

# Large-scale building extraction from open-source satellite imagery via super-resolution-based instance segmentation

Shenglong Chen, Yoshiki Ogawa, Yoshihide Sekimoto

## Background

Building footprint is a primary dataset of an urban geographic information system (GIS) database. However, the instance segmentation method based on the Mask R-CNN model ought to be improved toward extracting and fusing multi-scale features. Moreover, open-source satellite image datasets with wider spatial coverage and temporal resolution than high-resolution image may exhibit different coloration and resolution.

## Object

- We enhanced the multi-scale feature extraction ability of Mask R-CNN model via a Multi-path Vision Transformer backbone.
- We improved the quality of open-source satellite image by super-resolution and color normalization.

## Dataset

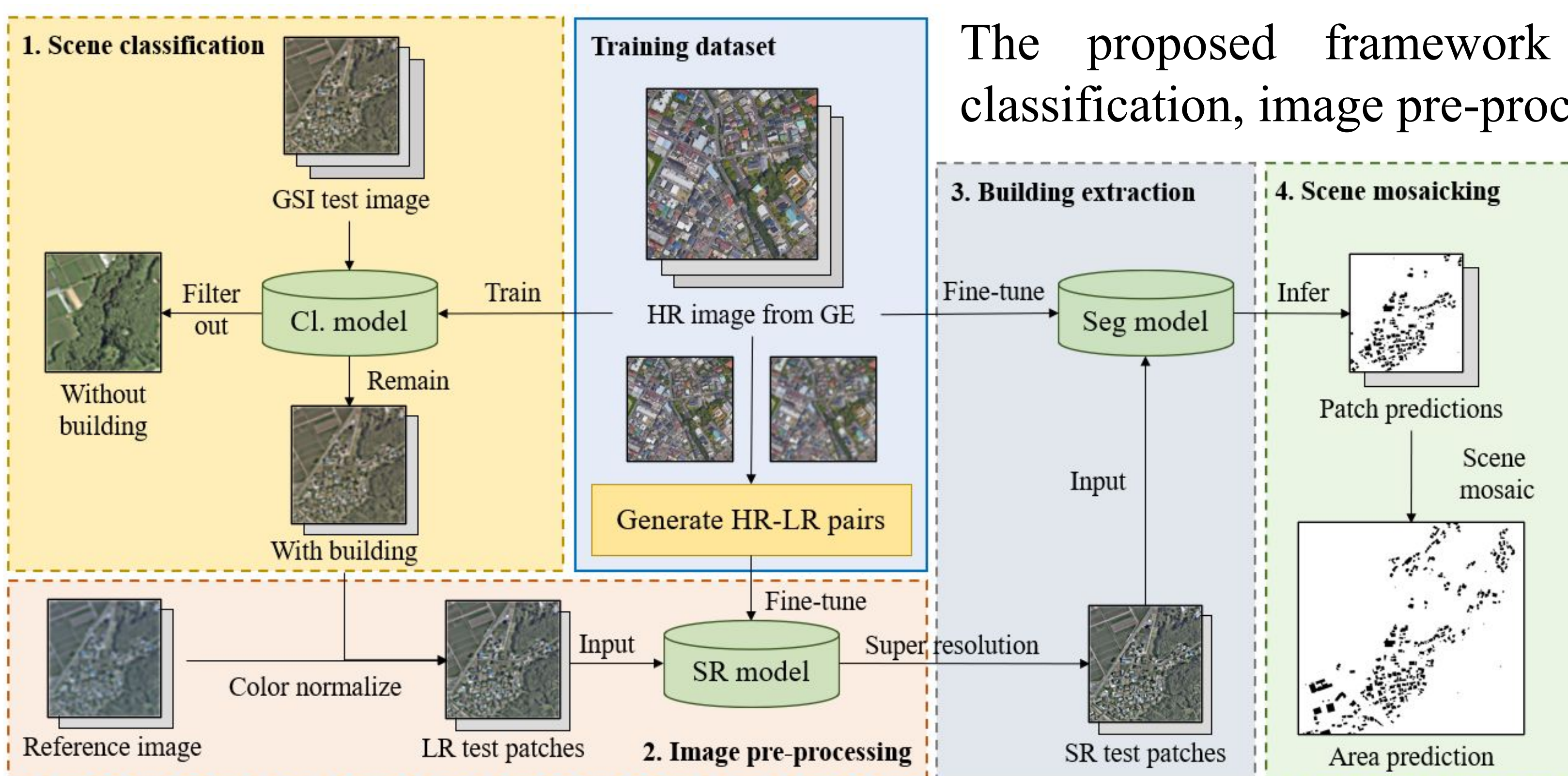
**Training data:** 1222 images from Google Earth (0.3 m) with 186000 annotated footprints.

**Test data:** 81348 patches from Geospatial Information Authority of Japan (GSI) (0.6~1 m).

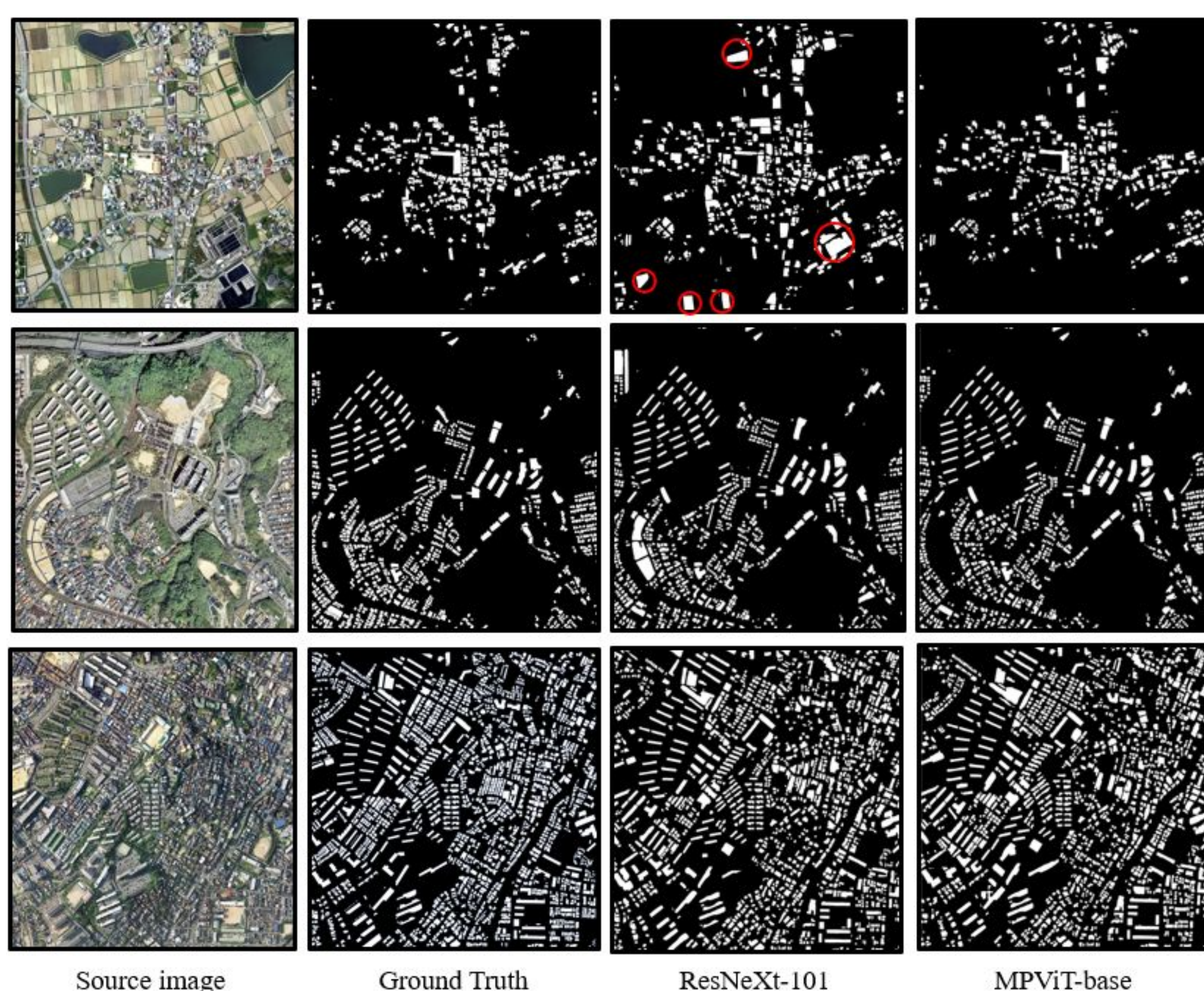
**Training area:** Shinjuku, Setagaya, Hachioji & Susono

**Test area:** Hyogo Prefecture, 19187 km<sup>2</sup>

## Methodology



## Result



Prediction results of MPViT and ResNeXt model

Object-wise metric of different areas				
Area	Backbone	Precision	Recall	F1
Rural	ResNeXt-10	0.71	<b>0.64</b>	0.67
	MPViT-b	<b>0.76</b>	0.63	<b>0.69</b>
Suburban	ResNeXt-10	0.82	0.67	0.74
	MPViT-b	<b>0.87</b>	<b>0.72</b>	<b>0.78</b>

Prediction results under different SR strategies