9. Sensor-Agnostic Soil Sealing Segmentation: Cross-Sensor Generalization of Transformer and Foundation Models on Very-High-Resolution Imagery

Julian Pajo1,2, Mauro Sylos Labini1, Nicolò Taggio1, and Giacomo Pavanello1

1 Planetek Italia, Bari, Italy 2 University of Bologna, Bologna, Italy pajo@planetek.it

Operational soil sealing monitoring must run on a heterogeneous and evolving set of satellite sensors, yet segmentation models are typically trained and selected on a single sensor. We study how well models trained on one sensor transfer to others without any target adaptation. Framing soil sealing as binary semantic segmentation (sealed versus unsealed), we train three Transformer-based architectures on four-band SPOT imagery at 1.5 m resolution: a specialized SegFormer extended with a spectral awareness embedding, the DINOv2-based foundation model CrossEarth, and the wavelength-aware foundation model DOFA. As a second architectural family and a strong reference we add a convolutional U-Net baseline trained under the same budget. All models are integrated into a unified, sensor-agnostic training, inference, and evaluation framework, and are evaluated cross-sensor on five Italian coastal cities acquired by different platforms at resolutions never seen in training. The in-domain ranking and the cross-sensor ranking are decoupled: the convolutional U-Net baseline leads in domain (mIoU 0.834), yet CrossEarth, last in domain, generalizes best across sensors (average mIoU 0.84), while U-Net, SegFormer-SAE, and DOFA drop sharply on the most dissimilar sensors, at times failing to detect the sealed class while retaining high aggregate accuracy driven by class imbalance. The results indicate that in-domain accuracy is not predictive of cross-sensor transfer, and that spectral-agnostic foundation models are a more reliable basis for operational, multi-sensor soil sealing monitoring.