

INTISARI

Sistem *instance segmentation* merupakan kombinasi deteksi objek dan segmentasi piksel yang memberikan informasi rinci mengenai objek-objek pada citra. Sistem *instance segmentation* dapat memberikan informasi yang lebih detail mengenai objek-objek yang dideteksi dikarenakan sistem *instance segmentation* melibatkan kombinasi *object detection* dan *image segmentation*. Mask R-CNN merupakan salah satu model yang populer untuk *instance segmentation*.

Pada penelitian sebelumnya Mask R-CNN menggunakan ResNet-50 sebagai backbone. Namun, penelitian ini mengusulkan penggantian ResNet-50 dengan EfficientNet-B3 dan EfficientNetV2-S untuk meningkatkan efisiensi komputasi. Penggantian ini bertujuan untuk mengurangi beban komputasi sambil tetap mempertahankan kinerja model. Pada penelitian ini data latih yang digunakan sebanyak 3833 gambar, dengan total 10 kelas objek *indoor* yang umum ditemukan sehari-hari. Terdapat pengujian untuk menganalisis performa segmentasi model dan performa komputasi model.

Hasil menunjukkan bahwa model Mask R-CNN dengan *backbone* EfficientNetV2-S mencapai kinerja segmentasi tertinggi dengan mAP sebesar 29,107. Kemudian, pada kinerja komputasi, model Mask R-CNN dengan ResNet-50 unggul dengan kecepatan 13,65 FPS pada uji coba video, menjadikannya yang tercepat di antara pilihan yang ada, kemudian diikuti oleh model dengan EfficientNet-B3, dan yang terakhir model dengan EfficientNetV2-S.

Kata Kunci: *Instance segmentation*, Mask R-CNN, EfficientNet, ResNet, *indoor object*

ABSTRACT

The instance segmentation system is a combination of object detection and pixel segmentation that provides detailed information about objects in an image. The instance segmentation system can provide more detailed information about detected objects because it involves a combination of object detection and image segmentation. Mask R-CNN is one of the popular models for instance segmentation.

In previous research, Mask R-CNN used ResNet-50 as its backbone. However, this research proposes replacing ResNet-50 with EfficientNet-B3 and EfficientNetV2-S to improve computational efficiency. This replacement aims to reduce the computational load while maintaining the model's performance. In this study, the training data used consisted of 3833 images, with a total of 10 common indoor object classes found in everyday life. There were tests conducted to analyze the segmentation performance of the model and the computational performance of the model.

The results show that the Mask R-CNN model with an EfficientNetV2-S backbone achieves the highest segmentation performance with a mAP of 29.107. Furthermore, in terms of computational performance, the ResNet-50-based Mask R-CNN model excels with a speed of 13.65 FPS in video tests, making it the fastest among the available options, followed by the model with EfficientNet-B3, and finally the model with EfficientNetV2-S.

Keywords: Instance segmentation, Mask R-CNN, EfficientNet, ResNet, Indoor object