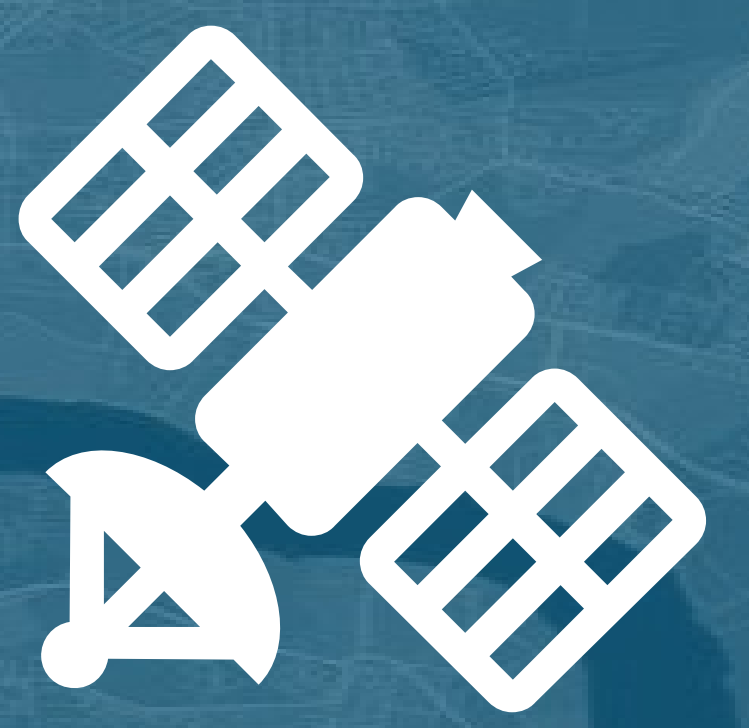


# BUILDING DATA FROM HIGH-RESOLUTION IMAGES



Francisco Caldeira & Cristina Gabriel,  
Statistics Portugal – Instituto Nacional de Estatística (INE)

## ! PROBLEM STATEMENT

- Statistics Portugal faces challenges in maintaining an up-to-date Geocoded Building Dataset (BGE), which is a foundation for the Portuguese Spatial Data Infrastructure, supporting surveys.
- Many buildings are outdated or missing, compromising data quality and spatial coverage. Conventional approaches are inefficient and lack scalability, so automated solutions are essential.

## 🎯 OBJECTIVE

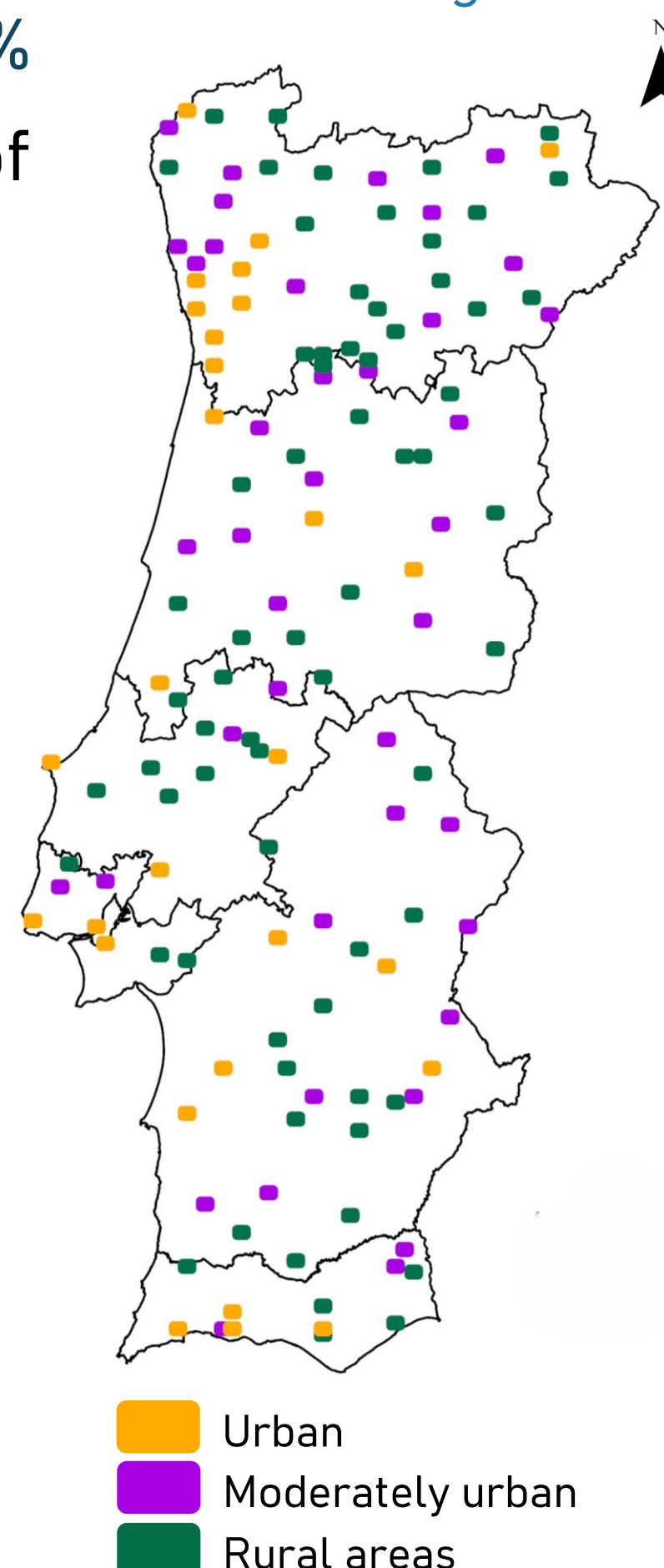
- Detect residential buildings from orthophotos using Deep Learning Models (DL): **Mask R-CNN** (M-RCNN) & **SAM**, and train them for Portugal's mainland.

## ⚙️ PROCESSING WORKFLOW

### 1. DATA PREPARATION

- ~130,000 buildings were manually georeferenced across all TIPAU classes (Figure 1): 50% urban, 29% rural, 21% moderately urban, and classified by roof colour (red, grey, white).

Figure 1. TIPAU Distribution of Georeferenced Buildings



### 2. EXPORT TRAINING SAMPLES

- The 'Export Training Data for Deep Learning' tool in ArcGIS Pro was used to generate image chips and metadata for model training.

### 3. TRAIN THE MODEL AND EVALUATE ACCURACY

- Two DL algorithms were trained to detect, classify and segment buildings using image chips and metadata.
- M-RCNN & SAM trained on 3,000–13,000 buildings; epochs 100–350 (Mask R-CNN), 200–500 (SAM).

### 4. RUN INFERENCE

- Inference applies trained DL models to classify, detect, or segment features. M-RCNN used 'Detect Objects Using Deep Learning' for building boxes and masks, while SAM used 'Classify Pixels Using Deep Learning' for pixel-wise semantic segmentation.

### 5. ACCURACY ASSESSMENT

- Test set: 10 images from diverse regions of mainland Portugal (8% of training data), covering all TIPAU classes; none used in training for unbiased evaluation.
- Ground truth: building footprints were manually georeferenced to assess their accuracy.
- Post-processing: excluded polygons <40 m<sup>2</sup>, removed overlaps with non-buildings, regularized shapes, and cleaned geometries.
- Model accuracy was assessed using Precision, Recall, F1, and IoU.

## 💡 CONCLUSIONS AND RECOMMENDATIONS

- The findings of this study demonstrated and confirmed the applicability and effectiveness of the proposed methodology for urban monitoring. Expanding the training dataset will further improve model performance. The building footprint detection and mapping process, combined with systematic integration into the BGE, will enhance and maintain a complete, up-to-date geospatial infrastructure, supporting statistical production more efficiently.

## 🏃 RESULTS

Inference applied the trained M-RCNN and SAM models to unseen data, comparing their ability to detect and segment building footprints across diverse geographic contexts (Figure 2).

Figure 2. Inference Results: Mask R-CNN vs SAM



Accuracy assessment results for the Precision metric ranged from 0.81–0.94 for M-RCNN and 0.79–0.93 for SAM. IoU, measuring footprint overlap quality, varied from 0.62–0.75 (M-RCNN) and 0.60–0.83 (SAM) across TIPAU classes (Figure 3).

Figure 3. Accuracy assessment Metrics

	PRECISION (0-1)		RECALL (0-1)		F1 (0-1)		IoU (0-1)	
	M-RCNN	SAM	M-RCNN	SAM	M-RCNN	SAM	M-RCNN	SAM
URBAN	0,82	0,93	0,86	0,78	0,84	0,85	0,71	0,72
	0,90	0,91	0,81	0,75	0,85	0,82	0,68	0,75
	0,92	0,93	0,85	0,73	0,88	0,82	0,75	0,75
	0,81	0,79	0,65	0,57	0,73	0,66	0,62	0,60
MODERATELY URBAN	0,82	0,91	0,83	0,78	0,83	0,84	0,72	0,74
	0,91	0,87	0,75	0,60	0,82	0,71	0,73	0,78
	0,89	0,93	0,81	0,67	0,85	0,78	0,75	0,82
RURAL	0,89	0,88	0,82	0,70	0,86	0,78	0,74	0,78
	0,88	0,89	0,84	0,71	0,86	0,79	0,75	0,81
	0,94	0,84	0,64	0,36	0,76	0,50	0,65	0,83