Synthetic image data generation for semantic understanding in everchanging scenes using BIM and Unreal Engine

Yujie Wei[1] and Burcu Akinci[2]

[1]Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA; e-mail: yujiew@andrew.cmu.edu
[2] Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA; e-mail: bakinci@cmu.edu

ABSTRACT

Scene understanding, such as object recognition and semantic segmentation on images, plays an important role in many existing information management workflows such as progress monitoring and facility management. Current studies on Architecture/Engineering/Construction (AEC) scene understanding often focus on using visual data captured for a particular building at project closeout, which has two primary limitations: 1) A scene understanding model trained based on static images collected at project closeout data often is not useful for understanding an ever-changing scene where facility components could have intermediate and incomplete states, such as construction sites; 2) Many facility components are occluded at project closeout, which creates many challenges when labeling or detecting them. By leveraging as-designed information present in building information models, this paper proposed an approach to generate synthetic data for training semantic understanding models reflecting the changes in site conditions through 4D-BIM and Unreal Engine. The paper contains two primary contributions: 1) The proposed workflow addresses issues with changing scenes by generating synthetic images with ground truth semantic segmentations during any stage of construction based on given schedule information; 2) The proposed method reduced the labeling effort by utilizing the semantically rich as-designed information that exists in a BIM. The proposed workflow was tested on an academic building in its ability to create a useful synthetic dataset using the Uniformat (2010) as semantic taxonomy. The experiment results showed that the proposed data augmentation can improve the ability of scene understanding for images captured in changing scenes.