



## INTISARI

Seiring dengan pertumbuhan populasi, pembangunan infrastruktur yang pesat, serta tingginya frekuensi bencana alam, kebutuhan akan data geospasial yang akurat semakin meningkat. Meskipun demikian, metode manual dalam pemetaan bangunan terbukti tidak efisien dan rentan terhadap kesalahan. Penelitian ini memanfaatkan teknologi deep learning dan visi komputer untuk mengatasi tantangan tersebut dengan mengembangkan model *instance segmentation* terbaru pada citra udara orthomosaic. Dengan mengimprovisasi Cascade Mask R-CNN dengan ResNeXt101 sebagai backbone untuk meningkatkan kemampuan ekstraksi fitur, model ini mampu mendeteksi detail yang lebih halus dalam lingkungan perkotaan yang kompleks. Penambahan *Deformable Convolutional Networks* (DCN) memberikan fleksibilitas dalam menangani objek dengan bentuk yang bervariasi, sehingga memungkinkan segmentasi yang lebih akurat pada bangunan yang memiliki bentuk tidak beraturan. Selain itu, penambahan *Soft Non-Maximum Suppression* (Soft-NMS) dapat memperbaiki deteksi objek dengan mengurangi kesalahan ketika menghadapi objek yang saling tumpang tindih, sehingga menghasilkan peningkatan performa segmentasi yang signifikan.

Dengan memanfaatkan arsitektur Cascade Mask R-CNN yang sudah diimprovisasi ini, model berhasil mencapai *mean Average Precision* (mAP) sebesar 0.574 dan mAP50 sebesar 0.790 untuk tugas *instance segmentation*. Hasil ini menunjukkan peningkatan performa yang signifikan dibandingkan metode konvensional. Pendekatan ini memberikan manfaat besar dalam perencanaan kota, manajemen bencana, dan pemantauan infrastruktur. Dengan otomatisasi proses segmentasi, pendekatan ini menawarkan solusi yang lebih efisien dan akurat untuk berbagai analisis geospasial.

Kata kunci : *Instance Segmentation* Bangunan, *Deep Learning*, Cascade Mask R-CNN, DCN, Soft-NMS



## ABSTRACT

*As rapid urban development and natural disasters increase the demand for accurate geospatial data, traditional manual methods for mapping buildings have proven inefficient and prone to error. This study leverages deep learning and computer vision to address this challenge by developing a cutting-edge instance segmentation model for orthomosaic aerial imagery. By adding Cascade Mask R-CNN with ResNeXt101, the model enhances feature extraction, enabling better detection of fine details in complex urban environments. Deformable Convolutional Networks (DCN) introduce flexibility in handling objects with deformable shapes, allowing for more accurate segmentation of buildings with irregular forms. Soft Non-Maximum Suppression (Soft-NMS) refines object detection by reducing errors when dealing with overlapping structures, leading to significantly improved segmentation performance.*

*Utilizing the Cascade Mask R-CNN architecture with these enhancements, the model achieved a mean Average Precision (mAP) of 0.574 and mAP50 of 0.790 for instance segmentation. This demonstrates a significant improvement over traditional methods. The approach offers substantial benefits in urban planning, disaster management, and infrastructure monitoring. By automating the segmentation process, this method provides greater efficiency and more precise solutions for geospatial analysis.*

**Keywords** : Building Instance Segmentation, Deep Learning, Cascade Mask R-CNN, DCN, Soft-NMS