

Comparative Analysis of Machine Learning Algorithms for Classification of UAV-based Photogrammetric Cultural Heritage Point Clouds

Mehmet Arkali, Muhammed Yahya Biyik, Muhammed Enes Atik

ITU, Civil Engineering Faculty, 34469 Maslak Istanbul, Türkiye – (markali, biyik16, atikm)@itu.edu.tr

Keywords: Photogrammetry, Point cloud, Machine learning, UAV, Classification, Cultural heritage

Abstract

Unmanned Aerial Vehicles (UAVs) are being increasingly utilized across different fields because of their ability to deliver quick, cost-effective, and precise spatial information. Innovations in photogrammetry and computer vision techniques, especially Structure from Motion (SfM) and Multi-View Stereo (MVS), have improved the generation of orthoimages, digital surface models, and dense point clouds, rendering UAVs highly efficient for documentation and three-dimensional reconstruction. In studies focused on cultural heritage, UAV-based photogrammetry has emerged as a crucial resource for accurately preserving and representing historical sites with great detail and resolution. In this context, the current study analyzes UAV-acquired point cloud data from the Temple of Hera in Italy and performs a comparative assessment of three machine learning algorithms, Support Vector Machines (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost), for the purpose of semantic segmentation tasks. According to our results, the XGBoost and Random Forests (RF) methods have reached to more than 90% F1 score for all classes, and the SVM method has reached 90% F1 score only for three classes.

1. Introduction

An unmanned aerial vehicle is an aircraft that operates without a pilot. It is known in the geomatics community as a Remotely Piloted Vehicle (RPV) or Unmanned Vehicle System (UVS), among other terms (Remondino et al., 2012). UAVs are currently employed across a diverse range of disciplines, including archaeology (Fiz et al., 2022), civil engineering (Guan et al., 2022), computer vision (Biyik et al., 2023), and geomatics engineering (Atik and Arkali, 2025). The adoption of UAVs for photogrammetric mapping and surveying has experienced significant growth, thanks to advancements in technology and algorithms. These enhancements have improved the quality of deliverables such as orthoimages and digital surface models (DSM), which are now widely utilized across various fields for accurate and efficient results. UAV applications can be categorized into several areas, including forestry and agriculture, archaeology and 3D reconstruction, environmental surveying, traffic monitoring, and cultural heritage (CH) (Uysal et al., 2015). CH constitutes a vital means of connecting historical contexts with future generations. The imperative to document and preserve cultural heritage is increasingly recognized as essential. Photogrammetry collects 3D information and textures of cultural heritage objects using overlapping images, determined by known camera positions and orientations (Yastikli, 2007). UAV-based photogrammetry offers a more efficient and effective approach to mapping and modeling cultural heritage sites than traditional terrestrial methods. UAV-based photogrammetry generates high-resolution models of large areas, utilizing low-cost sensors and providing high temporal resolution (Atik et al., 2025). UAV photogrammetry has advanced significantly due to the development of Structure from Motion (SfM) and Multi-View Stereo (MVS) algorithms. These technologies facilitate the generation of dense point clouds, which accurately and comprehensively represent three-dimensional objects and structures in a detailed manner. There is an increasing trend in UAV photogrammetry, which is progressively utilized for documenting, representing, and performing 3D reconstructions of locations and structures related to CH (Azzola et al., 2019; Frodella et al., 2020).

Point cloud classification for complex data requires complex discriminative rules. Machine learning is a powerful statistical tool that can be used to classify complex point clouds (Atik et al., 2021). Machine learning methods have recently been frequently used for the semantic annotation of cultural heritage models and the automated integration of HBIM environments (Croce et al., 2021). Attractive tools are offered for classifying cultural heritage point clouds and organizing semantic information. By exploiting the geometric features of the point cloud, machine learning methods can effectively interpret cultural heritage point cloud data and integrate it into HBIM processes.

This research study utilizes point cloud data obtained from previous investigations conducted by Pepe et al. (2022). A dataset has been gathered utilizing UAV for the 3D reconstruction of the Temple of Hera in Italy, employing a photogrammetric methodology. This study employs the Support Vector Machine (SVM), Random Forest (RF), and Extreme Gradient Boosting (XGBoost) algorithms for the classification of cultural heritage point clouds.

2. Material and Methods

2.1 Dataset

The dataset consists of 57 UAV images of archaeological remains located at Metaponto, Italy, captured on 01 February 2020. These images were processed using Structure from Motion and Multi-View Stereo (MVS) algorithms to create a georeferenced and scaled 3D point cloud. To ensure spatial accuracy, eight ground control points measured through GNSS survey were utilized. For classification, the point cloud was divided into training, validation, and test subsets. The training and validation subsets were manually labeled, while the test subset was reserved for algorithmic classification. Five semantic classes were established: architrave (0), capital (1), column (2), stylobate (3), and stereobate (4). This dataset facilitates the analysis and classification of cultural heritage structures in 3D, serving as a valuable resource for testing reconstruction algorithms, point cloud processing, and applications in virtual or augmented reality.

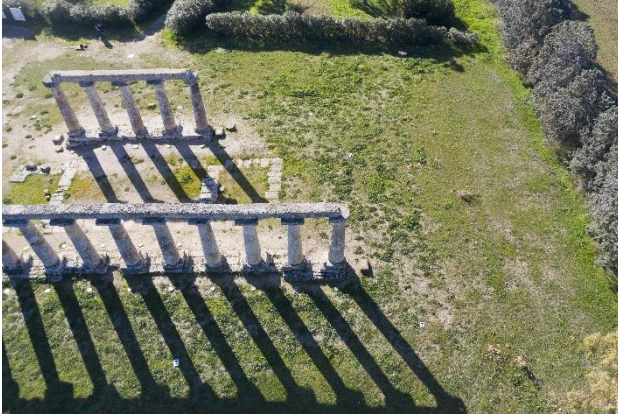


Figure 1. Study area and RGB view of heritages (Pepe et al., 2022)

2.2 Support Vector Machines (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm that is effective for both classification and regression tasks (Cortes and Vapnik, 1995). The optimal hyperplane is identified through the application of equation (3) for a specified set of samples represented as $x_i (i = 1, 2, \dots, N)$

$$f(x) = w^T x + b = \sum_{j=1}^N w_j x_j + b = 0 \quad (1)$$

In this context, w represents an N -dimensional vector and b signifies a scalar value. Together, these elements are employed to define the hyperplane.

2.3 Random Forest (RF)

Random Forest (RF) is an advanced bagging technique that generates a substantial number of independent trees and averages their outputs (Breiman, 2001). Each tree provides a class prediction, and the model selects the class that receives the most votes. For every tree, a training dataset is created using multiple bootstrap samples drawn from the original dataset. The RF classifier requires two key parameters: the number of variables to consider per node and the total number of trees for optimal splitting. Boot samples are taken from two-thirds of the training data, while the remaining one-third, known as out-of-bag (OOB) data, is used to assess errors, resulting in what is termed the generalized error. The error that arises is designated as the generalized error (Eq.2).

$$PE^* = P_{X,Y}(mg(X,Y) < 0) \quad (2)$$

where $mg()$ denotes the margin function. The margin quantifies the extent to which the average number of votes in (X, Y) for the correct class surpasses the average vote for any alternative class. A larger margin indicates a higher reliability in performing the classification.

2.4 Extreme Gradient Boosting (XGBoost)

An efficient and scalable algorithm, based on gradient boosting trees and known as XGBoost, has been developed as an effective approach for addressing both classification and regression problems (Chen and Guestrin, 2016). Given n samples and m features, the dataset $D = \{(x_i, y_i)\}$ is modeled using an ensemble of trees that utilizes K additive functions.

$$\hat{y}_i = \emptyset(x_i) = \sum_{k=1}^K f_k(x_i), \quad f_k \in \varphi, \quad (3)$$

$$\varphi = \{f(x) = \omega_{q(x)}\} (q: R^m \rightarrow T, \omega \in R^T, \quad (4)$$

In this context, $\hat{y}_i$ is the model's prediction from [Eq. (1)], x_i is an observation, and $f_k(x_i)$ is its predicted score. φ refers to the set of regression trees in [Eq. (2)] with an independent structure q . T is the number of leaves in the tree, and ω_q represents their weights.

2.5 Evaluation Metrics

The values of metrics derived from machine learning are computed using a confusion matrix. The formulas for these metrics are presented in Equations (5)–(7).

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$F1 \text{ score} = 2 \frac{Precision \cdot Recall}{Precision + Recall} \quad (7)$$

True positive (TP) refers to the count of points that share the same label in both the predicted results and the ground truth. False positive (FP) indicates the number of points that are predicted as positive, yet have a negative actual label. Conversely, false negative (FN) denotes the number of points that are predicted as negative while their true label is positive (Alpaydin, 2020).

3. Results and Discussion

The test area is at the Metaponto archaeological site in Italy and features the ruins of a 6th-century BC Doric temple dedicated to the Greek goddess Hera. Using three different ML techniques, the point cloud was classified into five distinct classes: lintel, capita, column, stylobate, and stereobate. Classification was applied according to ten parameters that describe the point cloud, and the point cloud was processed on the Cloud Compare software. The results of the point cloud classification, obtained using the XGBoost, RF, and SVM methods, are presented in Figures 2 – 4.

The classification process was implemented in Python using the Scikit-learn library. To compare these methods, all results were cross-checked using statistical parameters, including precision, recall and F1-score. According to the results obtained with the XGBoost method, classification can be achieved with high accuracy. Almost all of the architrave and column classes were correctly predicted. The stylobate class, however, performed less well than the other classes, with an F1-score of 0.93. The classification results and confusion matrix are presented in Table 1 and Table 2, respectively.

Class	Precision	Recall	F1-score	Support
0	0.99	1	0.99	17453
1	1	0.91	0.95	3015
2	0.98	1	0.99	16464
3	1	0.86	0.92	7581
4	0.95	1	0.98	17001

Table 1. XGBoost Methods results.

Class	Precision	Recall	F1-score	Support
0	0.99	0.89	0.94	17453
1	0.59	0.84	0.69	3015
2	0.96	0.96	0.96	16464
3	0.68	0.82	0.74	7581
4	0.93	0.87	0.9	17001

Table 5. SVM Method results.

Class	0	1	2	3	4
0	17448	5	0	0	0
1	224	2751	40	0	0
2	0	0	16464	0	0
3	0	0	272	6498	811
4	0	0	0	0	17001

Table 2. Confusion Matrix of XGBoost Method.

Class	0	1	2	3	4
0	15620	1758	75	0	0
1	201	2540	274	0	0
2	0	16	15762	685	1
3	0	0	248	6203	1130
4	0	0	13	2198	14790

Table 6. Confusion Matrix of SVM Method.

The RF algorithm also has a similar performance to XGBoost. The stylobate class has the lowest prediction metrics, while the architrave and column classes have the highest. Some points belonging to the stylobate class are labeled as columns. The classification results and confusion matrix are presented in Table 3 and Table 4, respectively.

Class	Precision	Recall	F1-score	Support
0	0.99	1	0.99	17453
1	1	0.93	0.96	3015
2	0.98	1	0.99	16464
3	1	0.85	0.92	7581
4	0.95	1	0.98	17001

Table 3. Random Forest Method results.

Class	0	1	2	3	4
0	17443	10	0	0	0
1	182	2799	34	0	0
2	0	0	16464	0	0
3	0	0	289	6468	824
4	0	0	0	0	17001

Table 4. Confusion Matrix of Random Forest Method

The SVM algorithm lags behind the RF and XBoost algorithms in performance. While the column class achieved high prediction metrics, the capital and stylobate classes achieved lower metrics with F1-scores of 0.69 and 0.74, respectively. While the capital class was mixed with the architrave and column classes, some points from the stylobate class were assigned to the stereobate class. The classification results and confusion matrix are presented in Table 5 and Table 6, respectively.

According to the results, RF and XGBoost have better performance than SVM thanks to their structure containing decision trees. RF and XGBoost better handle complex, noisy and large data in point clouds due to the flexible structure of decision trees. SVM has difficulties with imbalanced and noisy datasets. The confusion between geometrically similar structures is an expected result. This is because each point is defined by local geometric features, making it difficult to distinguish similar geometric structures. In this study, confusion occurred in the SVM algorithm, particularly between column, stylobate, and stereobate.

4. Conclusions

This study investigated the classification of cultural point findings using three different ML methods (RF, SVM, and XGBoost). The dataset used was generated photogrammetrically from UAV images. The results indicate that bagging and boosting algorithms are successful in classification. Classified point clouds automatically identify the geometries of cultural heritage, providing advantages in the use of these geometries in H-BIM type information systems. Furthermore, given that machine learning may be inadequate for large and complex datasets, future studies are planned to focus on deep learning methods. Furthermore, the use of explainable AI methods will be beneficial in understanding the black-box nature of ML algorithms.

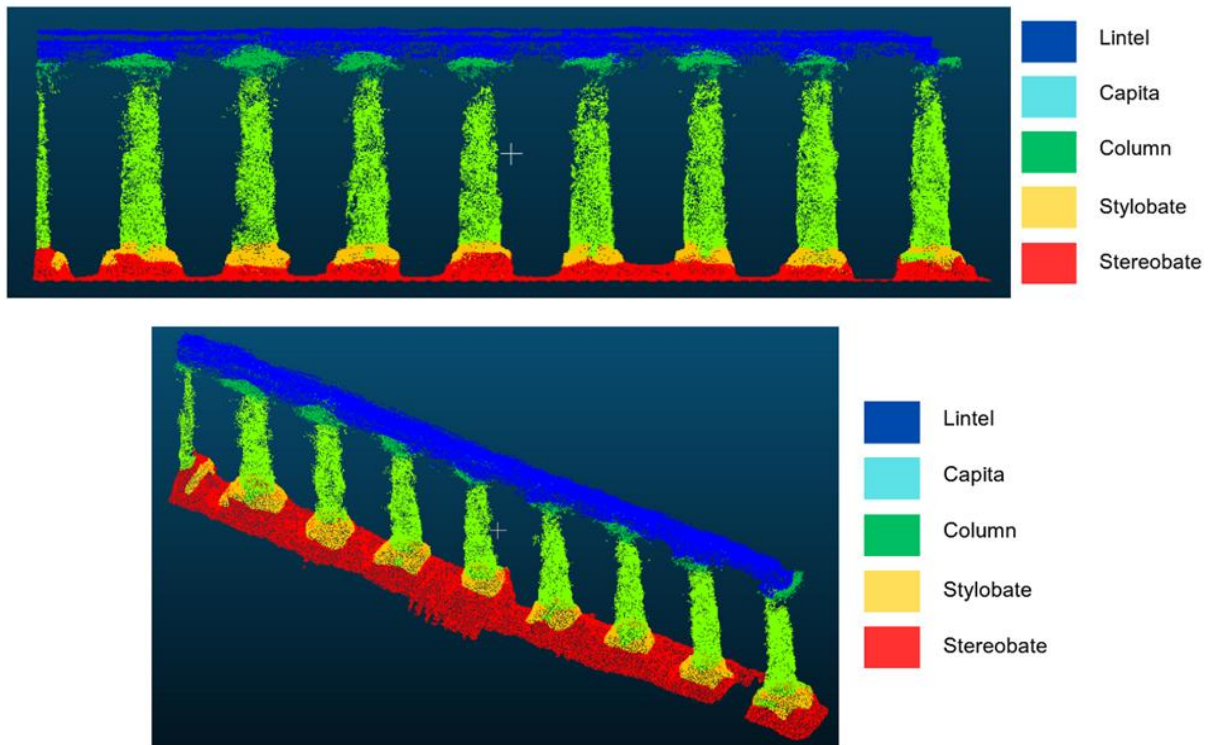


Figure 2. Classification of point cloud with XGBoost.

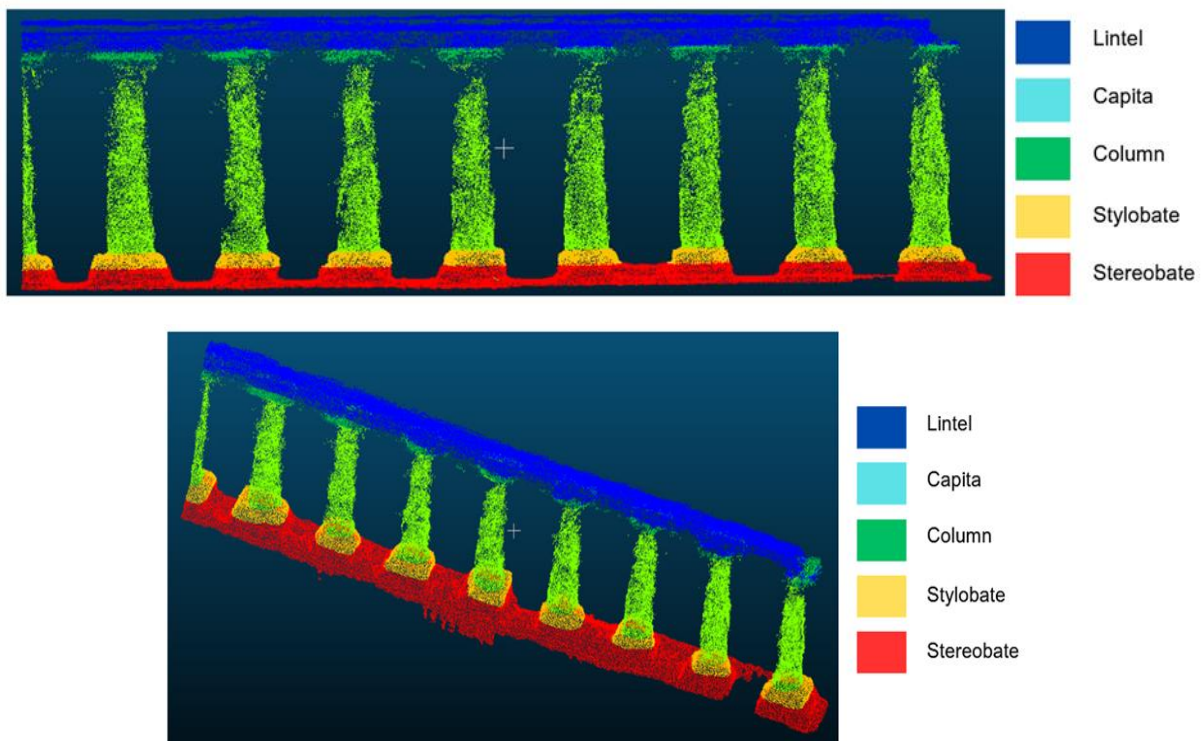


Figure 3. Classification of point cloud with RF.

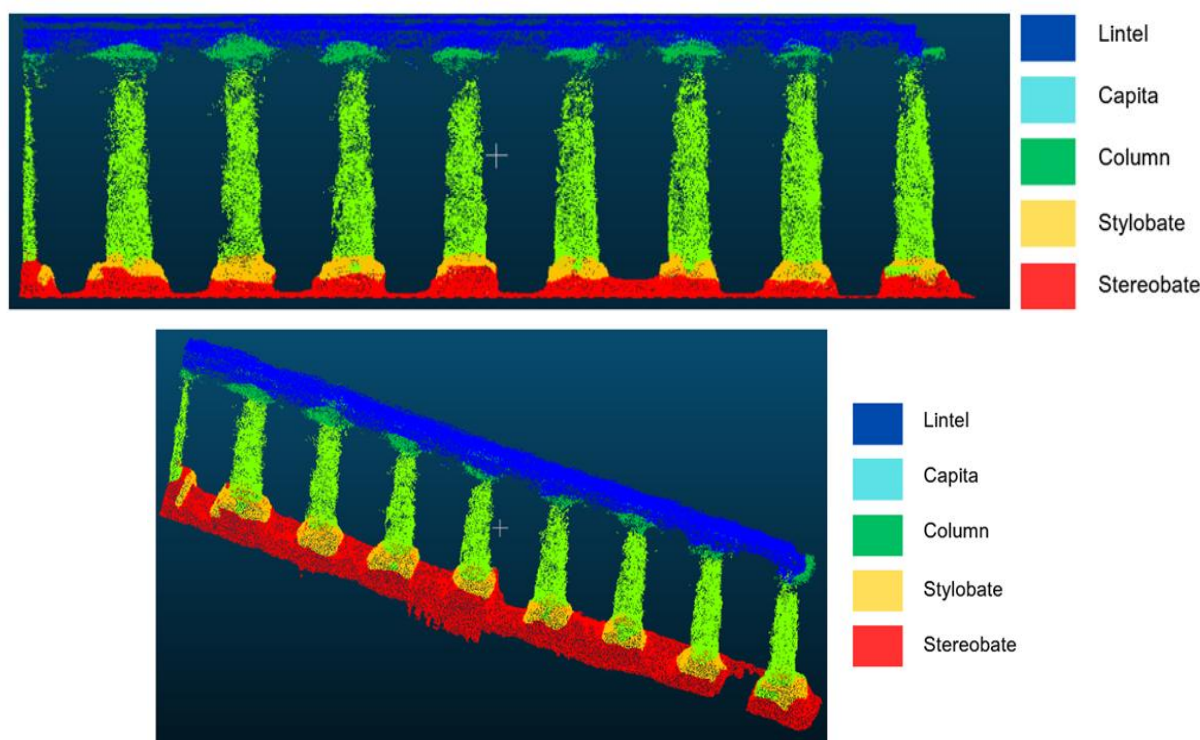


Figure 4. Classification of point cloud with SVM.

Acknowledgements

The dataset used in this study originates from “Data for 3D reconstruction and point cloud classification using machine learning in cultural heritage environment” by Massimiliano Pepe, Vincenzo Saverio Alfio, Domenica Costantino, and Daniele Scaringi. In the present work, the same dataset was employed with the application of different methods to explore alternative approaches to the original data.

References

- Alpaydin, E., 2020. *Introduction to Machine Learning*, 4th edition. MIT Press.
- Atik, M.E., Arkali, M., 2025. Comparative Assessment of the Effect of Positioning Techniques and Ground Control Point Distribution Models on the Accuracy of UAV-Based Photogrammetric Production. *Drones* 9, 15. <https://doi.org/10.3390/drones9010015>
- Atik, M.E., Arkali, M., Atik, S.O., 2025. Impact of UAV-Derived RTK/PPK Products on Geometric Correction of VHR Satellite Imagery. *Drones* 9, 291. <https://doi.org/10.3390/drones9040291>
- Atik, M. E., Duran, Z., Seker, D. Z. 2021. Machine Learning-Based Supervised Classification of Point Clouds Using Multiscale Geometric Features. *ISPRS International Journal of Geo-Information* 10(3), 187. <https://doi.org/10.3390/ijgi10030187>
- Azzola, P., Cardaci, A., Mirabella Roberti, G., Nannei, V.M., 2019. Uav Photogrammetry For Cultural Heritage Preservation Modeling and Mapping Venetian Walls of Bergamo. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2-W9, 45–50. <https://doi.org/10.5194/isprs-archives-XLII-2-W9-45-2019>
- Biyik, M.Y., Atik, M.E., Duran, Z., 2023. Deep learning-based vehicle detection from orthophoto and spatial accuracy analysis. *IJEG* 8, 138–145. <https://doi.org/10.26833/ijeg.1080624>
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>
- Croce, V., Caroti, G., De Luca, L., Jacquot, K., Piemonte, A., Véron, P. 2021. From the Semantic Point Cloud to Heritage-Building Information Modeling: A Semiautomatic Approach Exploiting Machine Learning. *Remote Sensing* 13(3), 461. <https://doi.org/10.3390/rs13030461>
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach Learn* 20, 273–297. <https://doi.org/10.1007/BF00994018>
- Fiz, J.I., Martín, P.M., Cuesta, R., Subías, E., Codina, D., Cartes, A., 2022. Examples and Results of Aerial Photogrammetry in Archeology with UAV: Geometric Documentation, High Resolution Multispectral Analysis, Models and 3D Printing. *Drones* 6, 59. <https://doi.org/10.3390/drones6030059>
- Frodella, W., Elashvili, M., Spizzichino, D., Gigli, G., Adikashvili, L., Vacheishvili, N., Kirkitadze, G., Nadaraia, A., Margottini, C., Casagli, N., 2020. Combining InfraRed Thermography and UAV Digital Photogrammetry for the Protection and Conservation of Rupestrian Cultural Heritage Sites in Georgia: A Methodological Application. *Remote Sensing* 12, 892. <https://doi.org/10.3390/rs12050892>

Guan, S., Zhu, Z., Wang, G., 2022. A Review on UAV-Based Remote Sensing Technologies for Construction and Civil Applications. *Drones* 6, 117.
<https://doi.org/10.3390/drones6050117>

Pepe, M., Alfio, V.S., Costantino, D., Scaringi, D., 2022. Data for 3D reconstruction and point cloud classification using machine learning in cultural heritage environment. *Data in Brief* 42, 108250. <https://doi.org/10.1016/j.dib.2022.108250>

Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., Sarazzi, D., 2012. UAV Photogrammetry for Mapping and 3D Modeling – Current Status and Future Perspectives. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXXVIII-1-C22, 25–31.
<https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-25-2011>

Uysal, M., Toprak, A.S., Polat, N., 2015. DEM generation with UAV Photogrammetry and accuracy analysis in Sahitler hill. *Measurement* 73, 539–543.
<https://doi.org/10.1016/j.measurement.2015.06.010>

Yastikli, N., 2007. Documentation of cultural heritage using digital photogrammetry and laser scanning. *Journal of Cultural Heritage* 8, 423–427.
<https://doi.org/10.1016/j.culher.2007.06.003>