Machine learning for large area archaeological feature detection

Applying transfer learning to airborne lidar data

Jürgen LANDAUER, Landauer Research, Germany
Ralf HESSE, State Office for Cultural Heritage Baden-Württemberg, Germany

Keywords: airborne lidar, prospection, automated detection, machine learning

Introduction

The application of airborne lidar data for archaeological prospection has so far largely been based on interpretation and manual mapping by humans. Increasingly, attempts are being made to apply algorithms to automatically detect specific types of archaeological features. Automation is hoped to allow the processing of very large area lidar data sets, the output of which would be useable for archaeological research as well as heritage management. The rapid progress in the field of machine learning and the increasing availability of algorithms and software as well as lidar data opens new opportunities for lidar-based archaeological prospection.

Materials and methods

Data set

In Baden-Württemberg, a state-wide airborne lidar data set is available. Since 2009, this data set has been used for full area coverage archaeological prospection, resulting in more than 900,000 manually mapped archaeological relief features. One of the limitations of machine learning approaches in the field of archaeology is the need for learning data sets containing large numbers of examples. The results of the manual lidar-based prospection in Baden-Württemberg with hundreds to thousands of mapped archaeological features per class are a good basis for such learning data sets. To develop and test the machine learning approach, a single type of archaeological features – charcoal burning platforms – is used, because they are morphologically well defined, numerous and relatively easy to recognize by humans. The manual mapping results largely lack quality control as field verification for tens of thousands of features is not feasible. Therefore, a first step was to establish a quantitative measure for the recognisability of each feature by a human expert. This is achieved by having each of the c. 29,000 charcoal burning platforms assessed multiple times by multiple users and assigning it to one of five recognisability classes from “certainly not” to “definitive”. The charcoal burning platform data set contains c. 29,000 snippets of 40 x 40 m resampled to 25 cm resolution. For each snippet, four different visualisations are supplied as as greyscale images, namely vertical Shaded Relief and Local Relief Model as well as greyscale averages of Sky-View Factor and Laplacian-of-Gaussian and Local Dominance and Laplacian-of-Gaussian, respectively. For each snippet, a recognisability score is provided. All geolocation information has been removed, absolute elevations have been replaced by relative elevations within the snippet, and snippets may be rotated randomly.

Transfer learning for archaeological feature detection

Machine learning has been adopted for the archaeological domain for a while now (e.g. see [1]): In general, artificial neural networks are trained with a large sample data set (typically ranging from thousands to millions) to detect features in data similar to these samples. Recent progress in this technology allows for developing this approach further: With so-called “Transfer Learning” methods, the sample data sets can be made substantially smaller (and thus more appropriate for relatively small archaeological data sets) without quality losses.

In this project the RESNET neural network architecture [1] is used, which has been pre-trained with several 100,000s of (non-archaeological) images. In order to apply it to the archaeological domain, a small set (merely several thousand) of lidar images of charcoal burning platforms are inserted into the learning pipeline of the pre-trained network. By doing so, its ability to recognize images is specialised (or “transferred”) into recognizing these charcoal platforms with relatively high accuracy: more than 85 percent of the approx. 30,000 of image snippets mentioned above are classified correctly within a few minutes. The

setup used is either a medium size PC with an Nvidia[tm] graphics card and 8 GB of memory or a cloud-based server from the Google Colab project [4] with similar hardware available to it.

The software is based on the highly efficient Fast.ai V1.0 library [2], which originally was developed by researchers from the University of San Francisco (released in October 2018)

Examples of the output are shown in Fig. 1: Along with a prediction of the type of feature possibly contained within a tile, a confidence rating is computed for each snippet. As a final quality enhancement step, an archaeologist can then review ratings with relatively low confidence and possibly adjust the prediction. This manual step further enhances the overall recognition rate.

With the degree of automation achieved here, large-scale detection projects have become feasible: In current experiments, automatic detection is tested for 100 km^2 areas. For this, the software splits the area into 40 x 40 m tiles (with 50 percent overlap, see Fig. 2) and each tile is then presented to the neural network for classification. Various factors influence the recognition rate, e.g. landscape form, but currently close to 80 percent recognition rate can be achieved even with this more challenging setup.

Conclusions and outlook

The application of machine learning for lidar-based archaeological prospection allows to rapidly process large area data sets and detect archaeological relief features with a recognition rate or c. 80% in the test case with charcoal burning platforms.

Very recent progress (2018 and 2019) in Transfer Learning technologies seemingly promises to further reduce the number of sample images needed for training. Manual effort is therefore greatly reduced, and it is planned to add further categories into the recognition pipeline, for example barrows, hollow ways, and ridge-and-furrow.

Moreover, the objective is to largely eliminate the need for manual quality control. This will require a stepwise improvement of the processing pipeline, but recognition rates above 90 percent seem to be possible.

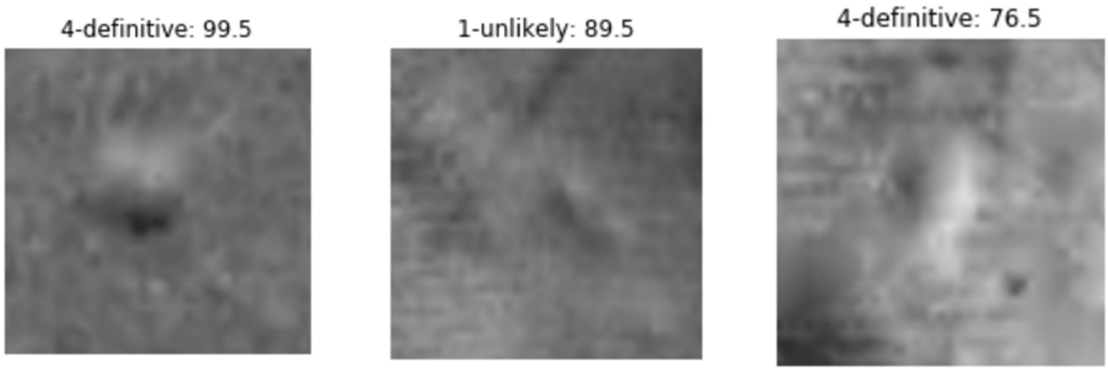

Fig. 1. Lidar snippets and its automatic classification: headlines show classification result (definitive, unlikely) and the confidence rating (in percent)

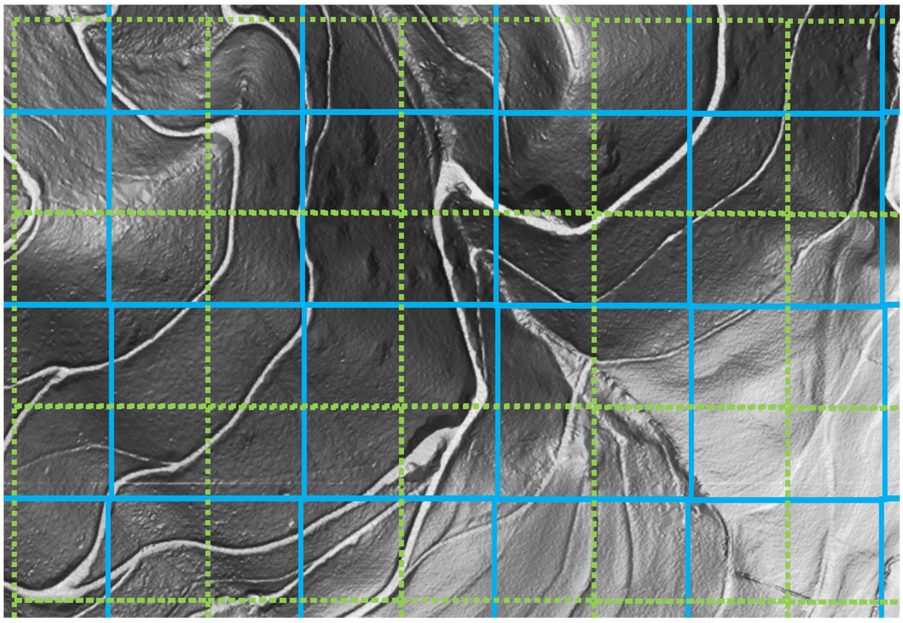

Fig. 2. Schematic grid with overlapping tiles (not to scale). The software will mark tiles with detected archaeological features (here: charcoal burning platforms)

References

Kramer, I. (2019). “Navigating the AI minefield: A step-by-step workflow to apply Deep Learning to detect Archaeological Sites”. In: Symposium on Interdisciplinary Machine Learning at Southampton University.

He, K: et al. (2016). “Deep Residual Learning for Image Recognition.” In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 770-778. https://arxiv.org/abs/1512.03385

https://www.fast.ai/ (Accessed: 18 July 2019)

https://colab.research.google.com (Accessed: 18 July 2019)