Appendix: Model Pipeline

Model's inherent predictive power | **Improving False Positives**

Data Preparation	Pre-Processing	Modeling	Prediction	Post-Processing

Data Preparation

> CapeTown
> Tiff images(`18~**`23**)
> 12500*12500 pixels
(8cm / pixels)
> CRS (ESRI:102562)
> Annotation
- 3 classes, GPKG file
- 5,162 labels

+ California Dataset
(1class, 32cm/pixels)
+ Germany Dataset
(1class, 15cm/pixels)

Pre-Processing

> Crop tiles into 320*320
patches (as-is)
> File naming scheme for
tiles restoration
> Stratified Data Split
(Train/Val/Test)

Modeling

> *Pixel-level segmentation*
Mask2Former(320*320)
Mask2Former(1024*1024)
NNUNet(1024*1024)

> Hyper Parameter Tuning
- (e.g.,) Learning Rate
> Data Augmentation
- (e.g.,) Geometric
transform, Noise + Blur

(Legacy Models)
UNet(ResNext50+FPN)
Yolo+Unet, UnetPlusPlus
Attention-UNet

Prediction

> Class(0: Background,
1: Solar Panel,
2: Water Heater,
3: Pool Heater)
> Pixel Coordinates
> JSON files

Post-Processing

> 1. Polygonization
- Stitch patches to tiles
- Compute Centroid X,Y
- Pixels → CRS → GPS
- Pixels into Polygons

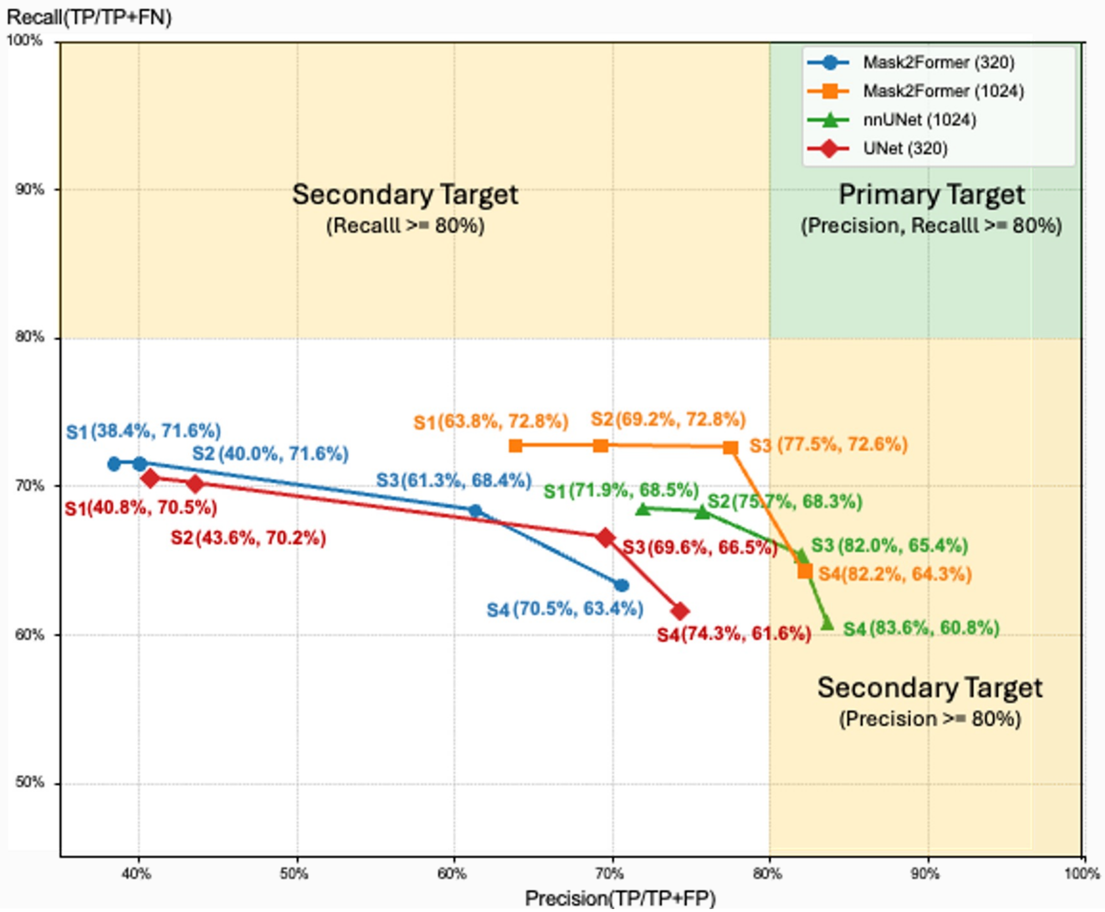
> 2. Grouping Polygons
- Merge polygons within
n pixels *(as-is: 2 pixels)*

> 3. Dropping Small objects
- Drop n m^2 *(as-is: 1.7m^2 for all classes)*

> 4. Building Footprint Filter
- Remove objects outside the building footprint

Appendix: Post-Processing & Evaluations

S1 Polygonization → S2 Grouping → S3 Remove small Polygons → S4 Building Footprint Filtering → **Evaluation**



Post-processing consistently improves Precision across all models

TP drops when applying building footprint filtering

Mask2Former (1024) shows the **best balance** of Precision and Recall **before Step 4**

nnUNet (1024) achieves the **highest Precision** after full post-processing (S1+2+3+4)

Solution

1. Re-evaluate Building Footprint filtering
2. Enhance Training Data Quality

Model Evaluation: Polygonization + Group neighboring polygons + Dropping small objects

Model	Class	TP	FP	FN	Precision	Recall
Mask2Former (1024)	**All**	**-**	**-**	**-**	**77.5%**	**72.6%**
	PV_normal	243	68	89	78%	73%
	PV_heater	220	62	83	78%	73%
	PV_pool	161	51	63	76%	72%
nnUNet (1024)	**All**	**-**	**-**	**-**	**81.9%**	**65.3%**
	PV_normal	248	72	88	78%	74%
	PV_heater	153	16	152	91%	50%
	PV_pool	167	37	61	82%	73%
Mask2Former (320)	**All**	**-**	**-**	**-**	**61.3%**	**68.3%**
	PV_normal	244	195	92	56%	73%
	PV_heater	168	68	112	71%	60%
	PV_pool	163	100	62	62%	73%
Unet(320)	**All**	-	-	-	69.6%	66.5%
	PV_normal	222	111	115	66%	65%
	PV_heater	189	57	111	76%	63%
	PV_pool	163	83	63	66%	72%