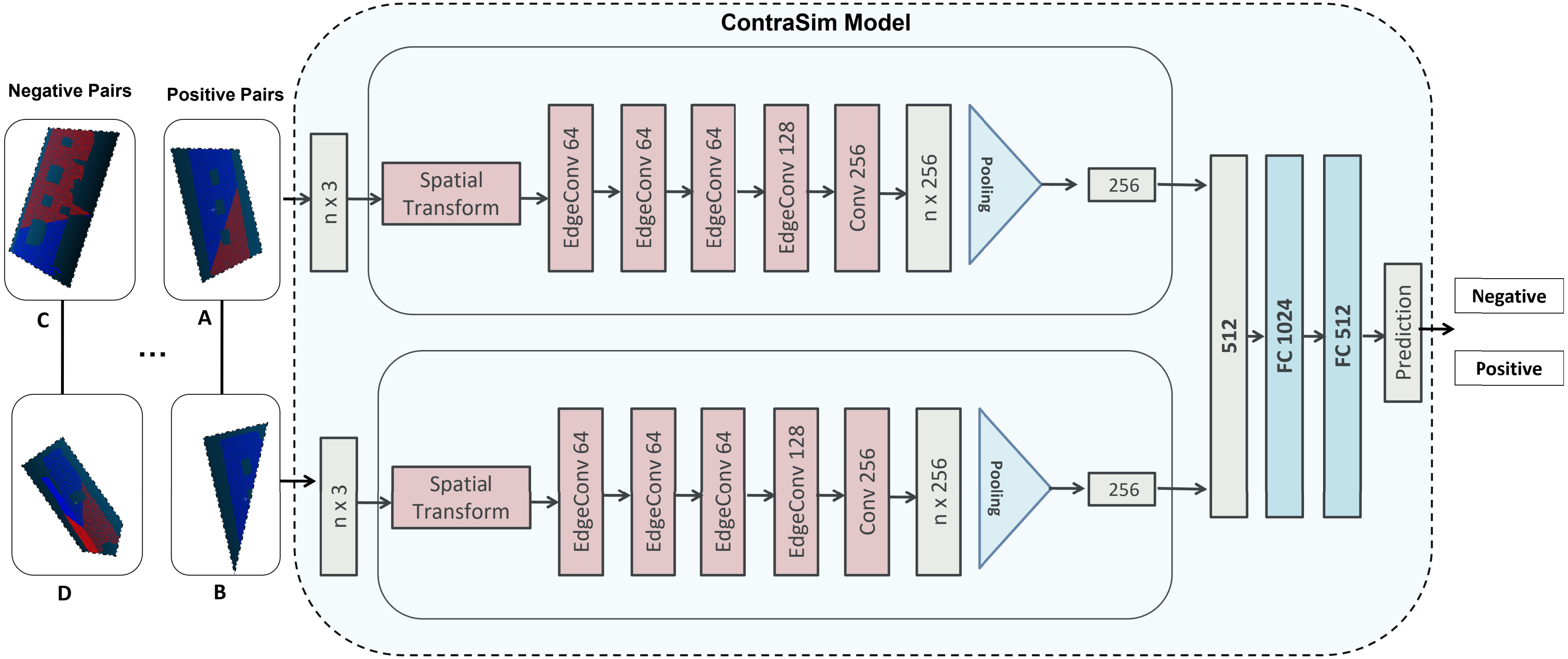


Figure 5.7 The architecture of ContraSim for part contrast learning

5.4.4.1 ContraSim Model

Contrastive learning is a technique in which features from the visual data are learned by learning the similarity and dissimilarity between the samples in the data. This is an efficient technique to learn the visual features in a self-supervised way with unlabelled data. Contrastive learning presents the similar and dissimilar parts of data points to each other to make the machine learning model learn the parts which belong to the same class and parts which do not.

In the proposed work, features are learnt from the point cloud segments of the building in a self-supervised fashion. The various objects from the building point clouds are segmented in the previous steps are split randomly into two parts as shown in Figure 5.6, with red boxed in point cloud objects. Two different parts of the same object are shown with blue and red coloured points. For the purpose of splitting the object point cloud into two parts, a random plane is found and points which are present on either side of that plane are considered to be two object splits.

Algorithm 5.3 describes the procedure *"Split Object"* for randomly splitting the point cloud object into two parts. In this algorithm, 3 random points, a, b, c are selected and multiplied with the X, Y, Z array of coordinates in point cloud object. It essentially tries to find a random plane in 3D, to split the point cloud objects into two parts, where the equation of the plane is $ax+by + cz= 0$. It sums the coordinates in (X, Y, Z) arrays and adds the points which has their Sum >= 0, to the first Split (Sp_1) and points which has their Sum < 0, to the second Split (Sp_2).

These pairs of segments of each object are compared with each other for positive and negative pair and hence features are learned for each object by learning similarity. Parts which belong to the same objects constitutes the positive pair, labelled as 1 and parts which belong to different object makes negative pair, labelled as 0. In the Figure 5.7,the process of learning features with contrastive learning is depicted. A sample positive pair input to the ContraSim model looks like, A-B of the same object and C-D belong to different object, which also makes A-C, A-D or B-C, B-D as negative pairs since they are coming from different objects. Hence, learning the features by learning the positive and negative pairs brings all the positive pairs and negative pairs close together respectively as a group.

Thus, this is a binary classification task with the positive and negative pairs as the input. This is a completely self-supervised learning process as no ground truth labels are used for feature extraction.

Algorithm: 5.3 – Algorithm to split a point cloud object

```
procedure SPLIT OBJECT(Object(X, Y, Z))
    a = (random(0, 1) − 0.5) ∗ 2.0
    b = (random(0, 1) − 0.5) ∗ 2.0
    c = (random(0, 1) − 0.5) ∗ 2.0
    X = X ∗ a
    Y = Y ∗ b
    Z = Z ∗ c
    Sum = X + Y + Z
    if Sum ≥ 0 then
        Sp_1 = X(idx(Sum)), Y(idx(Sum)), Z(idx(Sum))
    end if
    if Sum < 0 then
        Sp_2 = X(idx(Sum)), Y(idx(Sum)), Z(idx(Sum))
    end if
    return Sp_1, Sp_2
end procedure
```

ContraSim model utilizes DGCNN model as the backbone network for feature extraction purpose. DGCNN introduced a novel operation EdgeConv which aims towards recovering the topological information (local geometric features) in a point cloud. EdgeConv is applied directly on a dynamically generated graph in a Graph Neural Network (GCNN) to incorporate the local neighbourhood properties of a point cloud. In the Figure 5.7, the DGCNN backbone network is used for each part of the input object which is randomly split. Each branch of the network uses a spatial transformer network which allows a neural network to learn the spatial transformations to be applied on an input data (2D image or 3D point cloud), to enhance the geometric invariance of the network. After that, it uses a sequence of 4 EdgeConv layers with kernel sizes 64, 64, 64 and 128, followed by a convolutional layer to combine the feature embeddings generated from the previous EdgeConv layers. The final feature embedding is pooled into a $256 - D$ Max Pooling layer. The features generated from both the branches of the network are concatenated using a single vector by using 3 fully connected layers, where final classification is learned in the form of a binary classification for a positive or negative pairs.

The ContraSim model takes $numpoint$ (number of points) as an input parameter which indicates the number of points each cut part should have. As mentioned earlier, this model takes a pair as input i.e., two cut parts. These cut parts are converted into objects with $numpoint$ number of points. For objects with a smaller number of points, some random points are repeated to get the required number of points while for objects with higher number of points, points are removed to obtain objects with the desired number of points. For the purpose of experiments, various values for $numpoint$ like 256, 512, 1024, 2048 and 4096 were tried. Also, a batch size of 16 was used for training the model.

$$L_{cross-entropy}(\hat{y}, y) = -\sum_{i} y_i log(\hat{y}_i) \tag{5.1}$$

The loss function used for training the ContraSim model is cross entropy (see Eq. 5.1) which is optimized by stochastic gradient descent (SGD). ContraSim model is based on the binary classification task, and hence there are only two classes. The true labels are converted into one hot vector of size 2, and softmax activation is applied to the predicted values. Then, these are used for computing the cross-entropy loss.

5.4.4.2 Clustering

In this step clustering groups all the features learned from ContraSim model into the desired number of clusters, \$6\$ clusters in the current case. The features learned by the ContraSim model are given as the input to the K-means ++ clustering algorithm. Clustering algorithm starts by choosing an initial set of centroids and assigning each of the data points to one of these clusters. After every iteration, the centroids are updated based on the data points assigned to that particular cluster. The algorithm stops when the maximum number of iterations are completed which is kept as 300 for this problem by empirical observation. The Clustering algorithm takes K as an input, and K=6 was chosen as input for experiments as there are a total of six classes which are wall, door, window, slab, beam, and column and obtains 6 clusters for the features (see Figure 5.6).

Further, pseudo labels (cluster IDs) to each cluster are assigned which act as a label for the original point cloud data in a supervised 6-class classification.

5.4.4.3 Clusterify Model

In the above steps, features generated from ContraSim model are clustered and cluster ID as assigned as pseudo labels. Now another model Clusterify is trained which is a supervised n-class classification model using the original point cloud data and labels (pseudo labels) generated in the previous step. The Clusterify model (see Figure 5.6) aims at learning refined features of the building elements by making use of the cluster labels obtained in the Clustering step. Thus, in this step supervised learning is used to boost the feature learning process. In contrast with the ContraSim model where building element parts are considered as positive and negative pairs, in Clusterify model the entire building element is taken as input. Clusterify model tries to learn the mapping between the input building elements and the pseudo-labels and in this process where the potential learned features will help in differentiating building elements belonging to different classes. Clusterify model also used DGCNN network as a backbone network for learning the features as described in the previous section.

The Clusterify model also takes num_{point} number of points as input which indicates the number of points each object will have. All the building elements will be converted into objects with num_{point} number of points in the same way as done for the ContraSim model. However, in contrast with ContraSim model that takes cut parts of the building element as input, the Clusterify model takes the whole building element as the input. Therefore, for the experimentation purpose if the num_{point} is chosen as 256, 512, 1024, 2048 or 4096 for the ContraSim model training, then the num_{point} will be 512, 1024, 2048, 4096 or 8192 respectively for the Clusterify model that is twice the ContraSim model.

Similar to the ContraSim model, the loss function used by Clusterify model is softmax cross-entropy (see Eq. 5.2). Clusterify model deals with the task of assigning the building elements to one of the k-classes where the value of k is decided in the clustering step. The loss is computed in a similar way as discussed in the ContraSim model part.

$$L_{cross-entropy}(\hat{y}, y) = -\sum_i y_i log(\hat{y}_i) \tag{5.2}$$

Where, $\hat{y_i}$ is the vector of predicted scores and y_i is the corresponding true label in the form of one-hot encoding. The summation is done over all the classes.

Hand-crafted features (which will be discussed next) are augmented first with ContraSim features to generate pseudo labels and again with features learned by Clusterify model to be used for downstream tasks. These features capture local geometric properties of the point cloud objects and they are able to enhance the overall classification performance to great extent.

5.4.4.4 Augmentation of handcrafted features

In addition to the features that are automatically learnt, handcrafted features are additionally experimented with to fill the gaps in automatic feature extraction. The construction dataset in the consideration is largely different from the datasets used in computer vision for identifying elements from the point clouds. The datasets used in literature consists of general day to day elements such as table, chair, lamps. etc, which have a very distinguishable shape and structure and easy to differentiate between different classes if automatic feature extraction is done using contrastive learning. Hence, some of the very specific features from the building elements which make them distinct are identified and augmented with the automatically learnt features to improve the classification performance. These features are targeted towards capturing the local geometric properties of the components in point cloud which might be missed by the automatically learnt features using contrastive learning.

To begin the process of extracting handcrafted features (Figure 5.5) construction dataset was analysed, and various features were identified which makes one element different with another. Such features were: Average RGB value, Surface Area, Average z-coordinate, Average covariance features. These handcrafted features are independent of the domain of point cloud used and can be used for general construction data of point cloud data, which makes the system generalised for a vast variety of the construction data in contrast with the conventional rule-based systems, where rules are identified using the domain knowledge or the past trends.

• **Average RGB value:** The average RGB value for each point in a point cloud is an important feature which will help differentiating between the objects with different colours for example a red-brick wall from a column or beam which are grey in colour. Figure 5.8 (a) shows two objects of different classes in different colour, making the RGB information an important feature. It also captures the texture and roughness of the target surface which could be very different for all the object classes. The average RGB value for each object is calculated as:

$$AverageR = \frac{\sum R}{n_p}$$
$$AverageG = \frac{\sum G}{n_p}$$
$$AverageB = \frac{\sum B}{n_p} \quad (5.3)$$

Where np is the total number of points and $\sum R, \sum G, \sum B$ is the sum of all RGB values across the 3D object point cloud.

Figure 5.8 Examples of various handcrafted features

• **Element Surface Area**: For the purpose of finding the surface area feature for a point cloud, first a triangular mesh was generated for the unstructured point cloud to identify its dense 3D geometry. Triangular mesh is generated using Poisson surface reconstruction algorithm which uses a regularized optimization to generate a smooth surface in a point cloud. This mesh is used to find the surface area of the individual component in the point cloud. Algorithm 5.4 is used to calculate the surface area. In Figure 5.8 (b), it can be seen that the door has a smaller area when compared to the wall.

Algorithm: 5.4 – Algorithm for surface area calculation

procedure AREA OF TRIANGLE($p_{x1}, p_{y1}, p_{z1}, p_{x2}, p_{y2}, p_{z2}, p_{x3}, p_{y3}, p_{z3}$)
    $a_x = p_{x2} - p_{x1}$
    $a_y = p_{y2} - p_{y1}$
    $a_z = p_{z2} - p_{z1}$
    $b_x = p_{x3} - p_{x1}$
    $b_y = p_{y3} - p_{y1}$
    $b_z = p_{z3} - p_{z1}$
    $c_x = a_y \times b_z - a_z \times b_y$
    $c_y = a_z \times b_x - a_x \times b_z$
    $c_z = a_x \times b_y - a_y \times b_x$
    $area = 0.5 \times \sqrt{(c_x^2 + c_y^2 + c_z^2)}$
    return *area*
end procedure
procedure MESH AREA($X, Y, Z, num_{triangles}, V_1, V_2, V_3$)
    $Area = 0$
    for *i = 0 to* $num_{triangles}$ **do**
        $Area = Area +$ Area of Triangle($X[V_1(n)]$,
        $Y[V_1(n)], Z[V_1(n)]$,
        $X[V_2(n)], Y[V_2(n)], Z[V_2(n)]$,
        $X[V_3(n)], Y[V_3(n)], Z[V_3(n)]$)
    end for
    return *Area*
end procedure

- **Average z-coordinate**: Height of every object in the point cloud is an important feature, which can distinguish between objects with similar appearance but different functional roles. In the context of construction data, there are multiple objects which can have similar appearance taking the shape into consideration, for example a wall and a beam. The two rectangular objects have different functional roles and positions in construction data but could be ambiguous for a deep learning model to extract distinguishing features. Their placement with respect to the ground makes them different in a scene and hence a height feature can be important in this context. In Figure 5.8 (c) it can be seen that the height coordinate for various objects is separated apart, which can work as a good feature for classification.

- **Average covariance features**: For the purpose of capturing local geometric properties covariance features are widely used in the literature. Covariance features are calculated by first generating a covariance matrix of 3 for all the points for its neighbourhood. In this line, a covariance matrix for each point is generated with 50 nearest neighbour points found by K-nearest neighbour algorithm. After that the eigenvalues and eigenvectors from the covariance matrix are calculated. Since eigen values represent geometric properties of an object, features such as Linearity, Planarity, Scattering, and Verticality can be identified using them. These local geometric features are described below for the eigen values $\lambda - 0, 1, 2$, arranged in ascending order:

– Linearity(L): It suggests the linearity of a given point based on its surroundings, i.e., whether the surrounding is linear or not.

$$Linearity(L) = \frac{\lambda_2 - \lambda_1}{\lambda_2} \in [0, 1] \tag{5.4}$$

– Planarity(P): It provides the planarity of a given point based on its surroundings, i.e., whether the surrounding is a plane or not.

$$Planarity(P) = \frac{\lambda_1 - \lambda_0}{\lambda_2} \in [0, 1] \tag{5.5}$$

– Scattering(S): It suggests the sphericity of a given point based on its surroundings, i.e., whether the surrounding is curved or not.

$$Scattering(S) = \frac{\lambda_0}{\lambda_2} \in [0, 1] \tag{5.6}$$

– Verticality(V): This metric gives the verticality of a given point based on its surroundings, i.e., whether the surrounding is vertical or horizontal.

The unary vector of principal direction in R3+ is defined as the sum of the absolute values of the coordinate of the eigenvectors weighted by their eigenvalues.

$$[\vec{v}]_i \propto \sum_{j=1}^{3} \lambda_j |[v_j]_i|, wherei = 1, 2, 3et||\vec{v}|| = 1 \tag{5.7}$$

The vertical component of this vector is the verticality of the point. For each object, these covariance features are obtained, and average is taken of each object.

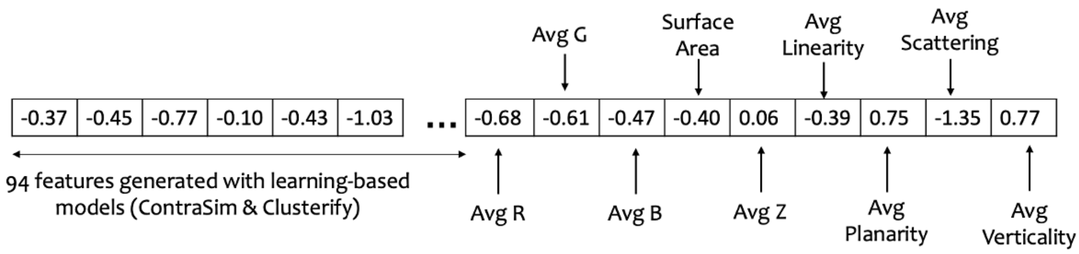


Figure 5.9 Feature engineering with handcrafted feature augmentation.

Figure 5.9 depicts the final feature vector after augmentation of handcrafted features in the automatically learnt features using contrastive learning.

5.4.4.5 Feature ranking and selection

The feature extraction process is described in the above sections, however out of all the extracted features using deep learning and hand-crafted, there are only a few features which are important and contributes to the performance of the machine learning model. Hence, some of the experiments were performed to analyse the feature importance and selected 103 important features from the pool of the features. Permutation importance is one such technique which is used for identifying important features in a large feature matrix. The process of important feature importance selection is described as:

Algorithm: 5.5 – Algorithm for permutation Importance

1: **procedure** Permutation Importance(M, D)
2: $Importance = \{\}$
3: $Score_{imp} = \text{Accuracy}_{prediction}(M, D)$
4: **for** $f_i : i = 1, ...n$ **do**
5: **for** **do**$K_j : j = 1, ...K$
6: $D_{i,j} = \text{RandomSuffle}(D_i)$
7: $Score_{imp_{i,j}} = \text{Accuracy}_{prediction}(M, D_{i,j})$
8: **end for**
9: $Score_{imp_i} = Score_{imp} - \frac{1}{K} \sum_{j=1}^{K} Score_{K,i}$
10: $Importance = Importance \bigcup Score_{imp_i}$
11: **end for**
12: **return** $Importance$
13: **end procedure**

Algorithm 5.5 demonstrates the process of finding important features from the pool of 1033 features with function *Permutation Importance*. The Alg. 5.5 takes a trained model M and a dataset D as parameters to the function *Permutation Importance* and tries to find the importance score for each feature f_i for randomly shuffling the rows for f_i in dataset D, K times (line 3,4). With every shuffle, it calculates its importance score $Score_{impi,\ j}$ which is accuracy of prediction in the current case for the shuffled dataset $D_{i,j}$ with Model M (line 5). After K iterations, it takes the average of importance scores obtained at each step $Score_{impi}$ (line 6), further adding the value to the list of all the importance scores (*Importance*) (line 7). The larger the value of $Score_{impi}$ for feature f_i, the more important is the feature to the trained model, pertaining to the fact that shuffling more important feature will cause more drop in the accuracy value.

The details of feature selection based on importance are discussed in section 5.4.1. The most important 103 features in the order of their ranking and importance score are listed as: Blue, Red, Green, Surface area, Average Z, Verticality, Planarity Linearity, Scattering and deep learning features. Here, Blue, Red, Green are the RGB colour values, and out of the 1033 features, top 100 features with the remaining three handcrafted features were taken to make a feature vector of size 103. In order to measure the sensitivity of the selection of vector dimension, experiments were performed with size, multiple dimensions (details in section 5.4.1). Hence, considering the highest performance, feature vector dimension 103 is selected for training the classifier. The feature vectors used for downstream training are scaled using the scaling function defined as $Z = \frac{(x - \mu)}{\sigma}$, where x is each feature, μ is mean of whole feature column and σ is the standard deviation of the feature column.

5.4.4.6 Classifier Training and Testing

In this step, the quality of feature extraction with model pre-training is tested with the remaining 20% S3DIS data for a classification task. The building elements are classified into one of the six classes (Wall, Door, Slab, Window, Beam and Column) with the remaining labelled dataset which is divided into train and test sets.

SVM classifier is used as the classifier which is trained on the train set which consists of features of the building elements extracted from both the ContraSim or Clusterify model as the input and the ground truth labels as the output. The final results of classification are presented on the test set in the results section. Equation below depicts the loss function used in training the SVM classifier.

$$\min_{w,b,\zeta} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \zeta_i$$
$$subject\ to\ y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i,$$
$$\zeta_i \geq 0, i = 1, \ldots, n$$

Here, w is the set of weights and b is the set of biases, C is the penalty when a sample is misclassified or within the margin boundary. yi refer to the training labels and xi is the feature map.

The testing of the pipeline is conducted on the test dataset, here the remaining unlabelled data from the S3DIS is used which have not been used till now at any stage. The experimental details, results and their interpretations are presented in the next section.

5.5 EXPERIMENTAL RESULTS AND DISCUSSION

In this section, all the experiments performed in the proposed work is described in detail. All the experiments were performed on GPU, NVIDIA Quadro RTX 6000 Passive with 24 GB memory, with 2.3 GHz clock speed and NVIDIA Tesla T4 with 16 GB memory with 2.2 GHz clock speed.

The experiments were performed on the test set created from the S3DIS dataset i.e., Area 2 and Area 4, which consist of 40 rooms and 50 rooms respectively. The pipeline shown in Figure 5.2 was executed till the Step 3 of unsupervised plane segmentation.

Next to select an appropriate merging threshold to form meaningful elements the method described in section 3.3.3 was used. For this, a case is demonstrated with Area 6 - Copy Room 1 as shown in Figure 5.11(a) from the S3DIS data set and applied the mentioned empirical method to get the range for the appropriate threshold.

CopyRoom1 has a total of 8 elements, with 1 beam, 2 slabs, 1 door and 4 walls. Executing unsupervised segmentation created a total of 11 planes for this room, with 3 for beam, 1 for ceiling, 1 for floor, 1 for door and 5 for walls segments. The distance between these planes is as shown in Figure 5.10.

Planes	**S1**	**B1**	**W1**	**D1**	**W2**	**S2**	**W3**	**W4**	**W5**	**B2**	**B3**
S1	0	2.02	2.27	1.53	1.75	2.59	2.33	1.62	2.19	1.82	1.70
B1		0	3.90	1.74	3.66	2.72	1.42	2.81	3.04	**0.81**	**0.64**
W1			0	2.68	1.15	2.45	3.57	2.14	1.96	3.63	3.64
D1				0	2.45	1.94	2.24	2.68	2.68	2.04	1.91
W2					0	3.04	3.73	2.28	2.57	3.48	3.40
S2						0	2.01	2.51	1.76	2.60	2.72
W3							0	2.23	2.11	**0.97**	1.25
W4								0	1.06	2.18	2.29
W5									0	2.48	2.66
B2										0	0.32
B3											0

Figure 5.10 The distance (in meters) between the centroid of the identified planes from the plane segmentation algorithm

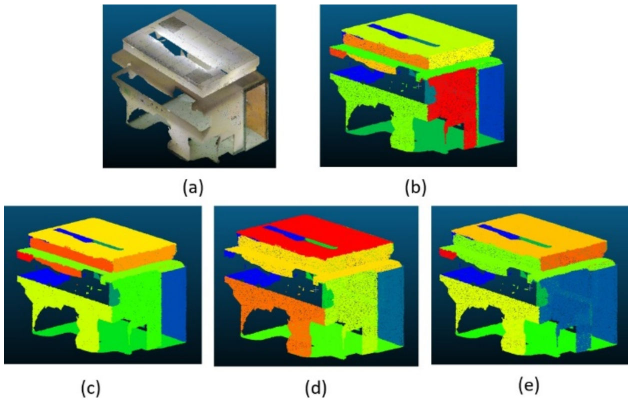


Figure 5.11 (a) Original Point cloud of CopyRoom1 from S3DISDataset (b) Segmented point cloud (c) Case 1: if $t \leq 0.81$ (d) Case 2: if $t > 0.81$ and $t < 0.97$ (e) Case 3: $if\ t \geq 0.97$

To appropriately club the planes into meaningful elements, the following three critical threshold ranges were identified by analysing the Figure 5.10.

- **if $t \leq 0.81\ m$**: Individual segments of a single beam element are not getting clubbed together.
- **if $t > 0.81\ m$ and $t < 0.97\ m$:** Two segments of the same wall are not getting clubbed together, but it is getting identified as a wall. Segments of the beam are getting clubbed together.
- **if $t \geq 0.97\ m$**: One segment of the beam (B2) and one segment of the wall (W4) will combine into one element, which will form a meaningless element, with parts of two different elements.

Therefore, it is meaningful to select a threshold as $0.81\ m < t < 0.97\ m$. Similarly, this empirical method can be applied to few other rooms for identifying the range of meaningful threshold for this data. But this range of the possible threshold will not vary much for a particular dataset as the elements will be the same as the case considered.

However, for a new data the same method will be used for selecting the appropriate merging threshold. Here to validate this method, experiments were performed by taking value of t as 0.8m, 0.9m and 1.0m and present the results.

$$IoU = \frac{|Object_{seg} \cap Object_{GT}|}{|Object_{seg} \cup Object_{GT}|} \tag{5.8}$$

Figure 5.13 depicts the IoU (Intersection over Union) values generated for various threshold values in meters across all the object classes. IoU (see Eq. 5.8) is a metric to measure the performance of a segmentation method, where $Objectseg$ is the segmented object for a specific class and $ObjectGT$ is the ground truth object taken from the original data. IoU measures the overlap in the areas of segmented object and the ground truth object. The threshold depicts the value taken for merging the different segmented planes for object classes to create one single object. However, lower the threshold value, the higher will be the IoU, because merging with a lower threshold value remains a very conservative way to merge the planes and create a consolidated object. This will eventually lead to a good overlap with the original planes but misses out merging all the points in an object, leading to higher miss-classification rate. However, the higher threshold value in merging will lead to over merging of the planes and lead to confusing the machine learning model in identifying the exact shape of the object.

Hence, the target at this stage is to select an optimal value of which gives a good IoU for segmentation and also leads to lower miss-classification rate. As empirically identifies in the previous step, threshold values (t) should be selected in the range of

$0.81\ m < t < 0.97\ m$. Hence in this study, threshold value is selected as 0.9 to facilitate the higher classification accuracy, which also aligns with the empirical observations. Figure 5.12 shows the trad-off between classification and threshold value and it can be seen that overall classification accuracy is higher with growing threshold value with highest on 0.9 for hybrid and hand- crafted features. Hence 0.9 threshold value is considered to be the most optimal value in this case. The threshold values considered for experiments are taken in the range of $0.1 to 1$, and median IoU is calculated for the respective classes in Figure 5.13. The median IoU value is lowest for door and window object class because these objects are built within the wall class object and segmentation model confuses the points belonging to door or window to the wall class object.

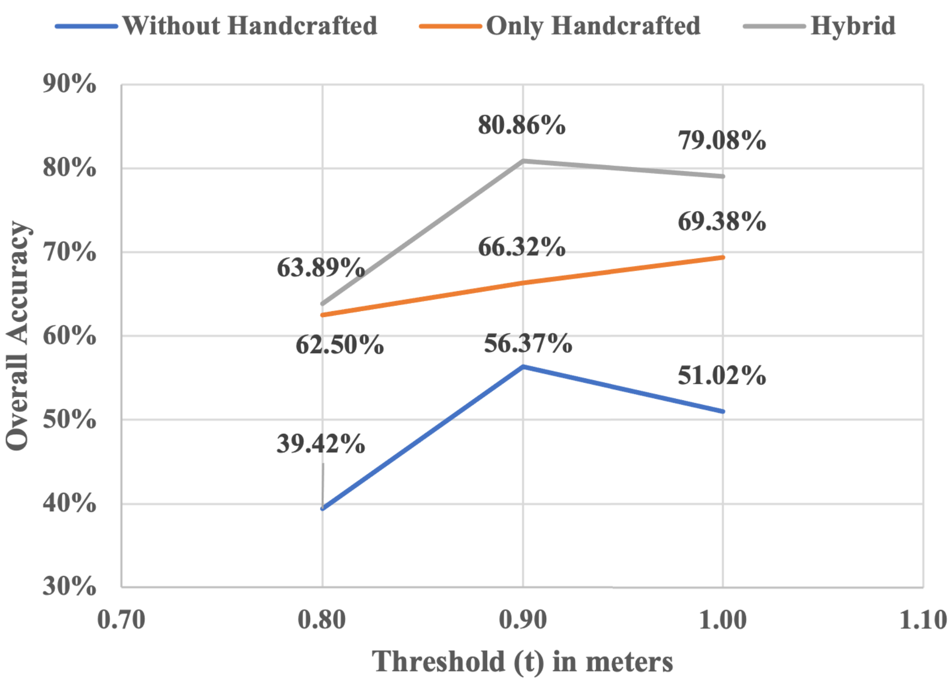


Figure 5.12 Overall Accuracy variation with the threshold and features used.

Object Class	Threshold 't' (in meters)										Median IoU
	0.10	**0.20**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**	**1.00**	
Wall	0.91	0.91	0.91	0.91	0.90	0.90	0.89	0.87	0.86	0.84	0.90
Door	0.71	0.70	0.70	0.70	0.69	0.67	0.64	0.61	0.58	0.54	0.68
Slab	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.98
Window	0.77	0.77	0.77	0.76	0.75	0.75	0.73	0.68	0.68	0.65	0.75
Beam	0.90	0.90	0.90	0.89	0.89	0.89	0.87	0.81	0.73	0.60	0.89
Column	0.87	0.87	0.86	0.83	0.80	0.76	0.72	0.67	0.65	0.59	0.78

Figure 5.13 Individual IoU values for different threshold and Median IoU of various object classes for unsupervised segmentation

The remaining pipeline as shown in Figure 5.2 was executed and the results obtained were evaluated based on the matrices provided below.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \qquad (5.9)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \qquad (5.10)$$

$$F1\ Score = 2 \times \frac{Precision * Recall}{Precision + Recall} \qquad (5.11)$$

$$Overall\ Accuracy = \frac{Predictions_{correct}}{Predictions_{Total}} \times 100 \qquad (5.12)$$

Precision is a measure of how many of the positive predictions made are correct (true positives). (Formula shown in Eq. 5.9)

Recall is a measure of how many of the positive cases the classifier correctly predicted, over all the positive cases in the data. It is sometimes also referred to as Sensitivity. (Formula shown in Eq. 5.10)

F1-Score is a measure combining both precision and recall. It is generally described as the harmonic mean of the two. Harmonic mean is just another way to calculate an "average" of values, generally described as more suitable for ratios (such as precision and recall) than the traditional arithmetic mean. (Formula shown in Eq. 5.11)

Accuracy: The base metric used for model evaluation is often Accuracy, describing the number of correct predictions over all predictions. Where, $Predictions_{correct}$ is the total number of correct predictions over all classes and $P\ redictions_{Total}$ is the total number of predictions. (Formula shown in Eq. 5.12)

5.5.1 Hybrid Feature ranking and analysis for selection

After performing the feature ranking algorithm for the overall 1033-dimensional hybrid feature vector, graded experiments were performed by selecting number of top 'n' features to select the relatively important features. Table 5.3 below shows the individual feature ranking obtained for various runs for the handcrafted features.

As it was expected, some of the handcrafted features contributed more than the deep learning features in object classification. It can be seen from the ranks in the table that the top important features are the R, G, B colour features, Surface area and Average Z. The covariance features rank slightly lower, with verticality contributing most, then linearity and scattering contributing almost equivalently and then scattering contributing the least.

Table 5.3 Individual Feature ranking obtained for various runs for the handcrafted features.

	Feature Rankings				
Features	**Run 1**	**Run 2**	**Run 3**	**Run 4**	**Run 5**
Blue	1	3	3	1	1
Red	2	1	1	2	2
Green	3	2	2	3	3
Surface Area	4	6	4	52	65
Average Z	5	46	6	4	5
Verticality	52	297	174	120	178
Planarity	681	696	698	698	697
Linearity	682	697	699	699	698
Scattering	859	851	856	852	854
Average Ranking	254.3	288.8	271.4	270.1	278.1

Slight deviations were obtained in the rankings as the contribution for each feature varied every time, however many features were ranked very closely to due equal importance score. Run 1 results were chosen to the documentation proceeded as it gave the lowest average ranking of the handcrafted features.

Below is the summary of two classes of experiments that were performed.

Experiment 1: Directly taking top features according to their rank.

The first experiment was performed by directly taking the top ‘n’ features and reporting the results for overall accuracy. Here values of ‘n’ are selected to capture the trend and sensitivity in overall accuracy. The values of n, that were chosen were 1, 2, 3, 4, 5, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300. These results have been shown in Figure 5.14. The graph shows that the overall accuracy increases till about the top 25 features are included, however after then it does not vary much and slightly decreases as more features are added.

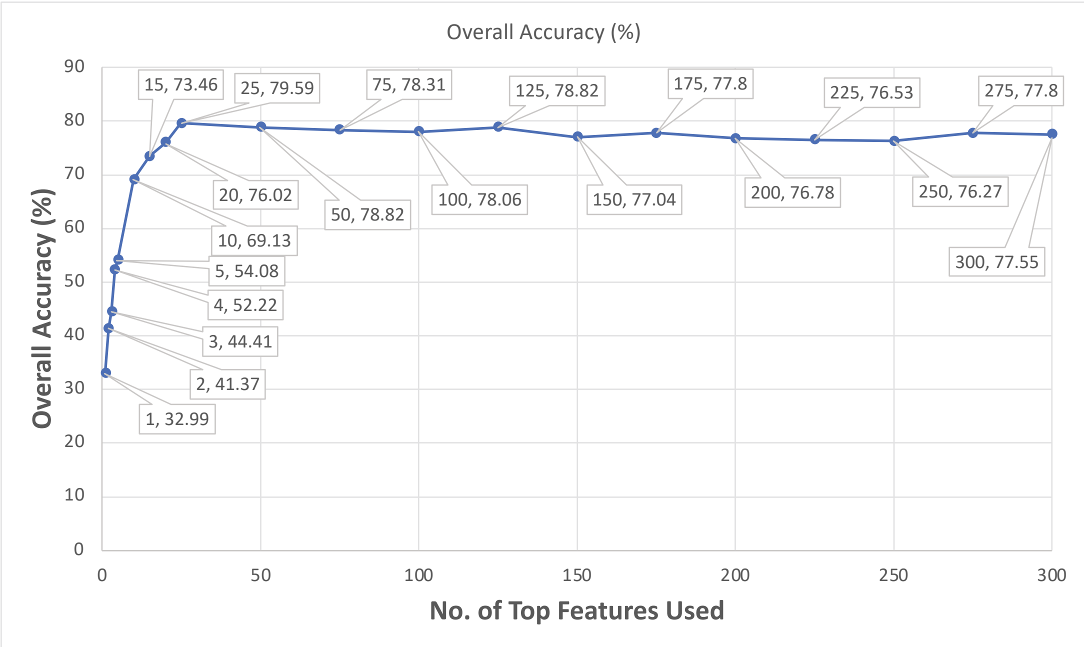


Figure 5.14 Overall Accuracy Vs number of top features used.

Note: The values in brackets (x, y) are (No. of top features, Overall Accuracy) for each point on the graph.

Table 5.4 The data in tabular format obtained for Experiment 1

Number of Top Features	Overall Accuracy (%)
1	32.99
2	41.37
3	44.41
4	52.22
5	54.08
10	69.13
15	73.46
20	76.02
25	79.59

Number of Top Features	Overall Accuracy (%)
50	78.82
75	78.31
100	78.06
125	78.82
150	77.04
175	77.8
200	76.78
225	76.53
250	76.27
275	77.8
300	77.55

It can be seen that the top 25 features contribute most to the accuracy as it reaches a maximum there at 79.59% and then slightly declines thereafter.

Experiment 2: Taking top features according to their rank and adding the remaining handcrafted features.

The second experiment was performed by directly taking the top 'n' features and the remaining handcrafted features which did not come in the top 'n' features. These results have been shown in Figure 5.15

It can be seen that the overall accuracy reached a maximum of 80.86% at 103 features, and slightly decreases thereafter as the number of features included were increased.

It is interesting to note that, as the number of features that are included passes a certain number, then the accuracy starts to slightly decline in both the experiments. This is because, all the features inclusion might not positively affect the accuracy and might result in overfitting. Therefore, if the features that are required for accurate classification are increased, then after a after a certain threshold the accuracy begins to decline.

For the thesis, the results corresponding to the 103 features were documented and reported. These contains the 100 top features with the remaining three handcrafted features as obtained in run 1.

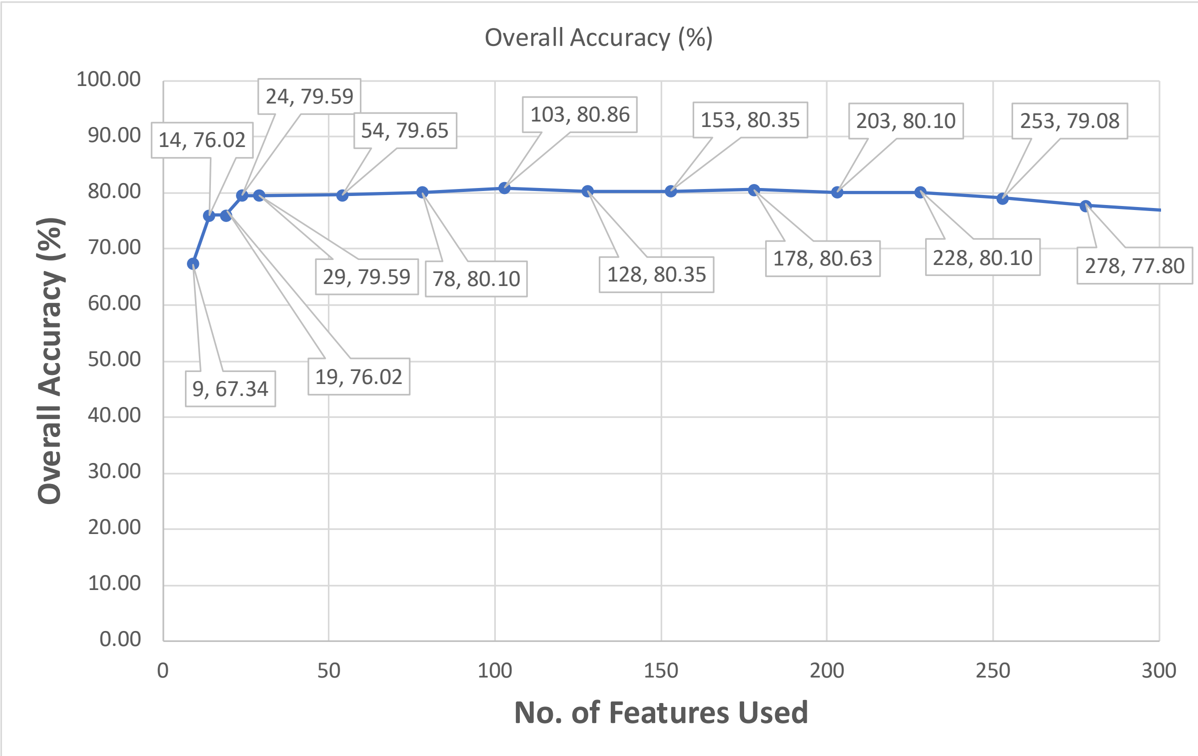


Figure 5.15 Overall Accuracy Vs number of total features used after considering the top 'n' features in addition to the remaining handcrafted features.

Note: The values in brackets (x, y) are (No. of features, Overall Accuracy) for each point on the graph.

Table 5.5 The data in tabular format obtained for experiment 2.

Number of Top Features Used	Remaining Handcrafted Features	Total Features Used	Overall Accuracy (%)
5	4	9	67.34
10	4	14	76.02
15	4	19	76.02
20	4	24	79.59
25	4	29	79.59
50	4	54	79.65
75	3	78	80.10
100	3	103	80.86
125	3	128	80.35
150	3	153	80.35
175	3	178	80.63
200	3	203	80.10
225	3	228	80.10
250	3	253	79.08
275	3	278	77.80
300	3	303	76.78

5.6 RESULTS AND INTERPRETATIONS

Table 5.6 shows Precision, Recall, F1-score for various object classes with 3 different feature engineering approaches proposed in this work, i.e., Deep learning-based features, handcrafted features, and Hybrid features. As it can be seen from Table 5.6 that overall accuracy for Hybrid method of feature engineering, i.e. augmenting the deep learning based features and handcrafted features, gives the highest overall classification accuracy for all the object classes which is 80.86 % as compared to overall accuracy with deep learning based features (56.37 %) and handcrafted features (66.32 %).

Table 5.6 Experimental results of classification pipeline with different sets of features on the S3DIS Dataset

Features	Without Handcrafted Features			Only Handcrafted Features			Hybrid Features		
Object Class	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
Wall	43.03%	44.73%	43.86%	58.44%	59.21%	58.82%	63.33%	75.00%	68.67%
Door	30.64%	31.66%	31.14%	73.33%	91.66%	81.48%	79.66%	78.33%	78.99%
Slab	57.89%	72.36%	64.32%	73.84%	63.15%	68.08%	85.88%	96.05%	90.68%
Window	64.40%	63.33%	63.86%	62.12%	68.33%	65.07%	83.33%	75.00%	78.94%
Beam	88.23%	50.84%	64.51%	52.00%	44.06%	47.70%	97.82%	76.27%	85.71%
Column	71.42%	73.77%	72.58%	76.27%	73.77%	75.00%	86.20%	84.03%	84.03%
Overall Accuracy	**56.37%**			**66.32%**			**80.86%**		

Color coding for Table 5.6

- 0% to 33.33% - Red (Low Score)
- 33.33% to 66.66% – Yellow (Medium Score)
- 66.66% to 100% – Green (High Score)

It can be seen from Table 5.6 that the individual F1-score is the highest for all the object classes for hybrid-based approach, specially Beam and Door object classes, which are nearly failed to be classified using deep learning-based features.

Table 5.7 depicts the confusion matrices for classification of objects with all 3 proposed feature engineering schemes. It can be seen that the Beam and Door object classes are highly miss-classified when only deep learning-based features are used. These numbers are improved and reach highest from handcrafted features to hybrid features by augmenting the deep learning based and hand-crafted features. The similar trend is observed for all the object classes. Hence, it can be concluded that hybrid features are performing much better as compared to deep learning-based features and handcrafted features used alone.

Table 5.7 Overall Confusion matrix for classification pipeline, without handcrafted features, with only handcrafted features and hybrid features for the S3DIS test data.

Without Handcrafted Features						
	Wall	Door	Slab	Window	Beam	Column
Wall	34	13	13	4	2	10
Door	20	19	9	8	1	3
Slab	7	8	55	5	0	1
Window	7	8	5	38	1	1
Beam	5	12	9	0	30	3
Column	6	2	4	4	0	45
Only Handcrafted Features						
	Wall	Door	Slab	Window	Beam	Column
Wall	45	7	4	0	13	7
Door	2	55	0	1	1	1
Slab	9	4	48	7	6	2
Window	10	3	4	41	2	0
Beam	3	3	7	16	26	4
Column	8	3	2	1	2	45
Hybrid Features						
	Wall	Door	Slab	Window	Beam	Column
Wall	57	7	0	5	0	7
Door	6	47	5	1	0	1
Slab	1	1	73	0	1	0
Window	13	1	1	45	0	0
Beam	3	3	6	2	45	0
Column	10	0	0	1	0	50

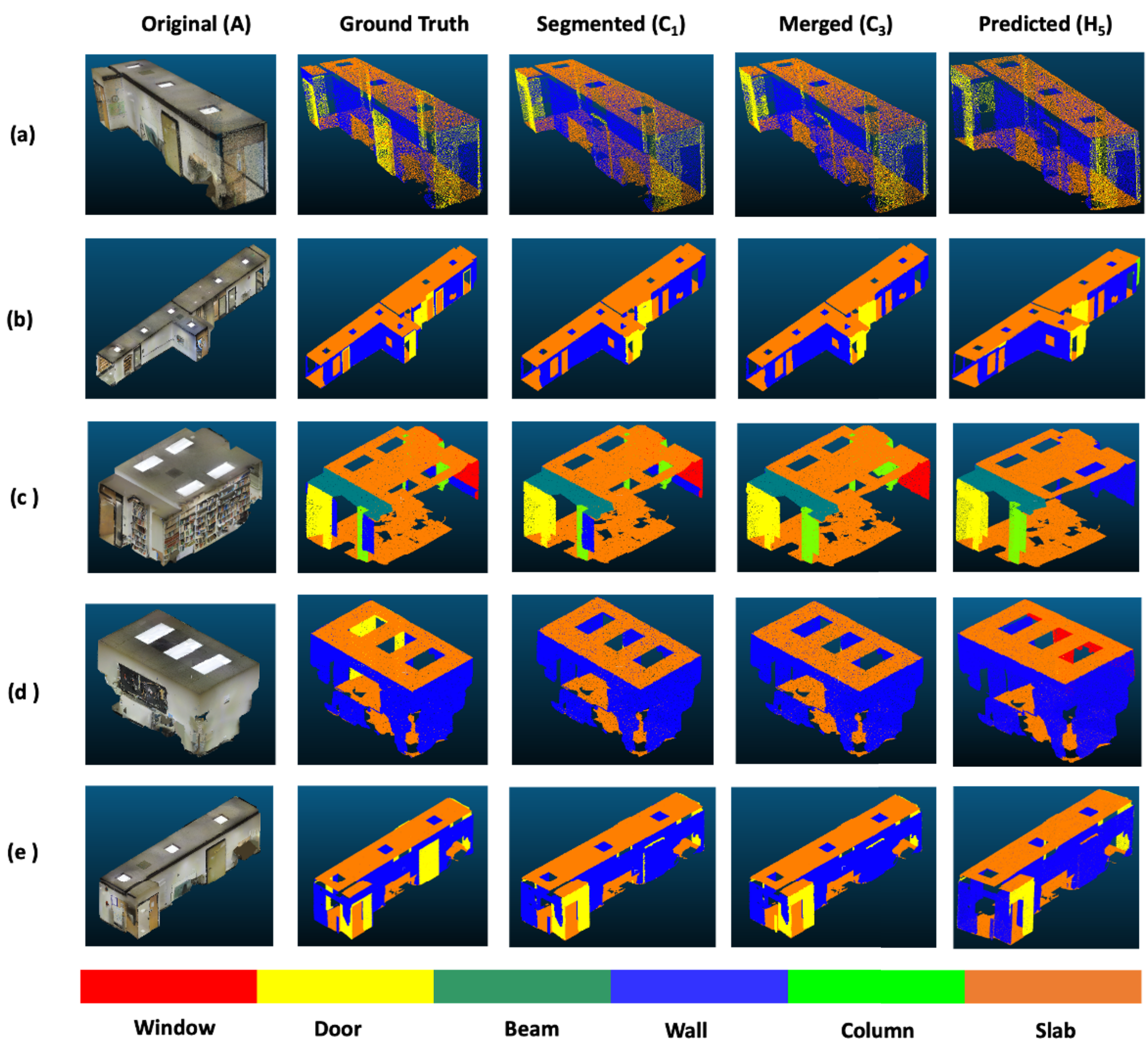


Figure 5.16 Visual results for various stages for S3DIS test Data

Figure 5.16 shows the visual results of the various stages for proposed framework for data samples taken from S3DIS dataset. This contains the original point cloud for a specific building, ground truth for the segmentation, segmented planes, merged planes for generating a single segmented object, and predicted object classes with the proposed pipeline. Legend in the lower panel shows the colour coding for all the object classes.

5.6.1 Comparison with other existing methods

Table 5.8 Class wise accuracy (percentages) comparison between different recognition methods

Class	Supervised Methods			Unsupervised Method
	(Armeni et al., 2016)	(F. Liu, Li, Zhang, & Zhou, 2017)	(Chen et al., 2019)	**Proposed (ConPro-NET)**
Beam	66.7	**78.6**	42.1	76.3
Ceiling	7.16	89.6	84.8	**96.1**
Floor	88.7	95	97.2	
Column	**91.8**	89.4	43.8	82
Door	54.1	33.4	55	**78.3**
Wall	72.9	60.1	52.4	**75**
Window	25.9	**75.3**	54.3	75
Average	58.1	74.48	61.37	**80.45**

Table 5.8 shows the comparison of class wise accuracy between the proposed method and existing approaches. Our methods outperform the existing supervised approaches for slab, door and wall classes. It can also be noted that the average accuracy is the highest for our method. However, our method requires an empirical threshold value to be selected based on the dataset or the project elements. A lot about the methods performance depends on how appropriately we can select the threshold.

5.7 CONCLUSION

This work presents ConPro-NET, a hybrid self-supervised method to detect and classify various components of a building point cloud, which is used for automatic construction progress monitoring. The key gap this chapter highlights is the incompetency of the supervised method for construction progress estimation domain where many building elements are project specific and cannot not be generalised. The labelling of point clouds acquired from construction site is a highly tedious and labour-intensive task.

Also, there are so many variations of the types of components used at construction site, therefore a project specific labelled dataset is difficult to obtain. Even the components on under-construction site are mostly grey and therefore traditional segmentation techniques cannot be applied.

The study contributes to the existing knowledge in the following ways. Firstly, a novel distance-threshold-based merging of planes approach is developed to conduct the distinctly challenging problem of segmenting various components in a construction environment. Secondly, the self-supervised approach is customized and applied to perform and work on the building construction dataset. Finally, construction-specific handcrafted features are identified and applied to the pipeline for improving its performance. The experiments showed that the hybrid features introduced and used in this paper improved the overall accuracy of the pipeline by achieving an 80.86% overall accuracy. The results also shows that the proposed method outperforms the existing approaches for the average class wise accuracy.

CHAPTER 6.

CASE STUDY EVALUATION

In the previous chapter, ConPro-NET, hybrid self-supervised approach was developed and experimented on the S3DIS dataset. This chapter present the results of the developed pipeline on a case study dataset of an under-construction project. This chapter is divided into seven sub-sections. The first sub-section is about project details. The second sub-section gives the details of the collected data. The third sub-section is on Revit modelling. The fourth sub-section is on annotating the dataset to prepare it for testing. The fifth sub-section is on implementing ConPro-NET for element detection on this dataset. The sixth sub-section shows the results obtained and their interpretation. The seventh sub-section discusses about the limitations of the method and future directions of research.

6.1 PROJECT DETAILS:

The case study was conducted on a local residential building project in Chennai, India. The existing building complex was built in 1990, and they were required to undergo an extension to have one bedroom and a bathroom with a balcony added to each house's existing plan. Figure 6.1 shows the previously existing quarter structure (left) and the currently built extension structure (right). The plinth beam stage was considered the zero-progress stage as no structural components were visible at this stage when the data collection started, as shown in Figure 6.1 (left). The current structure stage was considered the fifth stage, as shown in Figure 6.1 (right).

Figure 6.1 shows the image to reference the existing quarters being constructed to the existing structure.

Figure 6.1 Local Residential Project at Chennai

Figure 6.2 shows the drawing of the CAD plan for extension. The part shaded is the extension being developed on the existing structure.

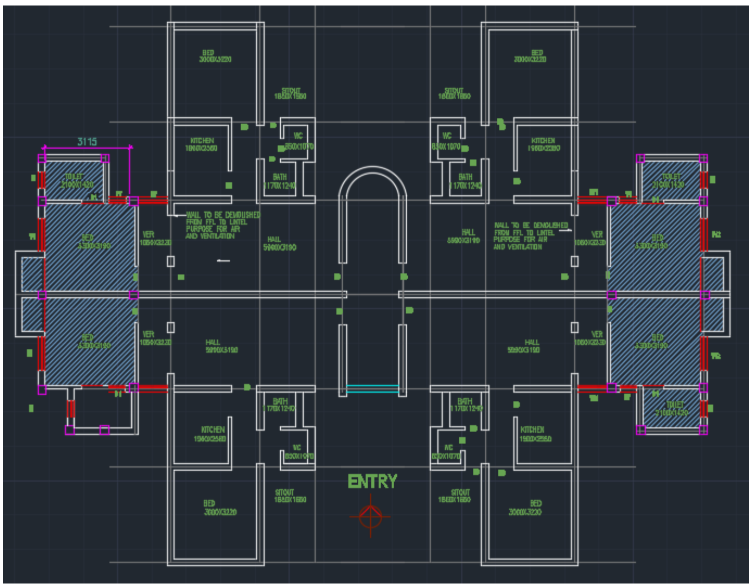


Figure 6.2 CAD Drawing for Extension Project

Figure 6.3 and Figure 6.4 shows the Plan and Elevation for the extension unit in detail. It consists of one Additional Bedroom with a window, a door to a balcony and a toilet.

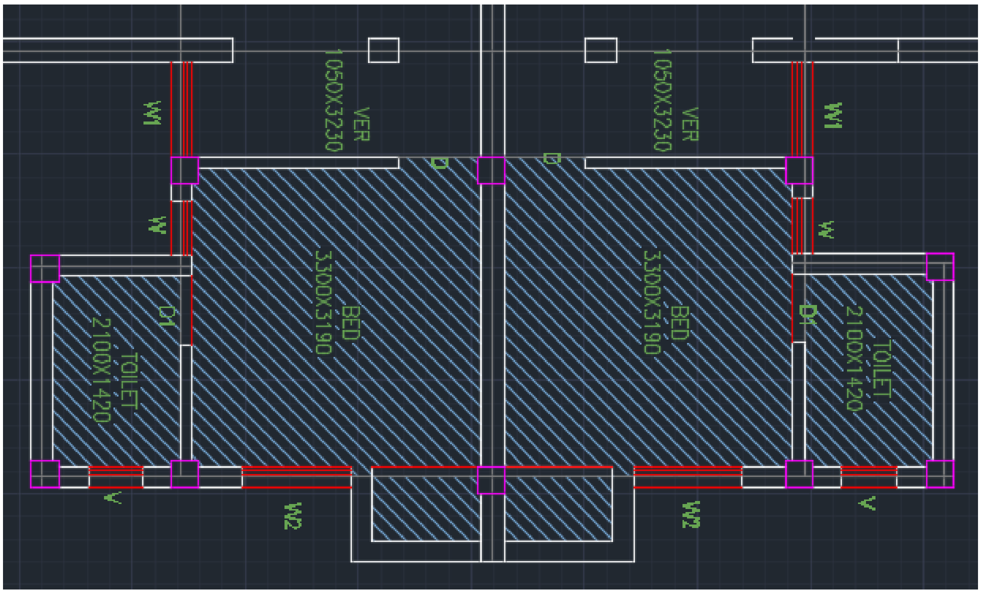


Figure 6.3 Plan for the Extension Unit in Detail

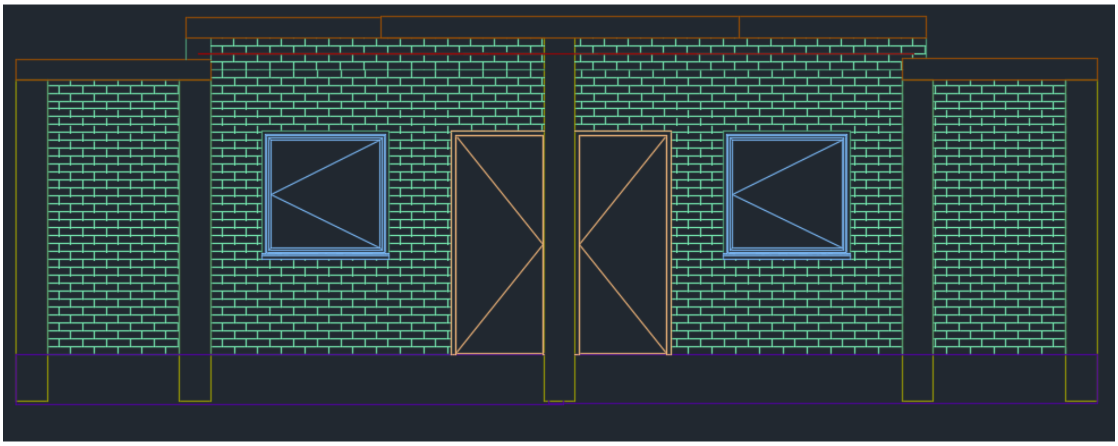


Figure 6.4 Elevation for One Floor

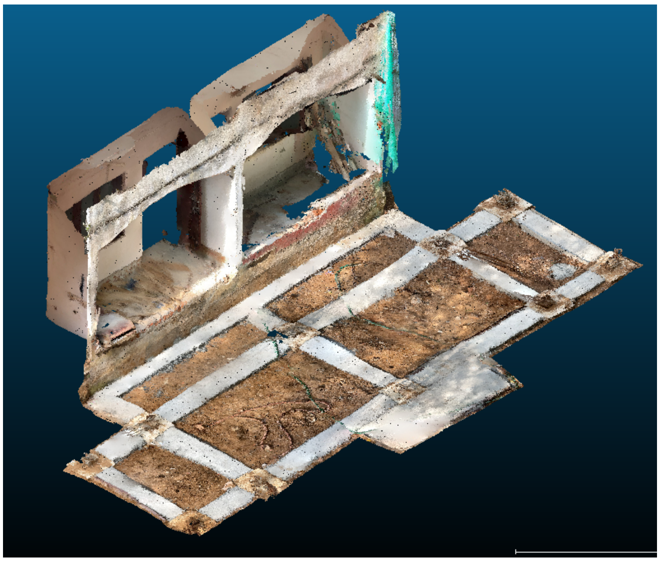
6.2 DATASET DETAILS:

For this study, only the ground floor structure was considered. The data was collected in the following five stages:

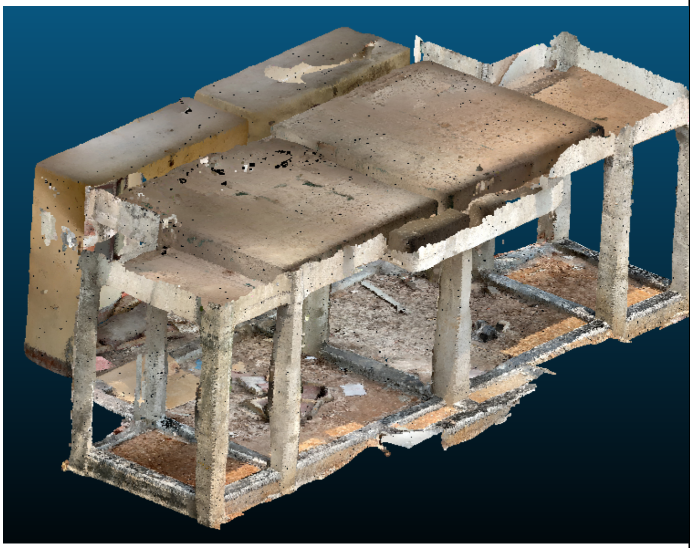
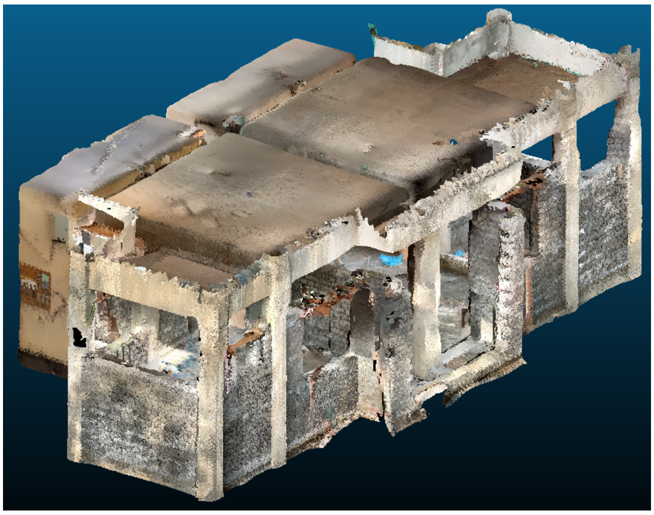
1. Plinth beam completion stage
2. Columns first lift completion stage
3. Columns second lift completion stage
4. Slab completion stage
5. Incomplete brickwork stage

These stages are shown in the Table 6.1 below.

Table 6.1 Point Cloud Data Captured at different Stages of the Construction Project

Sl. No.	**Stages**	**Point Cloud Collected**
1	**Zero Progress - Casted PPC for Plinth Beam**	

Sl. No.	Stages	Point Cloud Collected
2	Columns Erected First Lift	
3	Columns Erected Second Lift	

Sl. No.	Stages	Point Cloud Collected
4	Slab Casted	
5	Incomplete Brickwork	
6	Completed	NA

The data was collected using a hand-held device, iPad M1 Pro (Apple MHWC3HN/A 11' iPad Pro WIFI + Cellular, Space 1 TB). The LiDAR sensor in this iPad is used to capture the point clouds. This sensor works Direct Time of Flight (FTOF) ranging principle. Other details about the collected data are shown in Table 6.2.

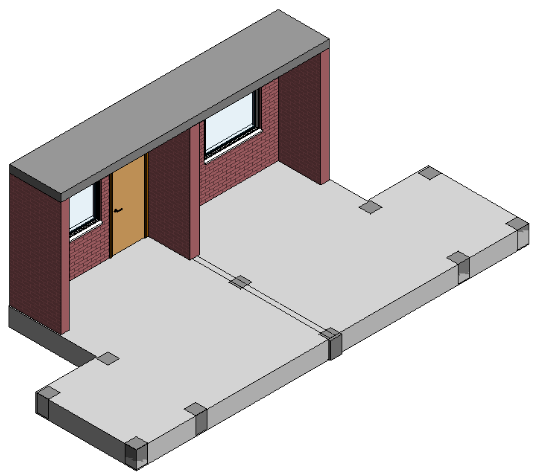
Table 6.2 Details about the data acquired

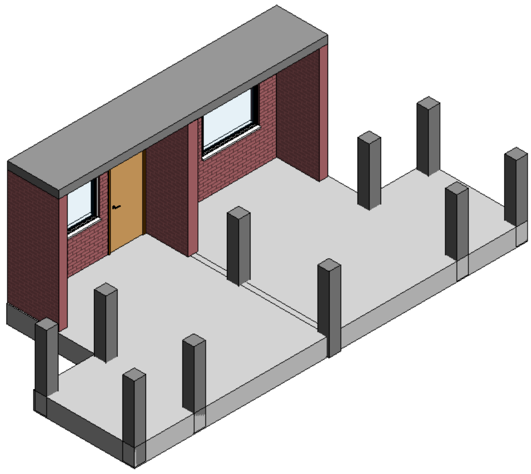
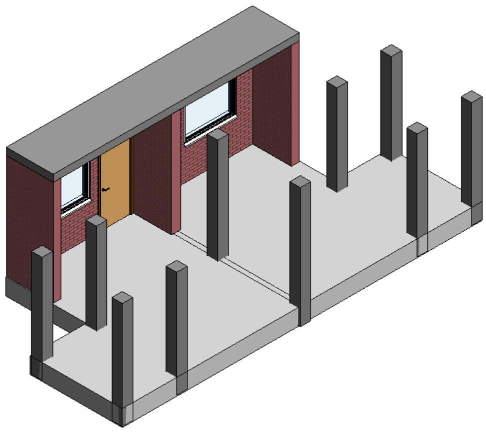
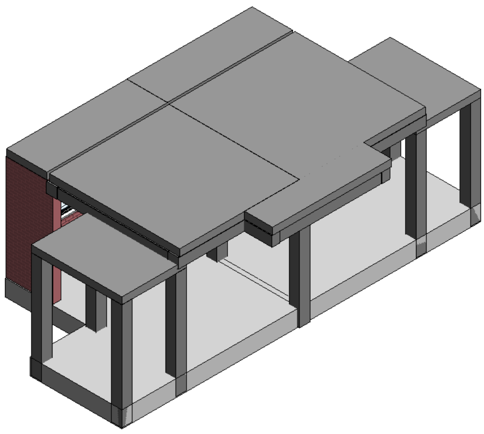
Description	Details
Acquisition Dates	19th March 2022, 27th March 2022, 2nd April 2022, 2nd May 2022, 15th May 2022
Acquisition Time	All the data was collected during 10AM to 11 AM to nullify the effect of variable sunlight.
Weather	The data was collected on clear sunny weather for all the five stages.
Data Volume	2.01 GB (Raw Point Clouds)
Application for Data Acquisition	3D Scanner (IOS)
Acquisition Device	iPad M1 Pro (Apple MHWC3HN/A 11' iPad Pro WIFI + Cellular, Space 1 TB)
Acquisition Settings	High Resolution
Application of Data Processing	Cloud Compare

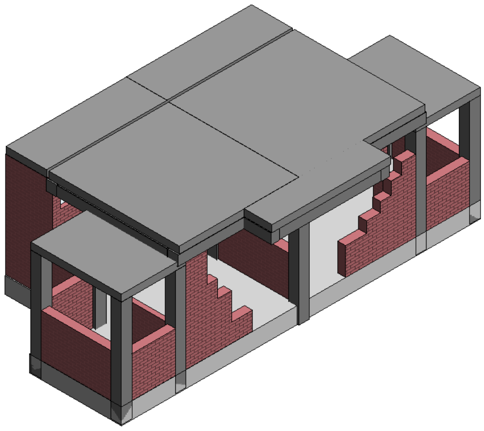
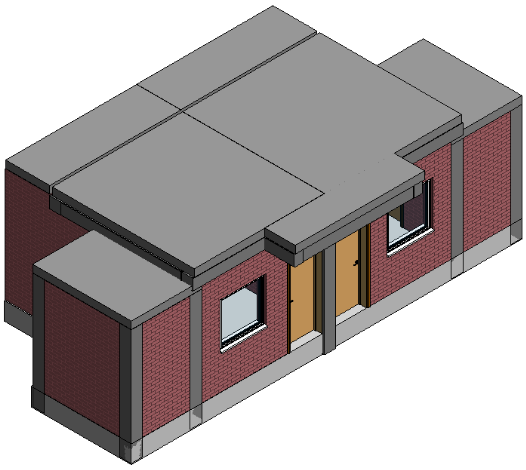
6.3 REVIT MODELLING

For reference and visualisation, Revit modelling was performed for the project and its various stages at which the site data was captured. The below Table 6.3 shows the stagewise models.

Table 6.3 As-built Revit BIM Model at different Stages of the Construction

Sl. No.	**Stages**	**As-Designed Revit Model**
1	**Zero Progress - Casted PPC for Plinth Beam**	

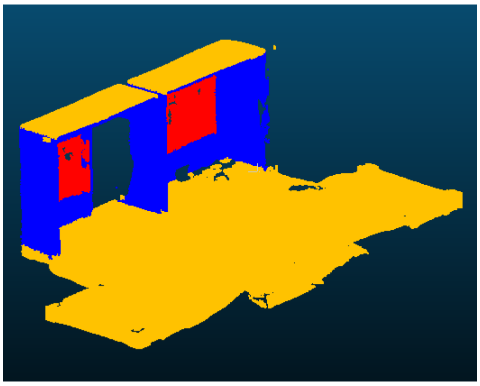
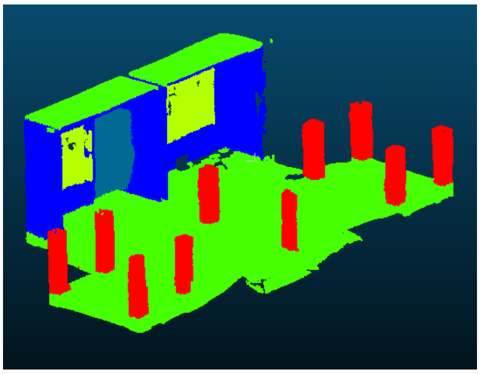
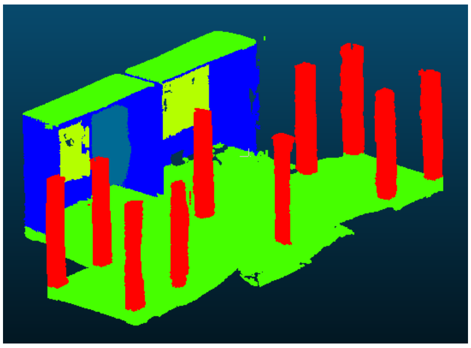
Sl. No.	Stages	As-Designed Revit Model
2	**Columns Erected First Lift**	
3	**Columns Erected Second Lift**	
4	**Slab Casted**	

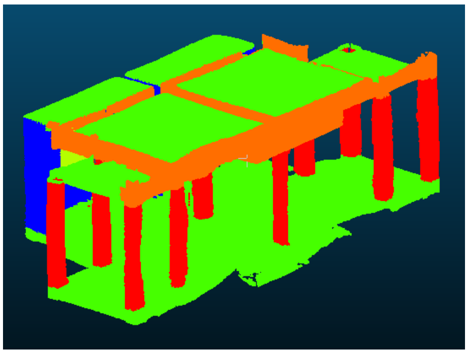
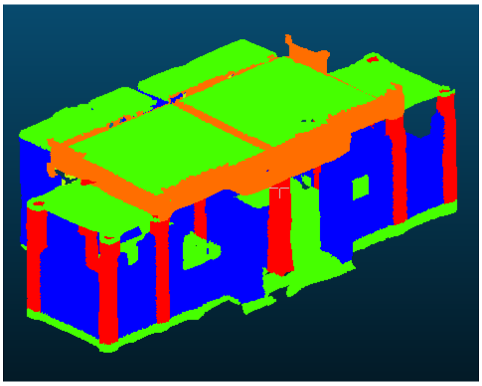
Sl. No.	Stages	As-Designed Revit Model
5	**Incomplete Brickwork**	
6	**Completed**	

6.4 ANNOTATING THE DATASET FOR TESTING

To validate the method the data needs to be annotated to compare the predicted results to the ground truth. The following Table 6.4 shows the annotated dataset.

Table 6.4 Annotated Point Cloud Data at different Stages of the Construction

Sl. No.	Stages	Annotated Model
1	Zero Progress - Casted PPC for Plinth Beam	
2	Columns Erected First Lift	
3	Columns Erected Second Lift	

Sl. No.	Stages	Annotated Model
4	**Slab Casted**	
5	**Incomplete Brickwork**	
6	**Completed**	**NA**

6.5 IMPLEMENTING CONPRO-NET

The stagewise unlabelled point cloud data as input, various steps of the developed method were performed sequentially as shown in Figure 6.5.

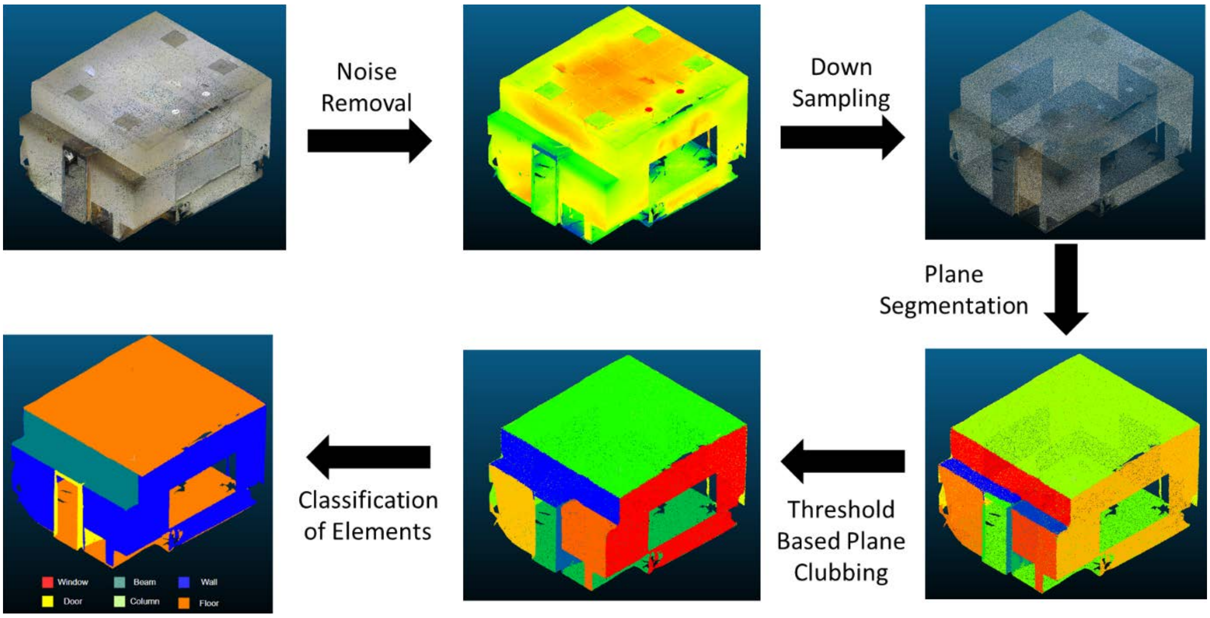


Figure 6.5 ConPro-NET - Hybrid Self-Supervised Pipeline Implementation

1. Noise Removal
2. Down Sampling
3. Plane Segmentation
4. Thresholding Based Plane Clubbing (threshold identified as 0.9m)
5. Classification and Labelling

6.6 RESULT AND INTERPRETATIONS

The developed hybrid-self supervised pipeline was executed using these five-point clouds as inputs and finally the corresponding five segmented and classified labelled point clouds were obtained as the outputs.

The best possible threshold of 0.9m was selected using the method described in section 3.3.3. However, by selecting this value, the merging of a few column elements with the slab and the merging of a few elements beams with the slab could not be avoided.

The developed hybrid-self supervised pipeline was executed using these five-point clouds as inputs and finally the corresponding five segmented and classified labelled point clouds were obtained as the outputs.

Figure 6.6 shows the visual results obtained by implementing the developed method using the hybrid features. Table 6.5 shows the combined confusion matrix for the case-study dataset with the three variations of, without handcrafted features, only handcrafted features, and hybrid features. Table 6.6, shows the results and accuracy obtained using each of these features.

From the results obtained it can be interpreted that:

- From Table 6.5 and Table 6.6 it can be observed that the results from the hybrid features outperforms the individual results by without handcrafted features and only handcrafted features.
- The overall accuracy obtained by Hybrid features is 80.95% which is slightly better than the accuracy obtained on the S3DIS dataset. This can be due to the size of the data being tested is small as compared to that of S3DIS data or the performance of the hybrid features is better in the mid-construction environment as they are designed for.
- From Figure 6.6 and Table 6.5, it can be seen that in Stage 4 and Stage 5, few

beam elements are not being recognized accurately as they are getting merged and embedded physically with the slab elements.

- In Figure 6.6 Stage 5, columns are not identified because all the columns are now physically embedded in walls that are getting built and therefore it is extremely hard for the algorithm to separate the column elements from the wall elements. Therefore, the algorithm predicts them as wall as they are now part of it.

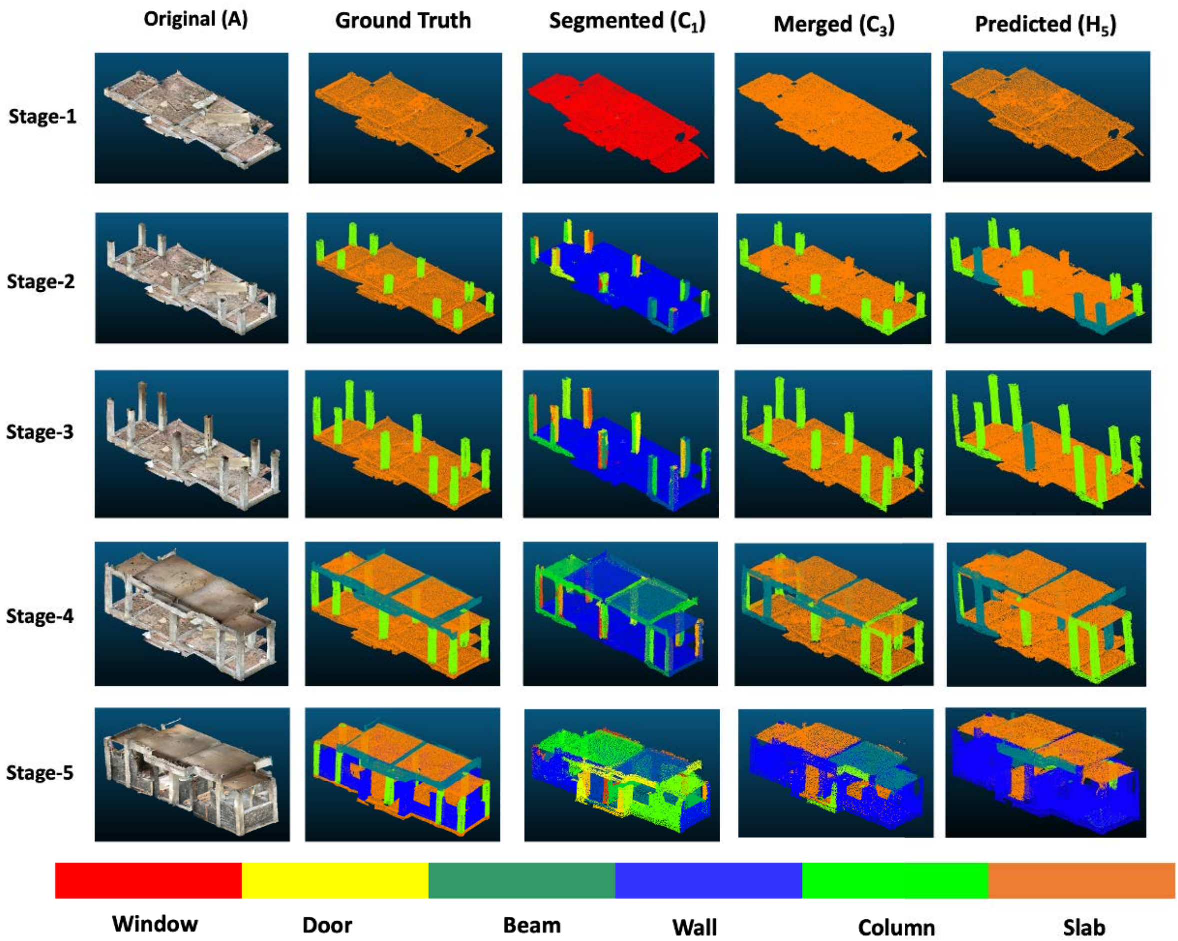


Figure 6.6 Visual results for various stages for the Residential Project Data

Table 6.5 Overall Confusion matrix for classification pipeline, with deep learning features, handcrafted features, and hybrid features for the entire Residential Project Data (Stage 1 to Stage 5)

Without Handcrafted Features				
	Wall	Slab	Beam	Column
Wall	10	0	1	1
Slab	1	14	2	0
Beam	0	2	5	0
Column	1	5	4	17

Only Handcrafted Features				
	Wall	Slab	Beam	Column
Wall	9	0	0	3
Slab	2	14	0	1
Beam	0	0	7	0
Column	0	0	12	15

Hybrid Features				
	Wall	Slab	Beam	Column
Wall	10	0	0	2
Slab	1	15	0	1
Beam	0	1	6	0
Column	0	1	6	20

Table 6.6 Combined results for deep learning, handcrafted features, and hybrid features for the entire Residential Project Data (Stage 1 to Stage 5)

Features	**Without Handcrafted Features**			**Only Handcrafted Features**			**Hybrid Features**		
Object Class	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
Wall	83.33%	83.33%	83.33%	81.81%	75.00%	78.26%	90.90%	83.33%	86.95%
Slab	66.66%	82.35%	73.68%	100.00%	82.35%	90.32%	88.23%	88.23%	88.23%
Beam	41.66%	71.42%	52.63%	36.84%	100.00%	53.84%	50.00%	85.71%	63.15%
Column	94.44%	62.96%	75.55%	78.94%	55.55%	65.21%	86.95%	74.07%	79.99%
Overall Accuracy	**73.01%**			**71.42%**			**80.95%**		

Color coding for Table 6.6

- 0% to 33.33% - Red (Low Score)

- 33.33% to 66.66% – Yellow (Medium Score)
- 66.66% to 100% – Green (High Score)

6.7 LIMITATIONS AND FUTURE WORK

The present study introduces a prospective hybrid self-supervised pipeline, called ConPro-NET, which can be employed for progress monitoring of building structures. However, there exist a few limitations in the proposed method, which need to be addressed. Below are the limitations alongside with potential future directions:

- Dataset: Currently, the dataset utilized for pre-training is the S3DIS dataset, which comprises completed elements. In future, to enhance accuracy in tracking incomplete components, a significant amount of mid-construction data will be captured and utilized. Also, the use of synthetic data which can be extracted from the as-planned BIM seems promising and will be explored in future.
- Process: The current methodology for merging planes in a new dataset involves manual threshold estimation. However, in future research, we aim to develop an automated algorithm that utilizes as-planned BIM to determine the optimal threshold value that maximizes overall accuracy.
- Learning Algorithm: Currently only two parts are used in learning. Objects can be split in multiple parts with multiple views and contrastive loss can be used for learning representations without having to label them. In real-world scenarios, point clouds are often captured alongside other modalities such as images or videos. Developing methods that can effectively learn from these

multi-modal sources could help improve the quality and robustness of unsupervised or self-supervised learning of point clouds.

- Loss Function: Contrastive loss function can be used in place of cross-entropy loss as contrastive loss considers the distance between the positive parts and negative part to learn the matching or non-matching pairs. Hence it can be efficiently used with unlabelled data in place of cross entropy loss.
- Feature Engineering: Furthermore, this study introduces a set of handcrafted features and utilizes them according to their calculated feature importance score. However, in the future, the individual contribution and computation time of these handcrafted features towards the overall accuracy will be evaluated and documented to ensure a balanced and optimized selection. In addition to this, other derived semantic features will be explored and their performance on the classification accuracy will be evaluated.
- Experimental Testbed: Here, the data was acquired using a low-cost hand-held device to show the method's applicability and robustness, it will be interesting to conduct experiments using high precision terrestrial laser scanners, which might result in better classification accuracy.

CHAPTER 7.

CONCLUSIONS

This chapter summarizes the research findings, presents the contributions, and highlights the future direction of research.

7.1 SUMMARY AND CONCLUSIONS

The main goal of this work is to explore the use of computer vision for construction progress monitoring. To achieve this goal, this study established three research objectives with sub-objectives and conducted several tasks to achieve them. Here, the research objectives have been translated in form of questions and conclusions drawn from these objectives are made.

Questions from Research Objective 1: What is the state of the art of progress monitoring in construction industry, in practice as well as in literature? What factors a construction firms should consider while selecting an appropriate progress monitoring technology for projects?

To seek the answer to this question, first a broad review of literature on progress monitoring methods available was conducted using a PRISMA methodology. Then these technologies were classified into six broad categories and evaluated for their advantages and disadvantages from an application perspective. Next, a systematic questionnaire survey was conducted with the participants from the industry (India and UAE). The questionnaire was designed to deduce the progress monitoring technologies being implemented on the projects and the challenges these technologies face for progress monitoring. Finally, after understanding the literature as well as practice, factors affecting selection of progress monitoring technologies were identified and reported. A method to use these factors as a basis for objectively selecting the technology for a specific project was also presented. A RII score was calculated for these factors based on a questionnaire survey. These factors with their importance can be evaluated and it can help in selecting a progress monitoring technology for a particular project.

Conclusions from Research Objective 1: Currently the state of art of progress monitoring at sites lags with the progress made in the literature. This is evident from the results of the survey-based study, which shows that construction sites still rely on excel-sheet based methods of manual progress estimation and other technologies are being sparsely used on sites. This gap is due to several reasons and challenges as highlighted for on-site implementation. As an attempt to explore the individual progress monitoring technologies were classified and analysed both of their advantages and disadvantages. Next, the method for progress monitoring technology selection at construction sites does not follow a scientific approach. Through a structured and systematic approach this study identifies the key factors for the selection and identifies their relative importance which will support in decision making. The most important factor was found to be the project type and characteristic which should be considered for progress monitoring technology selection. However, these importance values are the preliminary research findings, more in-depth work is required for their validation in future.

Question from Research Objective 2: What are the key components of a Computer-Vision Based Pipeline for Construction progress monitoring?

This question was answered by assembling an integrated CV-CPM framework that captures the process requirements of construction progress monitoring and enables the characterization and categorization of current and future work in the area. Next, this CV-CPM framework was utilized to position and compare various concepts and tools adopted by published research studies. After the framework was published, an attempt was made to position the new studies on CV-CPM along the framework. Finally, areas and strategies for future work using the framework were identified. A key area proposed was the need of benchmarking for the various tools, techniques, algorithms of the CV-CPM pipeline.

Conclusions from Research Objective 2: CV-CPM has the potential of creating an immense impact by providing real-time, accurate, reliable information to construction managers. Though a significant amount of work has been done in the last decade, specific challenges remain due to the construction industry's dynamic nature and complexities at sites. Some of these challenges have been addressed; however, significant gaps need to be filled to make pipelines accurate and automated to meet rising user expectations of real-time feedback.

It was found that the four levels of progress monitoring identified in this study strongly influence the technology selection for the pipeline at each stage. For implementing progress monitoring on a construction project, it is recommended that the pipeline and components required are selected based on the chosen level. Among several potential research areas, advancements in the as-built modelling stage are required to facilitate the quantification of progress. To enable this, exploring a hybrid approach that combines learning with heuristics is recommended. Addressing the requirements and following a well-developed roadmap in this area is essential to move CV-CPM research from laboratory studies to field applications.

Question from Research Objective 3: Can a hybrid approach increase the accuracy of un-supervised / self-supervised learning methods to be utilized to make CV-CPM feasible to construction progress monitoring?

For this, the research proposed a ConPro-NET, hybrid self-supervised approach for CV-CPM. This approach is the first application of unsupervised segmentation and self-supervised classification which is curated specifically to detect elements from construction point cloud data for progress quantification. This pipeline performed with high level of accuracy on the S3DIS dataset as well as an under-construction dataset as compared to the other supervised approaches from literature. A hybrid self-supervised approach was adopted which uses the mix of features obtained using a deep learning based contrastive approach as well as specific handcrafted features for construction elements. This method using hybrid features is a key to obtaining a reasonably good overall accuracy even from a self-supervised approach.

Conclusions from Research Objective 3: The key reasons for the existing learning based and heuristic based approaches of element identification not being used at construction sites is the ineffectiveness of these methods individually. The former requires large, labelled datasets as it is dependent on supervised models of learning and the later requires a significant amount of hard coding and domain knowledge. The labelling of point clouds acquired from construction site is a highly tedious and labour-intensive task. Also, there are so many variations of the types of components used at construction site, therefore a project specific labelled dataset is difficult to obtain. Even the components on under-construction site are mostly grey and therefore traditional segmentation techniques cannot be applied. Therefore, the direction of research is to utilize the unsupervised and self-supervised methods, which do not require large, labelled datasets. Also, the strengths of both learning and heuristics-based approaches were complimented by the use of hybrid feature vector. <u>It can be concluded that the hybrid features performed the best among the three different feature sets used in the pipeline.</u>

7.2 THESIS CONTRIBUTION

The detailed contribution of individual objectives has already been discussed at the end of each chapter. Here is a broad summary of the following scientific contributions that are achieved by this research:

1. **Evaluated the State of the art of construction progress monitoring in construction.**

Evaluated the state of the art of progress monitoring of construction, from literature and practice. Evaluated and classified various progress monitoring technologies available for construction. Identified various factors that affect selection of progress monitoring technologies for construction project. The relative importance index of these factors was obtained, and the critical factors were identified. These factors are initial steps in the direction of making the selection process objective and scientific and can be utilized by the decision makers to select the progress monitoring technology for a project.

2. **Developed a comprehensive framework capturing end to end process of computer vision-based construction progress monitoring from data acquisition to progress estimation.**

Developed an integrated framework for Computer Vision-Based Construction Progress Monitoring (CV-CPM) (The acronym CV-CPM was coined in this research). Introduced the four Levels of Progress monitoring (LPM) for categorization of on-site progress monitoring requirements and associating these with the various processes of CV-CPM framework. Through the framework several future research areas have been identified. The framework can be used to develop a roadmap for future work in CV-CPM.

3. **Developed ConPro-NET - a novel hybrid self-supervised approach utilizing the handcrafted features which performs equally par with supervised approach with less labelling effort requirements.**

The thesis proposed ConPro-NET, a novel hybrid self-supervised approach for computer vision-based construction progress monitoring. The pipeline consists of a customized unsupervised segmentation approach and a self-supervised contrastive learning for classification of construction elements. The proposed method for construction progress monitoring outperforms the existing approaches in terms of the effort put in as it is a self-supervised approach.

7.3 FUTURE RESEARCH DIRECTIONS

An extensive discussion on future works is presented appropriately in each chapter. This section highlights the summary of future research directions. The future research direction can be in the broad CV-CPM framework or in the niche area of hybrid self-supervised pipeline. These have been highlighted separately in the following sub-sections.

7.3.1 Future Research Directions: Progress Monitoring Technology Selection in Construction

There need to be separate framework for selection of progress monitoring technology based on the project type and it characteristic. Same technology cannot be used at every project. The framework can be a decision framework which will depend on many other factors. Also, there in future there should be a validation study of the ranking of factors that have been identified in this research.

7.3.2 Future Research Directions: CV-CPM

The CV-CPM research can be extended in the following directions. This section highlights potential research questions within the three stages of the CV-CPM framework. Several of these questions can be addressed through benchmarking studies, while others require exploration of new concepts.

7.3.2.1 Areas within the three stages of the framework

Stage 1: Data acquisition and 3D reconstruction

- How to objectively assess the factors affecting data acquisition technology and sensor mounting method shown in Figure 4.2?
- What are the appropriate sensing technology and mounting method for data acquisition based on project types and project characteristics (ex., Linear, underground, elevated etc.)?
- Based on factors identified in Figure 4, How to decide on using SfM and SLAM for a specific project?
- Vision-based data is usually enormous, and storage requirements are high. How sparse can data be without impacting the output?

Stage 2: As-built modelling

- How can the data be effectively pre-processed to meet the point cloud quality requirements for an as-built model generation?
- How can a hybrid of heuristics and learning-based approaches be deployed, such that it exploits the advantages and overcomes the shortcomings of each?
- How can cognitive computing be successfully implemented to recognise and measure complex geometries at reduced computational cost accurately?

Stage 3: Progress Monitoring

- How to objectively compare the computational requirements and output accuracy for the level of progress monitoring based on the type of as-built models as shown in Figure 4.4?

- How can virtual and physical environments be seamlessly connected to improve visualisation?
- How to select an appropriate as-built model and progress quantification technique based on project characteristics and the trade-off mentioned in Figure 4.4?
- How to generate automatic schedules update and control suggestions by applying cognitive computing techniques to the progress data produced by CV-CPM?

7.3.2.2 Digital Twins for CPM

In addition to the specific areas identified above, the CV-CPM framework is also relevant to supporting emerging areas of research and development, such as Digital Twin.

Digital Twin technologies enable the built facility's virtual representation to mirror and predict the state/behaviour of the physical facility. Key technologies for Digital Twin implementation include:

1. continuous streaming and processing of sensing data,
2. threaded models and simulations for current/future state/behaviour of the facility,
3. formulating/implementing interventions to limit future state/behaviour within specified limits.

Studies have envisioned the utility of Digital Twin in all phases of a construction project lifecycle. However, details on twinning an under-construction facility are limited. Commercial platforms such as Autodesk Tandem ("Autodesk Tandem," 2021) and Bentley iTwin (Bentley, 2021) are available for creating Digital Twin solutions. However, technological features available in these platforms are currently in the early stages and focus on the operations and maintenance phase of the built asset.

A key issue in creating a Digital Twin for construction progress monitoring is that the geometry of the facility keeps expanding; hence the corresponding sensor positioning is also dynamic. In this context, computer vision-based sensing would be the most appropriate input to capture geometrical attributes directly and rapidly. The proposed CV-CPM framework identifies the essential technologies to embed progress monitoring capabilities in a Digital Twin.

Pipelines that can stream and process the progress data accurately and rapidly to mirror the progress in the virtual model will support Digital Twinning for progress monitoring. For this, the process within and across the stages of the CV-CPM framework needs to be optimised for real-time updates.

In addition to mirroring the progress, a Digital Twin can forecast construction progress as well as simulate and evaluate control measures to bring the project back on track. As the usage of autonomous construction equipment increases, the control measures can be transmitted from the Digital to the physical world to be implemented by the equipment.

7.3.3 Future Research Directions: Hybrid Self-Supervised Pipeline

There are few current limitations of the existing pipeline, that should be researched in future.

7.3.3.1 Automatic algorithm for threshold selection

Currently, the method for threshold selection for clubbing the individual planes for the unsupervised segmentation is not automatic. In future, an automated algorithm to find the exact threshold value which gives the maxima for overall accuracy will be developed for simplifying the step for a new dataset.

7.3.3.2 Feature Engineering

A set of handcrafted features have been introduced in this paper and have been used according to the calculated feature importance score. In future, the individual contribution, and the computation time for the various handcrafted features to the overall accuracy will be evaluated and documented for a balanced and optimized selection.

There can be more innate features of the elements, for example, the topographic and morphological features which depends on the geometric shape and morphology of the object. Currently only the direct features were used, in future these derived features can be computed and tested to see how they contribute to the classification accuracy.

7.3.4 Future Research Directions: Benchmarking Studies

The integrated CV-CPM framework and the guidelines proposed for each stage is expected to assist in developing a strategy and a roadmap for benchmarking. Also, there are various parameters and hyperparameters which could be benchmarked in the developed hybrid-self supervised approach. As a wide range of studies is required, the technology roadmap needs to be developed collaboratively by the research community. A starting point for these studies can be to investigate the subjective ratings proposed in this study and quantify these ratings through controlled experimental testbeds as discussed in detail in section 4.5.1. The CV-CPM framework can help in structuring and prioritising the areas to be explored in developing a roadmap so that they form a standard reference for benchmarking studies.

APPENDIX A.

QUESTIONNAIRE SURVEY – PROGRESS MONITORING TECHNOLOGIES ON CONSTRUCTION PROJECTS

Progress Monitoring Technologies on Construction Projects

I am Varun Kumar Reja a Ph.D. Research Scholar at IIT Madras working on "Automated Progress Monitoring in Construction".

This Survey is for my research and is to get the state-of-art of Progress monitoring technologies used for Construction Projects. If You have worked on multiple projects you can fill this form again with the other project details. Kindly Share this form with anyone who you feel will be relevant for providing the information.
Thank You for Your Time.

[redacted]@smail.iitm.ac.in Switch account

* Required

Email *

Your email

Your Name (Not Mandatory to Mention)

Your answer

Organization Name *

Your answer

Project Name *

Your answer

Project Cost (Approximately in Crores)

Your answer

Technology in Specified Project Used for Progress Monitoring *

Kindly Check All The boxes That Apply

- [ ] Pen & Paper Based Quantity Calculation & DPR Reporting
- [ ] Excel Sheet Based Quantity Calculation & DPR Reporting
- [ ] QR Code / Bar-codes
- [ ] RFID Based / Sensor Based
- [ ] CCTV Based Surveillance
- [ ] Fixed Cameras (Time-lapse/Periodic Photos)
- [ ] Drone Based Monitoring (Images/Videos)
- [ ] GPS Based
- [ ] Laser Scanners
- [ ] Remote Sensing Based
- [ ] Still Cameras Photos / Mobile Photos or videos (For Only Documentation)
- [ ] Direct Quantity Capture From Images (Using Image Processing & 3D Reconstruction)
- [ ] Other:

Which Platform or Mobile Application did you use to Register Daily Progress? (Whatsapp Reporting / Procube / N-pulse etc.) Kindly Mention

Your answer

Challenges in the way progress were monitored in the project? *

Kindly Check All The boxes That Apply

- [ ] Low accuracy
- [ ] Time Consuming
- [ ] Ineffective as it was too infrequent to enable prompt control action
- [ ] Too much manual work Required to compute
- [ ] High Cost of Technology & Equipment
- [ ] Requirement of Skilled Staff
- [ ] Non-systematic & inconsistent because of human involvement
- [ ] Technological Implementation Issues
- [ ] Other:

Submit Clear form

APPENDIX B.

QUESTIONNAIRE SURVEY – RANKING FACTORS AFFECTING PROGRESS MONITORING TECHNOLOGY SELECTION ON CONSTRUCTION PROJECTS

Ranking Factors affecting progress monitoring technology selection

Kindly use the below definitions in case of any ambiguity or email to varunreja7@gmail.com

1. **Level of Automation:** The extent of manual effort required while using the technique.

2. **Time efficiency:** The speed of data acquisitions as well as data processing.

3. **Operating range:** The distance up to which the employed technology works.

4. **Utility:** Adaptability of the technology used in interior and exterior construction progress monitoring. In other words, whether the technology is a general case solution.

5. **Preparation required:** The level of preparation required while setting up the equipment or process at the deployment stage.

6. **Accuracy:** The reliability of the collected data along with precision.

7. **Training required:** The amount of training or knowledge a user requires before using a particular technique.

8. **Cost:** The amount of financial and computational costs incurred to adopt and implement the technology.

9. **Susceptibility in adverse weather:** The extent of use of the technology in harsh environmental conditions like low visibility

10. **Compatibility for use:** The level by which a particular technology can be integrated with other technologies or the existing Enterprise Resource Management (ERP) system.

11. **Statutory requirements:** The legal codes and procedures to be followed while using a technique enforced by the authorities.

12. **Mobility:** The ease, flexibility, and portability of the related equipment.

13. **Project type & characteristics**: The type of the project and characteristics of a particular construction project where the technology can be used.

Rate the Factors affecting progress monitoring technology selection in order of their relative importance (1 - Most Important, 13- Least Important) *

	1	2	3	4	5	6	7	8	9
Level of Automation	○	○	○	○	○	○	○	○	○
Time efficiency	○	○	○	○	○	○	○	○	○
Operating range	○	○	○	○	○	○	○	○	○
Utility	○	○	○	○	○	○	○	○	○
Preparation required	○	○	○	○	○	○	○	○	○
Accuracy	○	○	○	○	○	○	○	○	○
Training required	○	○	○	○	○	○	○	○	○
Cost	○	○	○	○	○	○	○	○	○
Susceptibility in adverse weather	○	○	○	○	○	○	○	○	○
Compatibility for use	○	○	○	○	○	○	○	○	○
Statutory requirements	○	○	○	○	○	○	○	○	○
Mobility	○	○	○	○	○	○	○	○	○
Project type & characteristics	○	○	○	○	○	○	○	○	○

Submit

Clear form

REFERENCES

1. Adán, A., Quintana, B., Prieto, S. A., & Bosché, F. (2020). An autonomous robotic platform for automatic extraction of detailed semantic models of buildings. *Automation in Construction*, *109*, 102963. https://doi.org/10.1016/j.autcon.2019.102963

2. Adán, Antonio, Quintana, B., Prieto, S. A., & Bosché, F. (2018). Scan-to-BIM for ‘secondary’ building components. *Advanced Engineering Informatics*, *37*(November 2017), 119–138. https://doi.org/10.1016/j.aei.2018.05.001

3. Alaloul, W. S., Qureshi, A. H., Musarat, M. A., & Saad, S. (2021). Evolution of close-range detection and data acquisition technologies towards automation in construction progress monitoring. *Journal of Building Engineering*, *43*, 102877. https://doi.org/10.1016/j.jobe.2021.102877

4. Ali, A. K., Lee, O. J., Lee, D., & Park, C. (2021). Remote indoor construction progress monitoring using extended reality. *Sustainability (Switzerland)*, *13*(4), 1–24. https://doi.org/10.3390/su13042290

5. ALICE Technologies. (2021). *AI for Construction Management*.

6. Alizadehsalehi, S., & Yitmen, I. (2016a). The Impact of Field Data Capturing Technologies on Automated Construction Project Progress Monitoring. *Procedia Engineering*, *161*, 97–103. https://doi.org/10.1016/j.proeng.2016.08.504

7. Alizadehsalehi, S., & Yitmen, I. (2016b). The Impact of Field Data Capturing Technologies on Automated Construction Project Progress Monitoring. *Procedia Engineering*, *161*, 97–103. https://doi.org/10.1016/j.proeng.2016.08.504

8. Alizadehsalehi, S., & Yitmen, I. (2019). A Concept for Automated Construction Progress Monitoring: Technologies Adoption for Benchmarking Project Performance Control. *Arabian Journal for Science and Engineering*, *44*(5), 4993–5008. https://doi.org/10.1007/s13369-018-3669-1

9. Alizadehsalehi, S., & Yitmen, I. (2021). Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX). *Smart and Sustainable Built Environment*. https://doi.org/10.1108/SASBE-01-2021-0016

10. Arif, F., & Khan, W. A. (2021a). Smart Progress Monitoring Framework for Building Construction Elements Using Videography–MATLAB–BIM Integration. *International Journal of Civil Engineering*. https://doi.org/10.1007/s40999-021-00601-3

11. Arif, F., & Khan, W. A. (2021b). Smart Progress Monitoring Framework for

Building Construction Elements Using Videography–MATLAB–BIM Integration. *International Journal of Civil Engineering*, *19*(6), 717–732. https://doi.org/10.1007/s40999-021-00601-3

12. Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., & Savarese, S. (2016). 3D Semantic Parsing of Large-Scale Indoor Spaces. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, *2016-Decem*, 1534–1543. IEEE. https://doi.org/10.1109/CVPR.2016.170

13. Arnaud, A., Christophe, J., Gouiffes, M., & Ammi, M. (2016). 3D reconstruction of indoor building environments with new generation of tablets. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, *02-04-Nove*, 187–190. New York, NY, USA: ACM. https://doi.org/10.1145/2993369.2993403

14. Asadi, K., Kalkunte Suresh, A., Ender, A., Gotad, S., Maniyar, S., Anand, S., … Wu, T. (2020). An integrated UGV-UAV system for construction site data collection. *Automation in Construction*, *112*(June 2019), 103068. https://doi.org/10.1016/j.autcon.2019.103068

15. Autodesk Tandem. (2021). Retrieved January 10, 2022, from Autodesk website: https://intandem.autodesk.com/

16. Bailey, T., & Durrant-Whyte, H. (2006). Simultaneous localization and mapping (SLAM): part II. *IEEE Robotics & Automation Magazine*, *13*(3), 108–117. https://doi.org/10.1109/MRA.2006.1678144

17. Barazzetti, L. (2016). Parametric as-built model generation of complex shapes from point clouds. *Advanced Engineering Informatics*, *30*(3), 298–311. https://doi.org/10.1016/j.aei.2016.03.005

18. Bassier, M., Vincke, S., Mattheuwsen, L., de Lima Hernandez, R., Derdaele, J., & Vergauwen, M. (2019). Percentage of completion of in-situ cast concrete walls using point cloud data and BIM. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-5/W2*(5/W2), 21–28. https://doi.org/10.5194/isprs-archives-XLII-5-W2-21-2019

19. Bassier, Maarten, Bonduel, M., Van Genechten, B., & Vergauwen, M. (2017). Segmentation of large unstructured point cloud using Octree-Based Region Growing and Conditional Random Fields. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W8, 2017 5th International Workshop LowCost 3D – Sensors, Algorithms, Applications*, *XLII*(November), 28–29.

20. Bentley. (2021). An Open Platform for Infrastructure Digital Twins : How Bentley Systems' iTwin Platform is enabling a digital twin ecosystem for the world's infrastructure. *A Frost & Sullivan White Paper*. Retrieved from https://prod-bentleycdn.azureedge.net/-/media/files/documents/white-

papers/wp_frost_sullivan_open_platform_for_infrastructure_digital_twins.pdf?la=pt-br&modified=00010101000000

21. Bognot, J. R., Candido, C. G., Blanco, A. C., & Montelibano, J. R. Y. (2018). Building construction progress monitoring using Unmanned Aerial System (UAS), low-cost photogrammetry, and Geographic Information System (GIS). *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *IV–2*(2), 41–47. https://doi.org/10.5194/isprs-annals-IV-2-41-2018

22. Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, *114*(March), 103179. https://doi.org/10.1016/j.autcon.2020.103179

23. Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., & Haas, R. (2015). The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components. *Automation in Construction*, *49*, 201–213. https://doi.org/10.1016/j.autcon.2014.05.014

24. Boulch, A., & Marlet, R. (2012). Fast and Robust Normal Estimation for Point Clouds with Sharp Features. *Computer Graphics Forum*, *31*(5), 1765–1774. https://doi.org/10.1111/j.1467-8659.2012.03181.x

25. Braun, A., Borrmann, A., Tuttas, S., & Stilla, U. (2014). Big data in smart building operation. In *eWork and eBusiness in Architecture, Engineering and Construction* (pp. 765–804). CRC Press. https://doi.org/10.1201/b17396-20

26. Braun, Alex, Tuttas, S., Borrmann, A., & Stilla, U. (2020). Improving progress monitoring by fusing point clouds, semantic data and computer vision. *Automation in Construction*, *116*(March), 103210. https://doi.org/10.1016/j.autcon.2020.103210

27. Bueno, M., González-Jorge, H., Martínez-Sánchez, J., & Lorenzo, H. (2017). Automatic point cloud coarse registration using geometric keypoint descriptors for indoor scenes. *Automation in Construction*, *81*(June), 134–148. https://doi.org/10.1016/j.autcon.2017.06.016

28. Bueno, Martín, Bosché, F., González-Jorge, H., Martínez-Sánchez, J., & Arias, P. (2018). 4-Plane congruent sets for automatic registration of as-is 3D point clouds with 3D BIM models. *Automation in Construction*, *89*(February), 120–134. https://doi.org/10.1016/j.autcon.2018.01.014

29. Chen, J., Fang, Y., & Cho, Y. K. (2017). Unsupervised Recognition of Volumetric Structural Components from Building Point Clouds. *Computing in Civil Engineering 2017*, 34–42. Reston, VA: American Society of Civil Engineers. https://doi.org/10.1061/9780784480823.005

30. Chen, J., Kira, Z., & Cho, Y. K. (2019). Deep Learning Approach to Point

Cloud Scene Understanding for Automated Scan to 3D Reconstruction. *Journal of Computing in Civil Engineering*, *33*(4), 04019027. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000842

31. Cheng, M. Y., & Chen, J. C. (2002). Integrating barcode and GIS for monitoring construction progress. *Automation in Construction*, *11*(1), 23–33. https://doi.org/10.1016/S0926-5805(01)00043-7

32. Dai, A., Chang, A. X., Savva, M., Halber, M., Funkhouser, T., & Niessner, M. (2017). ScanNet: Richly-Annotated 3D Reconstructions of Indoor Scenes. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, *2017-Janua*, 2432–2443. IEEE. https://doi.org/10.1109/CVPR.2017.261

33. Deng, Y., Gan, V. J. L., Das, M., Cheng, J. C. P., & Anumba, C. (2019). Integrating 4D BIM and GIS for Construction Supply Chain Management. *Journal of Construction Engineering and Management*, *145*(4), 04019016. https://doi.org/10.1061/(asce)co.1943-7862.0001633

34. Díaz-Vilariño, L., Khoshelham, K., Martínez-Sánchez, J., & Arias, P. (2015). 3D Modeling of Building Indoor Spaces and Closed Doors from Imagery and Point Clouds. *Sensors*, *15*(2), 3491–3512. https://doi.org/10.3390/s150203491

35. Dimitrov, A., & Golparvar-Fard, M. (2014). Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections. *Advanced Engineering Informatics*, *28*(1), 37–49. https://doi.org/10.1016/j.aei.2013.11.002

36. Ekanayake, B., Wong, J. K. W., Fini, A. A. F., & Smith, P. (2021a, July 1). Computer vision-based interior construction progress monitoring: A literature review and future research directions. *Automation in Construction*, Vol. 127. Elsevier B.V. https://doi.org/10.1016/j.autcon.2021.103705

37. Ekanayake, B., Wong, J. K.-W., Fini, A. A. F., & Smith, P. (2021b). Computer vision-based interior construction progress monitoring: A literature review and future research directions. *Automation in Construction*, *127*(April), 103705. https://doi.org/10.1016/j.autcon.2021.103705

38. Fathi, H., Dai, F., & Lourakis, M. (2015). Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Advanced Engineering Informatics*, *29*(2), 149–161. https://doi.org/10.1016/j.aei.2015.01.012

39. Franz, S., Irmler, R., & Rüppel, U. (2018). Real-time collaborative reconstruction of digital building models with mobile devices. *Advanced Engineering Informatics*, *38*(December 2017), 569–580. https://doi.org/10.1016/j.aei.2018.08.012

40. Furukawa, Y., Curless, B., Seitz, S. M., & Szeliski, R. (2009). Reconstructing building interiors from images. *2009 IEEE 12th International Conference on*

Computer Vision, 80–87. IEEE. https://doi.org/10.1109/ICCV.2009.5459145

41. Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2011). Integrated Sequential As-Built and As-Planned Representation with D4AR Tools in Support of Decision-Making Tasks in the AEC/FM Industry. *Journal of Construction Engineering and Management*, *137*(12), 1099–1116. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000371

42. Golparvar-Fard, M., Peña-Mora, F., & Savarese, S. (2015). Automated Progress Monitoring Using Unordered Daily Construction Photographs and IFC-Based Building Information Models. *Journal of Computing in Civil Engineering*, *29*(1), 04014025. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000205

43. Guven, G., & Ergen, E. (2021). Tracking major resources for automated progress monitoring of construction activities: masonry work case. *Construction Innovation*, *21*(4), 648–667. https://doi.org/10.1108/CI-05-2020-0081

44. Halder, S., Afsari, K., Serdakowski, J., DeVito, S., Ensafi, M., & Thabet, W. (2022). Real-Time and Remote Construction Progress Monitoring with a Quadruped Robot Using Augmented Reality. *Buildings*, *12*(11), 2027. https://doi.org/10.3390/buildings12112027

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

46. Hamledari, H., McCabe, B., Davari, S., & Shahi, A. (2017). Automated Schedule and Progress Updating of IFC-Based 4D BIMs. *Journal of Computing in Civil Engineering*, *31*(4), 04017012. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000660

47. Han, K., Degol, J., & Golparvar-Fard, M. (2018). Geometry- and Appearance-Based Reasoning of Construction Progress Monitoring. *Journal of Construction Engineering and Management*, *144*(2), 04017110. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001428

48. Han, K. K., Cline, D., & Golparvar-Fard, M. (2015). Formalized knowledge of construction sequencing for visual monitoring of work-in-progress via incomplete point clouds and low-LoD 4D BIMs. *Advanced Engineering Informatics*, *29*(4), 889–901. https://doi.org/10.1016/j.aei.2015.10.006

49. Han, K. K., & Golparvar-Fard, M. (2014). Automated Monitoring of Operation-level Construction Progress Using 4D BIM and Daily Site Photologs. *Construction Research Congress 2014*, *53*, 1033–1042. Reston, VA: American Society of Civil Engineers. https://doi.org/10.1061/9780784413517.106

50. Han, K. K., & Golparvar-Fard, M. (2015). Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs. *Automation in Construction*, *53*, 44–57. https://doi.org/10.1016/j.autcon.2015.02.007

51. Hannan Qureshi, A., Alaloul, W. S., Wing, W. K., Saad, S., Ammad, S., & Musarat, M. A. (2022). Factors impacting the implementation process of automated construction progress monitoring. *Ain Shams Engineering Journal*, *13*(6), 101808. https://doi.org/10.1016/j.asej.2022.101808

52. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Retrieved from http://www.springer.com/series/692

53. He, C., Zeng, H., Huang, J., Hua, X., & Zhang, L. (2020). Structure Aware Single-Stage 3D Object Detection From Point Cloud. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, *1*, 11870–11879. IEEE. https://doi.org/10.1109/CVPR42600.2020.01189

54. Hübner, P., Weinmann, M., & Wursthorn, S. (2020). Voxel-based indoor reconstruction from HoloLens triangle meshes. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *V-4–2020*(2018), 79–86. https://doi.org/10.5194/isprs-annals-V-4-2020-79-2020

55. Ibrahimkhil, M. H., Shen, X., & Barati, K. (2021). Enhanced Construction Progress Monitoring through Mobile Mapping and As-built Modeling. *Proceedings of the 38th International Symposium on Automation and Robotics in Construction (ISARC)*, (Isarc), 916–923. https://doi.org/10.22260/ISARC2021/0124

56. Iwaszczuk, D., Koppanyi, Z., Gard, N. A., Zha, B., Toth, C., & Yilmaz, A. (2018). Semantic labeling of structural elements in buildings by fusing RGB and depth images in an encoder-decoder CNN framework. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII–1*(1), 225–232. https://doi.org/10.5194/isprs-archives-XLII-1-225-2018

57. Jaiswal, A., Babu, A. R., Zadeh, M. Z., Banerjee, D., & Makedon, F. (2020). A Survey on Contrastive Self-Supervised Learning. *Technologies*, *9*(1), 2. https://doi.org/10.3390/technologies9010002

58. Jiang, L., Zhao, H., Shi, S., Liu, S., Fu, C. W., & Jia, J. (2020). PointGroup: Dual-set point grouping for 3D instance segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4866–4875. https://doi.org/10.1109/CVPR42600.2020.00492

59. Kang, T., Patil, S., Kang, K., Koo, D., & Kim, J. (2020). Rule-based scan-to-BIM mapping pipeline in the plumbing system. *Applied Sciences (Switzerland)*, *10*(21), 1–17. https://doi.org/10.3390/app10217422

60. Karsch, K., Golparvar-Fard, M., & Forsyth, D. (2014). ConstructAide:

Analyzing and visualizing construction sites through photographs and building models. *ACM Transactions on Graphics*, *33*(6), 1–11. https://doi.org/10.1145/2661229.2661256

61. Kavaliauskas, P., Fernandez, J. B., McGuinness, K., & Jurelionis, A. (2022). Automation of Construction Progress Monitoring by Integrating 3D Point Cloud Data with an IFC-Based BIM Model. *Buildings*, *12*(10). https://doi.org/10.3390/buildings12101754

62. Keyvanfar, A., Shafaghat, A., & Awanghamat, M. A. (2021). Optimization and Trajectory Analysis of Drone's Flying and Environmental Variables for 3D Modelling the Construction Progress Monitoring. *International Journal of Civil Engineering*. https://doi.org/10.1007/s40999-021-00665-1

63. Khairadeen Ali, A., Lee, O. J., Lee, D., & Park, C. (2021). Remote Indoor Construction Progress Monitoring Using Extended Reality. *Sustainability*, *13*(4), 2290. https://doi.org/10.3390/su13042290

64. Khaloo, A., & Lattanzi, D. (2017). Robust normal estimation and region growing segmentation of infrastructure 3D point cloud models. *Advanced Engineering Informatics*, *34*, 1–16. https://doi.org/10.1016/j.aei.2017.07.002

65. Kim, C. C., Kim, C. C., Son, H., & Kim, C. C. (2013). Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction*, *31*, 75–82. https://doi.org/10.1016/j.autcon.2012.11.041

66. Kim, C., Son, H., & Kim, C. (2013a). Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction*, *31*, 75–82. https://doi.org/10.1016/j.autcon.2012.11.041

67. Kim, C., Son, H., & Kim, C. (2013b). Fully automated registration of 3D data to a 3D CAD model for project progress monitoring. *Automation in Construction*, *35*, 587–594. https://doi.org/10.1016/j.autcon.2013.01.005

68. Kim, Hyunjoo, Anderson, K., Lee, S., & Hildreth, J. (2013). Generating construction schedules through automatic data extraction using open BIM (building information modeling) technology. *Automation in Construction*, *35*, 285–295. https://doi.org/10.1016/j.autcon.2013.05.020

69. Kim, Hyunsoo, & Kim, C. (2021). 3D as-built modeling from incomplete point clouds using connectivity relations. *Automation in Construction*, *130*(January), 103855. https://doi.org/10.1016/j.autcon.2021.103855

70. Kim, J., Chung, D., Kim, Y., & Kim, H. (2022). Deep learning-based 3D reconstruction of scaffolds using a robot dog. *Automation in Construction*, *134*(December 2021), 104092. https://doi.org/10.1016/j.autcon.2021.104092

71. Kim, S., Kim, S., & Lee, D.-E. (2020). Sustainable Application of Hybrid Point Cloud and BIM Method for Tracking Construction Progress. *Sustainability*,

12(10), 4106. https://doi.org/10.3390/su12104106

72. Kopsida, M., Brilakis, I., & Vela, P. A. (2015). A review of automated construction progress monitoring and inspection methods. *Proceedings of the 32nd CIB W78 Conference on Construction IT*. https://doi.org/https://doi.org/w78-2015-paper-044.

73. Kopsida, Marianna, & Brilakis, I. (2020). Real-Time Volume-to-Plane Comparison for Mixed Reality–Based Progress Monitoring. *Journal of Computing in Civil Engineering*, *34*(4), 04020016. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000896

74. Kopsida, Marianna, Brilakis, I., & Vela, P. (2015). A Review of Automated Construction Progress and Inspection Methods. *Proceedings of the 32nd CIB W78 Conference on Construction IT*, (January), 421–431.

75. Lee, J., Son, H., Kim, C., & Kim, C. (2013). Skeleton-based 3D reconstruction of as-built pipelines from laser-scan data. *Automation in Construction*, *35*, 199–207. https://doi.org/10.1016/j.autcon.2013.05.009

76. Lei, L., Zhou, Y., Luo, H., & Love, P. E. D. (2019). A CNN-based 3D patch registration approach for integrating sequential models in support of progress monitoring. *Advanced Engineering Informatics*, *41*(October 2018), 100923. https://doi.org/10.1016/j.aei.2019.100923

77. Li, Y., Li, W., Tang, S., Darwish, W., Hu, Y., & Chen, W. (2020). Automatic Indoor as-Built Building Information Models Generation by Using Low-Cost RGB-D Sensors. *Sensors*, *20*(1), 293. https://doi.org/10.3390/s20010293

78. Liu, F., Li, S., Zhang, L., & Zhou, C. (2017). 3DCNN-DQN-RNN : A Deep Reinforcement Learning Framework for Semantic. *IEE International Conference on Computer Vision(ICCV)2017*, (July), 5678–5687.

79. Liu, X., Zhang, Y., Ling, X., Wan, Y., Liu, L., & Li, Q. (2019). TopoLAP: Topology Recovery for Building Reconstruction by Deducing the Relationships between Linear and Planar Primitives. *Remote Sensing*, *11*(11), 1372. https://doi.org/10.3390/rs11111372

80. Liu, Y.-F., Cho, S., Spencer, B. F., & Fan, J.-S. (2016). Concrete Crack Assessment Using Digital Image Processing and 3D Scene Reconstruction. *Journal of Computing in Civil Engineering*, *30*(1), 04014124. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000446

81. Ma, J. W., & Leite, F. (2022). Performance boosting of conventional deep learning-based semantic segmentation leveraging unsupervised clustering. *Automation in Construction*, *136*(July 2021), 104167. https://doi.org/10.1016/j.autcon.2022.104167

82. Ma, Z., & Liu, S. (2018). A review of 3D reconstruction techniques in civil engineering and their applications. *Advanced Engineering Informatics*,

37(September 2017), 163–174. https://doi.org/10.1016/j.aei.2018.05.005

83. Maalek, R., Lichti, D. D., & Ruwanpura, J. Y. (2019). Automatic Recognition of Common Structural Elements from Point Clouds for Automated Progress Monitoring and Dimensional Quality Control in Reinforced Concrete Construction. *Remote Sensing*, *11*(9), 1102. https://doi.org/10.3390/rs11091102

84. Macher, H., Landes, T., & Grussenmeyer, P. (2017). From Point Clouds to Building Information Models: 3D Semi-Automatic Reconstruction of Indoors of Existing Buildings. *Applied Sciences*, *7*(10), 1030. https://doi.org/10.3390/app7101030

85. Mahami, H., Nasirzadeh, F., Hosseininaveh Ahmadabadian, A., & Nahavandi, S. (2019). Automated Progress Controlling and Monitoring Using Daily Site Images and Building Information Modelling. *Buildings*, *9*(3), 70. https://doi.org/10.3390/buildings9030070

86. Makantasis, K., Doulamis, A., Doulamis, N., & Ioannides, M. (2016). In the wild image retrieval and clustering for 3D cultural heritage landmarks reconstruction. *Multimedia Tools and Applications*, *75*(7), 3593–3629. https://doi.org/10.1007/s11042-014-2191-z

87. Mapillary. (2018). OpenSfM: Open source structure from motion pipeline. Retrieved July 19, 2020, from https://github.com/mapillary/OpenSfM

88. Masood, M. K., Aikala, A., Seppänen, O., & Singh, V. (2020). Multi-Building Extraction and Alignment for As-Built Point Clouds: A Case Study With Crane Cameras. *Frontiers in Built Environment*, *6*. https://doi.org/10.3389/fbuil.2020.581295

89. Mitsugami, I. (2011). Bundler: Structure from motion for unordered image collections. *The Journal of The Institute of Image Information and Television Engineers*, *65*(4), 479–482. https://doi.org/10.3169/itej.65.479

90. Mutis, I., Joshi, V. A., & Singh, A. (2021). Object Detectors for Construction Resources Using Unmanned Aerial Vehicles. *Practice Periodical on Structural Design and Construction*, *26*(4), 04021035. https://doi.org/10.1061/(ASCE)SC.1943-5576.0000598

91. Navon, R., & Sacks, R. (2007). Assessing research issues in Automated Project Performance Control (APPC). *Automation in Construction*, *16*(4), 474–484. https://doi.org/10.1016/j.autcon.2006.08.001

92. Nikoohemat, S., Diakité, A. A., Zlatanova, S., & Vosselman, G. (2020). Indoor 3D reconstruction from point clouds for optimal routing in complex buildings to support disaster management. *Automation in Construction*, *113*(May), 103109. https://doi.org/10.1016/j.autcon.2020.103109

93. Ochmann, S., Vock, R., & Klein, R. (2019). Automatic reconstruction of fully

volumetric 3D building models from oriented point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, *151*(October), 251–262. https://doi.org/10.1016/j.isprsjprs.2019.03.017

94. Omair, B., & Alturki, A. (2020). An Improved Method for Taxonomy Development in Information Systems. *International Journal of Advanced Computer Science and Applications*, *11*(4), 535–540. https://doi.org/10.14569/IJACSA.2020.0110470

95. Omar, H., Mahdjoubi, L., & Kheder, G. (2018). Towards an automated photogrammetry-based approach for monitoring and controlling construction site activities. *Computers in Industry*, *98*, 172–182. https://doi.org/10.1016/j.compind.2018.03.012

96. Omar, T., & Nehdi, M. L. (2016a). Data acquisition technologies for construction progress tracking. *Automation in Construction*, *70*, 143–155. https://doi.org/10.1016/j.autcon.2016.06.016

97. Omar, T., & Nehdi, M. L. (2016b). Data acquisition technologies for construction progress tracking. *Automation in Construction*, *70*, 143–155. https://doi.org/10.1016/j.autcon.2016.06.016

98. P. R., S., Raphael, B., & Vaidyanathan, K. (2019). Construction Progress Monitoring Using BIM and QR Code. *CIB W78 – Information Technology for Construction*, 455–466.

99. Paneru, S., & Jeelani, I. (2021a). Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction*, *132*, 103940. https://doi.org/10.1016/j.autcon.2021.103940

100. Paneru, S., & Jeelani, I. (2021b). Computer vision applications in construction: Current state, opportunities & challenges. *Automation in Construction*, *132*, 103940. https://doi.org/10.1016/j.autcon.2021.103940

101. Park, J., & Cho, Y. K. (2022). Point Cloud Information Modeling: Deep Learning–Based Automated Information Modeling Framework for Point Cloud Data. *Journal of Construction Engineering and Management*, *148*(2), 1–14. https://doi.org/10.1061/(asce)co.1943-7862.0002227

102. Patel, T., Guo, B. H. W., & Zou, Y. (2021). A scientometric review of construction progress monitoring studies. *Engineering, Construction and Architectural Management*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/ECAM-10-2020-0799

103. Pătrăucean, V., Armeni, I., Nahangi, M., Yeung, J., Brilakis, I., & Haas, C. (2015). State of research in automatic as-built modelling. *Advanced Engineering Informatics*, *29*(2), 162–171. https://doi.org/10.1016/j.aei.2015.01.001

104. Perez-Perez, Y., Golparvar-Fard, M., & El-Rayes, K. (2021).

Scan2BIM-NET: Deep Learning Method for Segmentation of Point Clouds for Scan-to-BIM. *Journal of Construction Engineering and Management*, *147*(9), 04021107. https://doi.org/10.1061/(asce)co.1943-7862.0002132

105. Pour Rahimian, F., Seyedzadeh, S., Oliver, S., Rodriguez, S., & Dawood, N. (2020). On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Automation in Construction*, *110*(October 2019), 103012. https://doi.org/10.1016/j.autcon.2019.103012

106. Poux, F., Mattes, C., Selman, Z., & Kobbelt, L. (2022). Automatic region-growing system for the segmentation of large point clouds. *Automation in Construction*, *138*(March), 104250. https://doi.org/10.1016/j.autcon.2022.104250

107. Previtali, M., Díaz-Vilariño, L., & Scaioni, M. (2018). Towards automatic reconstruction of indoor scenes from incomplete point clouds: Door and window detection and regularization. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *42*(4), 573–578. https://doi.org/10.5194/isprs-archives-XLII-4-507-2018

108. Pu, S., & Vosselman, G. (2009). Knowledge based reconstruction of building models from terrestrial laser scanning data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(6), 575–584. https://doi.org/10.1016/j.isprsjprs.2009.04.001

109. Pučko, Z., Šuman, N., & Rebolj, D. (2018). Automated continuous construction progress monitoring using multiple workplace real time 3D scans. *Advanced Engineering Informatics*, *38*(April), 27–40. https://doi.org/10.1016/j.aei.2018.06.001

110. Puri, N., & Turkan, Y. (2020). Bridge construction progress monitoring using lidar and 4D design models. *Automation in Construction*, *109*. https://doi.org/10.1016/j.autcon.2019.102961

111. Pushkar, A. (2018). *Development of framework for automated progress monitoring of construction project*.

112. Pushkar, A., Senthilvel, M., & Varghese, K. (2018). Automated Progress Monitoring of Masonry Activity Using Photogrammetric Point Cloud. *35th International Symposium on Automation and Robotics in Construction (ISARC2018)*, (Isarc). https://doi.org/10.22260/ISARC2018/0125

113. Quintana, B., Prieto, S. A., Adán, A., & Bosché, F. (2018). Door detection in 3D coloured point clouds of indoor environments. *Automation in Construction*, *85*, 146–166. https://doi.org/10.1016/j.autcon.2017.10.016

114. Rashidi, A., Brilakis, I., & Vela, P. (2015). Generating Absolute-Scale

Point Cloud Data of Built Infrastructure Scenes Using a Monocular Camera Setting. *Journal of Computing in Civil Engineering*, *29*(6), 04014089. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000414

115. Rausch, C., & Haas, C. (2021). Automated shape and pose updating of building information model elements from 3D point clouds. *Automation in Construction*, *124*(August 2020), 103561. https://doi.org/10.1016/j.autcon.2021.103561

116. Rebolj, D., Pučko, Z., Babič, N. Č., Bizjak, M., & Mongus, D. (2017). Point cloud quality requirements for Scan-vs-BIM based automated construction progress monitoring. *Automation in Construction*, *84*(August), 323–334. https://doi.org/10.1016/j.autcon.2017.09.021

117. Reja, V. K., & Varghese, K. (2019). Impact of 5G Technology on IoT Applications in Construction Project Management. *Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC 2019)*, (July), 209–217. Banff, Canada. https://doi.org/10.22260/ISARC2019/0029

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

119. Sbiti, M., Beddiar, K., Beladjine, D., Perrault, R., & Mazari, B. (2021, May 1). Toward BIM and LPS data integration for lean site project management: A state-of-the-art review and recommendations. *Buildings*, Vol. 11. MDPI AG. https://doi.org/10.3390/buildings11050196

120. Son, H., Kim, C., & Kim, C. (2015). Fully Automated As-Built 3D Pipeline Extraction Method from Laser-Scanned Data Based on Curvature Computation. *Journal of Computing in Civil Engineering*, *29*(4), 1–9. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000401

121. Son, H., Kim, C., & Kwon Cho, Y. (2017). Automated Schedule Updates Using As-Built Data and a 4D Building Information Model. *Journal of Management in Engineering*, *33*(4), 04017012. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000528

122. Sun, C.-Z., Zhang, B., Wang, J.-K., & Zhang, C.-S. (2021). A Review of Visual SLAM Based on Unmanned Systems. *2021 2nd International Conference on Artificial Intelligence and Education (ICAIE)*, 226–234. IEEE. https://doi.org/10.1109/ICAIE53562.2021.00055

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

124. Tran, H., Khoshelham, K., Kealy, A., & Díaz-Vilariño, L. (2019). Shape

Grammar Approach to 3D Modeling of Indoor Environments Using Point Clouds. *Journal of Computing in Civil Engineering*, *33*(1), 04018055. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000800

125. Tran, Ha, & Khoshelham, K. (2020). Procedural Reconstruction of 3D Indoor Models from Lidar Data Using Reversible Jump Markov Chain Monte Carlo. *Remote Sensing*, *12*(5), 838. https://doi.org/10.3390/rs12050838

126. Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W. (2000). Bundle Adjustment — A Modern Synthesis. In *International Workshop on Vision Algorithms* (pp. 298–372). Greece. https://doi.org/10.1007/3-540-44480-7_21

127. Unearth. (2023). *Construction Progress Monitoring and Reporting 101*. 1–22.

128. Vincke, S., de Lima Hernandez, R., Bassier, M., & Vergauwen, M. (2019). Immersive visualisation of construction site point cloud data, meshes and BIM models in a VR environment using a gaming engine. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-5/W2*(5/W2), 77–83. https://doi.org/10.5194/isprs-archives-XLII-5-W2-77-2019

129. Wang, B., Yin, C., Luo, H., Cheng, J. C. P., & Wang, Q. (2021). Fully automated generation of parametric BIM for MEP scenes based on terrestrial laser scanning data. *Automation in Construction*, *125*(February), 103615. https://doi.org/10.1016/j.autcon.2021.103615

130. Wang, C., Cho, Y. K., & Kim, C. (2015). Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction*, *56*, 1–13. https://doi.org/10.1016/j.autcon.2015.04.001

131. Wang, Q., Tan, Y., & Mei, Z. (2020). Computational Methods of Acquisition and Processing of 3D Point Cloud Data for Construction Applications. *Archives of Computational Methods in Engineering*, *27*(2), 479–499. https://doi.org/10.1007/s11831-019-09320-4

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

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

134. Wu, C. (2011). VisualSFM : A Visual Structure from Motion System. Retrieved July 19, 2020, from http://ccwu.me/vsfm/

135. Xiong, X., Adan, A., Akinci, B., & Huber, D. (2013). Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in Construction*, *31*, 325–337. https://doi.org/10.1016/j.autcon.2012.10.006

136. Xu, L., Oja, E., & Kultanen, P. (1990). A new curve detection method: Randomized Hough transform (RHT). *Pattern Recognition Letters*, *11*(5), 331–338. https://doi.org/10.1016/0167-8655(90)90042-Z

137. Xu, Y., Shen, X., & Lim, S. (2021). CorDet: Corner-Aware 3D Object Detection Networks for Automated Scan-to-BIM. *Journal of Computing in Civil Engineering*, *35*(3), 04021002. https://doi.org/10.1061/(asce)cp.1943-5487.0000962

138. Xu, Y., Shen, X., Lim, S., & Li, X. (2021). Three-Dimensional Object Detection with Deep Neural Networks for Automatic As-Built Reconstruction. *Journal of Construction Engineering and Management*, *147*(9), 04021098. https://doi.org/10.1061/(asce)co.1943-7862.0002003

139. Xue, J., Hou, X., & Zeng, Y. (2021). Review of Image-Based 3D Reconstruction of Building for Automated Construction Progress Monitoring. *Applied Sciences*, *11*(17), 7840. https://doi.org/10.3390/app11177840

140. Yang, L., Cheng, J. C. P., & Wang, Q. (2020). Semi-automated generation of parametric BIM for steel structures based on terrestrial laser scanning data. *Automation in Construction*, *112*(January), 103037. https://doi.org/10.1016/j.autcon.2019.103037

141. Zeng, S., Chen, J., & Cho, Y. K. (2020). User exemplar-based building element retrieval from raw point clouds using deep point-level features. *Automation in Construction*, *114*(September 2019), 103159. https://doi.org/10.1016/j.autcon.2020.103159

142. Zhirong Wu, Song, S., Khosla, A., Fisher Yu, Linguang Zhang, Xiaoou Tang, & Xiao, J. (2015). 3D ShapeNets: A deep representation for volumetric shapes. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, *07-12-June*, 1912–1920. IEEE. https://doi.org/10.1109/CVPR.2015.7298801

143. Zhu, X. (2005). *Semi-Supervised Learning Literature Survey*.

144. Zollmann, S., Hoppe, C., Kluckner, S., Poglitsch, C., Bischof, H., & Reitmayr, G. (2014). Augmented Reality for Construction Site Monitoring and Documentation. *Proceedings of the IEEE*, *102*(2), 137–154. https://doi.org/10.1109/JPROC.2013.2294314

CURRICULUM VITAE

1.	**NAME**	:	Varun Kumar Reja
2.	**DATE OF BIRTH**	:	21 October 1991
3.	**EDUCATIONAL QUALIFICATIONS**		
	Bachelor of Technology		**2010 – 2014**
	Institution	:	National Institution of Technology, Calicut
	Specialization	:	Civil Engineering
	Master of Technology		**2016 – 2018**
	Institution	:	Indian Institution of Technology, Madras
	Specialization	:	Civil Engineering – Construction Technology and Management
	Doctor of Philosophy		**2018 – 2023**
	Institution	:	Indian Institution of Technology, Madras
	Specialization	:	Civil Engineering – Construction Technology and Management
	Registration Date	:	16 July 2018
4.	**WORK EXPERIENCE**		**2014-2016**
	Organisation	:	L&T Construction
5	**FELLOWSHIPS / AWARDS**		

Build India Scholarship (BIS) – IIT Madras and L&T Construction

Prime Minister's Research Fellowship (PMRF) – Government of India

Collaborative Degree UTS President's Scholarship (COLUTSP) - UTS

Collaborative International Research Scholarship (COLIRS) - UTS

DOCTORAL COMMITTEE

CHAIRPERSON:	Dr. R. G. Robinson Professor & Head of the Department Department of Civil Engineering, IIT Madras
GUIDE(S):	Dr. Koshy Varghese Professor Department of Civil Engineering, IIT Madras
	Dr. Quang Ha Associate Professor School of Electrical and Data Engineering, UTS
MEMBERS:	Dr. Benny Raphael Professor Department of Civil Engineering, IIT Madras
	Dr. Ashwin Mahalingam Professor Department of Civil Engineering, IIT Madras
	Raghunathan Rengasamy Professor Department of Chemical Engineering, IIT Madras