# Mag-Net

## Improving magnetometer interpretation workflows with semantic segmentation

Julien WOLF, Magnitude Surveys, United Kingdom
Finnegan POPE-CARTER, Magnitude Surveys, United Kingdom
Paul JOHNSON, Magnitude Surveys, United Kingdom

***Keywords:*** *Machine Learning — Semantic Segmentation — Automatic Interpretation —Magnetometry — Commercial Geophysics*

**CHNT Reference:** add full reference here: Wolf, Julien, Pope-Carter, Finnegan, Johnson, Paul. 2020. Magnetising Neural Networks. Improving magnetometer interpretation workflows with semantic segmentation. Editorial team. Conference on Cultural Heritage and New Technologies (Proceedings of the 25th International Conference on Cutural Heritage and New Technologies 2020). DOI:xxxxxxx.

Commercial geophysics in Britain has met the increasing demands for large scale surveys with advances in motorised and modular survey systems. By rising to these physical challenges, however, a surge in data quantities poses new logistical challenges in keeping reporting turnover time low whilst focusing on producing qualitative interpretations and discussion. This paper presents preliminary findings of a development-led project addressing this issue. The aim is to develop a convolutional neural network (CNN) for the detection of anomalies in geophysical data intended for secondary (human) archaeological classification.

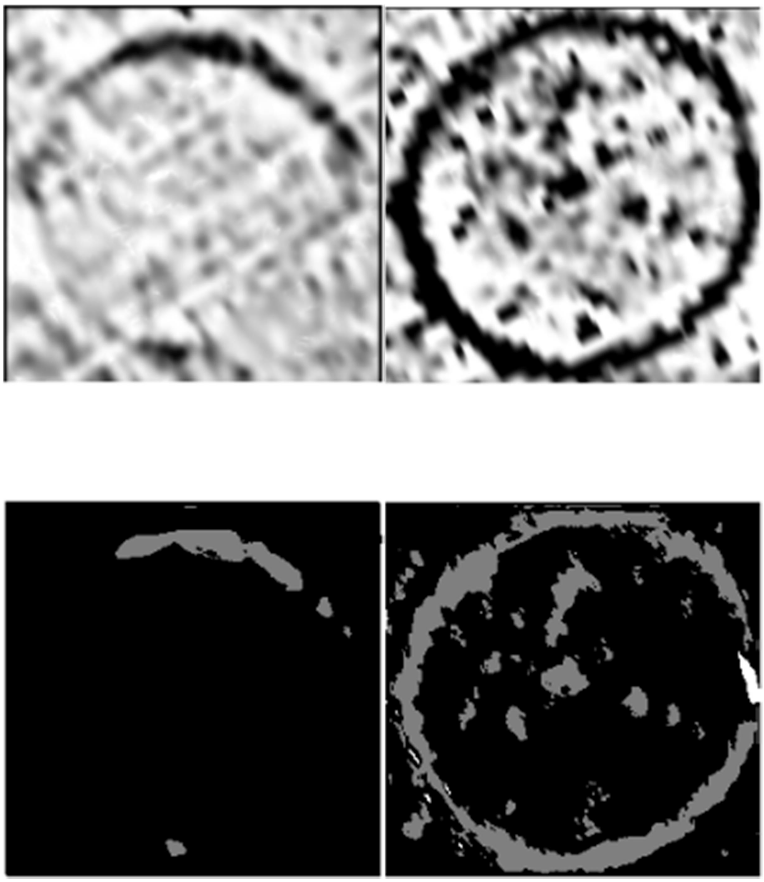
The project is currently exploring the limitations of machine learning in geophysics and complications arising from the nature of magnetometer data. Central to this topic are the choices in pre-processing methods, network architecture and hyper parameters and how they can affect and improve training results. As this is an ongoing project, this paper will present current methodology and results while focusing on the following questions:

How the quantity and quality of the data affects the development of CNN, and where this is affected specifically by the nature of geophysical data.

What types of network architectures and hyper parameter choices are particularly suited to detecting anomalies in magnetometer data?

How the development of a natively archaeological CNN, trained exclusively on geophysical data, may improve the algorithm's robustness in detecting anomalies of potentially archaeological nature in comparison with a transfer-learned network.

This pilot study assesses the feasibility of magnetometer data for semantic segmentation using a subset of data collected across numerous commercial surveys in Britain over the last five years. The current focus lies on training the network on ring ditch anomalies, which were chosen for having distinct recognisable features while being more complex than simple linear or point anomalies. The data varies in terms of scale, survey conditions and geological backgrounds, amongst other factors. Broad natural anomalies or anthropogenic activities in the background such as modern manuring practices can further complicate the data and reduce the likelihood of detection by machine algorithms. Initial results indicate that changes in the geological background which affect the clarity of the archaeological features can particularly impact the ability of the network to detect ring ditches, especially where these are ephemeral in nature (cf. Fig. 1).



*a.b.*

*Fig. 1. Gradient magnetometer data (above) with predictions (below) from a neural network based on Ronneberger et al.'s (2015)* Unet *architecture after 50 epochs. Black indicates background, grey archaeology, and white null-data. Ephemeral features such as a. are more difficult to pick up than strong features such as b.*

The machine learning algorithms used in this study are realised in Python using the Tensorflow and Keras libraries. These comprise of a Python API to integrate with Magnitude's bespoke PostGIS database and in-house digital archive. The API is managed using a Git version control system; version references are stored alongside parameter configurations with the results of each training iteration for reproducibility. As the desired output is a pixelwise classification, the network architectures investigated are based on variations of the networks introduced by Shelhamer et al. (2016) and Ronneberger et al. (2015) which both employ up-sampling to produce output labels with the same dimensions as the given inputs.

A key objective in this study is to assess the commercial viability of training networks only on archaeological data instead of using pre-trained models. Recent applications of CNN to archaeology have so far profited from the availability of large open datasets such as the ImageNet library, which reduce the requirements in terms of data quantity by pre-training a network on non-archaeological data. This allows for the application of CNN to smaller datasets, which are more common in archaeology. The authors however expect networks trained solely on archaeological data to be more robust in their predictions, as the nature of the data in libraries such as ImageNet can prove too divergent from archaeological data, both in terms of data formatting as well as complexity and distinctness (Trier et al. 2018, 227). Conversely, training on the above mentioned architectures can be timely and resource intensive, hence it is therefore also the aim of this project to find whether significant improvements can be made in this manner, or whether it is more cost effective to invest time in fine tuning network parameters instead. These observations not only apply to commercial applications but are also relevant in an academic climate afflicted by temporal and financial constraints.

In addition to the issues specific to geophysics, the project is also faced with issues common in archaeology pertaining to data collection. As the dataset comprises results from a few hundred surveys, a workflow was established as part of the above-mentioned API to process and organise the data in a standardised fashion. The final goal of this project is to integrate the training and testing of the network into Magnitude’s existing data workflow as an additional processing step prior to human interpretation.

## References


Ronneberger, O., Fischer, P. and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs]. [online]. Available at: http://arxiv.org/abs/1505.04597 (Accessed January 17, 2020).

Shelhamer, E., Long, J. and Darrell, T. (2016). Fully Convolutional Networks for Semantic Segmentation. arXiv:1605.06211 [cs]. [online]. Available at http://arxiv.org/abs/1605.06211 (Accessed January 22, 2020).

Trier, O.D., Salberg, A.-B. and Pilo, L.H. (2018). Semi-automatic mapping of charcoal kilns from airborne laser scanning data using deep learning. In M. Matusmoto & E. Uleberg, eds. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology. Oceans of Data - Computer Applications and Quantitative Methods in Archaeology. Oslo: Archaeopress Publishing Ltd, pp. 219–231. Available at: https://www.researchgate.net/publication/326771506_Semi-automatic_mapping_of_charcoal_kilns_from_airborne_laser_scanning_data_using_deep_learning (Accessed July 25, 2019).