

## **Transfer Learning Enabled Process Recognition for Module Installation of High-rise Modular Buildings**

Zhiqian Zhang<sup>1</sup>, Wei Pan<sup>2</sup>, and Zhen-Jie Zheng<sup>3\*</sup>

<sup>1</sup> *PhD Researcher, Department of Civil Engineering, University of Hong Kong*

<sup>2</sup> *Associate Professor, Department of Civil Engineering, University of Hong Kong*

<sup>3</sup> *PhD Researcher, Department of Civil Engineering, University of Hong Kong*

*\*Corresponding author's e-mail: zhengzj@hku.hk*

### **ABSTRACT**

High-rise modular buildings (HMB), based on the advanced approach of modular construction, have gained momentum in practice due to their offered benefits in accelerated construction, improved quality, reduced health and safety risks, and enhanced productivity. Modular construction with standard design of modules and repetitive processes of module installation is in favor of the development of construction automation.

As module installation is one of the critical activities in the delivery of HMBs, it is important to recognize the module installation process automatically so as to facilitate automation in modular construction. However, there is no detailed phase-division of module installation process. Also, little research has been carried out on intelligent process recognition for module installation due to the limited amount of images of real-life projects.

To fill in the knowledge gaps, this paper aims to build a transfer learning enabled process recognition model using convolutional neural network (CNN) for module installation of HMBs. The study first divided the module installation process into three stages: hooking, lifting and positioning, with a comprehensive literature review. Then the recognition model for module installation process was created and trained with the adoption of CNN-based transfer learning, and verified with images taken from real-life projects.

The results show that the three stages of module installation process are effectively recognized with the proposed model. The transfer learning enabled image recognition model for module installation process accelerates automation in the construction of HMBs for enhanced productivity and accuracy.

### **KEYWORDS**

High-rise modular buildings; Module installation; Process recognition; Transfer learning; Convolutional neural network

## INTRODUCTION

Modular construction (MC) represents the most advanced level of offsite construction with reference to Gibb's categorization of various offsite technologies (Gibb 1999), which is a novel approach by using the Design for Manufacture and Assembly (DfMA) theory and advanced manufacturing and logistics technologies (Pan & Hon, 2018). The modular units are typically fully factory finished and thus advance traditional prefabricated components and two-dimensional panel systems (Pan 2006). The intrinsic features of MC determine its significant advantages over conventional site-built construction methods, with demonstrated benefits in accelerated onsite construction, improved health and safety, enhanced productivity and quality, and minimized construction waste (Lawson et al. 2014; Pan et al. 2007).

High-rise modular buildings (HMB) refer to high-rise buildings which adopt MC technology. Internationally, MC has been adopted in many jurisdictions including UK, US, Australia and Singapore, etc. There have also emerged some remarkable projects of HMB, e.g. the 32-storey steel framed modular building in the Pacific Park of New York City, US, and the 40-storey residential building with concrete modules named Clement Canopy in Singapore.

Yang et al. (2018) pointed out that HMB with standard design of modules and repetitive processes of module installation is in favor of the development of construction automation. As module installation is one of the critical construction activities in the delivery of HMBs (Lawson et al. 2014), it is important to recognize the module installation process automatically so as to facilitate automation in the construction of HMB. Machine learning seems to be an efficient approach for the automated recognition of module installation via the trained models for image recognition. However, there is neither standard database nor enough information of the real-life HMB projects towards the module installation process, which severely hinders the model training to facilitate automated recognition.

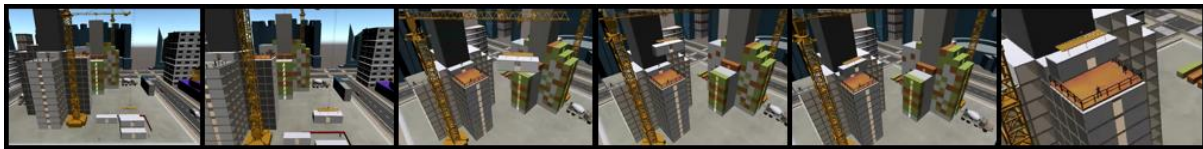
To overcome the problems of lacking enough data for the development of the recognition model, transfer learning was first introduced into this study. Transfer learning offers the ability of a system to recognize and apply knowledge learned in previous tasks to novel tasks, with the aim to extract the knowledge from one or more source tasks and applied the knowledge to a target task (Pan et al. 2010). Many studies have been done to address the effectiveness of transfer learning enabled image recognition when lacking enough data, such as deep learning based transfer learning to recognize emotion on small datasets (Ng et al. 2015), and transfer learning assisted prediction of wind speed in a newly-built farm (Hu et al. 2015).

Nevertheless, few studies have been done on the application of transfer learning in the field of construction, and little research has been carried out on transfer learning enabled image recognition for module installation. This paper aims to build a transfer learning enabled image recognition model using convolutional neural network (CNN) for automated recognition of module installation process of HMB. To achieve the aim three steps of research were followed: 1) conduct the analysis of the module installation process; 2) develop the transfer learning enabled recognition model; and 3) conduct experiment to train and test the proposed model with the collected images.

## METHODS

### Module installation process analysis

Compared with the adoption of conventional site-built construction techniques, high-rise building with MC is unique in its ability on simultaneous construction and manufacturing as most of the onsite works are transformed into the factory while with little works remaining on site. Site activities of HMB normally include foundation and lift core construction, module installation and connection, and the remaining fit-out works (Lawson et al. 2014). Among the site works of HMB, module installation (Figure 1) is one of the most critical activities to ensure the successful project delivery (Gibb 1999). Module installation usually occupies heavy-duty tower cranes which is in the critical path of the construction programme (Olearczyk et al. 2014), and it always concerns much of the construction safety (Fard et al. 2017).

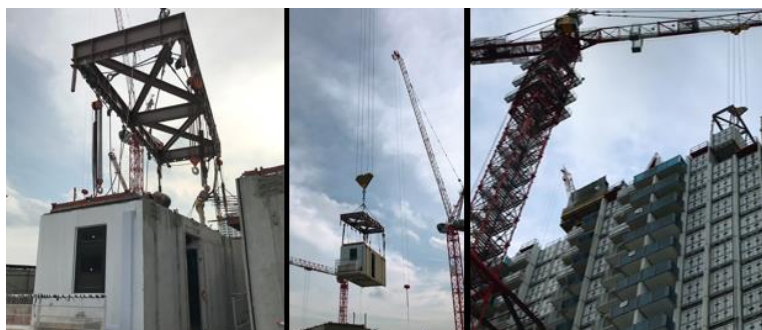


**Figure 1.** Simulation of module installation of HMB

The process of module installation normally consists of hoisting/lifting, positioning (setting in place and levelling), connecting and waterproofing (connecting to adjacent modules structurally and to make a watertight seal) (Gibb 1999; Staib et al. 2008). In this paper, module installation is divided into three stages mainly, including hooking, lifting, and positioning based on the literature review and site observation.

- *Hooking*: the process of which the module is hooked from the truck or from the storage yard by the crane via a lifting frame;
- *Lifting*: the process of which the module is elevated or rotated after hooking up before reaching the designated position;
- *Positioning*: the process of which the orientation of the module is slightly adjusted in order to be set in place within the designed tolerance.

To form the database in order to develop the recognition model, images of the module installation of HMB were collected via the internet searching on Google Images, site records, and video clips from the modular suppliers. Sample of the collected images are shown in Figure 2, which are categorized with the three installation stages.



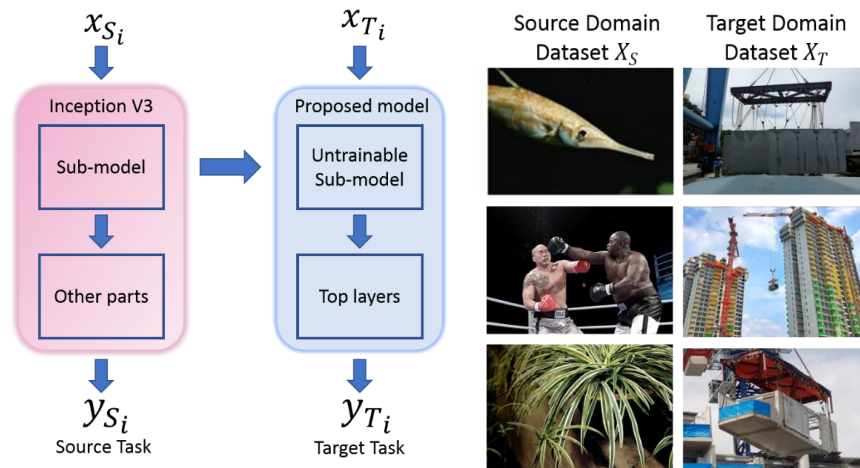
**Figure 2.** From left to right: hooking, lifting, and positioning (Courtesy of Teambuild)

## Model development

### Transfer learning

The framework of transfer learning used in this paper is showed in Figure 3. Let  $D_s = \{(x_{S_1}, y_{S_1}), (x_{S_2}, y_{S_2}), \dots, (x_{S_n}, y_{S_n})\}$  and  $D_T = \{(x_{T_1}, y_{T_1}), (x_{T_2}, y_{T_2}), \dots, (x_{T_i}, y_{T_i})\}$  denote the source and target domain data respectively,  $x_{S_i} \in X_S$  and  $x_{T_i} \in X_T$  denote the data instance of the source and the target domain respectively,  $y_{S_i} \in Y_S$  and  $y_{T_i} \in Y_T$  denote the corresponding class label of the source and the target domain respectively.

Though transfer learning has shown its powerful adaptability in many task, it is not easy to apply in the context of module installation due to the great difference between the source domain dataset  $D_S$  (the ImageNet Dataset) and the target domain dataset  $D_T$  (the dataset of the images of module installation), as showed in Figure 3 (Right). To address such challenges, two strategies are adopted in this paper (Figure 3. Left). First, the Untrainable Sub-model extracted from Inception-v3 (Szegedy et al. 2016), which is trained based on ImageNet Dataset, is used for features extraction. This is motivated by the effectiveness of Inception-v3 in transfer learning tasks (Mednikov et al. 2018, Mittal et al. 2018). Second, features extracted from the Untrainable Sub-model are directly input into top layers of the network architecture to avoid overfitting as the target domain dataset  $D_T$  is small.



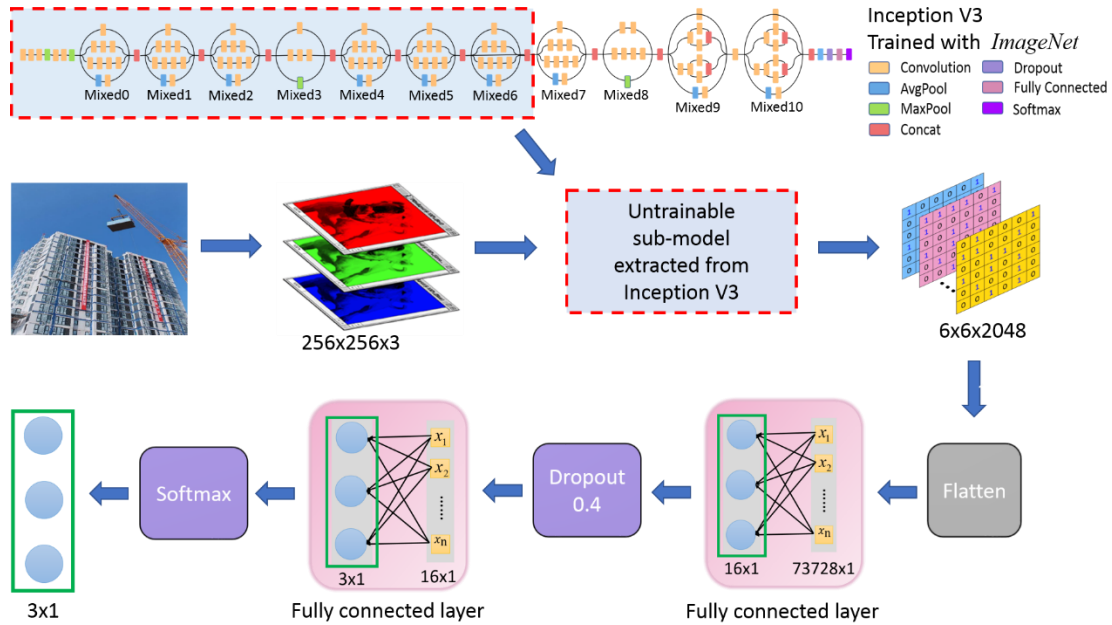
**Figure 3.** Left: Model adopted for transfer learning; Right: Data instances of the source ( $X_S$ ) and the target ( $X_T$ ) task

### Convolutional neural network design

The Inception-v3 model is used in the Source Task, the structure of which is illustrated by Szegedy (2016). The proposed network takes a  $256 \times 256$ -pixel RGB image as the input and produces a distribution over the ImageNet object classes, and it consists of 11 mixed convolutional layers (Mixed0, Mixed1, ..., and Mixed10), followed by top layers as showed in Figure 4.

The Untrainable Sub-model in Target Task is extracted from the Inception-v3, specifically from its input layer to the mixed9 layer. The reason is that the layer of mixed6 has one of the lowest

dimensional outputs as  $6 \times 6 \times 2048$  features, which can reduce the number of parameters to avoid overfitting. The features extracted from the Untrainable Sub-model are flattened into  $73728 \times 1$  dimensions, followed by a  $16 \times 1$  fully connected layer and ReLU activations. The dropout layer with a rate of 0.4 is added to avoid overfitting, and the output layer is  $3 \times 1$  fully connected layer with SoftMax activation. The proposed network is shown in Figure 4.



**Figure 4.** The proposed network architecture

## Experiment

### Dataset

In line with the procedures of the module installation analyzed in the former section, the images of module installation were classified into three categories, which are hooking, lifting, and positioning. 249 images of the module installation in total had been collected as the Target Task Dataset ( $X_T$ ), which were randomly split into 70% (174 images) training samples and 30% (75 images) test samples. Image distribution of the three categories is shown in Figure 5.



**Figure 5.** Image distribution

### Network Training

All experiments were conducted in the Google cloud platform named Google Collaboratory, using Keras library with TensorFlow backend. First, the Inception-v3 model in Keras library (Reference) was adopted as the Sourcing Task. Then, for the knowledge transferring part, the Sub-model illustrated in Figure 3 (Left), was extracted from the trained Inception-v3 model, and its parameters were set to be untrainable when fit into in the proposed recognition model. Finally, the category cross-entropy was selected as the loss function for the Target Task, so that the objective function  $L(\theta)$  became a convex function to converge (Boer 2005). The objective function  $L(\theta)$  is represented as:

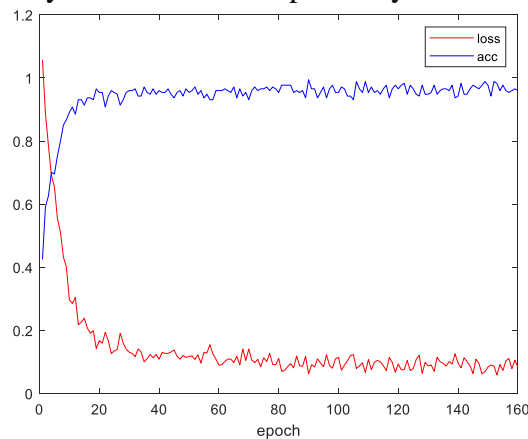
$$L(\theta) = \arg \min_{\theta} \left\{ - \sum_{j=1}^m y'_{T_j}(\theta) y_{T_j} \log(y_{T_j}) \right\} \quad j = 1, 2, \dots, m \quad (1),$$

where  $\theta$  is the parameter vector,  $y'_{T_j}$  and  $y_{T_j}$  are the predicted label and ground true label in the target task respectively, and  $m$  is the total number of samples in the target task.

Besides, Stochastic Gradient Descent (SGD) was applied in the training process, as it is one of the most effective optimization algorithms in convex optimization (Busto 2015). The SGD was set up with a learning rate of 0.01 for the first 32 epochs and 0.001 for the remaining 128 epochs, and a batch size of 16.

## RESULTS AND DISCUSSION

It is shown in Figure 6 that both the loss and the accuracy of the training samples get converged after around 100 epochs. The overall test accuracy reached a quite high level considering such a small dataset, which is up to 0.8533 as is shown in Table 1. Also, the test accuracy is smaller than the training accuracy (0.9713), which indicates that the proposed recognition model does not over fit the dataset. However, the test accuracies of the hooking and positioning are lower than the overall result, with only 0.78 and 0.72 respectively.



**Figure 6.** Results of the network training – loss and accuracy

**Table 1.** Results of the network training – training and test accuracies

	Hooking	Lifting	Positioning	Overall
Train. Acc.	1	1	0.9242	0.9713
Test Acc.	0.7778	0.9318	0.7273	0.8533

Quite a high accuracy (85.33%) is achievable with the proposed recognition model considering such a small scale of target domain Dataset, which is mainly due to the reasonable segmentation of modular installation and the application of transfer learning in the model training process. The three stages of module installation process (hooking, lifting and positioning) well reflect the features of on-site activities of HMB, which facilitates to the automated recognition with the proposed model. In addition, the adoption of transfer learning offers a chance for features extraction and overfitting avoidance with the mature Sub-model, which in turn greatly improves the performance of the proposed model. Due to the small scare of the hooking and positioning images available, with only 88 for the former and 33 for the latter (Figure 5), it is explainable that the test accuracies of these two categories are less than 80%. Since only a small batch of images was used in this study, it is believed that the model could be further improved when more data are provided.

## CONCLUSIONS

This paper has proposed a transfer learning enabled CNN model to recognize the module installation process of HMB. First, the process of module installation is divided into three stages: hooking, lifting and positioning. Second, the proposed recognition model is developed based on transfer learning. Last, the results of the experiment show that the proposed recognition model can well recognize the stages of module installation with an accuracy of 85.33%, which could be increased with more images available.

The developed model is believed to be helpful for the planning and control of HMB construction. Except for the automated recognition of the module installation process, the model can be further used to calculate the cycle time of the installation with relevant monitoring devices. In summary, the proposed recognition model accelerates automation in the construction of HMBs for enhanced productivity and accuracy.

## REFERENCES

- Boer, P. D. E. (2005). "A Tutorial on the Cross-Entropy Method." 19–67.
- Busto, P. P. (2015). "Adaptation of Synthetic Data for Coarse-to-Fine Viewpoint Refinement." *BMVC*, 2.
- Fard, M. M., Terouhid, S. A., Kibert, C. J., and Hakim, H. (2017). "Safety concerns related to modular/prefabricated building construction." *International Journal of Injury Control and Safety Promotion*, 24(1), 10–23.
- Gibb, A. G. F. (1999). *Off-Site Fabrication: prefabrication, preassembly, and modularisation*. Whittles Publishing.
- Lawson, M., Ogden, R., and Goodier, C. (2014). *Design in Modular Construction*. CRC Press.
- Mednikov, Y., Nehemia, S., Zheng, B., Benzaquen, O., and Lederman, D. (2018). "Transfer Representation Learning using Inception - V3 for the Detection of Masses in Mammography." *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2587–2590.
- Mittal, S., Kaushik, P., Hashmi, S., and Kumar, K. (2018). "Robust Real Time Breaking of Image CAPTCHAs Using Inception-v3 Model." *2018 Eleventh International Conference on Contemporary Computing (IC3)*, IEEE, 1–5.

- Ng, H., Nguyen, V. D., Vonikakis, V., and Winkler, S. (2015).“Deep Learning for Emotion Recognition on Small Datasets Using Transfer Deep Learning for Emotion Recognition on Small Datasets Using Transfer Learning.” *In Proceedings of the 2015 ACM on international conference on multimodal interaction*, (November), 443–449.
- Olearczyk, J., Al-Hussein, M., and Bouferguène, A. (2014).“Evolution of the crane selection and on-site utilization process for modular construction multilifts.” *Automation in Construction*, Elsevier B.V., 43, 59–72.
- Pan, S. J., and Yang, Q. (2010).“A Survey on Transfer Learning.” *IEEE Transactions on knowledge and data engineering*, 22(10).
- Pan, W. (2006).“A Decision Support Tool for Optimising the Use of Offsite Technologies in Housebuilding.”
- Pan, W., Gibb, A. F., and Dainty, A. R. J. (2007).“Perspective of UK housebuilders on the use of offsite modern methods of construction.” *Construction Management and Economics*, 25(2), 183–194.
- Pan, W., and Hon, C. K. (2018).“Modular integrated construction for high-rise buildings.” *Proceedings of the Institution of Civil Engineers-Municipal Engineer*, 1–12.
- Staib, Dörrhöfer, and Rosenthal. (2008).*Modular Construction Design and Structure New Technologies*. Institut für internationale Architektur-Dokumentation GmbH & Co. KG München, Munich, Germany.
- Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. (2016).“the Impact of Residual Connections on Learning.” *In Proceedings of the IEEE conference on computer vision and pattern recognitio*, 2818–2826.
- Torre, M. L. D.La, Sause, R., Slaughter, S., and Hendricks, R. H. (1994).*Review and Analysis of Modular Construction Practices*. Lehigh University.
- Yang, Y., Pan, M., and Pan, W. (2018).“Co-evolution of innovative building technologies through interaction : The case of modular integrated construction and robotics.” *CIOB (HK) Outstanding Paper Awards*, (1).