

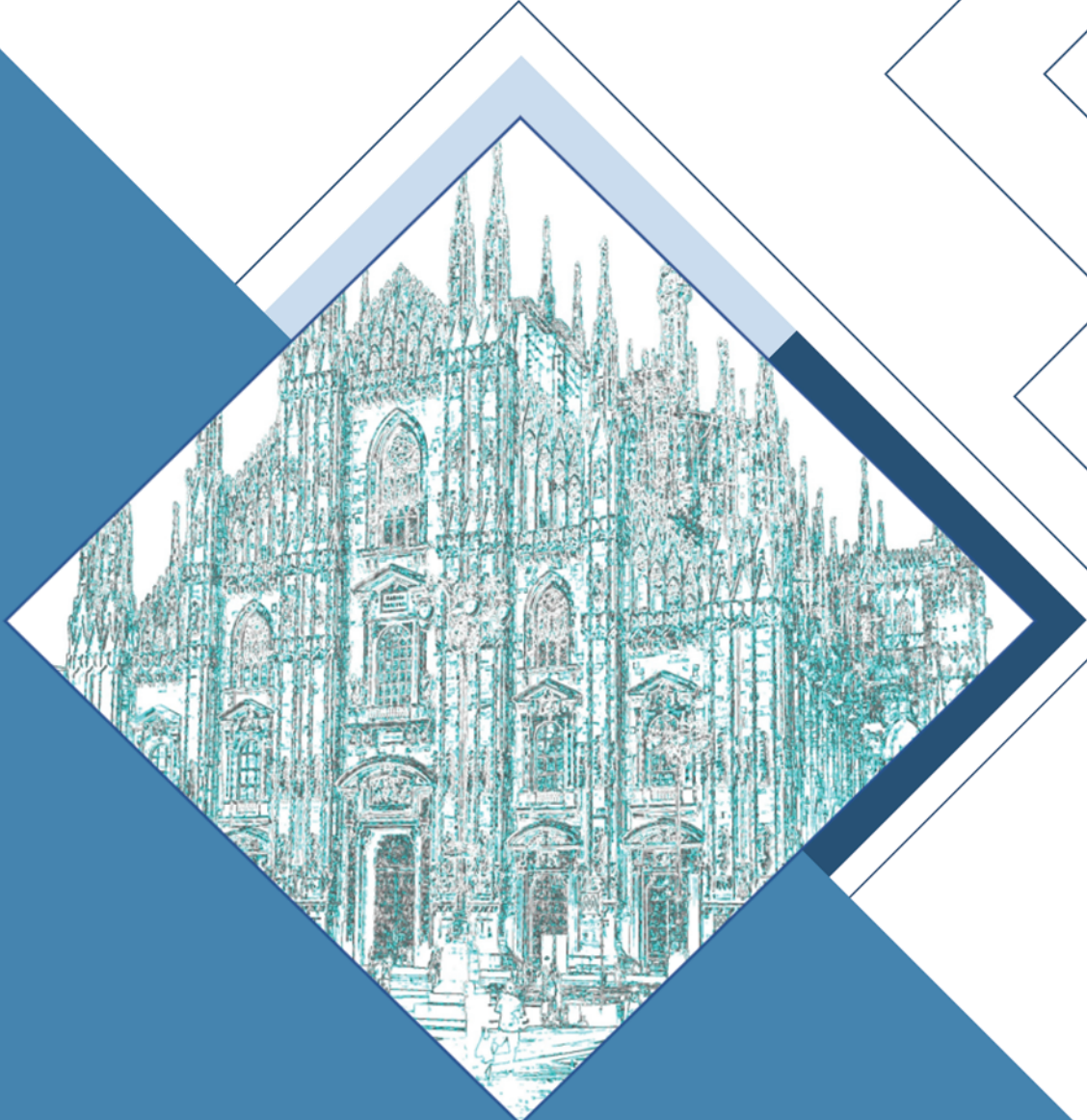


XII AIT



istituto per il rilevamento
elettromagnetico dell'ambiente

INTERNATIONAL **CONFERENCE**



*Smart Earth Observation
for a Sustainable Future*

ABSTRACT BOOK

12 – 14 NOVEMBER 2025 | MILAN

Analytical evaluation of Deep Learning models applied to remotely sensed imagery for object detection and segmentation tasks.

A. Demartis,^{a*} F. Giulio Tonolo,^a

^aPolitecnico di Torino, Department of Architecture and Design (DAD), Laboratory of Geomatics for Cultural Heritage (G4CH), Viale Pier Andrea Mattioli 39, 10125, Torino (TO);

*corresponding author e-mail: andrea.demartis@polito.it

Remote sensing tasks—such as agro-forestry monitoring, land classification, environmental surveillance, and border cross supervision—increasingly rely on robust and accurate automated solutions to reduce the manual workload traditionally required by human operators. In this context, artificial intelligence (AI), and particularly deep learning (DL) applied to image-based approaches, has become the state-of-the-art in spatial data analysis. This technological evolution has led to the emergence of GeoAI, a growing interdisciplinary field combining geospatial analysis and AI that allows the extraction of information from data.

Satellite remote sensing remains a reference point of this domain, offering broad spatial coverage, high temporal resolution, and widely available open-access datasets - such as the one from the Copernicus Sentinel constellations - that enable monitoring applications. However, some applications such as detailed post disaster damage estimation or detection of objects related to environmental crimes, require higher spatial resolution and tailored data acquisition, which are not always achievable with open satellite datasets. In these cases, Remotely Piloted Aircraft Systems (RPASs) have become increasingly valuable due to their flexibility in the acquisition of very high-resolution data with different types of sensors.

Despite the growing accessibility of Earth Observation (EO) data, deep learning workflows remain complex and often inaccessible to non-experts due to the technical skills required for coding, dataset preparation, and model training. To overcome this barrier, a range of pre-trained DL models have been developed, including models specifically adapted to imagery with a ground sample distance of few centimeters (i.e. aerial/RPAS imagery), along with commercial software solutions that integrate them into user-friendly Graphical User Interfaces, broadening their usability.

Pre-trained DL models for image-based Object Detection (OD) can be divided into two categories: traditional models, trained to recognize a fixed set of object classes, and prompt-based (PB) models, which integrate Large Language Models (LLMs) to support zero-shot detection. This means they can identify objects specified by a non-expert user through textual prompts, even if those objects were not present in the training dataset. This capability is valuable in remote sensing contexts, allowing for greater flexibility and reducing dependence on task-specific training where annotated datasets are scarce.

The proposed study presents an analytical evaluation of both categories, focusing mainly on PB models. Experiments are conducted on UAV imagery in OD applications - within the scope of the EU-funded “EMERITUS” Horizon Europe project focused on fighting environmental crimes - and on open Sentinel-2 satellite imagery in both OD and Semantic Segmentation tasks for land cover classification.

With the support of:



N|V|5



XII AIT INTERNATIONAL
CONFERENCE

12 – 14 NOVEMBER 2025 | MILAN

ISBN 979-12-985355-1-0