

Sensitivity Analysis of Excavator Activity Recognition Performance based on Surveillance Camera Locations

Yejin SHIN¹, Seungwon SEO², and Choongwan KOO^{3*}

¹ Department of Architecture, Incheon National University, Korea, E-mail address: yj1385@inu.ac.kr

² Department of Architecture, Incheon National University, Korea, E-mail address: ssw9541@inu.ac.kr

³ Division of Architecture & Urban Design, Incheon National University, Korea, E-mail address: cwkoo@inu.ac.kr

Abstract

Given the widespread use of intelligent surveillance cameras at construction sites, recent studies have introduced vision-based deep learning approaches. These studies have focused on enhancing the performance of vision-based excavator activity recognition to automatically monitor productivity metrics such as activity time and work cycle. However, acquiring a large amount of training data, i.e., videos captured from actual construction sites, is necessary for developing a vision-based excavator activity recognition model. Yet, complexities of dynamic working environments and security concerns at construction sites pose limitations on obtaining such videos from various surveillance camera locations. Consequently, this leads to performance degradation in excavator activity recognition models, reducing the accuracy and efficiency of heavy equipment productivity analysis.

To address these limitations, this study aimed to conduct sensitivity analysis of excavator activity recognition performance based on surveillance camera location, utilizing synthetic videos generated from a game-engine-based virtual environment (*Unreal Engine*). Various scenarios for surveillance camera placement were devised, considering horizontal distance (20m, 30m, and 50m), vertical height (3m, 6m, and 10m), and horizontal angle (0° for front view, 90° for side view, and 180° for backside view). Performance analysis employed a 3D ResNet-18 model with transfer learning, yielding approximately 90.6% accuracy.

Main findings revealed that horizontal distance significantly impacted model performance. Overall accuracy decreased with increasing distance (76.8% for 20m, 60.6% for 30m, and 35.3% for 50m). Particularly, videos with a 20m horizontal distance (close distance) exhibited accuracy above 80% in most scenarios. Moreover, accuracy trends in scenarios varied with vertical height and horizontal angle. At 0° (front view), accuracy mostly decreased with increasing height, while accuracy increased at 90° (side view) with increasing height. In addition, limited feature extraction for excavator activity recognition was found at 180° (backside view) due to occlusion of the excavator's bucket and arm.

Based on these results, future studies should focus on enhancing the performance of vision-based recognition models by determining optimal surveillance camera locations at construction sites, utilizing deep learning algorithms for video super resolution, and establishing large training datasets using synthetic videos generated from game-engine-based virtual environments.

Key words: excavator; activity recognition; location of surveillance camera; synthetic video