# Detection of archaeological surface ceramics using very high-resolution remotely sensed data

Argyro Argyrou[1][0000-0001-6134-5799], Apostolos Papakonstantinou[1][0000-0002-6464-2008], Michalakis Christoforou[2][0000-0002-6217-5987], Dimitrios D. Alexakis[3][0000-0003-4070-4654] and Athos Agapiou[1][0000-0001-9106-6766]

[1] Department of Civil Engineering and Geomatics, Faculty of Engineering and Technology, Cyprus University of Technology, Saripolou 2-8, 3036 Achilleos 1 Building, 2nd Floor, Lemesos, Cyprus, P.O Box. 50329, 3603
[2] Department of Agricultural Sciences, Biotechnology and Food Science, Faculty of Geotenical Sciences and Environmental Management, Cyprus University of Technology, Archbishop Kyprianou Str. 30, 3036 Limassol, Cyprus, PO Box 50329, 3603
[3] Laboratory of Geophysical—Satellite Remote Sensing & Archaeo-environment (GeoSatRSeArch Lab), Institute for Mediterranean Studies, Foundation for Research and Technology—Hellas (FORTH), 130 Nikiforou Foka Str. 74100 Rethymno, Crete, Greece;
`ac.argyrou@edu.cut.ac.cy`

**Abstract.** This study explores using Artificial Intelligence (AI) to enhance the detection of archaeological features from images taken by Unmanned Aerial Vehicles (UAVs). The aim is to complement traditional archaeological survey methods. With the advancement of Remote Sensing (RS), the use of low-altitude systems for documenting archaeological sites has become popular due to cost-effectiveness. Additionally, the application of Machine Learning (ML) and Deep Learning (DL) techniques has shown promise in improving the efficiency of archaeological feature detection. However, there are still challenges in accurately interpreting results, especially when methodologies are inadequate. This study addresses the need to detect non-uniformly better-distributed surface ceramics in high-resolution aerial images, a challenge characterized by imbalanced data distribution. The authors propose an innovative AI-driven methodology; the research aims to fill a gap in the current literature, improving the precision and reliability of archaeological research through advanced technological approaches. Upon reviewing the overall results alongside the in-situ survey records, it has been determined that using high-resolution UAV imagery combined with automated AI methods is an effective preliminary approach for detecting surface archaeological ceramics.

**Keywords:** Artificial Intelligence (AI), Object detection, Multispectral UAV, Surface ceramic detection.

# 1 Introduction

## 1.1 A Subsection Sample

This study aims to unveil the vast potential of Artificial Intelligence (AI) in Unmanned Aerial Vehicle (UAV) imagery for automated archaeological feature detection, with the noble aim of bolstering traditional fieldwalking archaeological surveys. Over the years, Remote Sensing science has proven its mettle in bolstering archaeological research [1]. Drones and other low-altitude systems have been widely embraced for documentation purposes of archaeological sites due to their cost-effectiveness [2]. Moreover, analytical tools such as machine learning automatic detection have revolutionized the efficiency of archaeological feature detection [3]. Researchers have recently harnessed computer vision techniques and deep learning with artificial intelligence to elevate archaeological research to new heights. In 2020, [4] underscored the surge in the use of these advanced technologies. The success of remote sensing approaches is intricately woven with the methodology used during the research process. However, there are instances when the methodology used is insufficient, making it challenging to assess the outcomes accurately. Consequently, it becomes difficult to interpret the results in a way that aligns with the research objectives. According to a recent study conducted by [5], although Machine Learning (ML) and Deep Learning (DL) algorithms are widely used for the classification and identification of ceramics, the results from the detection of archaeological structures using the DL algorithm (especially with aerial imaging) still require improvement.

This study aims to fill a gap in the current literature by utilising AI methods for detecting non-uniformly distributed surface ceramics in high-resolution remote imagery. This can be viewed as an "imbalanced data distribution" problem that requires further examination. To overcome this challenge, novel approaches need to be developed.

# 2 Methods and materials

Our methodology combines RGB and multispectral images with artificial intelligence (AI) techniques to identify surface archaeological ceramics in archaeological and controlled environments. The images were captured through low-altitude sensor cameras. Two flight campaigns were performed using the DJI Phantom 4 Pro system (spectral bands: Blue (B): 468 nm ± 47 nm; Green (G): 532 nm ± 58 nm; Red (R): 594 nm ± 32.5 nm), and the DJI P4 multi-spectral systems (spectral bands: Blue (B): 450 nm ± 16 nm; Green (G): 560 nm ± 16 nm; Red (R): 650 nm ± 16 nm; Red edge (RE): 730 nm ± 16 nm and Near-infrared (NIR): 840 nm ± 26 nm). We employed traditional photogrammetric techniques, which involve obtaining measurements from photographs, to generate an orthophoto-mosaic for each specific set of images. Initially, we applied supervised machine learning (ML) classifiers to the high-resolution derivatives of the images. These classifiers were trained to recognize and categorize specific features within the images. We then used various evaluation metrics, such as accuracy, to assess the performance of the classification. These metrics helped us to measure the effectiveness of the classifiers and guide the modeling process. The overall process is visually presented in Figure 1 below, which illustrates the workflow of the classification and evaluation techniques.

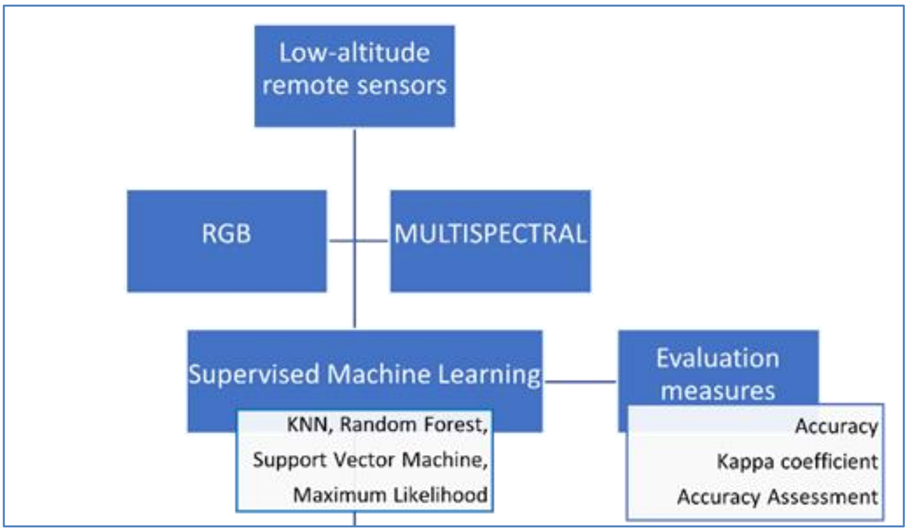

**Fig. 1.** Using UAV images in simulated and real case studies, research methodology.

Many standard classifier learning algorithms assume equal class distribution and misclassification costs, which may not perform well on imbalanced data sets. Imbalanced data sets have skewed class proportions, with the majority class having a more significant proportion of samples than the minority class, such as archaeological surface ceramics. Small sample sizes, separability, and sub-concepts within the minority class can influence modeling. Small sample sizes lead to poor representation of the minority class, making accurate classification difficult. Separability refers to how distinct the minority class is from the majority class, affecting misclassification rates. Sub-concepts within the minority class may require different models for accurate classification [6].

Based on the results of our experiment, it seems that the problem arises from an imbalance between the surface ceramics and the surrounding environment, which includes the soil and crops. To investigate this issue further, we need more precise classification results and effective tools to handle imbalanced data or update learning algorithms. The literature provides various solutions, such as rebalancing the class distribution by resampling the data space at the data level or modifying existing classifier learning algorithms to better understand the small class of ceramics at the algorithmic level. Dealing with imbalanced data can challenge boosting algorithms in binary classification problems. Nonetheless, boosting algorithms are commonly used to enhance the predictive abilities of weak learners and turn them into strong learners [7]. Among the three main boosting algorithms, AdaBoost, is often used as an Ensemble Method in Machine Learning. The fundamental concept behind AdaBoost is to construct a model and assign equal weights to all data points. It then assigns higher weights to incorrectly classified points, and the subsequent model emphasizes those points with higher weights. This training process continues until a lower error rate is achieved [8].

## 3 Results

This study focuses on applying classifiers trained on image samples categorized into three distinct classes: 'ceramics', 'soil', and 'crops'. The efficacy of these classifiers was assessed through the analysis of overall accuracy, determined by evaluating randomly distributed testing pixels. Accuracy is the ratio of correctly predicted samples within the test set to the total predictions made.

$$\textbf{Accuracy = Correct Predictions / Total Predictions.} \tag{1}$$

For the soil and crops classes, the accuracy was estimated to be around 80%. In contrast, ceramics demonstrated lower accuracy across all four classifiers. This inconsistency raised questions regarding the optimal number of testing pixels required to ensure that the assessed accuracy effectively represents true accuracy. The question was whether increasing the sample size of testing pixels would improve classifier performance.

Referencing [9] it is suggested that to achieve an accuracy level of 90%, a sample size of 225 is required, while a 95% accuracy level necessitates 119 testing pixels. These recommendations presuppose that the classes adhere to a normal distribution characterized by a symmetric spread of measurement values centered around the mean. Following this guidance, the classifiers were re-evaluated with 225 testing samples. This sampling employed ArcGIS Pro's capabilities to randomly generate 225 points for post-classification accuracy analysis, utilizing the Image Analyst Toolbox. The methodology ensured an equitable distribution of points across each class, proceeding to create and manually update the 'Ground Truth' and 'Classified' fields within the final attribute table. This process culminated in the utilization of the Compute Confusion Matrix tool to compare these fields.

The ceramics class displayed considerable variability in accuracy, with RGB images ranging from 12% to 24% and multispectral images from 23% to 61% (Table 1). This variability highlights the challenge of misclassification, particularly of minority classes, where classifiers demonstrate high predictive accuracy for predominant classes but falter with minority classes. To confront this limitation, our research introduces a methodology to enhance surface ceramics' detection. This approach leverages low-altitude multispectral and RGB cameras, underpinned by the application of weak learners, to optimize predictive accuracy.

**Table 1.** Ceramics Accuracy after supervised classification.

| | KNN | MAX_Likelihood | SVM | RF |
|---|---|---|---|---|
| RGB | 13% | 12% | 24% | 15% |
| Multispectral | 23% | 52% | 61% | 31% |

## 4 Accuracy Improvements

Research has shown that using low-altitude sensors to detect surface ceramics can yield significant and valuable results. However, this approach has challenges, particularly in accurately detecting the minority class of ceramics, which can significantly impact the overall detection accuracy. We are framing ceramic surface detection as an "imbalanced data distribution" problem to address this issue. Previous studies have identified the problem of misclassifying minority classes, such as archaeological ceramics. Despite achieving high overall accuracy, there remains a low actual detection rate for the ceramic class. This is because classifiers tend to predict classes with abundant data more accurately than those with limited data. Therefore, addressing these imbalances in data distribution is crucial for improving detection accuracy, particularly for the minority ceramics class.

## 5 Conclusion

This research evaluated the potential use of artificial intelligence techniques to automatically identify archaeological ceramics from high-resolution images of unmanned aerial vehicles (UAVs). Additionally, we aimed to develop a methodology that could produce results in terms of time and accuracy comparable to, if not better than, those obtained through traditional archaeological field surveys. To accomplish this, we utilized supervised machine learning algorithms to analyze RGB and multispectral images captured by a UAV.

The comprehensive research revealed that the classifiers utilized in the study demonstrated a high level of effectiveness in accurately predicting soil types and crop varieties with commendable precision. However, when it came to predicting surface ceramics, which represents a minority class, the classifiers were found to be less successful. To address this disparity, authors introduced a novel methodology to enhance the accuracy of detecting surface ceramics through drone imagery. This innovative approach involves implementing a boosting technique designed for weak learners, encompassing RGB and multispectral images. Combining these two types of imagery, the newly proposed method is anticipated to yield significantly improved results, providing a more reliable and precise means of identifying surface ceramics than the previous technique.

Moreover, the authors plan to broaden their research on ceramics to encompass a variety of types from different historical periods, each exhibiting unique spectral behaviors within the same geographic area. They intend to perform meticulous lab-based spectral measurements to ensure a statistically significant differentiation among the ceramics during the same flight. Additionally, the team has meticulously scheduled drone surveys to enrich the dataset for algorithm training and comprehensive outcome assessment. Their overarching goal is to mitigate noise, elevate distinguishability, and thoroughly assess imbalanced ceramics data using established measures such as F-measure, G-mean, and ROC analysis.

**Acknowledgements:** This project has received funding from the European Union's Horizon Europe Framework Programme (HORIZON-WIDERA-2021-ACCESS-03, Twinning Call) under the grant agreement No 101079377 and the UKRI under project number 10050486. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the UKRI. Neither the European Union nor the UKRI can be held responsible for them. This paper is a part of the PhD of the first author on the same topic.

## References

1. Traviglia, A., Torsello, A. Landscape pattern detection in archaeological remote sensing. Geosciences,7, 128, (2017). [doi:10.3390/geosciences7040128]
2. Materazzi, F. and Pacifici, M. Archaeological crop marks detection through drone multispectral remote sensing and vegetation indices: A new approach tested on the Italian pre-Roman city of Veii. Journal of Archaeological Science: Reports 41, 103235, (2022). [https://doi.org/10.1016/j.jasrep.2021.103235]
3. Orengo H.A., & Garcia-Molsosa, A., "A brave new world for archaeological survey: Automated machine learning-based potsherd detection using high-resolution drone imagery" Journal of Archaeological Science, 112, 105013, (2019), [https://doi.org/10.1016/j.jas.2019.105013]
4. Manuel Domínguez-Rodrigo M., Gabriel Cifuentes-Alcobendas, G., Jiménez-García, B., Abellán, N., Pizarro-Monzo, M., Organista, E., Baquedano, E. Artificial intelligence provides greater accuracy in the classification of modern and ancient bone surface modifications. Scientific Reports, 10, 18862, (2020). [https://doi.org/10.1038/s41598-020-75994-7]
5. Jamil A. H., Yakub F., Azizan A., Roslan S. A., Zaki S. A., Ahmad S. A., A Review on Deep Learning Application for Detection of Archaeological Structures. Journal of Advanced Research in Applied Sciences and Engineering Technology 26, Issue 1, 7-14, (2022).
6. Y. Sun, A.K.C. Wong, M.S. Kamel, "Classification of imbalanced data: a review", International Journal of Pattern Recognition and Artificial Intelligence 23(04), (2009), DOI: 10.1142/S0218001409007326, Sourcedx.doi.org

7. Freund, Y., “Boosting a Weak Learning Algorithm by Majority”. In: Fulk MA, Case J, editors. Proceedings of the Third Annual Workshop on Computational Learning Theory, COLT 1990, University of Rochester, Rochester, NY, USA, August 6-8, p. 202{216 (1990)
8. Schapire, R.E., “The Strength of Weak Learnability”. *Machine Learning*. (1990);5(2):197{227.
9. Richards, J.A., Remote Sensing Digital Image Analysis, Springer-Verlag Berlin Heidelberg, ISBN, 978-3-030-82327-6 (eBook), (2022), https://doi.org/10.1007/978-3-030-82327-6