

# Training a Visual Scene Understanding Model Only with Synthetic Construction Images

Jinwoo Kim,<sup>1</sup> Daeho Kim,<sup>2</sup> Julianne Shah,<sup>3</sup> and SangHyun Lee, M.ASCE<sup>4</sup>

<sup>1</sup>Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109; e-mail: [jinwooki@umich.edu](mailto:jinwooki@umich.edu)

<sup>2</sup>Department of Civil and Environmental Engineering, University of Toronto, Toronto, ON M5S 1A4; e-mail: [daeho@umich.edu](mailto:daeho@umich.edu)

<sup>3</sup>Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109; e-mail: [jeshah@umich.edu](mailto:jeshah@umich.edu)

<sup>4</sup>Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI 48109; e-mail: [shdpm@umich.edu](mailto:shdpm@umich.edu)

## ABSTRACT

While the use of deep neural networks (DNN) for computer vision is increasing in the construction domain, the shortage of training datasets prevents such models from achieving their maximum potential. To address this issue, we investigate the potential of using synthetic data for vision model development. Specifically, we synthesize construction images and train a DNN model only with the synthetic data. We then evaluate the performance of the synthetic data-trained model on a worker detection task, and the results demonstrate the great potential of synthetic images: 97.3% of mean Average Precision. Given the benefits of synthetic data—it is possible to automatically create an unlimited number of images without manual labeling—this finding is promising. Moreover, this approach can be readily applied to other computer vision tasks, without requiring the manual labeling. This finding will enable the creation of more accurate and scalable DNN models for construction applications.