東京大学　関本研究室 / Sekimoto Lab. IIS, the University of Tokyo.

Revamping Building Extraction with YOLOv8: A Multi-Domain Assessment in Bangkok's Diverse Land-Use Types

Bhanu Prasad CHINTAKINDI, Shenglong CHEN, Yoshiki OGAWA, Yoshihide SEKIMOTO

Background

- Building footprint data is crucial for urban planning, crisis management, navigation, database creation, change detection, and 3D modeling.
- YOLO models (You Only Look Once), especially YOLOv7 and YOLOv8, are prominent in object detection due to their speed and accuracy.
- Existing models face limitations in rural regions or developing countries due to intricate building patterns across diverse land-use types.

Objective

- Develop seven specialized YOLOv8 and YOLOv7 models trained from scratch to accommodate distinct land-use types and their unique building patterns.
- Conduct a multi-domain assessment across various land-use types to evaluate the accuracy, applicability, and adaptability of these models.

Methodology

1. Model Development:

a. Seven specialized building extraction models were developed using instance segmentation techniques.

b. YOLOv8x-seg model was chosen as the baseline and trained from scratch to create models for Rural, BuiltUp, SubUrban, Urban, Commercial, and ALUT (All Land-Use-Types).

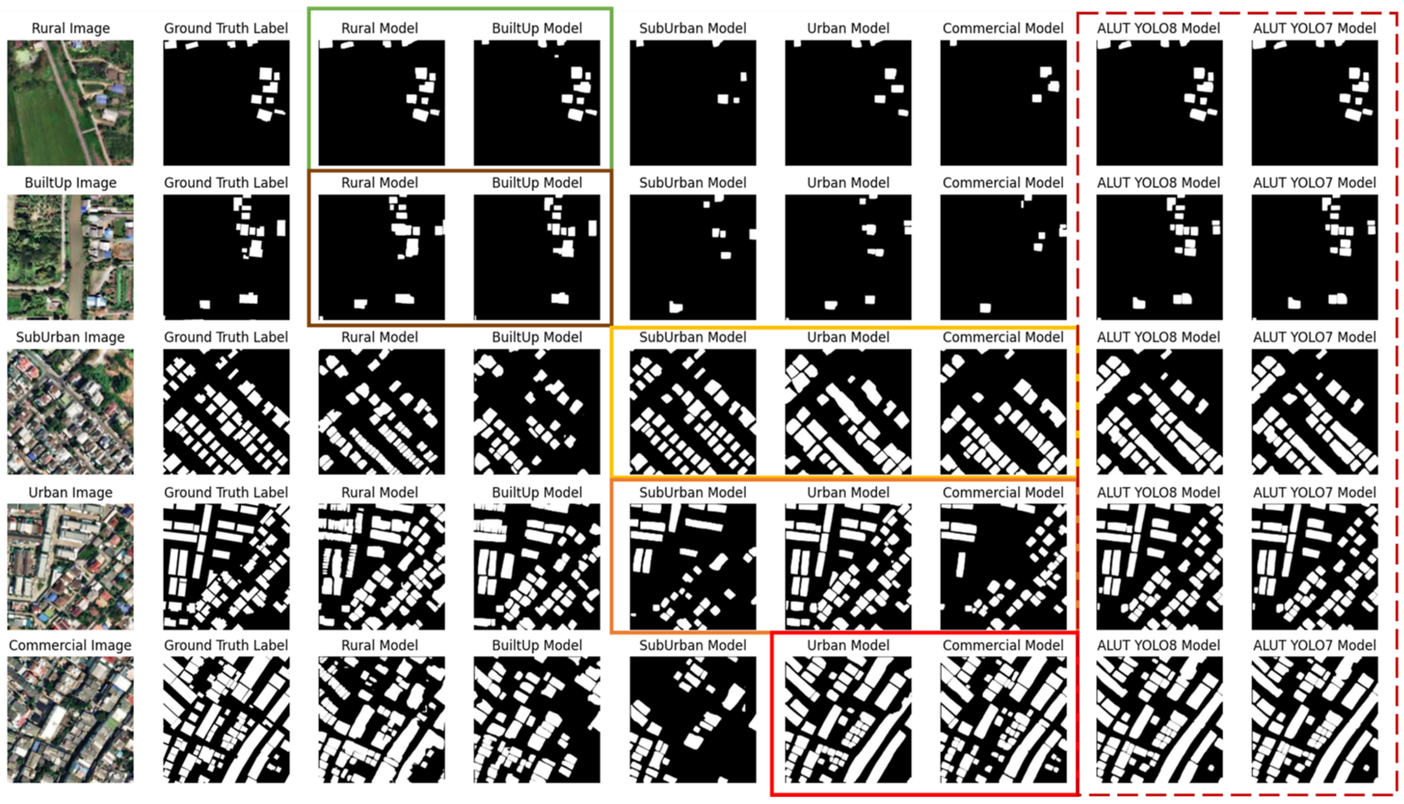
c. YOLOv7-seg model was used as an additional baseline, trained with the ALUT dataset.

2. Multi-Domain Assessment:

a. Evaluated and compared model performance within and across land-use types using metrics such as F1 scores, visual predictions, and radar plots.

Dataset

- **Source:** Open-source 0.3m resolution satellite imagery from Mapbox and 2 million building footprints from the Bangkok Metropolitan Administration GIS Portal.
- **Land-Use Types**: Categorized into Commercial, Urban, SubUrban, BuiltUp, and Rural based on Bangkok's 2013 land-use zoning plan.
- **Data Preparation:** Cropped 0.3m resolution satellite image patches (640x640 pixels) with a 30% overlap were used to create training, validation, and testing datasets for each land-use category.

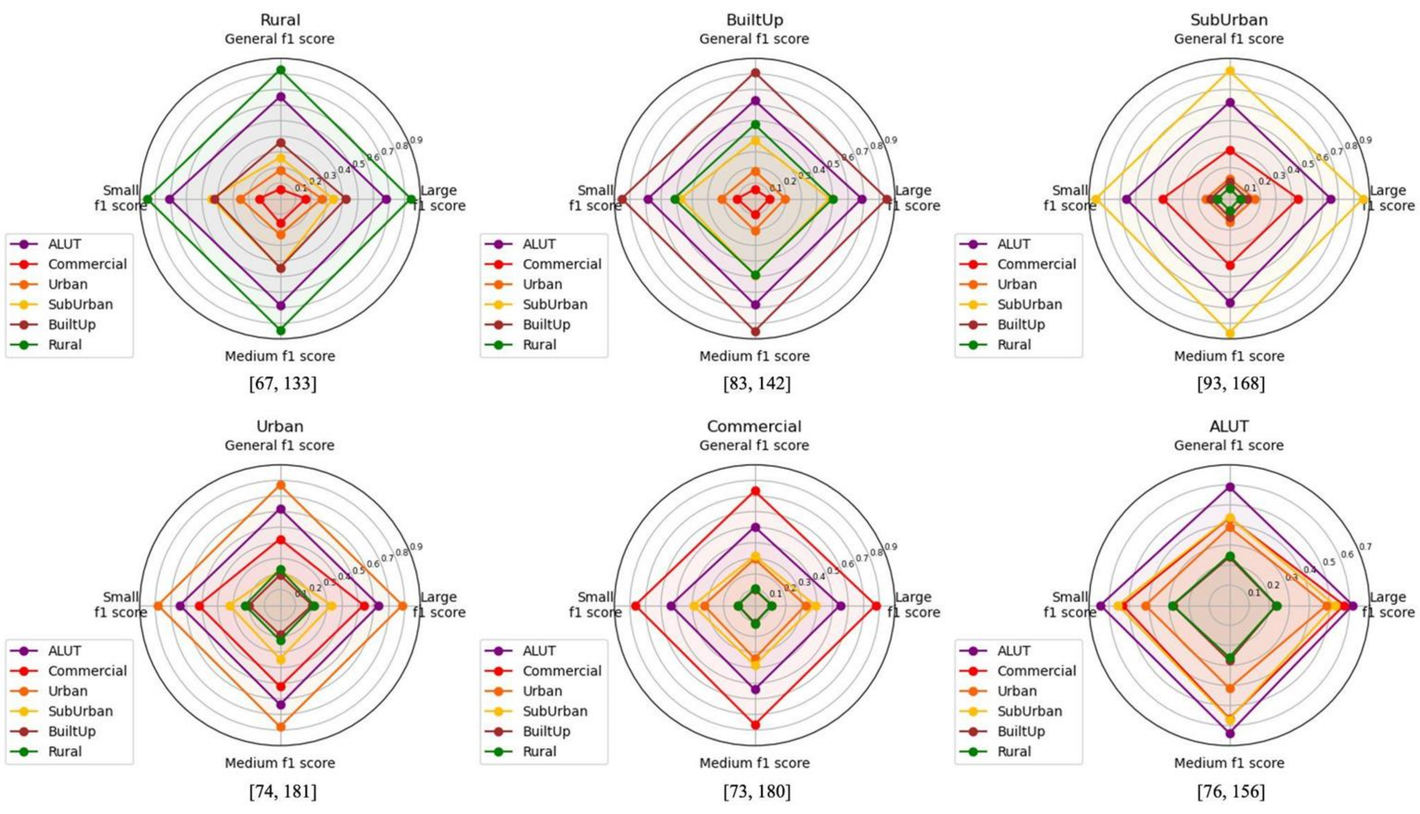


Predicted results by Individual Land-Use-Type Models and corresponding source image and ground truth label in (a) Rural Area, (b) BuiltUp Area, (c) SubUrban Area, (d) Urban Area, and (e) Commercial Area.

Results

Accuracy Analysis:

- Models showed remarkable adaptability, aligning closely with ground truth buildings.
- Performance varied depending on land use; for instance, the rural model performed best in rural zones, and the commercial model excelled in commercial zones.
- The ALUT model consistently ranked second, demonstrating strong generalization capabilities.



Radar Plots illustrating the performance of the six YOLOv8: Individual Land-Use-Type Models and ALUT Model across different building-size categories in various test areas, including (a) Rural, (b) BuiltUp, (c) SubUrban, (d) Urban, (e) Commercial, and (f) ALUT. The quantile limit values are also included for reference.

Conclusion

- The study highlights the effectiveness of specialized YOLOv8 models in accurately extracting building footprints across diverse land-use types.
- Future research involves creating a global building extraction model by ensembling individual models for specific land-use cases, providing a template for accurate building extraction in various international contexts.