

ISASeg: 2025 Multimodal Semantic Segmentation Challenge

Bellaterra, Spain Nov. 12 – 14, 2025

Datasets:

In this contest, we present the ISASeg dataset, a multimodal resource for 1-meter resolution semantic segmentation of impervious surfaces (ISA). Known formally as the Yangtze River Economic Belt (YREB, China) Multimodal Super-Resolution Segmentation Dataset, it integrates Sentinel-2 multispectral (MSI) imagery, Sentinel-1 synthetic aperture radar (SAR) data, and 1-meter RGB imagery generated via a GAN applied to Sentinel-2 RGB data. Ground-truth labels—produced by a semi-automated process with manual refinement—distinguish Non-ISA (0) from ISA (1), covering about 4.26 billion ISA pixels and 25.33 billion Non-ISA pixels. Spanning 36 major YREB cities, including Wuhan, Nanjing, Wuhu, Yueyang, and Jiujiang, the dataset provides spatially aligned multimodal data acquired from January 1 to December 31, 2021, with temporal mean synthesis ensuring complete spatial coverage. All data were preprocessed using Google Earth Engine and advanced GAN techniques.

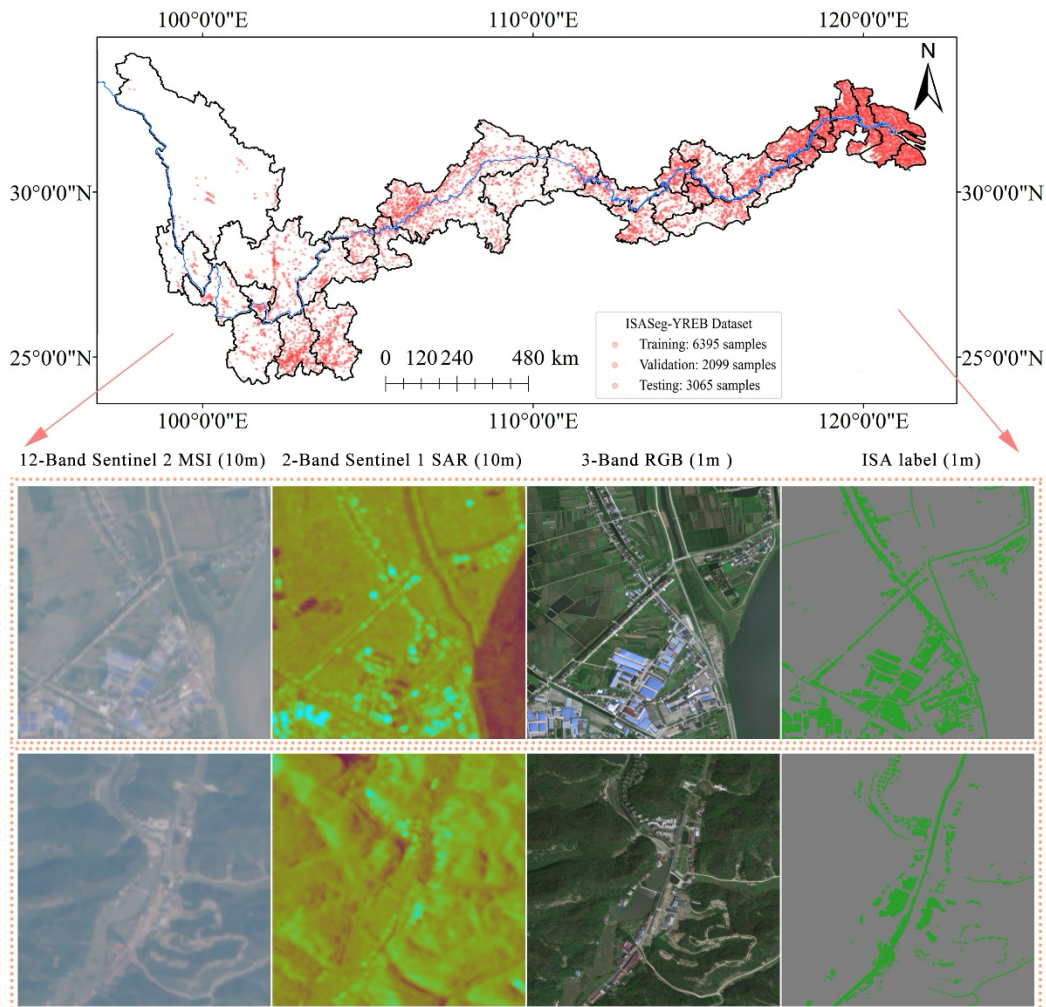


Figure1: The ISASeg dataset

For the entire YREB, a total of 11,559 image pairs—comprising both MSI and SAR data—

were collected to construct the training, validation, and test datasets. These samples comprehensively cover the upper, middle, and lower reaches of the YREB. Each tile has a spatial size of 160×160 pixels, corresponding to 1-meter resolution RGB imagery with a size of 1600×1600 pixels. As depicted in Fig. 1, the dataset comprises Sentinel-2 MSI imagery, Sentinel-1 SAR data, 1-meter-RGB images, and corresponding 1-meter label files. For easy portability and distribution, the dataset together with the ground-truth segmentation masks (only for training tiles) are prepared in .tiff format in separate folders. Fig. 2 shows the class distributions and spectral statistics for train/ validation/test sets. Moreover, the following table summarizes the main characteristics of the **ISASeg** dataset:

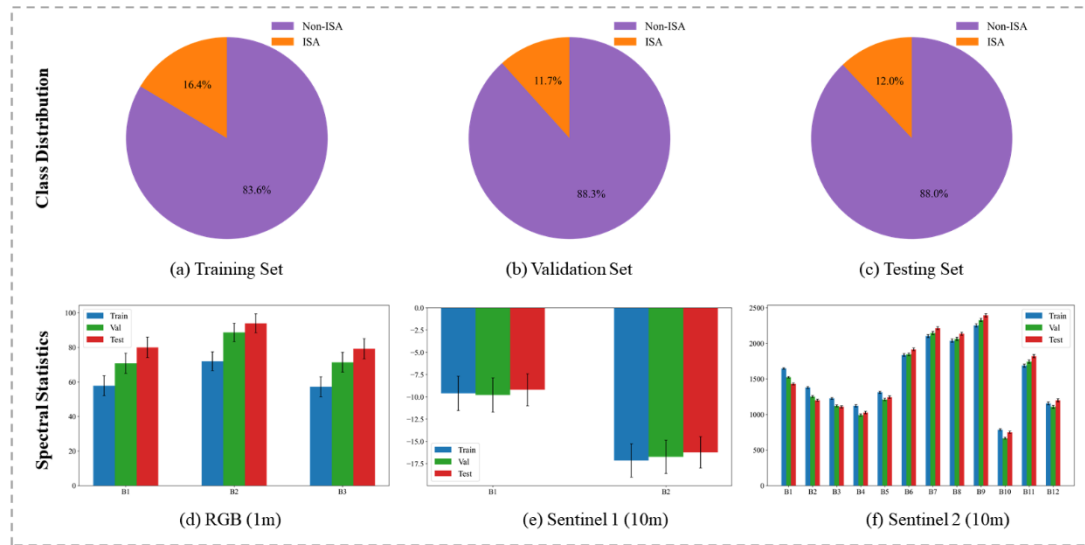


Figure2: Statistics for the **ISASeg** dataset

Name	Size	Resolution	Number of bands	Number of training /validation/testing tiles
ISASeg	$1600 \times 1600/160 \times 160$	1m/10m	RGB with 3 bands Sentinel-2 MSI with 12 bands Sentinel-1 SAR with 2 bands (VH/VV)	6395/2099/3065

Evaluation Protocols:

To evaluate the model performance with respect to the **ISASeg** task, three standard and widely used semantic segmentation metrics are consider, namely, the overall accuracy (OA), the mean intersection over union (mIoU), and the mean F1 score (mF1). More specifically, the OA, also known as pixel accuracy (PA), collects each pixel prediction, thus it can be defined as follows:

$$OA = \frac{\sum_{i=1}^l p_{ii}}{\sum_{i=1}^l \sum_{j=1}^l p_{ij}}$$

where i , j , and l represent the real value, the predicted value, and the total number of classes, respectively, and the p_{ij} denotes the number of pixels that predict the i -th class as the j -th class.

Similarly, the mIoU computes the intersection and union of two sets, which is defined as:

$$mIoU = \frac{1}{l} \sum_{i=1}^l \frac{p_{ii}}{p_{ij} + \sum_{j=1}^l p_{ji} - p_{ii}}.$$

Instead, the last score, mF1, is the harmonic mean of precision (P) and recall (R), which is given by:

$$mF1 = \frac{1}{l} \frac{2 \times P \times R}{P + R},$$

where

$$P = \sum_{i=1}^l \frac{p_{ii}}{\sum_{j=1}^l p_{ij} + p_{ii}}, R = \sum_{i=1}^l \frac{p_{ii}}{\sum_{j=1}^l p_{ji} + p_{ii}}.$$

As a result, the participants, to join the contest, must submit their predicted segmentation masks for the **ISASeg** validation dataset and test dataset with 2 classes. Afterwards, the team who will provide the model with the best performance (considering the average of the ranks obtained for the above-mentioned three quality metrics) will be announced as the winner of the contest.

Download Dataset for Contest:

The participants can resort to either Google Drive or Baidu Netdisk to acquire the training set. The corresponding links are shown below:

Google Drive:

https://drive.google.com/file/d/1-5ScGUHbpud4qxx4SZPypkpdK-gt_jQM/view?usp=sharing

Baidu Netdisk:

<https://pan.baidu.com/s/1Hf2sOE6prWi6hiwW0IEWxw?pwd=ua84> key: ua84

Reference:

- [1] Jie Deng, Danfeng Hong, Chenyu Li, Naoto Yokoya. "Joint Super-Resolution and Segmentation for 1-m Impervious Surface Area Mapping in China's Yangtze River Economic Belt." arXiv preprint arXiv: 2505.05367 (2025).