2025 IMEKO TC26 International Conference on
Metrology for Archaeology and Cultural Heritage
Bergamo, Italy, October 15-17, 2025

From Visual Perception to Automated Detection: Integrating NDVI, SCP, and AI for the Identification of Archaeological Surface Markers

Filomena Papaleo[1], Simone Giosuè Madeo[2],

[1] *Università degli Studi di Genova,* filomenapapaleo@gmail.com
[2] *Università degli Studi di Genova,* simonegiosue.madeo@gmail.com

***Abstract* – This paper presents a comparative analysis between traditional archaeological landscape interpretation and outputs derived from artificial intelligence and multispectral remote sensing. The study focuses on the middle Tanagro Valley, where NDVI (Normalized Difference Vegetation Index) and the Semi-Automatic Classification Plugin (SCP) for QGIS were applied to satellite and drone imagery to detect cropmarks and soilmarks indicative of buried structures. These results were compared with expert archaeological interpretations based on photointerpretation and historical data.**
The objective is to evaluate the predictive validity of automated processing versus manual interpretation and to assess the effectiveness of automated classification in reconstructing settlement patterns. Results highlight the potential of integrated approaches in enhancing archaeological detection, while also acknowledging risks of overfitting and false positives in complex landscapes.

Keywords: NDVI; Remote sensing; SCP; AI; Archaeological survey; Predictive modelling

I. INTRODUCTION

In recent years, landscape archaeology has increasingly integrated predictive technologies based on artificial intelligence (AI) to support and complement traditional survey methods [1]. This paper presents the results of an experimental study conducted in the middle Tanagro Valley (southern Italy), where outputs from machine learning algorithms – such as Random Forest models rained to recognize archaeological patterns were critically compared with data collected through traditional methods. This approach reflects a confrontation between two distinct epistemological paradigms: manual interpretation grounded in archaeological expertise versus automated detection based on computational routines.

Using multispectral imagery acquired from drones and high-resolution satellites, predictive models were implemented to suggest the potential location of archaeological sites, based on parameters such as NDVI (Normalized Difference Vegetation Index), variations in the green spectral channel, terrain morphology anomalies, and integration with digital terrain models (DTMs) [2].

The experiment focused on selected areas of the middle Tanagro Valley and aimed to evaluate the effectiveness of automated analysis of survival traces and vegetational anomalies in reconstructing settlement dynamics, with a particular emphasis on the role of vegetational patterns in detecting buried anthropogenic features.

Through the generation of thematic maps using NDVI and the Semi-Automatic Classification Plugin (SCP) and their integration into GIS workflows [3], [4], the paper critically discusses the shift from intuitive recognition by the archaeologist to automated segmentation and classification. The comparison revealed a partial but significant overlap between AI outputs and material evidence recorded in the field, as well as notable discrepancies that raise methodological questions about the validity of predictive models. In particular, machine learning models showed high sensitivity to environmental anomalies but limited ability to distinguish between natural and archaeological signals.

This case study emphasizes the interplay between prediction and empirical validation, and between algorithmic efficiency and contextual archaeological knowledge. The paper concludes with a proposal for a structured integration of both approaches within a broader methodological framework for landscape archaeology.

II. TECHNICAL WORKFLOW AND STUDY AREA DEFINITION

The research was conducted in the Vallo di Diano, an intramontane basin in southern Italy that has historically served as a natural corridor connecting the Tyrrhenian coast with the inland regions of Basilicata and Campania. The area is characterized by a rich and multilayered archaeological landscape, ranging from prehistoric cave occupations to Roman rural systems and medieval fortified sites. This variety of archaeological contexts made the Vallo di Diano an ideal testing ground for the integration of AI-based predictive models, remote sensing, and traditional archaeological interpretation.

Scholarly interest in the Vallo di Diano dates back to the late 19th century, when railway construction and early speleological explorations brought to light Iron Age necropolises, prehistoric cave sites, and Bronze Age settlements [5]. The necropolis of Sala Consilina, with its cremation and inhumation burials accompanied by rich grave goods, underscored the region's role in the socio-cultural dynamics of pre-Roman Italy. Later systematic excavations and surveys in Padula, Atena Lucana, and Sant'Arsenio confirmed the complexity of the local settlement systems, revealing both indigenous practices and external cultural influences, particularly from Magna Graecia.

For the purpose of this study, multiple areas within the Vallo di Diano were selected to evaluate the predictive effectiveness of AI-based analyses and to ensure the reliability of archaeological correlations. For the detection of survival traces, three specific areas were investigated:

1) Mattine (Auletta), where the remains of a Roman rural complex are preserved;
2) Tempio (Polla), with the remains of a funerary mausoleum and a broader necropolis sector that has remained largely unaffected by modern urban development;
3) Calvanello (San Rufo), where the ruins of a medieval castle are still visible and provide a well-defined stratigraphic context.

These sites were selected for their variation in size, elevation, and typology, providing a diverse dataset to support the training and evaluation of machine learning models. Moreover, the three contexts exhibit different levels of interpretability based on the legibility of survival traces in the imagery. As an exploratory step, we used a large multimodal language model (ChatGPT-5 Plus) to assess the detection of survival traces in orthophotos of known archaeological contexts [6], [7]. The outcomes were partially positive. In the case of the Roman villa at Auletta (Mattine), the model correctly recognized the structural typology within the orthophoto set and plausibly hypothesized its chronology; however, it only marginally succeeded in outlining the perimeter of the survival traces (Fig. 1).

Fig. 1. Mattine (Auletta, SA). AI-based detection of archaeological traces on 2006–2016 orthophotos (Google Earth): detected structures shown in red.

For the medieval remains in San Rufo and the Roman mausoleum in Polla, the model identified the archaeological nature of the former and inferred a plausible medieval chronology, but it did not recover the functional attribution; performance at Polla was weaker and did not yield a reliable delineation of the mausoleum.

These results underline that a general-purpose multimodal LLM (e.g., ChatGPT) can identify the archaeological character of some contexts, yet it struggles to delineate survival traces with cartographic precision. Future experiments with task-specific Convolutional Neural Networks (CNNs) trained on local orthophotos—for example, segmentation models such as Mask R-CNN or YOLO-Seg and detectors tailored to linear features—are likely to yield more satisfactory results. CNNs can be tuned to the task, generate masks/polygons readily usable in GIS, and support quantitative evaluation (precision/recall, IoU) with improved reproducibility. While they require an initial investment in annotation, we anticipate higher recall and boundary accuracy, and fewer false positives, than with a general-purpose LLM [8], [9].

In addition, two areas were selected for the identification of cropmarks and soilmarks: the first area, located between Teggiano and Sala Consilina, is well-known for the presence of a Roman centuriation system; the second one, between Sala Consilina, Padula, and Montesano sulla Marcellana, is characterized by the presence of numerous dispersed Roman rural sites (*villae rusticae*). This selection, spanning distinct archaeological and geomorphological profiles, was critical for assessing the generalizability of AI-assisted detection across diverse settlement types and environmental settings.

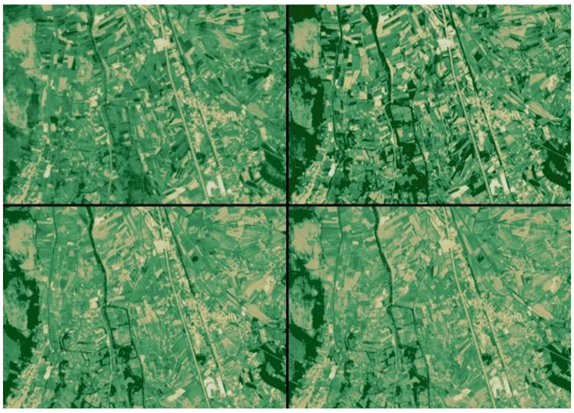

Fig. 2. Multi-temporal comparison of false-color (NDVI/near-infrared) satellite imagery for the rural and urban sectors of the study area (2019–2022). The composites highlight vegetation distribution and density (green), built-up and other anthropogenic surfaces (light tones), and water bodies or unvegetated areas (dark tones). Inter-annual chromatic variation reveals changes in land use, vegetation cover, and urban expansion, supporting environmental, landscape, and archaeological analyses.

The methodological framework relies on open-source tools, particularly QGIS 3.40.2, and combines visual interpretation, statistical classification, and geospatial modeling. NDVI (Normalized Difference Vegetation Index) was used as a key indicator to detect vegetative anomalies potentially linked to buried features [10], while the SCP (Semi-Automatic Classification Plugin) supported supervised and unsupervised classification of satellite and drone imagery across multiple spectral ranges, especially infrared and near-infrared bands [11]. These tools were essential in identifying soilmarks and cropmarks—subtle variations in soil color or plant growth caused by subsurface structures such as foundations, walls, and ditches [12], [13].

Satellite imagery from multiple years (2019 to 2023) was processed to capture temporal variations in vegetation (Fig. 2). Random Forest classifiers [14], [15], trained to detect minimal vegetational shifts often linked to anthropogenic alterations in soil composition and moisture retention, were applied to NDVI differentials (Fig. 3). Significant anomalies were recorded between 2020 and 2022, with peaks of vegetative density followed by normalization in 2023. These patterns were mapped and interpreted as potential archaeological features, including linear anomalies possibly corresponding to centuriation grids or ancient roadways.

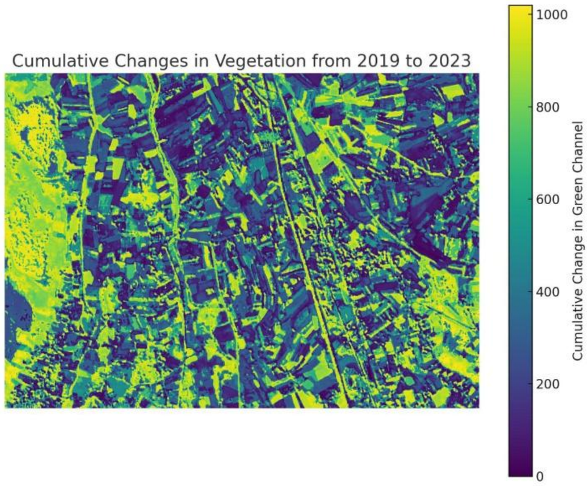

Fig. 3. Cumulative inter-annual change in vegetation greenness (2019–2023), derived from the green spectral channel.

To enhance the predictive capabilities of the workflow, the research employed machine learning algorithms capable of ingesting large datasets and detecting complex patterns. These models were tested on both sectors of the Vallo di Diano and validated through comparison with existing archaeological maps and field verification campaigns. The preliminary results show a promising overlap between AI-detected anomalies and known sites, but also reveal discrepancies that call for a critical reassessment of purely automated approaches. AI tends to over-detect in areas with intense agricultural use or complex geology, leading to false positives that can only be filtered through expert archaeological knowledge.

One of the most innovative aspects of the methodology is the use of viewshed analysis as a complementary tool. Based on high-resolution Digital Terrain Models (DTMs) [16], visibility calculations were performed from known or hypothesized site locations to assess their strategic positioning in terms of intervisibility and landscape control. The results suggest that many of the sites identified by AI lie along visual corridors that connect ancient routes, agricultural zones, and natural vantage points. This supports the hypothesis that visibility played a key role in site selection, not only for defensive purposes but also for symbolic and territorial claims.

The research also explored the use of AI for semantic recognition tasks. By feeding the model with georeferenced examples of Roman centuriation and prehistoric settlements, it was possible to experiment with object detection techniques. These are still in a preliminary phase, but early results indicate that AI can be trained to recognize structured patterns even in non-uniform terrains, particularly when cross-validated with historical cartography and survey data.

Finally, all spatial data were integrated into a comprehensive GIS platform, allowing for dynamic querying, layer comparison, and multi-temporal analysis. The system is designed to be open and replicable, promoting a new standard for archaeological research that balances automation with human expertise.

III. CHALLENGES IN AUTOMATED DETECTION AND PREDICTIVE MODELING

While the integration of AI, multispectral analysis, and GIS-based methodologies has shown considerable promise in archaeological prospection, several challenges remain. A primary issue concerns the risk of false positives, particularly in areas affected by intensive agricultural practices, fluctuating vegetation, or heterogeneous soils. Even well-trained classification algorithms such as Random Forest may misinterpret seasonal or anthropogenic vegetation changes as archaeological anomalies, especially where ground-truth validation is limited.

Another critical issue is the temporal variability of satellite imagery. NDVI values are highly sensitive to climatic fluctuations, irrigation cycles, and crop rotations, all of which may distort the interpretation of vegetational anomalies if not controlled by rigorous multi-temporal analysis. The use of multi-year imagery (2019–2023) mitigated some of these issues, but several patterns remain ambiguous and require field verification.

A further limitation involves the so-called "black box" nature of some machine learning models, particularly deep learning approaches. While such models can achieve high accuracy in anomaly detection, their decision-making processes often remain opaque, reducing interpretability and scientific transparency. This raises concerns about reproducibility and the possibility of overfitting when applied to complex landscapes.

The absence of standardized protocols for applying AI in archaeology hinders comparability across case studies. As digital workflows become increasingly complex, there is a pressing need for shared methodological frameworks that define best practices in data preprocessing, classification, and validation.

Despite these challenges, AI-based approaches remain powerful tools when combined with expert archaeological knowledge. Their effectiveness is greatly enhanced by hybrid workflows that balance automation with interpretative reasoning, field verification, and contextual awareness.

IV. CONCLUSIONS

This research highlights the significant potential of integrating artificial intelligence, multispectral remote sensing, and GIS-based analysis in archaeological prospection. By combining NDVI differentials, automated classification tools such as the SCP plugin, and experimental AI models, it was possible to detect and interpret subtle surface anomalies across selected sectors of the Vallo di Diano. The results demonstrate that digital and automated approaches can enhance the archaeologist's ability to identify buried features, especially in morphologically and historically complex landscapes.

The choice of the Vallo di Diano as a case study reflects both its historical significance and its suitability for advanced spatial analysis. The integration of AI, NDVI, SCP, and GIS tools provides a powerful methodological framework for detecting, interpreting, and validating archaeological markers. At the same time, limitations remain—particularly in reducing false positives, ensuring interpretive transparency, and establishing shared protocols for validation. However, the results confirm the potential of this hybrid approach to transform archaeological prospection and contribute to a deeper understanding of ancient landscapes.

This study also underlines that AI is not a substitute for archaeological reasoning but rather a complementary tool that can extend the scope of traditional research when combined with contextual knowledge and field verification. Future developments will focus on refining predictive models, expanding training datasets with confirmed ground-truth data, and developing open, replicable workflows.

V. REFERENCES

[1] A. Garcia-Molsosa, H.A. Orengo, D. Lawrence, G. Philip, K. Hopper, C.A. Petrie, "Potential of deep learning segmentation for the extraction of archaeological features from historical map series", Archaeol Prospect, 28(2), Apr-Jun 2021, pp. 187-199.

[2] F. Soula, P. Fallavollita, L. Zappino, M. Balsi, S. Esposito, M.G. Melis, "Remote sensing in Archaeology. The state of the art and presentation of METAdAtA research project's preliminary results", Revue Française de Photogrammétrie et de Télédétection, 216, March 2018, pp. 61-79.

[3] K.M.S. Allen, S.W. Green, E.B.W. Zubrow, "Interpreting Space: GIS and Archaeology. Applications

of Geographic Information Systems", London, 1990.

[4] H. Rua, "Geographic information systems in archaeological analysis: a predictive model in the detection of rural Roman villae," *Journal of Archaeological Science*, vol. 35, no. 2, pp. 224–236, 2008.

[5] F. Papaleo, "Scoperte archeologiche nel Vallo di Diano", in AA.VV., "Il Vallo di Diano. Dalla preistoria all'età contemporanea", Agropoli, 2024, pp. 13-22.

[6] A. Agapiou, V. Lysandrou, "Interacting with the Artificial Intelligence (AI) Language Model ChatGPT: A Synopsis of Earth Observation and Remote Sensing in Archaeology", Heritage 2023, 6, pp. 4072–4085.

[7] G. Ciccone, "ChatGPT as a Digital Assistant for Archaeology: Insights from the Smart Anomaly Detection Assistant Development", Heritage, 7, 2024, pp. 5428–5445.

[8] W.B. Verschoof-van der Vaart, K. Lambers, "Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands", Journal of Computer Applications in Archaeology, 2(1), 2019, pp. 31–40.

[9] G. Caspari, P. Crespo, "Convolutional Neural Networks for Archaeological Site Detection Finding 'Princely' Tombs", Journal of Archaeological Science, vol. 110, October 2019.

[10] N. Masini and R. Lasaponara, "Investigating the spectral capability of QuickBird data to detect archaeological remains buried under vegetated and not vegetated areas," Journal of Cultural Heritage, vol. 8, no. 1, 2007, pp. 53–60.

[11] L. Congedo, "Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS" , Journal of Open Source Software, 6(64), 3172, August 2021.

[12] G.L.A. Fry, B. Skar, G. Jerpåsen, V. Bakkestuen, L. Erikstad, "Locating archaeological sites in the landscape: a hierarchical approach based on landscape indicators," Landscape and Urban Planning, vol. 67, no. 1–4, 2004, pp. 97–107.

[13] O. Belvedere, A. Burgio, R. M. Cucco, D. Lauro, "Relazioni tra geomorfologia, processi post-deposizionali e visibilità del suolo nella lettura dei dati di prospezione archeologica," Archeologia e Calcolatori, vol. 16, 2005, pp. 129–152.

[14] H.A. Orengo, F.C. Conesa, A. Garcia-Molsosa, A. Lobo, A.S. Green, M. Madella, C.A. Petrie, "Automated detection of archaeological mounds using machine-learning classification of multisensor and multitemporal satellite data", Proceedings of the National Academy of Sciences, 117(31), 4 August 2020, pp. 18240-18250.

[15] M.E. Castiello, M. Tonini, "An Explorative Application of Random Forest Algorithm for Archaeological Predictive Modeling. A Swiss Case Study", Journal of Computer Applications in Archaeology, 4(1), 2021, pp. 110–125.

[16] G. Vizzino, "GIS and DTM for the analysis of the archaeological data in Vaste (Southern Apulia)," SCIRES-IT, vol. 10, 2020, pp. 53–64.