

ABSTRACT

Amid the increasingly intense competition between online and offline shopping, physical retail stores remain the primary choice for most Indonesian consumers. To maintain this competitive edge, store layout planning becomes a crucial strategic aspect, as it can influence customer flow, visit duration, and purchase decisions. However, optimizing store layouts is often hindered by the limited availability of customer behavior data, resulting in less targeted layout strategies. This study aims to develop a computer vision system that integrates object detection and pose estimation to detect crowd patterns and customer interactions in retail stores based on visual data that provides an indication and validation of the effectiveness of retail store layout arrangements. This study uses the CRISP-DM methodology approach as a framework, which includes stages of business understanding, to system deployment. The method involves training a YOLOv11 model for crowd density detection and a classification algorithm based on keypoint coordinates to detect product-picking interactions. Visual data from both approaches are analyzed using Rank Spearman Correlation to evaluate the relationship between crowd levels and interaction intensity. The results show that the YOLOv11 model achieved an mAP50 of 94% and an F1-score of 89.9%, while the classification model using the Random Forest Classifier algorithm achieved an accuracy of 93.3% and an F1-score of 92%. From the results of the implementation of the test video, it was found that in Area-01 there was a significant relationship with the high relationship strength category, while in Area-02 the relationship was low and insignificant. This shows that high-density areas can significantly encourage the intensity of customer interaction, although it does not apply evenly across the area. This finding confirms that a visual-based integrative approach can provide a concrete strategic basis for decision making in optimizing retail store layouts.

Keywords— YOLOv11, Object detection, Pose estimation, Crowded detection, Interaction detection, Retail, Store Layout, Spearman Rank