Transfer learning with ADAF: extending archaeological detection beyond its training ground

Nejc Čož, Luka Škerjanec, Žiga Kokalj

Introduction

Automatic Detection of Archaeological Features (ADAF) is a deep learning tool for detecting archaeological features from airborne laser scanning (ALS, lidar) data (Čož et al. 2025). It applies convolutional neural networks (CNNs) for semantic segmentation of ALS derived digital elevation models (DEM) to detect archaeological sites in the landscape. Developed through collaboration between ZRC SAZU, the Discovery Programme (Ireland), Bias Variance Labs, and Transport Infrastructure Ireland, ADAF was designed as a user-friendly application that lowers the technical barrier for archaeologists.

Originally trained and validated on Irish and UK ALS data, ADAF's models demonstrated strong performance in recognising barrows, ringforts, and enclosures across Ireland. The present study explores how well a model trained on Irish and UK data performs in a fundamentally different environment, i.e. southern Herzegovina in Bosnia and Herzegovina, and what this reveals about model transferability across contrasting landscapes.

Methods and Materials

The original ADAF training dataset was built from extensive ALS archives supplied by Transport Infrastructure Ireland. Because of low number of barrows, the dataset was enriched with ALS data from the UK because of morphological similarities of barrows in the two regions. The data were processed using the simple local relief model (SLRM) visualisation, divided into tiles, and paired with manually digitised masks of known archaeological features based on the Sites and Monuments Record. The HRNet architecture was employed for semantic segmentation, implemented through the AiTLAS framework.

To test transferability, the pretrained Irish ADAF model was applied to ALS data from southern Herzegovina, acquired within the ERC project "Unde venis? Unravelling the Enigma of the Stećci Tombstones (STONE, 101089123)". The project is conducting a comprehensive investigation into the prehistoric and medieval landscape dynamics of the historic region of Herzegovina. In the absence of publicly available ALS data for this area, a dedicated ALS survey was commissioned. The survey covers an area of approximately 3,800 km² in eastern Herzegovina, extending from the vicinity of Mostar in the west to the Morine Plateau in the north, the Bileća Basin in the east, and Popovo Polje in the south.

The Herzegovinian landscape, part of the Dinaric karst, contains thousands of prehistoric barrows identified through manual inspection. Prehistoric barrows or burial mounds (locally known as gomile, or humke) are among the most numerous and archaeologically significant archaeological structures in this region, ranging in size from a few meters to up to 30 m in diameter, with both circular and irregular base shapes, they tend to occupy prominent but approachable places in the landscape, low ridge crests, hill spurs above valley floors, and terraces that command long views along river corridors and ancient routeways. In limestone karst zones, barrows are often stone-built cairn mounds set on rocky knolls, while in the more soil rich karst fields they tend to be built from earth. Compared to their Irish and British counterparts, these barrows are smaller, without vegetation cover and densely clustered, while former tend to appear as larger, isolated mounds and covered by grass in gentler terrain. These morphological and environmental differences provide a strong basis for assessing how well an Irish-trained model can generalise to a new archaeological and geomorphological context.

Results from preliminary tests

Applying the original ADAF models to the Herzegovinian ALS data produced promising results. The Irish-trained model successfully detected roughly half of the manually recorded barrows, maintaining a notably low false-positive rate. This suggests that, despite environmental and cultural differences, the morphological signal of barrow-type monuments is sufficiently consistent for partial model transfer.

Performance was lower in areas with dense barrow clusters or pronounced karst microrelief, where the model struggled to distinguish archaeological mounds from natural features such as dolines and rock outcrops. Nonetheless, the quality of the detections exceeded expectations given that no Herzegovinian data were included in the training process. The model's conservative predictions, yielding few false positives, indicate that it generalises cautiously and relies on robust morphological cues.

Discussion

These preliminary results demonstrate both the potential and the limitations of transfer learning in archaeological feature detection. While ADAF's Irish-trained model recognises many Herzegovinian barrows, differences in scale, spacing, and landscape texture reduce recall. Such limitations highlight the availability (or rather lack) of training data as a key challenge in deep learning for archaeology, especially for generalisation across different domains.

The next phase will involve retraining ADAF with a curated Herzegovinian dataset derived from the ERC STONE ALS data. By fine-tuning the original model with expertly labelled examples from Herzegovina, we aim to quantify improvements in detection accuracy and better understand how model adaptation can enhance the results. The study will also analyse which features of the Herzegovinian dataset (e.g. morphology, slope, elevation, or context) most affect transfer performance.

Beyond technical refinement, this research contributes to broader discussions on AI transferability in archaeological prospection. Developing adaptable models across regions is essential for scaling up automated detection methods beyond local datasets. By systematically evaluating performance in distinct geomorphological settings, we aim to outline best practices for reusing pretrained archaeological models and sharing training data between projects.

Conclusion

Testing ADAF in southern Herzegovina demonstrates that deep learning models trained on one region can successfully identify similar archaeological features in another, even across substantial environmental contrasts. The results underscore both the robustness of CNN-based detection and the necessity of region-specific retraining.

Future work within ERC STONE will refine ADAF through domain adaptation, creating a model tailored to the karstic environment. This will provide valuable insights into the scalability of archaeological AI and its potential for cross-regional applications. Ultimately, this study contributes to developing a reproducible, transferable framework for automated archaeological detection, integrating deep learning, ALS data analysis, and landscape archaeology within complex terrains.

References

Čož, N., Kokalj, Ž., and Kostovska, A. 2025. *ADAF - Automatic Detection of Archaeological Features Library v1.0.0*. V. v1.0.0. Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU), released July 11. https://doi.org/10.5281/ZENODO.15861628.

Forenbaher, S., 2023. Rane grobne gomile na Jadranu i raznolikost pogrebnih običaja u 3. Tisućljeću pr. Kr. / Early Burial Mounds in the Adriatic and the Diversity of Mortuary Practice in the 3rd Millennium B.C. Monografije Instituta za Arheologiju 15. Zagreb.

Gavranović, M., 2011. Die Spätbronze- und Früheisenzeit in Bosnien. Universitätsforschungen zur Prähistorischen Archäologie, Band 195, Verlag Dr. Rudolf Habelt GmbH, Bonn.