AI-Driven Detection of Unauthorized Buildings to Protect Hamlet Heritage in Fragile Territories in Italy

Marco Seccaroni[1], Domenico D'Uva[1]

[1] Dept. of Architecture, Built Environment, and Construction Engineering, Politecnico di Milano
Via Giuseppe Ponzio, 31, 20133 Milano - (marco.seccaroni, domenico.duva)@polimi.it

Keywords: AI-assisted Segmentation, Cadastral Discrepancy, Unauthorized Buildings, Heritage Documentation, GIS-based Analysis, Segment Anything Model.

Abstract

Unauthorised construction continues to endanger the integrity of historic hamlets and vernacular heritage in fragile Italian territories. This research presents an automated methodology for detecting unregistered buildings through the integration of cadastral datasets, high-resolution orthophotos, and the Segment Anything Model (SAM), a foundation model for image segmentation. By prompting SAM with cadastral centroids, building footprints were extracted and compared to official records to identify spatial discrepancies and undocumented structures. The method proved effective in detecting both geometric anomalies and potential building code violations. Despite some limitations related to material reflectivity and complex roof morphologies, the workflow is reproducible, scalable, and open source. By combining AI-assisted segmentation and GIS-based spatial analysis, the approach contributes to the development of digital tools for heritage documentation, territorial monitoring, and planning control.

1. Introduction

The preservation of Italy's historic built environment, particularly its rural hamlets, is increasingly threatened by two interrelated phenomena: unauthorised construction and unsustainable land consumption. These phenomena compromise the integrity of urban fabrics that have developed through incremental and historically coherent processes. Within these contexts, numerous building insertions have been carried out in violation of planning regulations and often without adherence to traditional typologies.

The issue of building abuse is particularly pervasive. Although slightly decreasing at the national scale (CRESME, 2024), the incidence of unauthorised construction remains alarmingly high in southern regions. Recent data report that Calabria and Basilicata exhibit rates of unauthorised building reaching 54.1%, followed by Campania (50.4%) and Sicily (48.2%) (Corrado et al., 2024). These figures contrast sharply with the national average of 15.1%, and with northern regions such as Trentino-Alto Adige, where enforcement mechanisms and a consolidated culture of legality have produced significantly lower incidence rates.

In parallel, Italy continues to face an unsustainable rate of land consumption, exacerbated by successive legislative amnesties that have weakened planning enforcement. According to the Italian Institute for Environmental Protection and Research, land take is advancing at an average rate of 20 hectares per day (ISPRA, 2023), a trend that is particularly detrimental in a country characterised by limited territorial extension and high environmental and cultural vulnerability. In this context, the detection and documentation of unauthorised constructions are critical. State-of-the-art methods rely increasingly on geospatial technologies. Manual inspection of GIS data involves the comparison of high-resolution satellite imagery with cadastral datasets to identify discrepancies and verify undocumented buildings (Li et al., 2020; not in the reference list). The proliferation of unlawful constructions poses significant threats to protected areas, including historical hamlets and archaeological zones, leading to the irreversible alteration of cultural landscapes and systematic violation of urban planning frameworks.

Consequently, there has been a growing turn toward digital and remote-sensing techniques for monitoring and early detection of unauthorized structures. Recent advances have greatly improved the ability to automatically identify recent unauthorized building modifications using spatial data. Continuous monitoring via satellite or aerial imagery, combined with modern image analysis, enables timely intervention against illegal urban sprawl and encroachments on protected habitats. In fact, early identification of illegal constructions from satellite imagery can prevent the deterioration of cultural heritage sites by allowing prompt countermeasures; often ground inspections are infeasible, so remote sensing provides a feasible alternative for accurate information. International agencies and researchers alike recognise that remote observation is a valuable tool to catch suspicious building changes in areas where illicit construction is common (Jovanović et al., 2021).

Remote sensing provides the foundational data for most digital detection of unlawful construction. Very-high-resolution (VHR) optical satellite images (and aerial photographs or UAV imagery) can reveal fine details of the built environment, enabling analysts or algorithms to spot new structures that were not present before. By comparing current imagery with existing geographic information system (GIS) data, such as cadastral maps or approved development plans, one can flag discrepancies that indicate unpermitted buildings. For example, Varol, Gürbüz, and Alparslan (2019) used a combination of airborne LiDAR point clouds, stereo satellite images, and official development plan maps to successfully identify illegal constructions in an urban district.

Before the machine learning era, many detection workflows relied on rule-based algorithms and expert-defined criteria to isolate potential illegal structures. In these approaches, domain knowledge is encoded as logical rules or thresholds applied to imagery. For instance, analysts might segment an image into objects (using techniques like multi-resolution segmentation) and then apply rule sets based on shape, size, spectral signature, or context (e.g. proximity to roads) to classify which segments are buildings. This object-based image analysis (OBIA) can be

https://doi.org/10.5194/isprs-archives-XLVIII-M-9-2025-1349-2025

effective in separating man-made structures from natural features using human-crafted rules (Hossain & Chen, 2019).

OBIA frameworks rose to prominence in the 2000s and 2010s to incorporate both spectral and spatial information in classification. The chief limitation is the need for expert input to craft rules or configure segmentation parameters, which can be time-consuming. Nonetheless, in scenarios with limited training data or very specific local building characteristics (as is often the case in heritage villages), a carefully tuned OBIA system can be a practical solution (Hossain & Chen, 2019).

The advent of machine learning has brought classification techniques based on large datasets. In the context of remote sensing, these methods became popular in the 2010s for tasks such as building detection. These approaches have been used to map urban growth and can be adapted to flag new constructions by comparing classification results over time. One advantage is that they can work with relatively small training sets and are computationally lighter than deep neural networks. In heritage contexts, where illegal constructions might have distinct colour or texture, even simple classifiers can sometimes pick up these differences if properly trained.

However, the limitations of traditional ML are increasingly apparent. These models rely heavily on the quality of input features; a poor choice of features can limit performance. They generally cannot match the accuracy of modern deep learning on complex image recognition tasks, especially when the scene is heterogeneous.

Deep learning, particularly convolutional neural networks (CNNs), has revolutionized the detection of buildings from remote sensing imagery in the past five years. CNN-based models automatically learn hierarchical features directly from raw pixel data. These models have demonstrated a step-change in accuracy, often detecting even small structures with high reliability, given sufficient training data. For instance, Ostankovich and Afanasyev (2018) used a pre-trained GoogLeNet CNN to classify satellite image patches and cross-referenced the outputs with cadastral maps. Similarly, Liu et al. (2024) proposed a lightweight detector derived from YOLO, called YEMNet, tailored to capture illegal construction objects in real time.

Unlawful construction detection is inherently a change detection problem: one needs to find new structures (or expansions of existing ones) that have appeared over time without authorization. Multi-temporal analysis techniques leverage imagery from multiple dates to highlight changes in the built environment. Traditionally, change detection in remote sensing involves methods like image differencing, change vector analysis, or principal component analysis. However, these methods are sensitive to noise and require careful calibration (Holail et al., 2025).

Recent studies have proposed deep learning models that take images from two or more dates and directly output a change map. These networks can ignore irrelevant differences and focus on meaningful changes. Lu et al. (2023) proposed a sliding-window cross-attention mechanism for near-real-time building change detection, achieving high accuracy even for small constructions.

For the purpose of this work, a methodology whose flexibility and reliability have been demonstrated has been chosen: Meta's Segment Anything Model (SAM). It is a general-purpose segmentation model trained on 11 million images with over 1 billion masks. SAM has shown strong zero-shot generalization to geospatial domains, enabling effective segmentation of features such as buildings in aerial and satellite imagery (Guo et al., 2024; Osco et al., 2023).

2. Methodology

Recent studies have demonstrated the effectiveness of the Segment Anything Model (SAM) for the automatic segmentation of complex objects from satellite and UAV imagery, particularly when guided by point prompts. These approaches have proven particularly effective in identifying urban structures and articulated forms in visually dense contexts, outperforming conventional supervised models in both segmentation accuracy and generalization capability (Guo et al., 2024; Liu et al., 2024).

In this study, SAM was used to automatically segment buildings visible in high-resolution orthophotos (Figure 1), with the aim of comparing the resulting building footprints with the official Italian cadastral map (Figure 2), in order to identify geometric discrepancies or unregistered constructions.

The orthophotos, georeferenced according to the national spatial reference system (EPSG: 6706), served as the raster base for analysis. The cadastral building footprints were obtained via the WMS service provided by the Italian Revenue Agency, which delivers an up-to-date vector layer of officially registered buildings.

For each cadastral polygon, the geometric centroid was calculated and used as a point prompt to guide the segmentation with SAM (Figure 3). A spatial validity check was performed to ensure that the centroid actually lay within the building polygon; in some cases, such as C-shaped buildings, the centroid may fall outside the structure, compromising segmentation quality. In such cases, the centroid was replaced with an alternative interior point to ensure correct prompt placement.

The interaction between georeferenced imagery, point prompts, and SAM was managed using the Samgeo Python package, an open-source tool specifically designed to apply SAM to geospatial raster data. Samgeo enables the generation of segmentation masks (Figure 4) that retain geographic reference consistency and can be exported in standard raster or vector GIS formats, facilitating integration into spatial analysis workflows (Wu & Osco, 2023).

The segmentation masks produced by SAM were initially post-processed to close internal holes and then filtered to remove areas smaller than 5 m² (Figure 5). This threshold is indicative and may vary by municipality, as the minimum surface area requiring building authorization depends on local regulations. The resulting masks were then compared with the original cadastral polygons using established spatial metrics. Based on this comparison, each case was classified into one of the following four categories:

- match (overlap between cadastral footprint and segmentation mask greater than 75%);
- discrepancy (overlap below 75%);
- unregistered building (present in imagery but absent in the cadastral map);
- non-existent building (present in the cadastral map but absent in imagery, typically considered a segmentation error).

This methodology enables a fully automated, scalable, and reproducible workflow for assessing consistency between the built environment and cadastral records. By integrating general-purpose AI-based segmentation with authoritative geospatial data, the approach supports operational use cases such as urban monitoring, cadastral verification, and territorial control.

A parametric filter was applied to exclude segmentations with a surface area below the minimum threshold established at the municipal level, in accordance with local regulations.

3. Result

The segmentation process based on SAM produced georeferenced building masks that, once post-processed, were compared with the official cadastral data of the selected study area (Figure 6). The case study was selected due to its numerous morpho-typological features, including isolated houses, aggregated buildings, rooftops with solar panels, roofs made from different materials, and a variety of building shapes ranging from simple and regular to complex and articulated structures. This diversity allowed for a robust assessment of the model's performance under heterogeneous conditions. The comparison enabled the verification of the model's reliability in detecting existing buildings and its potential use for cadastral verification and updating activities.

The analysis led to the identification of two buildings clearly visible in the orthophoto but not recorded in the cadastral database (Figure 7). These structures, consistent in shape, orientation, and materials with adjacent, officially registered buildings, lacked corresponding cadastral polygons. This suggests either unreported constructions or outdated cadastral records. Both buildings exhibit characteristics typical of permanent residential structures.

In addition to the unregistered buildings, several geometric anomalies were observed. Specifically, the SAM-generated masks sometimes showed greater alignment with the actual contours visible in the orthophoto than the cadastral polygons, which occasionally appeared simplified, misaligned, or partially incomplete. Some buildings were represented in the cadastral dataset with regular geometries that did not reflect their actual planimetric articulation (Figure 8).

Nevertheless, some limitations of the model also emerged. In several cases, rooftops equipped with photovoltaic panels were segmented inconsistently: the model tended to identify the panels as separate entities from the underlying roof, resulting in fragmented masks that split the building unit. Similarly, in the presence of heterogeneous roofing materials (e.g., part in terracotta tiles and part in light-coloured sheet metal within the same dwelling), SAM segmented the surfaces discontinuously, sometimes omitting the brighter or more reflective portions. This resulted in partial masks that did not correctly represent the entire building footprint.

Overall, the results confirm the potential of the model for large-scale urban applications, while highlighting the need for further refinements to effectively handle material variation and reflectivity in building roofs.

Figure 1. High-resolution orthophoto (20 cm/pixel) used as the base layer for automatic building segmentation and comparison with cadastral data.

Figure 2. Overlay of the high-resolution orthophoto with official cadastral building footprints (in orange), used to assess spatial correspondence and identify discrepancies.

Figure 3. Initial SAM segmentation output overlaid on the orthophoto. The masks still contain internal holes and unfiltered small segments.

Figure 4. Refined SAM segmentation used for comparison, with internal holes removed and small segments filtered out.

Figure 5. Elements identified by the SAM mask but absent from the cadastral map, highlighted in red as potential unregistered buildings.

Figure 6. Comparison between satellite imagery and cadastral footprints, illustrating alignment issues due to perspective distortion in off-nadir satellite acquisition

4. Conclusion

This study demonstrates that integrating generalist segmentation models such as SAM with official cadastral data constitutes an innovative and high-potential approach for the automated detection of spatial inconsistencies and unregistered buildings in urban environments. The method proved effective in identifying geometric anomalies and undocumented structures through a fully automated workflow, positioning itself as a scalable solution for territorial monitoring and urban documentation. These features are fully aligned with the objectives of CIPA, particularly concerning the application of artificial intelligence in the analysis, management, and digital documentation of the built heritage.

Despite the promising results, the method presents certain limitations. The accuracy of the segmentation masks is sensitive to building morphology and roof materials, especially in cases of reflective or heterogeneous rooftops. In such contexts, the traditional approach based on a single starting point may be insufficient. The adoption of a grid of distributed prompts within the building footprint enables more complete segmentation but increases the computational load, requiring further studies to identify an optimal balance between precision and efficiency.

Another critical aspect concerns the dependence on high-resolution nadir orthophotos, which are the only images capable of ensuring geometric consistency for comparisons. Satellite imagery, due to perspective distortions, is not suitable for precise comparisons, but it may still serve as a viable tool for the preliminary identification of unregistered buildings in large and hard-to-reach areas.

From the perspective of land management and heritage protection, the ability to automatically and periodically monitor the evolution of the built environment represents a strategic opportunity for urban planning, transformation control, and the prevention of building code violations, including in protected areas. Integration with orthophoto time series and optimization of input mechanisms are promising development directions to improve the robustness of the approach. Since all data used are vector-based and georeferenced, comparison between different time frames is technically straightforward and immediate: the only operational constraint lies in the periodic availability of updated orthophotos.

In conclusion, the study highlights the value of integrating open-source artificial intelligence tools with institutional geospatial data for the digitalization and governance of the built environment. If further developed, the proposed method could serve as an effective operational support tool for public authorities, researchers, and professionals involved in the documentation and sustainable management of the built environment.

Figure 7. Comparison between SAM segmentation (in red) and cadastral footprints (in black), highlighting spatial agreements and geometric discrepancies.

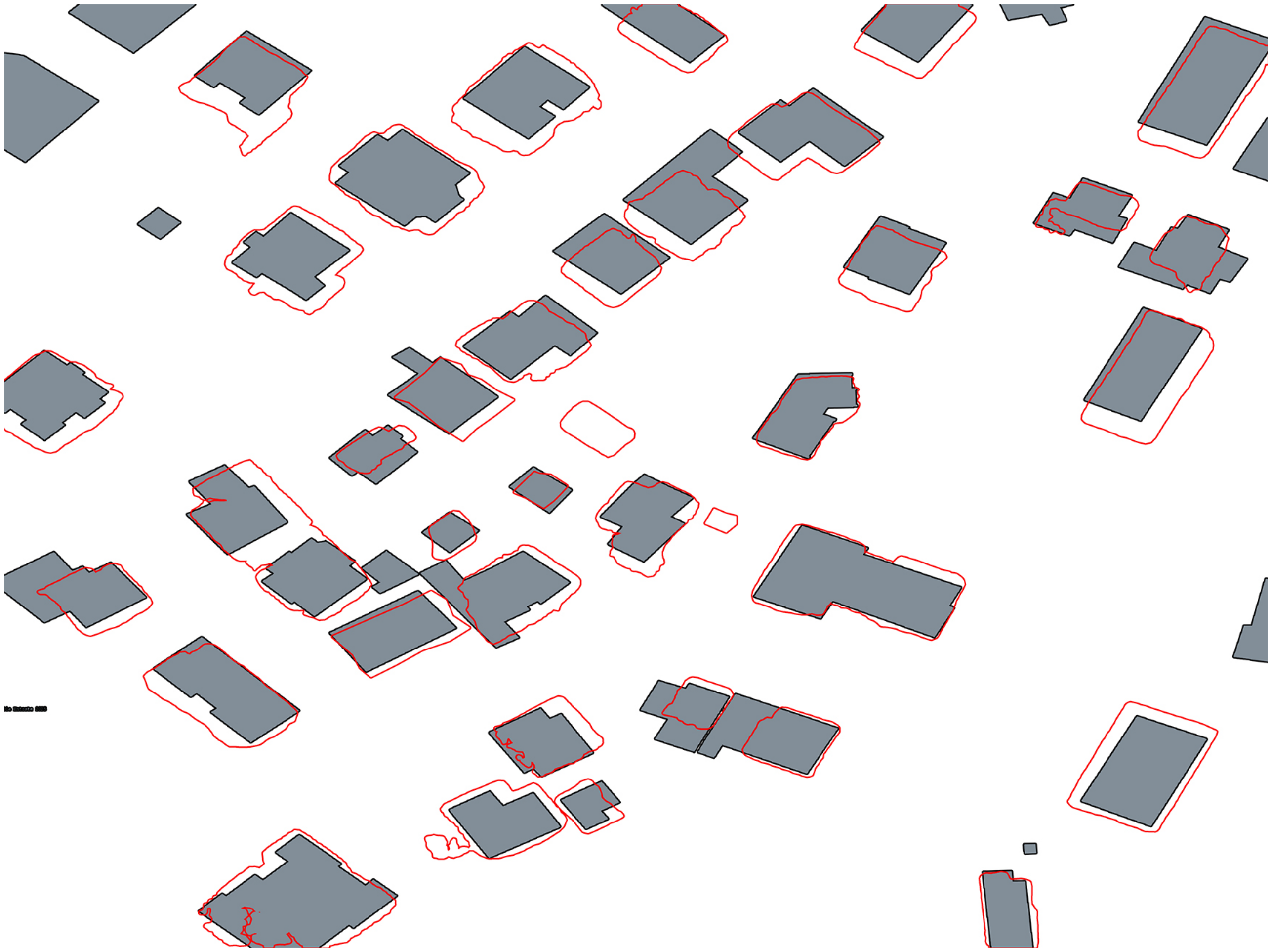

Figure 8. Comparison between SAM segmentation (in red) and cadastral footprints (in black).
The background is a cadastral map extract to enhance visibility.

References

Corrado, R., Trogu, D., Loru, M., & Nonnis, O. (2024). Illegal post-wildfire constructions detected using satellite time series: A case study in Sardinia, Italy. *Remote Sensing of Environment*, 297, 113645. https://doi.org/10.1016/j.rse.2023.113645

Guo, C., Hao, K., & Zuo, C. (2024). A Few-Shot Road Extraction Method Using Customized Segment Anything Model. 2024 9th International Conference on Intelligent Computing and Signal Processing (ICSP), 1036–1039. https://doi.org/10.1109/ICSP62122.2024.10743708

Guo, C., Hao, K., & Zuo, C. (2024). A few-shot road extraction method using customized Segment Anything Model. In 2024 9th International Conference on Intelligent Computing and Signal Processing (ICSP) (pp. 1036–1039). IEEE. https://doi.org/10.1109/ICSP62122.2024.10743708

Holail, H., Ismail, M. A., & Aly, H. A. (2025). Edge-CVT: A Siamese edge-guided transformer for building change detection. *ISPRS Journal of Photogrammetry and Remote Sensing*, 205, 45–58. https://doi.org/10.1016/j.isprsjprs.2024.12.010

Hossain, M. D., & Chen, D. (2019). Segmentation for object-based image analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150, 115–134. https://doi.org/10.1016/j.isprsjprs.2019.02.009

Jovanović, D., Gavrilović, M., Sladić, D., Radulović, A., & Govedarica, M. (2021). Building change detection method to support register of identified changes on buildings. *Remote Sensing*, 13(16), 3150. https://doi.org/10.3390/rs13163150

Liu, N., Xu, X., Su, Y., Zhang, H., & Li, H.-C. (2025). PointSAM: Pointly-Supervised Segment Anything Model for Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, 63, 1–15. https://doi.org/10.1109/TGRS.2025.3529031

Liu, W., Zhou, L., Zhang, S., Luo, N., & Xu, M. (2024). A new high-precision and lightweight detection model for illegal construction objects based on deep learning. *Tsinghua Science and Technology*, 29(4), 1002–1022. https://doi.org/10.26599/TST.2023.9010090

Lu, W. et al. (2023). Sliding-window dissimilarity cross-attention for near-real-time building change detection. *Remote Sensing*, 17(1): 135. https://doi.org/10.3390/rs17010135

Osco, L. P., Wu, Q., de Lemos, E. L., Gonçalves, W. N., Ramos, A. P. M., Li, J., & Marcato Junior, J. (2023). The Segment Anything Model (SAM) for remote sensing applications: From zero to one shot. *International Journal of Applied Earth Observation and Geoinformation*, 124, Article 103540. https://doi.org/10.1016/j.jag.2023.103540

Ostankovich, V. & Afanasyev, I. (2018). Illegal buildings detection from satellite images using GoogLeNet and cadastral maps. In Proc. 2018 Int. Conf. on Intelligent Systems (IEEE). https://doi.org/10.1109/IS.2018.8710565

Varol, B., Yılmaz, E. Ö., Maktav, D., Bayburt, S., & Gürdal, S. (2019). Detection of illegal constructions in urban cities: Comparing LiDAR data and stereo KOMPSAT-3 images with development plans. *European Journal of Remote Sensing*, 52(1),335–344.https://doi.org/10.1080/22797254.2019.1604082

Varol, B., Gürbüz, S. Z., & Alparslan, A. (2019). Detection of illegal constructions in urban cities: Comparing LiDAR data and stereo imagery with development plans. European Journal of Remote Sensing, 52(sup1), 41–50. https://doi.org/10.1080/22797254.2019.1585351

Wu, Q., & Osco, L. (2023). samgeo: A Python package for segmenting geospatial data with the Segment Anything Model (SAM). Journal of Open Source Software, 8(89), 5663. https://doi.org/10.21105/joss.05663

CRESME. (2024). Rapporto congiunturale e previsionale sull'industria delle costruzioni. Roma: CRESME.

ISPRA. (2023). Rapporto sul consumo di suolo in Italia 2023. Roma: Istituto Superiore per la Protezione e la Ricerca Ambientale.