ABSTRACT

Crowd phenomena frequently occur in public spaces such as stations, shopping centers, and tourist destinations, necessating surveillance systems to ensure public safety and order. Conventional surveillance systems that rely on CCTV are still dependent on human monitoring, making them inefficient and prone to human error. Therefore, this study proposes a crowd counting method based on deep learning using the YOLO11 (You Only Look Once version 11) algorithm. The model is trained and evaluated using the JHU-CROWD++ dataset, which contains various crowd densities, lighting conditions, camera angles, and weather degradations. Model optimization is conducted through transfer learning and genetic algorithms for hyperparameter tuning. The experiment results show that the YOLO11s variant achieves the best performance, with a Mean Absolute Error (MAE) of 64.61 and Root Mean Square Error (RMSE) of 107.21, as well as an inference time of 410.57 miliseconds, which outperforms other models such as DeTR and Faster R-CNN. These findings demonstrate that YOLO11 is a suitable approach for real-time crowd surveillance systems to enhance public safety.

Keyword: YOLO11, Object Detection, Crowd Counting, JHU-CROWD++, Real-Time