Discovering the Urbanity of Nomad Mongolia

Using Locational Modelling and Deep Learning for Space-Borne Large-Scale Survey

Tom Carl GRENDA, HTW Dresden, Germany
Hendrik ROHLAND, DAI Bonn / HTW Dresden, Germany
Marco BLOCK-BERLITZ, HTW Dresden, Germany
Christina FRANKEN, DAI Bonn, Germany

Keywords: *Remote Sensing—Deep Learning—Survey—Nomadic Urbanism—Locational Modelling*

CHNT Reference: Grenda, T. C., Rohland, H., Block-Berlitz, M. and Franken, C. (2022). 'Discovering the Urbanity of Nomad Mongolia. Combining Deep Learning with Space-Borne Large-Scale Survey', in CHNT Editorial board. *Proceedings of the 27th International Conference on Cultural Heritage and New Technologies, November 2022.* Heidelberg: Propylaeum.
DOI: xxxxxxx.

1 Motivation and Introduction

The German Archaeological Institute (DAI) and its Partners (Mongolian Academy of Sciences, National University of Mongolia and HTW Dresden since 2018) are researching the development of urban settlements in early medieval and medieval Mongolia. However, to further understand the process of urbanisation on the steppe over time, data on supra-regional scale is necessary, that allows for the analysis of the system of urban and proto-urban settlements. To get closer towards such a dataset as the basis for further research, we combine a GIS-based site preference modelling and machine learning-enabled analysis of remote sensing data.

2 Aims and Methods of the Study

In Mongolia, there is no large coverage of lidar-data available and data. At the same time, the arid conditions, the sparse vegetation, and the extensive land-use in most parts of Mongolia lead to an excellent preservation of archaeological remains of urban and proto-urban settlements, such as walls and building platforms (Fig. 1). These conditions also provide a generally good ground visibility. Furthermore, the structures we are interested in are usually quite large, in most cases measuring several tens or hundreds of metres, and they often feature long linear structures such as rectangular walls. This means that they can be detected using mid-resolution (6–12 m ground sampling distance (GSD)) satellite imagery. These provide sufficient detail for our aims and can still be handled, even if large tracts of land are to be surveyed. Even then, Mongolia is still a very large country which makes it necessary to narrow down the study areas. Therefore, we employed site preference modelling to identify the regions where the occurrence of urbanised settlements is possible, given the environmental conditions in which known settlements occur, and defined areas of interest for the remote sensing application.

Fig. 1. In Mongolia, many traces of ancient walled sites and cities can be found, but only few of them are as well preserved and easily recognizable as the site of Tsagaan Sumyn Balgas in the Khotont Sum of Arkhangaj Province. (Photo: © Hendrik Rohland)

2.1 Aims of the Study

With our contribution we want to present preliminary results and discuss the methodology of the pilot study for our further work. Our study aims to test the viability of different types of mid-resolution satellite data for large scale archaeological surveys, using Deep Learning to scale up the surveyed areas. Image data from platforms such as Sentinel-2 and PlanetScope will be compared to assess how Deep Learning Algorithms perform on detecting larger archaeological features using different types of input data, including RGB-data and different vegetation indices. Given a high performance of the model, it shall be used for the further archaeological exploration and monitoring of large swaths of land, especially in Mongolia, contributing to the building of a comprehensive dataset of urban sites of the country.

2.2 Site Preference Modelling and Test Site Selection

To narrow down, simple, rule-based site-preference modelling has been conducted. More than 100 known proto-urban and urban sites from all over Mongolia have been collected from the literature and mapped using a GIS. Basic, descriptive statistics have been used to identify the range of several environmental variables, within which the known settlements occur. The acquired information has been used to identify the areas, in which all considered variables are within the ranges of the dataset of known sites. The resulting area has been defined as the overall survey area, from which we selected a smaller part with sufficient training-and test data for the pilot study. The selected test site is an area in the Archangaj, Övörchangaj and Bulgan provinces of Mongolia, where a rich archaeological heritage of the historical nomad empires such as the Hunnu, Turk, Uyghur, Mongol and Manju empires are located.

2.3 Progress of Deep Learning and their Applications

Within the last decade multiple new Deep Learning architectures have emerged, among them Convolutional Neural Networks (CNN). Within the field of Computer Vision (CV), this architecture are is frequently employed and solves complex image processing tasks, which could previously only be handled by humans on a much smaller scale.

In recent years, Deep Learning models became more complex by integrating more layers and algorithms in order to improve overall prediction accuracy, sometimes even surpassing human accuracy. Especially when remote sensing data with high spatial and temporal resolution is to be analysed, a Deep Learning approach offers a valid alternative to manual analysis.

3. Deep Learning: Net Architecture and Training

Based on the archaeological sites identified in the literature and from manual survey, satellite imagery with a GSD of 6 m was acquired from Planet Labs[1]. The images contain data on multiple spectral bands, ranging from the visible light to near infrared spectrum.

As part of the image preprocessing various vegetation indices such as the Normalized Difference Vegetation Index (NDVI) are calculated using the different image bands, resulting in new output images. Those output images are then annotated using the Computer Vision Annotation Tool (CVAT)[2] maintained by Intel. The images along with the corresponding bounding boxes of the archaeological sites are exported and split into training, testing and validation subsets and propagated through the Faster R-CNN model.

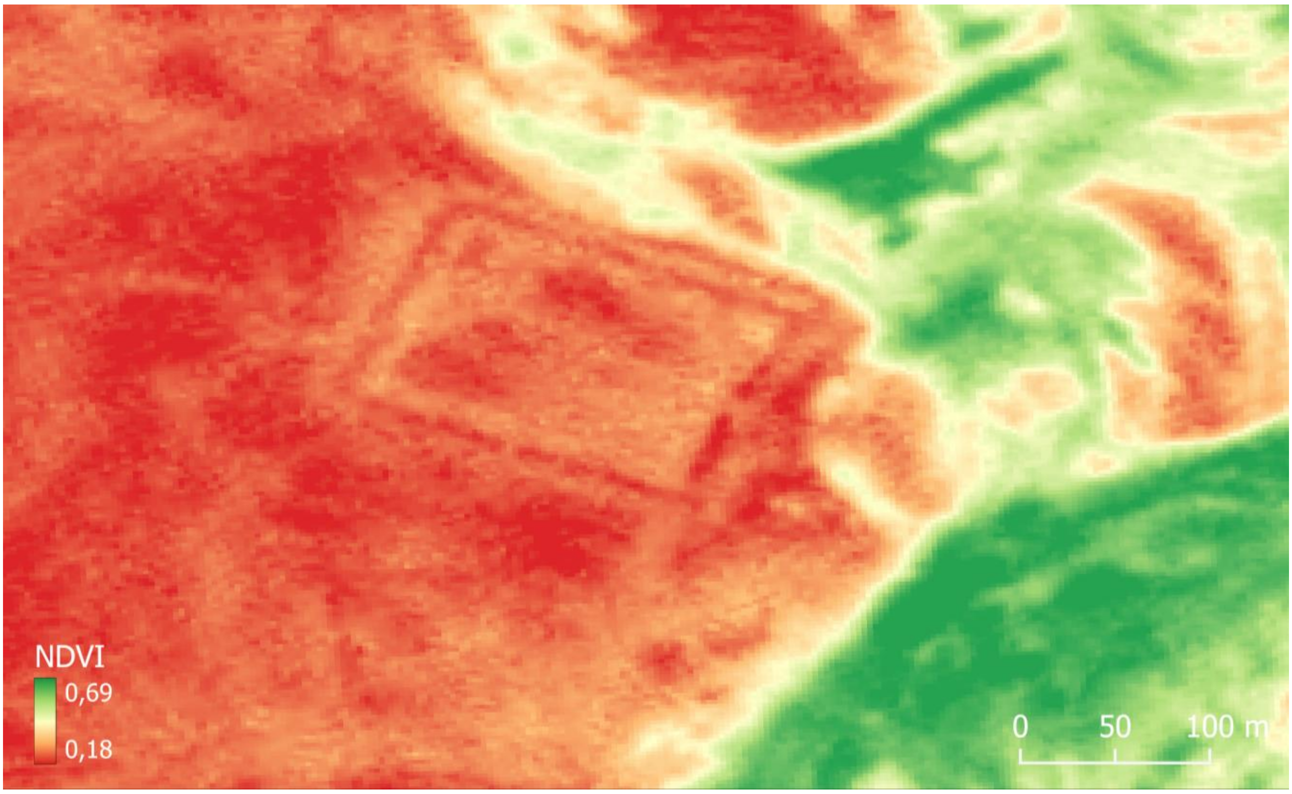

Fig. 2. The calculation of vegetation indices such as NDVI can help to enhance the visibility of archaeological structures, which have an influence on the plant health because of compression or disturbance of the soil. (© Image Data Planet.com; Visualization T C. Grenda)

The training of the model is based on the Faster R-CNN architecture proposed by Ren et al. 2017. This represents a further development of the Fast R-CNN architecture, which was introduced by Girshick 2016 and builds directly on the original R-CNN model architecture published by Girshick et

[1] https://www.planet.com/

[2] https://openvinotoolkit.github.io/cvat/v2.0.0/about/

al. 2014. R-CNN is one representation of a so-called two-stage classifier in which the object boundaries and classes are determined in two successive process steps. To further improve training accuracy a ResNet 50 backbone as originally presented by He et al. 2016 is implemented.

5. Conclusion and Future Work

After first testing and results, the approach seems promising, but there surely will be room for improvement. Currently the site location model is very generalised and the data was drawn from sites ranging from Xiong-nu to Manju times. For analytical purposes, it will be interesting to look at differing site preferences at different times. However, many of the known sites are barely studied and the dating is often-times uncertain. This underlines the need for further studies on the ground, obtaining high-resolution survey data such as DEMs and artefact mappings. This will open up the path to a comparative study of such sites within Inner Asia.

On the part of the Machine Learning application, the application of a One-Stage Classifier can be tested to speed-up the detection process. Further potential lies in the realm of digital image processing of remote sensing data. The calculation of different indices, not only focusing on vegetation marks, but also properties of the soil such as moisture and composition, seems to be a promising approach. Another interesting possibility is the application of image enhancement using Super-Resolution by either applying conventional approaches as described by Nasrollahi 2014 or more recent approaches leveraging Deep Convolutional Networks as introduced by Dong et al. 2016 to upscale low-resolution images in order to detect more features on the surface.

Bibliography

Dong, C., Loy, C. C., He, K. and Tang, X. (2016). 'Image Super-Resolution Using Deep Convolutional Networks', in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, doi: 10.1109/TPAMI.2015.2439281.

Girshick, R., Donahue, J., Darrell, T. and Malik, J. (2014). 'Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation', *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580-587, doi: 10.1109/CVPR.2014.81.

Girshick, R. (2015). 'Fast R-CNN', *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

He, K., Zhang, X., Ren, S. and Sun, J. (2016). 'Deep Residual Learning for Image Recognition', *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, doi: 10.1109/CVPR.2016.90.

Nasrollahi K., Moeslund T. B. (2014). 'Super-resolution: A comprehensive survey', *Machine Vision and Applications*, Volume 25, Issue 6, pp. 1423–1468, doi: 10.1007/s00138-014-0623-4.

Ren, S., He, K., Girshick, R. and Sun, J. (2017). 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks', in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, doi: 10.1109/TPAMI.2016.2577031.