

# **Improving automation in computer vision based indoor construction progress monitoring: A deep learning approach**

**Biyanka Jayangani Ekanayake**

Thesis submitted in fulfilment of the requirements for the degree of

**Doctor of Philosophy**

under the supervision of

Associate Professor Johnny Wong

Associate Professor Alireza Fini

Associate Professor Peter Smith

University of Technology Sydney

Faculty of Design Architecture and Building

December 2022

## **Certificate of original authorship**

I, Biyanka Jayangani Ekanayake, declare that this thesis is submitted in fulfilment of the requirements for the award of the degree of Doctor of Philosophy, in the School of Built Environment of the Faculty of Design, Architecture and Building 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.

This document has not been submitted for qualifications at any other academic institution.

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

Production Note:

**Signature:** Signature removed prior to publication.

**Date:** 29 December 2022

## **Note on the thesis format**

This thesis was prepared to fulfil the criteria for Doctor of Philosophy and is in the format of a conventional thesis. It adheres to the guidelines set forth in “Graduate Research Candidature Management, Thesis Preparation and Submission Procedures 2022 - University of Technology Sydney”. This thesis used the American Psychological Association (APA) 7<sup>th</sup> edition referencing style. In Chapter 2: Literature Review, the source of the figures created by the thesis author (Biyanka Jayangani Ekanayake) is cited as “The Author”.

## **Acknowledgement**

First and foremost, I would like to express my gratitude to my amazing supervisors. My principal supervisor, Associate Professor Johnny Wong gave me freedom and valuable insights for innovation in my research with constant encouragement. I sincerely thank him for being there for me not only to provide constructive feedback for my write-ups but also to listen to my typical PhD despairs, always with a kind smile. I am grateful to my supervisor, Associate Professor Alireza Fini for giving valuable insights, which made me work harder and think more critically. He even accompanied me to a construction site to ensure that everything went smoothly for my data collection during the COVID pandemic. My supervisor, Associate Professor Peter Smith kindly guided me to see my research from different perspectives, especially when I was too invested in the algorithmic part. My heartfelt gratitude goes to him for inspiring me to resolve issues with a calm approach. I have been blessed to be guided by such kind, supportive and understanding supervisors.

I am grateful to the building contractor organisation and Benjamin Lunn for their great support during the data collection process. Without Ben's support, the access to the construction sites and setting up of cameras would have been impossible. I express my special thanks to UTS 5D BIM Lab for employing me as a research assistant and also providing additional financial support for the research equipment I needed. My gratitude is extended to School of Built Environment and Faculty of Design, Architecture and Building of UTS for providing the research space and great facilities to complete my thesis. I would also like to thank the DAB HDR Director Associate Professor Toby Slade, former HDR Director Associate Professor Michael Er, DAB HDR Support team, Georgina Donovan, and Vanessa Setiadi and the Graduate Research School for all the support during my candidature.

My heartfelt thanks goes to many wonderful people I met in Sydney and UTS. Thank you, Adi for being the friend I needed when I first came to Sydney. My dearest friends, Mia and Tazy, thank you for making me part of your families and taking me on many trips in Sydney. Big thanks to Mona for her valuable advice. Since I started my PhD journey, I have been blessed to have incredible friends like Elham, Mukesh, Megan, Bingyang, Mani, Mirka, Javid, Enya, Caroline, Fiona, Samuel, Xiaohang, Juanwen and

rest of the DAB PhD Girls et al. group. I never felt alone in my PhD journey and in Sydney because of these fellow PhD friends.

My biggest gratitude goes to my parents for the unconditional love, support, wonderful upbringing, and opportunities they have given me. I could not have done any of this without their blessing. I am forever indebted to my beloved brother Thilina for taking care of everything and everyone while I was in Sydney for 3 years working on my PhD. With great love, I want to thank my boyfriend Ashan for his patience, understanding, care and support.

## Table of contents

Certificate of original authorship .....	i
Note on the thesis format .....	ii
Acknowledgement.....	iii
Table of contents .....	v
List of tables.....	viii
List of figures .....	viii
Abbreviations .....	x
Peer reviewed publications from this research.....	xii
Abstract .....	xiii
1 Introduction.....	1
1.1 Background .....	1
1.2 Problem statement .....	4
1.3 Research questions .....	5
1.4 Research aim and objectives .....	5
1.5 Overview of the research methodology.....	6
1.6 Scope and limitations .....	6
1.7 Outline of the thesis.....	8
2 Literature review.....	10
2.1 Challenges of traditional construction progress monitoring .....	10
2.2 State-of-the-art progress monitoring approaches .....	12
2.2.1 Laser scanning based technologies .....	14
2.2.2 Radio wave based technologies .....	15
2.2.3 Geospatial based technologies .....	17
2.2.4 Computer vision based technologies.....	18
2.3 Algorithmic approaches to analyse the as-built status from image data .....	24
2.3.1 Conventional machine learning approaches.....	24
2.3.2 Deep learning approaches .....	27
2.3.3 Convolutional neural networks .....	29
2.4 Indoor construction progress monitoring .....	31
2.4.1 Advances in indoor image collection and visualisation.....	35
2.4.2 Automated activity state recognition of indoor works.....	37
2.4.3 Adoption of deep learning in indoor construction progress monitoring ...	39
2.5 Research gaps .....	42
2.5.1 Technical challenges for as-built indoor elements visual recognition .....	42

2.5.2	The deep learning model to overcome the challenges .....	48
2.5.3	Calculating work-in-progress using the deep learning model.....	50
2.6	Chapter summary .....	52
3	Research design and methods .....	54
3.1	Determining the methodological framework .....	54
3.1.1	Design science research .....	55
3.1.2	Research process mapped with the design science research framework ..	56
3.2	Data collection from the case projects .....	59
3.2.1	Case projects .....	59
3.2.2	Time-lapse cameras, related accessories, and camera settings .....	61
3.2.3	Developing the training images dataset .....	66
3.3	Research design for developing the deep learning approach .....	70
3.3.1	Mask R-CNN network architecture .....	70
3.3.2	Building the deep learning model: Transfer learning and fine-tuning .....	72
3.3.3	Training platform for the deep learning model: Colab virtual machine ...	74
3.3.4	Metrics to evaluate the performance of deep learning models .....	76
3.4	Chapter summary .....	77
4	Results and discussion: Part I .....	80
4.1	Empirical evidence from the case projects.....	80
4.1.1	Indoor objects related challenges .....	82
4.1.2	Lighting conditions related challenges.....	85
4.1.3	Camera positioning related challenges.....	88
4.2	Pre-processing techniques to reduce the impacts of the technical challenges..	92
4.2.1	Perspective transformation in image processing.....	93
4.2.2	Low illumination enhancement in image processing.....	94
4.3	Discussion .....	98
4.4	Chapter summary .....	101
5	Results and discussion: Part II.....	103
5.1	Automated visual recognition of indoor as-built elements.....	103
5.1.1	The workflow of modifying Mask R-CNN and training on Colab.....	103
5.1.2	Performance evaluation of the optimised Mask R-CNN model .....	111
5.2	Calculating the work-in-progress of indoor elements .....	117
5.3	Discussion .....	123
5.4	Chapter summary .....	126
6	Conclusions and recommendations .....	128

6.1	Achieving the research objectives and summary of the outcomes.....	128
6.1.1	Investigating the technical challenges and their impacts for indoor as-built elements visual recognition .....	129
6.1.2	Developing an optimised deep learning algorithm for automated visual recognition of indoor as-built elements .....	130
6.1.3	Automatically calculating the percentage work completion of indoor elements by post-processing the optimised deep learning algorithm .....	132
6.2	Theoretical contributions and practical implications .....	133
6.2.1	Contributions to the body of knowledge .....	133
6.2.2	Implications for the construction industry .....	134
6.3	Limitations of the current study and future research directions .....	136
	References .....	139
	Appendix 1 Research ethics application approval .....	156

## List of tables

Table 2.1: Data acquisition technologies for progress monitoring .....	23
Table 2.2: Overview of computer vision based indoor construction monitoring .....	32
Table 2.3: Technical challenges for visual recognition of as-built indoor elements .....	46
Table 3.1: Information about the case projects .....	61
Table 3.2: Time-lapse camera settings.....	65
Table 4.1: Technical challenges for as-built indoor elements visual recognition.....	99
Table 5.1: Customising certain hyperparameters.....	110

## List of figures

Figure 1.1: Mapping of research objectives, methods and corresponding chapters .....	8
Figure 2.1: Process of automated progress monitoring in construction projects .....	12
Figure 2.2: Trimble GX 3D laser scanner .....	14
Figure 2.3: 3D as-built reconstruction of a building using point cloud .....	15
Figure 2.4: RFID and barcode technologies .....	16
Figure 2.5: GPS data logger .....	18
Figure 2.6: A construction site a) 2D image; b) 3D photogrammetric model .....	20
Figure 2.7: 2D site images integrated with 4D BIM.....	21
Figure 2.8: Feature extraction: a) Original image; b) Subjected to Canny edge detection .....	25
Figure 2.9: Image segmentation: a) Original image; b) Binary image .....	26
Figure 2.10: Mechanisms: a) Traditional machine learning; b) Deep learning .....	27
Figure 2.11: Mechanism behind deep neural networks .....	29
Figure 2.12: Typical structure of a CNN and the functions of the hidden layers .....	30
Figure 2.13: Timeline of the computer vision based indoor progress monitoring studies .....	42
Figure 3.1: Process of design science research .....	57
Figure 3.2: Case projects: a) Project 1; b) Project 2; c) Project 3 .....	60
Figure 3.3: Brinno TLC200 Pro time-lapse camera.....	62
Figure 3.4: Installing cameras: a) Using a tripod; b) Taping to scaffolding .....	63
Figure 3.5: Camera accessories: a) Gorillapod; b) Weather resistant housing .....	64
Figure 3.6: Indoor location with a) Natural lighting; b) Artificial lighting .....	64
Figure 3.7: a) Image captured from Project 1; b) Image sourced from the internet .....	67
Figure 3.8: Sample of Python code for image augmentation.....	68
Figure 3.9: a) Original image; b) Image flipped vertically and contrast enhanced.....	69
Figure 3.10: Annotated image.....	70
Figure 3.11: Mask R-CNN network architecture .....	71
Figure 3.12: Illustrating transfer learning .....	73
Figure 3.13: Working mechanism of Colab.....	75
Figure 3.14: Summary of the research process .....	78
Figure 4.1: Sample of Python code for Canny edge detection.....	81
Figure 4.2: Construction equipment creating occlusions.....	82
Figure 4.3: Construction personnel and material creating occlusions .....	83
Figure 4.4: Movement of the drywall sheets.....	84

Figure 4.5: a) Original image; b) Canny edge detector extracting insulation and drywall states.....	84
Figure 4.6: Low illumination in the indoor site .....	85
Figure 4.7: Backlight caused by natural light entering the site.....	86
Figure 4.8: Extracting the region of interest affected by a bright artificial lighting .....	87
Figure 4.9: a) Original image; b) Subjected to Canny edge detection.....	87
Figure 4.10: The slightly inclined region of interest captured by the camera.....	88
Figure 4.11: Field of view is affected by camera positioning.....	89
Figure 4.12: Camera location is far from the walls.....	89
Figure 4.13: Obstruction to the field of view of the camera .....	90
Figure 4.14: Camera being accidentally turned away .....	90
Figure 4.15: a) Original image; b) Original subjected to Canny edge detection .....	91
Figure 4.16: Region of interest subjected to Canny edge detection.....	92
Figure 4.17: Python code for perspective transformation.....	93
Figure 4.18: Perspective transformation: a) Original image; b) Corrected image.....	94
Figure 4.19: Python code for low light enhancement: Part I .....	96
Figure 4.20: Python code for low light enhancement: Part II.....	97
Figure 4.21: Low illumination enhancement: a) Original image; b) Enhanced image... ..	98
Figure 5.1: The technical algorithm .....	104
Figure 5.2: Mounting Google Drive to Colab.....	105
Figure 5.3: The specifications of the Colab virtual machine .....	106
Figure 5.4: Installing Detectron2 .....	107
Figure 5.5: Importing some common libraries and PyTorch related utilities .....	107
Figure 5.6: Loading the dataset.....	108
Figure 5.7: Loading the backbone and fine-tuning the hyperparameters.....	109
Figure 5.8: Setting up the directory to save the trained weights.....	110
Figure 5.9: Training the model.....	110
Figure 5.10: Comparison of mAPs with different learning rates .....	112
Figure 5.11: Comparison of training loss with different learning rates .....	113
Figure 5.12: Comparison of validation loss with different learning rates.....	113
Figure 5.13: Training loss and validation loss of the model.....	114
Figure 5.14: Loss: a) Bounding box loss; b) Class loss; c) Mask loss; d) Total loss....	115
Figure 5.15: Inferring using the trained model .....	116
Figure 5.16: Examples of indoor as-built state recognition using Mask R-CNN .....	116
Figure 5.17: a) Image of an indoor wall; b) Segmented mask area calculations .....	118
Figure 5.18: Mounting Google Drive .....	119
Figure 5.19: Importing the libraries .....	120
Figure 5.20: Test images and output directories .....	120
Figure 5.21: Pixel masks area calculation.....	121
Figure 5.22: a) Work completion; b) As-planned wall area from CAD model .....	122
Figure 5.23: Indoor partition wall is fully installed .....	123
Figure 6.1: Contributions of the PhD study .....	136

## Abbreviations

ALS	Airborne laser scanner
AP	Average precision
AVI	Audio Video Interleave
AU\$	Australian dollars
BIM	Building information model
CAD	Computer aided design
CDF	Cumulative distribution function
CLAHE	Contrast-limited adaptive histogram equalisation
CNN	Convolutional neural network
COCO	Common objects in context
CUDA	Compute unified device architecture
cuDNN	CUDA based deep neural networks
CV	Computer vision
FC	Fully connected
FCN	Fully convolutional net
FN	False negatives
FP	False positives
FPN	Feature pyramid network
FPS	Frames per second
GIS	Geographic information system
GNSS	Global navigation satellite system
GPS	Global positioning system
GPU	Graphic processing unit
HD	High definition
HDR	High dynamic range
HOG	Histogram of oriented gradients
IMU	Initial measurement unit
IoU	Intersection over union
JPG	Joint Photographic Experts Group
LBP	Local binary patterns
LiDAR	Light detection and ranging
mAP	Mean average precision
Mask R-CNN	Mask regional convolutional neural network
MLS	Mobile laser scanner
MVS	Multi view stereo
NSW	New South Wales
OpenCV	Open source computer vision
PDA	Personal digital assistant
PPHT	Progressive probabilistic Hough Transform
PSPNet	Pyramid scene parsing network
RANSAC	Random sample consensus
R-CNN	Region-based convolutional neural networks

ReLU	Rectified linear unit
ResNet	Residual network
RFID	Radio frequency identification
RGB-D	Red, green, and blue colour channels
RPN	Region proposal network
SFM	Structure from motion
SGD	Stochastic gradient descent
SLAM	Simultaneous localisation and mapping
SSD	Single shot multi-box detector
SVM	Support vector machine
TLS	Terrestrial laser scanner
TN	True negative
TP	True positive
TPU	Tensor processing unit
UAS	Unmanned aerial system
UAV	Unmanned aerial vehicle
UGV	Unmanned ground vehicle
US\$	United States dollars
UWB	Ultra-wide band
VO	Visual odometry
VSLAM	Visual simultaneous localisation and mapping
YOLO	You Only Look Once
YOLOCT++	You Only Look at CoefficientT++

## **Peer reviewed publications from this research**

**Ekanayake, B.**, Wong, J. K. W., Fini, A. A. F., and Smith, P. (2021), Computer vision based interior construction progress monitoring: A literature review and future research directions. *Automation in Construction*, 127, 103705. (Q1 journal).

**Ekanayake, B.**, Fini, A. A. F., and Wong, J. K. W. (2021), Technical challenges for automated indoor construction progress monitoring. *44th AUBEA Conference*, Deakin University, Melbourne, Australia, 28–29 October 2021, pp 152–162.

**Ekanayake, B.**, Fini, A. A. F., Wong, J. K. W., and Smith, P. (2022), A deep learning based approach to facilitate the as-built state recognition of indoor construction. *Construction Innovation*. (Q1 journal).

## **Abstract**

Inefficient and inaccurate monitoring of work-in-progress significantly contributes to schedule and budget overruns in construction projects. The adoption of computer vision technologies has expedited automation, accuracy, and efficiency in the process of monitoring construction progress by overcoming the challenges of laborious, costly and error prone manual methods. The majority of computer vision based progress monitoring studies have focused on the exterior construction environment and few studies have been conducted for indoor sites. The literature review found that robust object recognition from indoor site images is inhibited by technical challenges related to indoor objects, lighting conditions and camera positioning. Compared to conventional machine learning algorithms, deep learning models are robust in automated visual recognition of as-built elements and calculation of work completion. Therefore, this study aimed to develop a deep learning based approach for improving automation in computer vision based indoor construction progress monitoring. The design science research framework was employed as the underpinning methodology since this study focused on addressing a practical problem by developing a technology based artefact.

Three indoor sites were used as case projects and time-lapse cameras were used to capture progress images. The framing, insulation and drywall installation of indoor walls was selected as the progress monitoring scenario. The images captured were first used to demonstrate the technical challenges and their impacts on feature extraction. Using transfer learning and hyperparameter fine-tuning, a Mask Recurrent Convolutional Neural Network (Mask R-CNN) model was optimised. Google Colab was used to train the deep learning model on the images captured to reduce computational complexities. It was found that visual recognition of indoor as-built elements of the Mask R-CNN model was improved by optimising the learning rate with Adam optimiser and cosine annealing. Training efficiency improvements were made through the batch size, network resolution and number of iterations. The model reported a mean average precision of 88.02%, and training and validation losses of 0.19 and 0.20 respectively. This model was post-processed to automatically calculate the work-in-progress of indoor elements by extracting the areas of the segmented pixel masks of the as-built states detected.

The key contribution of this research is the development of an optimised and post-processed model of Mask R-CNN, which is capable of automated as-built visual

recognition and work-in-progress measurement of indoor construction elements. This study also contributes to offering a simplified, efficient, and shareable training workflow on Colab that can be replicated in similar studies. The novelty of this research is in the optimisation process of the Mask R-CNN model and the Colab based training workflow. Upon providing training images, the Mask R-CNN model and the training workflow can be extended to other as-built scenes to improve automation in the progress monitoring process. In order to achieve this, the study provides an easy-to-follow technical flow diagram and further explanations on the workflow of the deep learning model, which can be understood by construction professionals with a non-programming background.