Mask-RCNN Performance in Building Footprint Extraction in Formal and Informal Areas of the City of Cape Town, South Africa

Khumeleni Makungo (South Africa)

Key words: Informal settlements; Photogrammetry; Building footprint extraction; detection; segmentation; Mask R-CNN; high-resolution aerial imagery; LiDAR-derived nDSM

SUMMARY

In the City of Cape Town Metropolitan (CoCT), South Africa, GIS analysts manually digitize building footprints in aerial and stereo-aerial images. This process is time-consuming, expensive, and inefficient. Recent studies have explored automatic and semi-automatic methods for extracting building footprints, but none consider the unique characteristics of the landscape, including different types of residential zones, industrial zones, and informal settlements. This study employed Mask R-CNN, a deep learning model, to tackle this challenge. Separate models were trained on data from the Blaauwberg district within CoCT, which encompasses all three zones. The trained models were then tested individually on datasets representing formal residential, industrial zones, and informal settlements. Evaluation metrics such as precision, recall, F1-score, and Average Precision (AP) were calculated to assess model performance. The Mask R-CNN algorithm demonstrated successful extraction of building footprints in formal residential and industrial zones from aerial imagery and LiDAR-derived nDSM, achieving satisfactory AP and F1-scores. However, extracting footprints in densely populated informal settlements proved challenging, with poor algorithm performance. However, by combining footprints extracted from both aerial imagery and LiDAR-derived nDSM, improved results were observed in informal settlements. The study revealed that the Mask R-CNN algorithm is highly effective in extracting formal residential building footprints using both aerial imagery and LiDAR-derived nDSM, as well as industrial footprints using LiDAR-derived nDSM. To optimize performance in informal settlements, a fusion of footprints from aerial imagery and LiDAR-derived nDSM is necessary. These trained models can be utilized to enhance the existing City of Cape Town's 2D building footprint layer, providing a more comprehensive and up-to-date representation.

Mask-RCNN Performance in Building Footprint Extraction in Formal and Informal Areas of the City of Cape Town, South Africa (13609)
Khumeleni Makungo (South Africa)

FIG Congress 2026
The Future We Want - The SDGs and Beyond
Cape Town, South Africa, 24–29 May 2026