ANALISIS SEGMENTASI BANGUNAN PERKOTAAN MENGGUNAKAN CNN KONVENSIONAL, U-NET BERBASIS VGG16, DAN DEEPLABV3+ DENGAN RESNET-50 PADA CITRA SATELIT RESOLUSI TINGGI

Oleh

Putu Haryaka Seta Dewa, NIM 2215101033

Jurusan Teknik Informatika

Abstrak

Pertumbuhan urbanisasi di Indonesia mendorong kebutuhan pemetaan bangunan yang akurat sebagai dasar perencanaan tata ruang, mitigasi bencana, pemantauan perkembangan wilayah, dan pengelolaan infrastruktur perkotaan. Penelitian ini membandingkan performa tiga model segmentasi bangunan pada citra satelit resolusi tinggi, yaitu CNN konvensional (*CNN-K*), U-Net berbasis VGG16 (*U-VGG*), dan DeepLabV3+ dengan ResNet-50 (*DL-ResNet*). Dataset utama berasal dari kawasan Bali Selatan dengan 1.216 patch citra yang telah melalui proses labeling dan augmentasi untuk memperkaya variasi pola bangunan. *CNN-K* digunakan sebagai baseline dengan arsitektur encoder–decoder sederhana, *U-VGG* memanfaatkan VGG16 pretrained untuk menangkap tekstur dan kontur bangunan secara detail, sedangkan *DL-ResNet* menggunakan modul ASPP untuk memahami objek dengan berbagai skala pada kawasan padat. Evaluasi dilakukan menggunakan metrik akurasi, IoU, *Dice Coefficient*, P*recision*, *Recall*, dan *F1-score* pada data pelatihan, validasi, dan pengujian. Hasil menunjukkan *U-VGG* unggul dengan dice 89%, precision 87%, dan IoU 81%, memberikan keseimbangan antara presisi, efisiensi memori, dan kemampuan generalisasi. *DL-ResNet* mendekati performa *U-VGG* (dice 85%, precision 91%, IoU 80%) dengan kualitas detail visual segmentasi yang lebih tinggi, namun membutuhkan sumber daya komputasi lebih besar. *CNN-K* mencatat performa terendah pada data uji (dice 40%, precision 69%, IoU 14%) akibat overfitting. Temuan ini menegaskan bahwa *U-VGG* cocok untuk aplikasi pemetaan cepat dan efisien di kawasan perkotaan, sedangkan *DL-ResNet* lebih sesuai untuk analisis yang memerlukan ketelitian tinggi, seperti perencanaan infrastruktur detail atau monitoring perubahan bangunan secara presisi pada citra satelit resolusi tinggi.

Kata-kata kunci: segmentasi bangunan, citra satelit resolusi tinggi, CNN konvensional, U-Net, VGG16, DeepLabV3+, ResNet-50

URBAN BUILDING SEGMENTATION ANALYSIS USING CONVENTIONAL CNN, VGG16-BASED U-NET, AND DEEPLABV3+ WITH RESNET-50 ON HIGH-RESOLUTION SATELLITE IMAGERY

By

Putu Haryaka Seta Dewa, NIM 2215101033

Department of Informatics Engineering

Abstract

The rapid growth of urbanization in Indonesia has increased the demand for accurate building mapping to support spatial planning, disaster mitigation, regional development monitoring, and urban infrastructure management. This study compares the performance of three building segmentation models on high-resolution satellite imagery: a conventional CNN (CNN-K), a VGG16-based U-Net (U-VGG), and DeepLabV3+ with a ResNet-50 backbone (DL-ResNet). The dataset was collected from the South Bali region and consists of 1,216 image patches that were manually labeled and augmented to increase building pattern diversity. CNN-K was used as a baseline model with a simple encoder–decoder architecture. U-VGG leverages a pretrained VGG16 to capture detailed building textures and contours, while DL-ResNet employs an Atrous Spatial Pyramid Pooling (ASPP) module to handle multi-scale objects in dense urban areas. Model performance was evaluated using accuracy, Intersection over Union (IoU), Dice coefficient, precision, recall, and F1-score on training, validation, and testing datasets. The results show that U-VGG achieved the best overall performance, with a Dice score of 89%, precision of 87%, and IoU of 81%, indicating a strong balance between accuracy, memory efficiency, and generalization capability. DL-ResNet demonstrated comparable performance (Dice 85%, precision 91%, IoU 80%) and produced more detailed visual segmentation results, but required higher computational resources. In contrast, CNN-K achieved the lowest performance on the test dataset (Dice 40%, IoU 14%) due to overfitting. These findings suggest that U-VGG is suitable for efficient urban building mapping, while DL-ResNet is more appropriate for applications requiring higher segmentation precision.

Keywords: building segmentation, high-resolution satellite imagery, conventional CNN, U-Net, VGG16, DeepLabV3+, ResNet-50