The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-1/W5-2025
13th International Conference on Mobile Mapping Technology (MMT 2025)
"Mobile Mapping for Autonomous Systems and Spatial Intelligence", 20–22 June 2025, Xiamen, China

Implementation of an automated georeferencing workflow for architectural elements in GIS using ML and Cloud Computing

Elisabetta Doria[1], Luca Carcano[2]

[1] DICAr, Department of Civil Engineering and Architecture, University of Pavia – Pavia, Italy – elisabetta.doria@unipv.it
[2] luca.carcano@gmail.com

Keywords: Urban maintenance, 3D GIS, 3D models, Deep Learning, Cloud Computing, Drones.

Abstract

This research presents a scalable, cloud-based workflow integrating Machine Learning (ML) and 3D Geographic Information Systems (GIS) to support the automated detection of architectural elements and urban management. Via Unmanned Aerial Vehicle (UAV) georeferenced images, the system enables an automated and scheduled detection, geolocation, and import of architectural elements (e.g., domes, photovoltaics panels, tanks) data and metadata into a 3D GIS environment. A validated urban case study was conducted using UAV-acquired georeferenced images processed through a Structure-from-Motion (SfM) pipeline. Orthoimage chunks and dataset were uploaded to Google Cloud Storage, triggering an event-driven architecture built on a Cloud Computing Infrastructure. The pipeline leverages Vertex AI object detection via AutoML, the predictions of which are subsequently enriched with geospatial metadata. The output data is stored in BigQuery and Cloud Storage for urban GIS integration and analysis. Results confirm the viability of the pipeline for repeatable, and automated urban monitoring, reducing manual labour and improving safety for building maintenance workers. This approach is focused on the use of mobile mapping data processing, 3D reconstruction of urban areas, AI process for detection and urban maintenance and to develop smart city applications.

1. Introduction

The management and preservation of urban built heritage, particularly in historical centres, require innovative approaches combining traditional architectural knowledge and advanced technological tools. The integration of Cloud Computing, Machine Learning (ML), and 3D Geographic Information Systems (GIS) offers pathways to the automation of the detection and monitoring of architectural elements. This research develops a scalable, cloud-based workflow that leverages ML and 3D GIS to support the automated detection of architectural elements, such as roof construction technologies and hydraulic systems (e.g. water tanks), using UAV georeferenced images. The aim of the research is to integrate Cloud Computing, in particular Machine Learning platforms, with 3D GIS platforms to support planned maintenance and urban management with automated object detection and post processing of geospatial data (Doria et al, 2022). Namely, the goal is to automatically identify technical urban elements (such as domes, tanks etc.) using Object Detection from UAV images, infer and export the coordinates of such objects and import them in a 3D GIS. Repeatability of UAV flight plans, and the automated recognition of elements according to a fixed methodology, allow for a scheduled update of the GIS. By automating the identification, geolocation, and integration of these elements into a 3D GIS environment, the system facilitates scheduled updates and inspections, reducing manual labour and improving safety for maintenance personnel.

2. Pipeline overview and aims

The process was tested on a case study providing a large dataset of UAV images acquired for the construction of an existing and validated urban SfM model and a 3D GIS model, that is already deployed and in use (Parrinello and Picchio, 2019; Doria, 2022).

Georeferenced aerial images of roofs were acquired with UAVs, and used to build Structure from Motion (SfM) models and integrating dimensional data using a TLS survey as reference. Ortho images were derived from SfM, by dividing in multiple chunks, uploaded to an enterprise cloud platform (Google Cloud Platform, GCP) and fed to a computer vision model, previously trained to identify a set of architectural elements. The dimensions of image chunks were chosen according to platform, setting a fixed pixel size. When scaling the use case, this process leads to a variable number of chunks, according to the spatial extent of the starting SfM model. The coordinates of each architectural element recognised by the algorithm were estimated and stored in a format the GIS 3D models can handle. The goal is to develop an automated workflow for monitoring, structured as follows: acquisition of roofs images with flight-plane UAV instruments; processing of orthoimages; automatic recognition of architecture elements for maintenance purposes; inference of the coordinates of each identified element; import of the coordinates into a GIS; repetition of the process for each new acquisition and scheduled inspection.

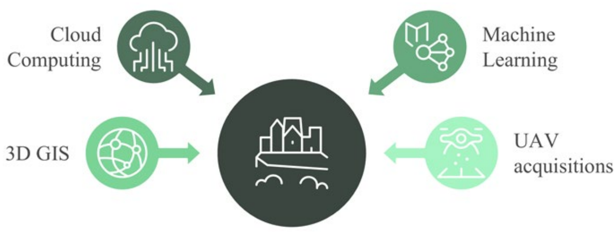

Figure 1. Automated urban Heritage management systems

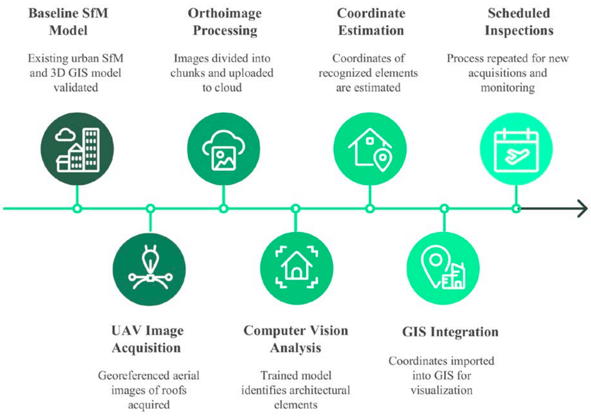

Figure 2. Workflow for the automated roof/urban area monitoring developed in the research

3. State of the art

The analytical processes aimed at understanding urban contexts have increasingly relied on the development of integrated and reliable databases, for structured data (e.g., tabular data and numbers), and data stores for unstructured data (e.g., natural language, images,). These data sets facilitate strategies for managing built heritage and enable the communication and virtual experience of digitised spaces to a wide audit (Miceli et al, 2020; Morandotti and Doria, 2023; De Marco and Bocconcino, 2025). In historic city centres, information about events that have modified the urban fabric and morphology must be preserved over time, forming a complex ensemble of data that necessitates digitisation for effective management (Bocconcino and Manzone, 2019). The surge in information involved in public decision-making processes has necessitated the adoption of digital technologies and techniques for information management. Information and Communication Technologies (ICT), rapidly evolving due to technological advancements, have significantly influenced decision-making frameworks in public management, offering advanced monitoring tools (Kouziokas and Perakis, 2017). To handle large amount of data and metadata, which are interconnected to create extensive datasets, advancements in data acquisition methods have simplified and expedited these processes with fast survey tools. However, data collection does not inherently constitute knowledge. A process of correlation and analysis is required to transform raw data into analysable insights. Computational systems can process vast amounts of unstructured information, identifying patterns and correlations that may not be immediately apparent, thereby generating new knowledge for the data analysis technician. Artificial Intelligence (AI) (McCarthy, 2007) has become pivotal in automating cognitive tasks (Corea, 2019), processing and reinterpreting external stimuli based on active research trajectories (Buratti et al, 2021). ML, Neural networks and Deep Learning (DL) models have enabled the management of increasingly complex use cases by scaling computational complexity and network interconnections (Mishra, 2021; Mishra et al 2024). The scalability of DL models allows them to process growing amounts of both structured and unstructured datasets. Computer Vision intersects with AI in applications where learning from unstructured data addresses complex problems like segmentation and object detection.

The automated identification of technological elements in urban environments, as proposed in this research, leverages object detection (OD). OD is a subset of computer vision that develops and builds techniques to identify and classify objects within images; Convolutional Neural Networks (CNNs), since their inception (LeCun and Bengio, 1998), have become the state-of-the-art approach in this domain due to their high performance (Redmon et al., 2016). This study explores AutoML, a suite of models that utilise ensemble and transfer learning techniques to expand the range of identifiable objects. The availability of AutoML models in Cloud Computing Platforms, often in code-less solutions, expedite the development of baseline ML models, promoting the democratization of AI (He et al, 2021). Managing the potentially increasing size of photographic dataset and the consequent computational complexity of object detection imposes significant constraints. To overcome the limitations of on-premises data management, this research builds on top of a Cloud Computing infrastructure, leveraging Data Storage, the Integrated Machine Learning Platform and the seamless integration offered across the suite of software provided by the cloud vendor. This approach abstracts from the computational power available locally, avoiding the constraints imposed by physical hardware and infrastructures (Rivera, 2020).

3.1 Related researches

Object Detection techniques have been developed to identify entities based on existing data, enabling predictions on new, unseen data. From image data, three primary objectives emerge: classification, detection, and segmentation (Xiao et al., 2020; La Placa and Doria, 2024).

Figure 3. Three-dimensional model obtained with SfM from UAV of the entire historic centre of the case study. The model, georeferenced both by the photographs' own metadata and by the use of external GPS, was used for the export of orthoimages for the detection of the elements of interest in the roofs. The database was realised according to “3D Bethlehem” research (Parrinello and Picchio, 2019)

These approaches have been extensively documented as support tools for managing and documenting cultural heritage and urban planning (Li et al., 2020). In urban environments, such techniques have been applied to monitor infrastructure criticalities and construction sites (Dandabathula et al., 2019; Wang et al., 2022), manage urban waste (Majchrowska et al., 2022), enforce compliance with building regulations (Cunha et al., 2021), and conduct environmental and agricultural monitoring (Carbone et al., 2022). High-performance applications of these techniques are also documented in the field of built cultural heritage, aiding in the identification and close-range imaging (Due Trier et al, 2021) and supporting decision-making for interventions on documented heritage (Monna et al., 2021; Zou et al., 2019; Hatir and Ince, 2021). Machine learning models and pipelines, based on open-source technologies such as Python libraries and Kubeflow, can be exported and shared with collaborators outside the platform. Recent studies have demonstrated the efficacy of integrating AI and computer vision in urban analysis.

For instance, the URBAN-i model utilizes deep learning and computer vision to detect informality and slums in urban scenes, aiding in urban modelling and understanding city dynamics (Ibrahim et al, 2018).

Similarly, the UV-SAM model adapts the Segment Anything Model for urban village segmentation, providing accurate boundary identification from satellite images (Zhang et al., 2024). This process can facilitate high degrees of cooperation across multiple entities, including universities, companies, and municipalities. Such collaboration can be organised during both the development and final usage phases, with pre-trained models accessible via dedicated Application Programming Interfaces (APIs).

4. Methodologies, tools and technologies

As described in the case study of Bethlehem’s historic centre (Doria, 2022), the automation of the technological census of urban architectural elements (e.g., antennas, tanks, domes) is vital for efficient monitoring and planning. The research proposes an operational model combining UAV-based image acquisition (Parrinello and Picchio, 2019) with automated classification techniques to identify urban infrastructure elements and integrate them into structured, usable datasets. This approach avoids labour-intensive manual mapping, offering a model applicable in other urban contexts.

However, one of the major constraints in earlier approaches, such as those involving on-premises processing and storage, is the limited scalability of both computing power and dataset management. Eventarc, event management software, routes events to Cloud functions, forming the core of this event-driven computing architecture; as soon as UAV orthoimage chunks are uploaded to a Google Cloud Storage (GCS) bucket, Eventarc triggers the sequential steps of the pipeline without manual intervention.

This not only reduces the time between image acquisition and data availability for analysis, but automates the interaction between different phases of the workload.

The processing pipeline continues with Vertex AI, the integrated ML platform of GCP, that allows for the training, deployment, and prediction (both batch and real time) of deep learning models, including AutoML models designed specifically for image object detection. A model (previously trained on urban datasets) processes the image files, outputting bounding box coordinates and confidence scores in a JSON Lines format; these predictions are stored temporarily in GCS. The next step, geospatial enrichment, links detected urban elements to their actual geolocation. A secondary Cloud function reads the original orthoimages and associates detection results with spatial metadata extracted from the images themselves (e.g., camera orientation, UAV altitude). This enriched data becomes useful not only for inventorying urban assets, but also for the integration with 3D GIS platforms; the enriched results are persisted in Big Query (Google Cloud’s data warehouse solution), which supports both manual and programmatic queries.

The output datasets can be visualized in GIS environments, shared with municipal stakeholders and collaborators for further analysis. Recent studies have demonstrated the effectiveness of ML and DL approaches in similar contexts. For instance, Fiorucci et al. (2022) utilized DL for archaeological object detection on LiDAR data, providing new evaluation measures and insights. Similarly, Jiang et al. (2023) developed a model for infrastructure detection along highways based on remote sensing images from UAVs,

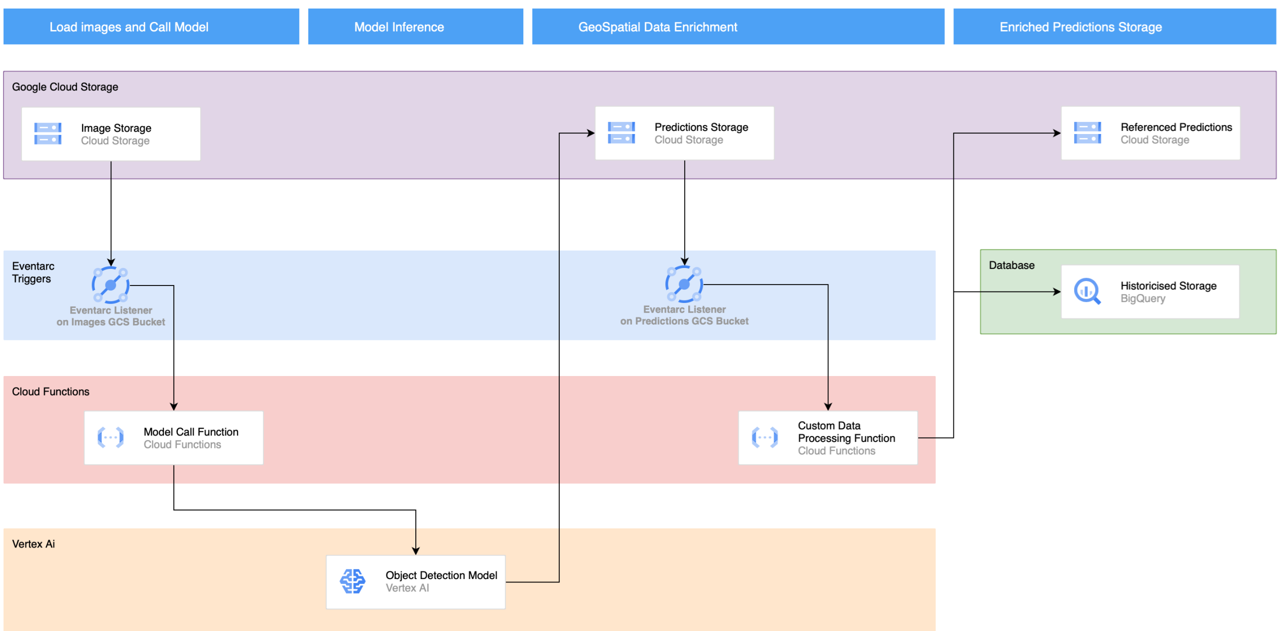

Figure 4. High level design of the underlying Cloud Architecture

showcasing the applicability of such techniques in infrastructure monitoring. In the sector of built heritage conservation, ML techniques have been applied for monitoring and planning purposes. A study by Tucci et al. (2024) presented a fully automated approach for the segmentation and classification of architectural elements, aiming to provide tools for a more detailed assessment of the risk of detachment of parts of the sandstone elements.
The different phases are described in the following subparagraph in more details:

4.1 Model Inference

The first phase is that of Data Ingestion and starts with the upload of ortho image chunks to a designated Google Cloud Storage (GCS) bucket. This triggers Eventarc, Google Cloud's solution for building event-driven architectures (Stopford, 2018). Upon detection, the event is routed to a Cloud Function responsible for initiating the processing pipeline (Lakshmanan, 2022). This function constructs a request file containing the image storage URI on GCS, as well as the Vertex AI model name and version to be used for inference. The Cloud Function then triggers a batch prediction job on Vertex AI, a programmatic way to get predictions from a model deployed on the Model Registry.

Figure 5. Preliminary images of the detection process. Above, raw image from drone with integrated location metadata thanks to the GPS systems of the aircraft used. Chuncks for SfM photogrammetry were made for urban districts and georeferenced according to GPS targets. Object detection training took place on the raw imams, and recognition and identification was initiated on the orthoimages to make the object georeferencing processes effective This process also makes it possible to develop systems of scheduled surface inspections, planned over time.

The batch prediction request is executed using the Vertex AI Python SDK. Vertex AI hosts an AutoML deep learning model (He, 2021) designed for computer vision tasks and trained on UAV images to identify the architectural elements of this study. Upon execution, the model processes the image files and outputs predictions in the form of a JSON Lines (JSONL) file, a standard interchange format used across multiple components of Vertex AI. Each record includes the list of detected objects, bounding box coordinates, and a confidence score corresponding per each class. As a managed service, Vertex AI ensures high scalability, allowing to adjust the number of nodes according to workload demand. The resulting predictions are written to a designated Prediction Storage Bucket on GCS.

4.2 Geospatial Data Enrichment

To enrich the predictions with geospatial context, another Eventarc trigger monitors the Prediction Storage Bucket. When new output files are detected, this trigger activates a secondary Cloud Function that performs geospatial data enrichment. This function loads both the original ortho images and their associated predictions, and enriches the detection data by linking each object to its geographic coordinates. This is achieved by extracting metadata embedded in the source images, which typically includes location, scale, and orientation parameters. The enriched data is then exported into a structured dataset, enabling downstream spatial analyses and integration with GIS systems.

4.3 Enriched Predictions Storage

The final stage involves persisting enriched data for long-term access and analysis. Results are written to BigQuery, enabling querying, analytics, and programmatic data access. Additionally, both raw predictions and enriched, georeferenced outputs are stored in Cloud Storage. Results are historicised on the data warehouse, associating the detected objects with a timestamp, referencing the presence of objects in a specific moment therefore allowing to detect additions and subtractions to built elements. This dual storage strategy ensures flexibility, supporting both tabular exploration via BigQuery and spatial visualization in GIS platforms or custom visualization tools.

5. Results

The implemented workflow demonstrates a robust and scalable system for the automated detection and georeferencing of architectural elements using UAV imagery, cloud computing, and machine learning. During testing on the historic centre of Bethlehem, the system achieved over 85% average precision (AP) in detecting and localizing objects such as water tanks, and antennas across a heterogeneous urban morphology. Confidence thresholds could be flexibly adjusted to prioritize either recall (for more exhaustive identification) or precision (for cleaner detections), depending on the maintenance strategy.
Batch inference using Vertex AI significantly reduced processing times compared to traditional desktop or on-premise methods, completing prediction cycles for hundreds of orthoimage chunks within hours. Moreover, the event-driven architecture allows fully automated processing once new images are uploaded, requiring no human intervention from the upload of images to the database storage. This allows to create a feedback loop for continuous monitoring, since historical predictions are available for comparison and

enrichment through time, highlighting modifications in the built environment.

The pipeline effectively maintained metadata integrity, linking object IDs with geospatial coordinates, timestamp, and detection confidence, which can be queried or visualized within BigQuery or any GIS environment.

Furthermore, the modularity of the pipeline allows users to retrain models with additional classes or fine-tune the system to different urban contexts.

Key highlights of the results include:

- Training process is done once, then model is deployed and leveraged;
- Time consuming tasks are automated, allowing human technicians to focus on more specialised tasks;
- Model performance is monitored, model can be retrained in case of performance degradation over time or the necessity to detect new unseen elements;
- The system processed 150 images in 4 hours; batch prediction jobs can be parametrised to run on larger machines or a higher number of computing nodes, to comply with heavier workloads or more strict time requirements;
- GIS integration: Seamless import of spatially enriched detection data into 3D GIS environments.

Feedback from stakeholders (urban planners, engineers, heritage conservationists) highlighted increased usability and reduced dependency on specialist operators, enabling broader adoption within interdisciplinary teams.

6. Conclusions

The study successfully demonstrates a repeatable and scalable pipeline for the automated georeferencing of architectural elements using UAV imagery and AI. The research not only confirms the feasibility of integrating machine learning with 3D GIS but also shows significant potential for improving urban asset management by enabling data-driven inspections, proactive maintenance, and enhanced heritage documentation. This methodology empowers municipalities and technicians with an automated and adaptable tool, fostering a proactive approach to urban maintenance. The cloud-native design supports event-driven scalability, allowing the same system to be applied to both small pilot studies and extensive urban areas without significant reconfiguration.

The workflow's modular nature means it can be extended in several directions:

- Detection class expansion training models on new urban or infrastructural elements (e.g., heat systems, materials, chimneys, photovoltaics panels);
- Temporal analysis using historical UAV datasets to detect changes over time and flag anomalies or maintenance needs creating a map of risk;
- Semantic enrichment linking detected elements to building IDs, maintenance records, or regulatory compliance databases for more integrated lifecycle and urban management.

In the longer term, this work sets a starting point for smart city applications based on automated spatial intelligence. The approach is particularly relevant for historic city centres, where maintenance must be balanced with preservation, and where safety concerns limit physical inspections. This workflow contributes to both theoretical research in digital urban and built maintenance and practical advances in the fields of architectural conservation and infrastructure monitoring.

Future development may include user-friendly dashboards, broader dataset interoperability, and integration with municipal decision-support systems.

The process reinforces the idea that AI and cloud-based geospatial workflows are not just technical solutions, but enablers of a new paradigm in urban governance.

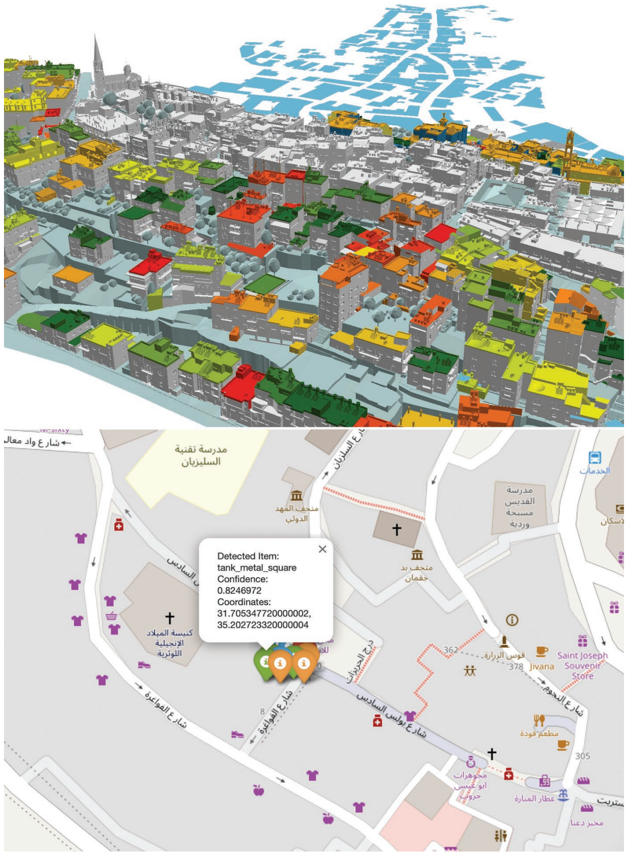

Figure 6. Above, 3D GIS model created during the 3D Bethlehem project (Doria et al, 2022). The model was used to validate the location extracted from the orthoimages of the elements (tanks, below) automatically identified with object detection. The location thus obtained can be integrated into GIS to develop monitoring systems over time.

Acknowledgements

The research was conducted thanks to the dataset acquired during the International Cooperation project "Management and control of urban growth for the development of heritage and improvement of life in the city of Bethlehem", called "3D Bethlehem". It is a cooperation project co-financed by the Italian Agency for Cooperation and Development (AICS). The project was coordinated by the Municipality of Pavia, with a partnership composed of Municipality of Bethlehem, University of Pavia (scientific coordination), Bethlehem University, Province of Pavia, Pavia Province Order of Engineers, SISTERR Pavia Territorial System for International Cooperation, ANCI Lombardia, VIS International Voluntary Service for Development NGO and Engineers Association of Jerusalem Centre. The project was scientifically coordinated by Prof. Sandro Parrinello (Parrinello et al, 2024) and the research laboratory DAda-LAB of DICAr, Department of Civil Engineering and Architecture, University of Pavia.

Author Contributions: Conceptualization, E.D.; Methodology, E.D. and L.C.; Investigation, E.D.; Data Curation, E.D. and L.C., ; Software, L.C.; Formal Analysis: L.C.; Validation: E.D. and L.C.; Visualisation: E. D.; Writing – Original Draft Preparation, E.D.; Writing – Review & Editing, E.D.; Project Administration, E.D.

References

Bocconcino, M. M. and Manzone, F. (2019), Sistemi informativi e strumenti grafici per la manutenzione di manufatti complessi | Information systems and graphic tools for the maintenance of complex buildings, in Garda, E., Mele,C. and Piantanida, P. (eds), Colloqui.AT.e 2019 – Ingegno ecostruzione nell'epoca della complessità – Atti del Convegno – Torino, 25-28 Settembre 2019, Edizioni Politecnico di Torino, Torino, pp. 678-688.

Buratti, G., Conte, S. and Rossi, M. (2021), Artificial Intelligency, Big Data and Cultural Heritage, in Giordano, A., Russo, M. and Spallone, R. (eds), Representation challenges – Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain, Franco Angeli, Milano, pp. 29-33.

Carbone, C., Fawakherji, M., Trianni, V. and Nardi, D. (2022), Photorealistic simulations of crop fields for remote Sensing with UAV swarms, in Parrinello, S., Barba, S., Dell'Amico, A. and di Filippo, A. (eds), D-SITE. Drones – Systems of Information on Cultural Heritage for a spatial and social investigation, Pavia University Press, Pavia, pp. 576-581.

Corea, F. (2019), An Introduction to Data – Studies inBig Data, vol. 50, Springer, Cham. doi.org/10.1007/978-3-030-04468-8

Cunha, H. S., Sclauser, B. S., Wildemberg, P. F., Fernandes, E. A. M., dos Santos, J. A., de Oliveira Lage, M., Lorenz, C., Barbosa, G. L., Quintanilha, J. A. and Chiaravalloti-Neto, F. (2021), Water tank and swimming pool detection based on remote sensing and deep learning – Relationship with socioeconomic level and applications in dengue control, in PLoS ONE, vol. 16, issue 12, pp. 1-24. doi.org/10.1371/journal.pone.0258681

Dandabathula, G. B., Saini, S., Parikh, D., Sharma, P., Khandelwal, S. and Rao, S. S. (2019), Automatic detection of overhead water tanks from satellite images using faster-RCNN, in International Journal of Advanced Research in Computer Science, vol. 10, n. 5, pp. 34-37. [Online] Available at: dx.doi.org/10.26483/ijarcs.v10i5.6466

De Marco, R., Bocconcino, M. M. (2025). Digital Intelligences and Urban InfoSystems in Territorial Re-Education. In Encyclopedia of Information Science and Technology, Sixth Edition (pp. 1-36). IGI Global.

Doria, E. (2022), Automation of urban technological census. The historical centre of Bethlehem. AGATHÓN | International Journal of Architecture, Art and Design, 12(online), 178–189. https://doi.org/10.19229/2464-9309/12162022

Doria, E., La Placa, S., Miceli, A. (2022), Il GIS come strumento di analisi urbana per la gestione del Life Cycle della città/GIS as an urban analysis tool for the management of the city's Life Cycle. DISEGNO, RILIEVO E PROGETTAZIONE, 3, 138-165.

Due Trier, Ø., Reksten, J. H. and Løseth, K. (2021), Automated mapping of cultural heritage in Norway from airborne lidar data using faster R-CNN, in International Journal of Applied Earth Observation and Geoinformation, vol. 95, article 102241, pp. 1-19. doi.org/10.1016/j.jag.2020.102241

Fiorucci, M., Verschoof-van der Vaart, W.B., Soleni, P., Le Saux, B., and Traviglia, A. (2022), Deep Learning for Archaeological Object Detection on LiDAR: New Evaluation Measures and Insights. Remote Sensing, 14(7), 1694. https://doi.org/10.3390/rs14071694

Hatir, M. E. and İnce, I. (2021), Lithology mapping of stone heritage via state-of-the-art computer vision, in Journal of Building Engineering, vol. 34, article 101921, pp. 1-12. doi.org/10.1016/j.jobe.2020.101921

He, X., Zhao, K., Chu, X. (2021), AutoML: A survey of the stateof-the-art. In Knowledge-Based Systems, 212, 2021, 106622.

Ibrahim, M. R., Haworth, J., and Cheng, T. (2018), URBAN-i: From urban scenes to mapping slums, transport modes, and pedestrians in cities using deep learning and computer vision. arXiv preprint arXiv:1809.0360

Jiang, X., Cui, Q., Wang, C., Wang, F., Zhao, Y., Hou, Y., Zhuang, R., Mei, Y., and Shi, G. (2023), A Model for Infrastructure Detection along Highways Based on Remote Sensing Images from UAVs. Sensors, 23(8), 3847. https://doi.org/10.3390/s23083847

Kouziokas, G. N. and Perakis, K. (2017), Decision Sup-port System Based on Artificial Intelligence, GIS and Re-mote Sensing for Sustainable Public and Judicial Management, in European Journal of Sustainable Development, vol. 6, issue 3, pp. 397-404.

La Placa, S., Doria, E. (2024), digital documentation and fast census for monitoring the university's built heritage, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2/W4-2024, 271–278, https://doi.org/10.5194/isprs-archives-XLVIII-2-W4-2024-271-2024

Lakshmanan, V. (2022), Data Science on the Google Cloud Platform. O'Reilly Media, Inc.

Lecun, Y. and Bengio, Y. (1998), Convolutional net-works for Images, Speech, and Time-Series, in Arbib, M.A. (ed.), The Handbook of Brain Theory and Neural Net-works, MIT Press, pp. 255-258.

Li, K., Wan, G., Cheng, G., Meng, L. and Han, J. (2020), Object detection in optical remote sensing images – A survey and a new benchmark, in ISPRS Journal of Photogrammetry and Remote Sensing, vol. 159, pp. 296-307. doi.org/10.1016/j.isprsjprs.2019.11.023

Majchrowska, S., Mikołajczyk, A., Ferlin, M., Klawikowska, Z., Plantykow, M. A., Kwasigroch, A. and Majek, K. (2022), Deep learning-based waste detection in natural and urban environments, in Waste Management, vol. 138, pp. 274-284. doi.org/10.1016/j.wasman.2021.12.001

McCarthy, J. (2007), From here to human-level AI. Artificial Intelligence, 171(18), 1174-1182.

Miceli, A., Morandotti, M., & Parrinello, S. (2020). 3D survey and semantic analysis for the documentation of built heritage. The case study of Palazzo Centrale of Pavia

University. Vitruvio: International Journal of Architectural Technology & Sustainability, 5(1).

Mishra, M. (2021), Machine learning techniques for structural health monitoring of heritage buildings – A state-of-the-art review and case studies, in Journal of Cultural Heritage, vol. 47, pp. 227-245.

Mishra, M., Zhang, K., Mea, C., Barazzetti, L., Fassi, F., Fiorillo, F., & Previtali, M. (2024), Deep Learning-Based AI-Assisted Visual Inspection Systems for Historic Buildings and their Comparative Performance with ChatGPT-4O. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 327-334.

Monna, F., Rolland, T., Denaire, A., Navarro, N., Granjon, L., Barbé, R. and Chateau-Smith, C. (2021), Deep learning to detect built cultural heritage from satellite imagery – Spatial distribution and size of vernacular houses in Sumba, Indonesia, in Journal of Cultural Heritage, vol. 52, pp. 171-183. doi.org/10.1016/j.culher.2021.10.004

Morandotti, M., Doria, E. (2023), information system as tool for cultural heritage documentation and preservation. Protocol structuring and testing on a case study, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-M-2-2023, 1081–1088, https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-1081-2023, 2023.

Parrinello, S. and Picchio, F. (2019), Integration and comparison of close-range sfm methodologies for the analysis and the development of the historical city center of bethlehem, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W9, 589–595, https://doi.org/10.5194/isprs-archives-XLII-2-W9-589-2019

Parrinello, S., Picchio, F., La Placa, S. (2024). The Construction of an Informative 3D Model for the Monitoring of City Heritage Risk. In Reviving Aleppo: Urban, Legal and Digital Approaches for Post-War Recovery (pp. 243-274). Cham: Springer International Publishing.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016), You Only Look Once – Unified, Real-Time ObjectDetection, in IEEE Conference on Computer Vision andPattern Recognition (CVPR) – Las Vegas, NV, USA, June27-30, 2016, IEEE Computer Society, pp. 779-788.

Rivera, J. (2020), Object detection with a model trained in Google Cloud AutoML. In Practical TensorFlow.js. Berkeley: Apress, pp. 163-184.

Stopford, B. (2018), Designing event-driven systems. O'Reilly Media, Incorporated.

Tucci, G., Bonora, V., and Conti, A. (2024), Machine Learning-Based Monitoring for Planning Climate-Resilient Conservation of Built Heritage. Drones, 8(6), 249. https://doi.org/10.3390/drones8060249

Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S. and Lan, S. (2020), A review of object detection based on deep learning, in Multimedia Tools Application, vol. 79, pp. 23729-23791. doi.org/10.1007/s11042-020-08976-6.

Zhang, X., Liu, Y., Lin, Y., Liao, Q., and Li, Y. (2024), UV-SAM: Adapting Segment Anything Model for Urban Village Identification. arXiv preprint arXiv:2401.08083

Zou, Z., Zhao, X., Zhao, P., Qi, F. and Wang, N. (2019), CNN-based statistics and location estimation of missing components in routine inspection of historic buildings, in Journal of Cultural Heritage, vol. 38, pp. 221-230doi.org/10.1016/j.culher.2019.02

Appendix

Across the development of this project multiple open source libraries developed and maintained by the community were used, both for image manipulation and geospatial data analysis.

Open source for Geospatial Foundation:
https://www.osgeo.org
GDAL - Library used to manage geospatial data in Python:
https://pypi.org/project/GDAL/
OpenCV - Open source computer vision library:
https://pypi.org/project/opencv-python/
Folium – Open source library for visualising map data:
https://pypi.org/project/folium/