Investigating an automatic segmentation and classification of archaeological magnetic features through Artificial Intelligence (AI) Techniques.

Ilias-Marios Sarris, Department of Applied Physics, Chalmers University of Technology, Sweden, eliassarris98@gmail.com
Dimitris Chalkiadakis, Instituto de Física Interdisciplinar y Sistemas Complejos CSIC-UIB, Campus Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain, jim.chalkia@gmail.com
Apostolos Sarris, Sylvia Ioannou Chair on Digital Humanities, Department of History and Archaeology, University of Cyprus, Cyprus, asarri01@ucy.ac.cy

Keywords: Archaeology — Magnetic survey — Artificial Intelligence — Machine Learning — Segmentation — Classification

CHNT Reference: Sarris, I-M, Chalkiadakis, D. and Sarris, A. (2022). 'Investigating an automatic segmentation and classification of archaeological magnetic features through Machine Learning Techniques.', in CHNT Editorial board. *Proceedings of the 27th International Conference on Cultural Heritage and New Technologies, November 2022.* Heidelberg: Propylaeum.
DOI: xxxxxxx.

Introduction

Due to the recent advances of automated cart-based or tractor-pulled multi-sensor collection of geophysical data (mainly through magnetic and ground penetrating (GPR) techniques), extensive cultural landscapes are scanned creating a massive volume of data. Unavoidably, despite the need of the final interpretation of the data, namely giving meaning to the actual geophysical features, a preliminary assessment of the geophysical anomalies is required to have an initial evaluation of the quality of the geophysical anomalies and a preliminary interpretation of the geophysical anomalies. Thus, the archaeo-geophysical community has started to investigate ways for the automatic recognition of the geophysical anomalies using either object-oriented approaches (Pregesbauer, M., Trinks, I., and Neubauer, W., 2014) or employing Artificial Intelligence (AI) techniques. The later have been mainly used to GPR depth slices using Machine Learning (ML) and Deep Learning (DL) techniques (Küçükdemirci and Sarris, 2019, 2020, 2022, Green and Cheetham, 2019, Green 2020, Manataki, Vafidis and Sarris A. 2021).

The current presentation deals with the experimentation of the application of different AI algorithms on images produced through the processing of magnetic data from various archaeological sites of Greece. The aim of the project was to experiment with different AI algorithms to propose different segmentation maps that could depict archaeological features in compliance to the suggested magnetic anomalies.

Methods and Tools

Input data consisted of magnetic images resulting from the magnetic prospection of a number of Neolithic settlements (*magoules*) in the area of Thessaly, Central Greece, that have been prospected under the IGEAN project (Sarris et al., 2017). Magnetic surveys covered large extents of the interior of the settlements and their surroundings. Data were processed according to the pipeline that is presented by Sarris (2020). Different anthropogenic features were recognized from the particular surveys including ditches, burnt daub based structures and stone based structures (unburnt houses).

Experiments for the detection of archaeomagnetic features were carried out using different algorithms. Initial efforts were based on Convolution Neural Networks (CNN), based on the U-net architecture and using 25 different archaeological sites as input data. The advantage of the U-net encoder is that it is designed to operate with small number of datasets (Ronneberger, 2015). Due to the imbalance of classes in the available images, in this part of the experiments, it was considered essential to merge all the three suggested magnetic anomalies (ditches, daub made and stone made houses) to one category, creating a Boolean mask of the features (annotated data) contained within the Neolithic settlements. The input magnetic images were grayscale .jpg files (namely 1 channels raster images) with a resolution of about 4000x4000 pixels each. At the same time, in order to increase the number of the available images, data augmentation was carried out through cropping of the images and the corresponding masks to smaller tiles, rotation, contrast and noise techniques.

Further trials were carried out through the use of the Random Forest classifier as a supervised ML classifier technique which is usually applied mainly to segmented single band or multi band images and based on the training features it classifies the rest of the image avoiding an overestimation of the classified features. ArcGIS Pro was employed for the particular process. The algorithm is based on a number of decision trees (the higher the number of decision trees (maximum number of trees) corresponds to a higher degree of classification accuracy) of different order of importance that derive for each pixel from the training dataset. By selecting a random subset of the training data and running the algorithm a number of times, a different decision tree is created and the final classification will be based on the significance degree of each decision tree avoiding overfitting of the data. In order to apply the classification scheme, magnetic data were treated as elevation measurements and a pseudo local relief models (LRM) (Bofinger and Hesse, 2011, Novák, 2014). of the magnetic data was generated which was subsequently segmented. Classification was based on both the segmented images together with the original images and the results of the supervised Random Forest classification were compared to the results derived by an unsupervised Isocluster classification (Isocluster of 3, 5, 10 categories).

Discussion of Results

The above processes proved very successful for the segmentation of the geomagnetic features. U-net experiments were sufficient in delineating anomalies that are correlated to archeological features in close correlation to the annotated data (Fig. 1). Random Forest classifier refined the classification of the anomalies. The supervised classification with maximum number of trees 25 and maximum tree depth of 60 was able to identify the different categories of the anthropogenic features and some further subtle magnetic anomalies in a much better way than the unsupervised isocluster classification (Fig. 2). It is obvious that further experimentation is needs with different AI frameworks (Küçükdemirci and Sarris 2022).

The importance of annotating the geophysical anomalies with an archaeological meaning has been stressed by Verdonck, et al. (2019). It is a crucial step that will lead towards a much better improvement of the application of AI techniques for the segmentation and the preliminary interpretation of the geophysical data. Working in this direction, the Digital Humanities GeoInformatics Lab of the University of Cyprus has created an open access digital depository of annotated datasets (different categories of features) resulting from various geophysical techniques that can further be populated with geophysical images and annotated data to be used to the particular processes (geosignatures.ucy.ac.cy).

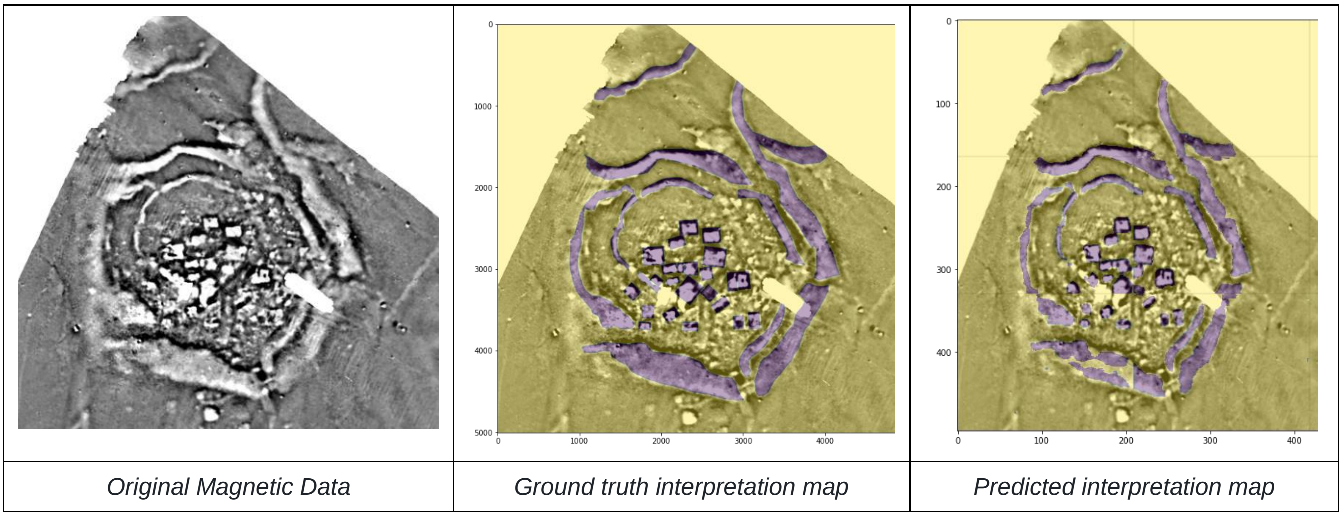

Fig.1: Results of U-net architecture for the classification of the geomagnetic anomalies from the Almyros 2 magoula.

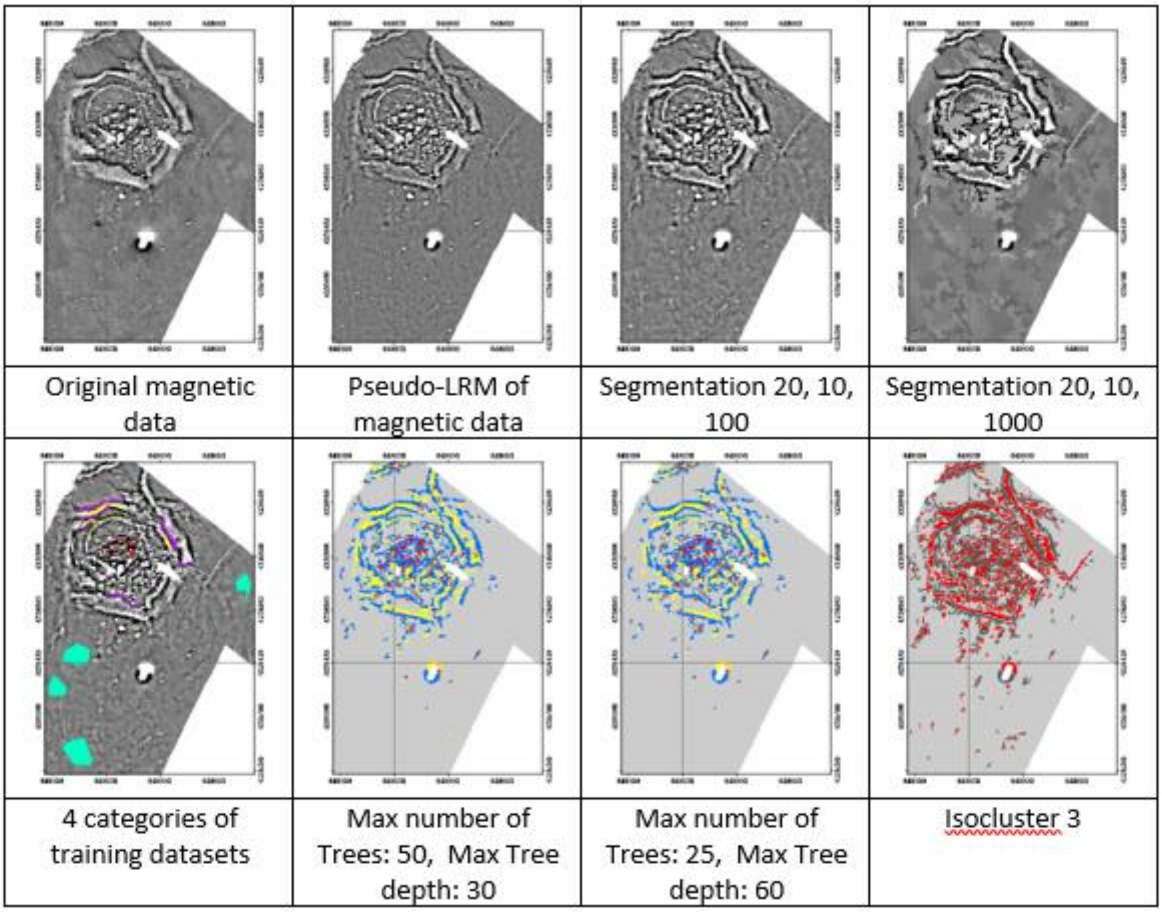

Fig.2: Different stages of the application of Random Forest Classifier to the magnetic data from the Almyros 2 magoula.

Acknowledgement and Funding: Part of the research was supported by a visitors' travel grant of the Dutch Research Council (NWO) at the University of Leiden.

Conflict of Interests Disclosure

There are no any financial or personal relationships with other individuals or organisations that could make the results of this research to appear biased or influenced.

Author Contributions:

Conceptualization: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Data curation: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Formal Analysis: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Funding acquisition: A. Sarris
Investigation: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Methodology: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Project Administration: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Resources: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Software: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Supervision: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Validation: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Visualization: N. Chetogiannaki, A. Sarris
Writing – original draft: I.-M. Sarris, D. Chalkiadakis, A. Sarris
Writing – review & editing: I.-M. Sarris, D. Chalkiadakis, A. Sarris

References:

Bofinger, J. and Hesse, R. (2011). As far as the laser can reach… Laminar analysis of LiDAR detected structures as a powerful instrument for archaeological heritage management in Baden-Württemberg, Germany. In: Cowley (ed), Re-mote Sensing for Archaeological Heritage Management, EAC Occasional Paper No. 5, 163-173.

Green, A. and Cheetham, P. (2019). Rise of the Machines: Improving the identification of possible graves in GPR data with interactive survey guidance and machine learning,*13th International Conference on Archaeological Prospection*, Sligo, Ire-land, New Global Perspectives on Archaeological Prospection ISBN 978-1-78969-306-, pp.300-304.

Green, A. (2020). Detecting Graves in GPR Data: Assessing the viability of machine learning for the interpretation of graves in B-scan data using medieval Irish case studies. Phd Thesis. Bournemouth University.

Küçükdemirci M. and Sarris A. (2019) Automated segmentation of archaeo-geophysical images by convolutional neural networks, *13th International Conference on Archaeological Prospection*, Sligo, Ireland, New Global Perspectives on Ar-chaeological Prospection ISBN 978-1-78969-306-, pp.295-299.

Küçükdemirci M. and Sarris A. (2020). Deep learning based automated analysis of archaeo-geophysical images. *Archae-ological Prospection*. 27:107–118. https://doi.org/10.1002/arp.1763.

Küçükdemirci, M. and Sarris, A. (2022). GPR Data Processing and Interpretation Based on Artificial Intelligence Ap-proaches: Future Perspectives for Archaeological Prospection. *Remote Sensing*, 14, 3377, https://doi.org/10.3390/rs14143377

Manataki, M., Vafidis, A. and Sarris, A. (2021). GPR Data Interpretation Approaches in Archaeological Prospection. Appl. Sci. 2021, 11, 7531.https://doi.org/10.3390/app11167531.

Novák, D. (2014). Local Relief Model (LRM) Toolbox for ArcGIS. https://onedrive.live.com/re-dir?resid=4A4965781C37FFAC!90606&authkey=!AKcM9WetyjhaEAA&ithint=folder%2czip

Pregesbauer, M., Trinks, I., and Neubauer, W. (2014). An object-oriented approach to automatic classification of archae-ological features in magnetic prospection data. *Near Surface Geophysics* 12(5), 651–656. https://doi.org/10.3997/1873-0604.2014014.

Ronneberger, O., Fischer, P. and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, 234-241. Springer, Cham.

Sarris, A., Kalogiropoulou, E., Kalayci, T. & Karimali, L. (eds.) (2017). Communities, Landscapes, and Interaction in Neolithic Greece. Proceedings of International Conference, Rethymno 29-30 May 2015. Ann Arbor, MI: International Monographs in Prehistory.

Sarris. A. (2020). Processing and analyzing Geophysical Data. In Archaeological Spatial Analysis: A Methodological Guide, ed. by Mark Gillings, Piraye Hacıgüzeller and Gary Lock, 1st Edition (January 29, 2020), Routledge, pp. 376-407, ISBN 9780815373230 - CAT# K338400.

Verdonck, L., De Smedt, P. and Verhegge. J. (2019). Making sense of anomalies: practices and challenges in the archaeological interpretation of geophysical data. In R. Persico, S. Piro and N. Linford (eds.), Innovation in Near-Surface Geophysics. Instrumentation, application, and data processing methods, 151–194. Amsterdam.