



Universiteit  
Leiden  
The Netherlands

## **Erfassung von Meilerpodien in Lidar-Daten: von der manuellen Kartierung über Crowdsourcing zum maschinellen Lernen**

Hesse, R.; Verschoof-van der Vaart, W.B.

### **Citation**

Hesse, R., & Verschoof-van der Vaart, W. B. (2019). *Erfassung von Meilerpodien in Lidar-Daten: von der manuellen Kartierung über Crowdsourcing zum maschinellen Lernen*.

Retrieved from <https://hdl.handle.net/1887/3217470>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3217470>

**Note:** To cite this publication please use the final published version (if applicable).

# Erfassung von Meilerpodien in Lidar-Daten:

## Von der manuellen Kartierung über Crowdsourcing zum maschinellen Lernen

Ralf Hesse, Landesamt für Denkmalpflege Baden-Württemberg

Wouter B. Verschoof-van der Vaart, Universität Leiden

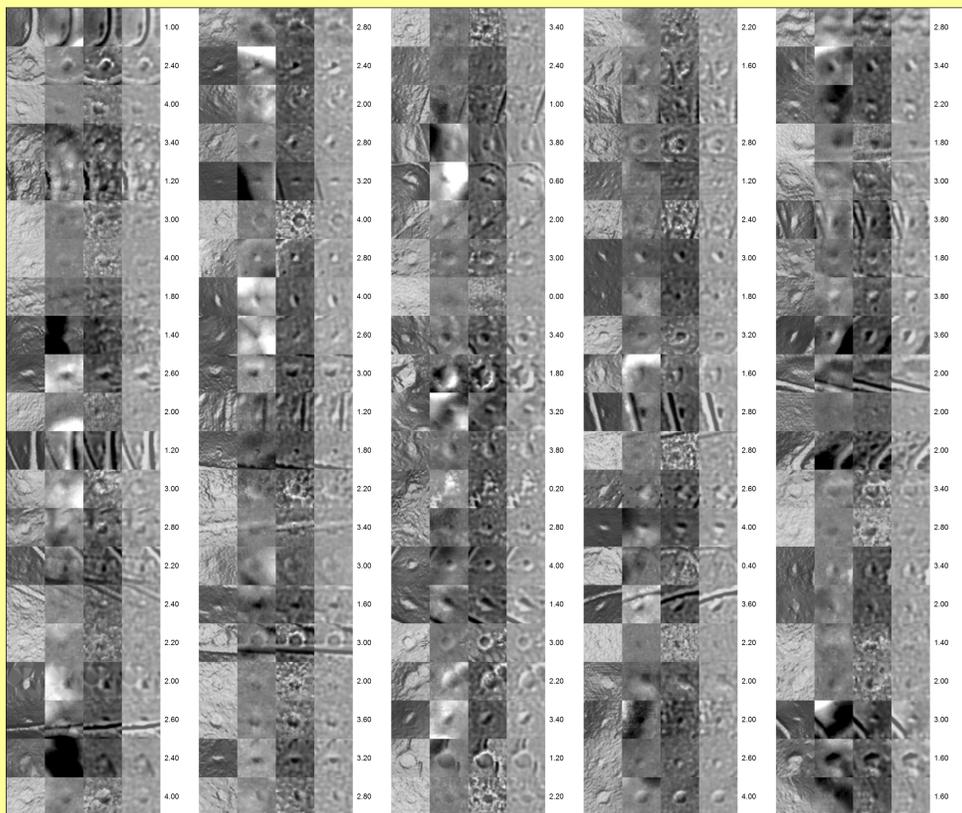
### Einleitung

Seit dem Jahr 2009 werden Lidar-Daten in Baden-Württemberg zur flächendeckenden archäologischen Prospektion verwendet. Insgesamt konnten bisher nahezu eine Million archäologische Reliefmerkmale kartiert werden, darunter ca. 30.000 Meilerpodien. Zwar können durch die Nutzung verschiedener Lidar-Visualisierungen viele nur reliktsch erhaltene Strukturen sichtbar gemacht und erfasst werden, eine qualitative Bewertung der bisherigen manuellen Kartierung steht aber aus. Zukünftig werden die Ergebnisse der flächendeckenden Lidar-Neubefliegung Baden-Württembergs mit höherer Auflösung verfügbar werden. Da eine erneute manuelle Bearbeitung sehr zeitaufwendig wäre, ist eine zumindest teilweise maschinelle Auswertung wünschenswert.

Beide Zielstellungen – die Bewertung der bisherigen Kartierung und die automatisierte Auswertung weiterer Lidar-Daten – sind mittels maschinellem Lernen erreichbar. Eine wesentliche Voraussetzung hierfür ist die Erstellung umfangreicher und qualitativ hochwertiger Trainingsdatensätze.



Je nach Topographie und Ausprägung sind Meilerpodien in verschiedenen Lidar-Visualisierungen unterschiedlich deutlich zu erkennen



Meilerpodien in verschiedenen Lidar-Visualisierungen mit Mittelwert aus fünf Bewertungen

### Ziel: Trainingsdatensätze für maschinelles Lernen

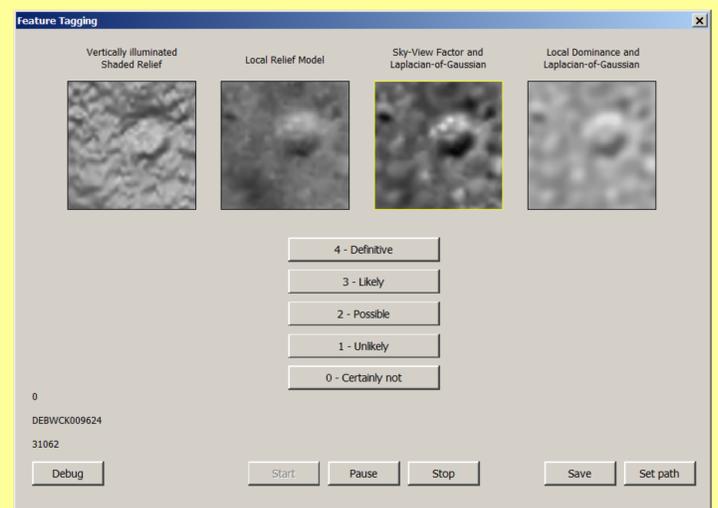
Ein grundlegendes Problem für Entwicklung und Vergleich von automatisierten Verfahren zur Detektion archäologischer Objekte in Lidar-Daten ist häufig der Mangel an ausreichend großen Trainingsdatensätzen. Dieser Mangel wird insbesondere im Vergleich mit dem 20,000 Kategorien und 14 Millionen Bilder umfassenden ImageNet-Datensatz deutlich. Aus den Ergebnissen der flächendeckenden Lidar-basierten Prospektion in Baden-Württemberg können derartige Trainingsdatensätze erstellt werden. Da Meilerpodien morphologisch gut beschreibbar und für Menschen gut erkennbar sind, sind sie für Entwicklung, Test und Vergleich von Algorithmen zur Mustererkennung und maschinelles Lernen besonders geeignet und wurden daher für die Entwicklung eines ersten Trainingsdatensatzes herangezogen.

Die Trainingsdatensätze bestehen aus Zusammenstellungen von kleinen (40x40 m), nicht geographisch lokalisierbaren Ausschnitten in Form von Rasterdaten (Graustufenbilder verschiedener Visualisierungen und relative Höhen). So soll ermöglicht werden, automatische Detektionsalgorithmen sowohl auf der Basis von Bilddaten als auch auf der Basis von Höhendaten zu entwickeln. Da Lidar-Daten in vielen Fällen nicht weitergegeben werden können oder sollen, beinhaltet der Datensatz quasi-synthetische, nicht geographisch lokalisierbare Ausschnitte. Hierzu werden (a) die Geokoordinaten und (b) die absolute Höhe entfernt sowie (c) die Ausschnitte beliebig gedreht und/oder gespiegelt. Eine Lokalisierung oder eine über die Nutzung als Trainingsdatensatz hinausgehende Verwendung sind damit unmöglich.

### Crowdsourcing

Da die vorliegenden Kartierungsergebnisse durch einen einzigen Bearbeiter erstellt wurden, die Erkennbarkeit von Meilerpodien starken Schwankungen unterliegt und eine Qualitätskontrolle bisher nicht stattgefunden hat, ist eine qualitative oder quantitative Bewertung der morphologischen Ausprägung und der Erkennbarkeit für menschliche Bearbeiter wünschenswert.

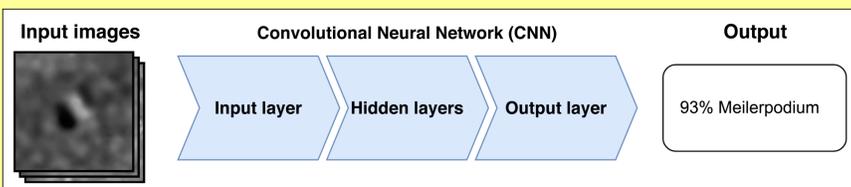
Der Zeitaufwand für die mehrfache Bewertung von nahezu 30.000 Objekten ist sehr hoch. Daher wird ein Crowdsourcing-Ansatz verfolgt, um alle Objekte durch mehrere Bearbeiter mehrfach bewerten zu lassen. Quantitativ auswertbare Daten werden durch eine Bewertung in fünf Stufen von „definitiv erkennbar“ bis „definitiv nicht erkennbar“ erzeugt. Die hierfür entwickelte Benutzeroberfläche erlaubt eine schnelle und unkomplizierte Bearbeitung. Nach Abschluss der mehrfachen Bewertung ist eine Konsolidierung und Homogenisierung der gewonnenen Daten erforderlich.



Benutzeroberfläche für die Bewertung durch Crowdsourcing

### Ausblick

Der für Meilerpodien erstellte Trainingsdatensatz soll für Entwicklung, Test und Vergleich von automatisierten Verfahren zur Erkennung von Meilerpodien angewendet werden. Besonders vielversprechend erscheinen derzeit *Convolutional Neural Networks*. Erste Ergebnisse sollen auf der CAA-Tagung in Krakow im Mai 2019 vorgestellt werden. Schrittweise Erweiterungen des Datensatzes um Beispiele aus anderen Untersuchungsgebieten sowie um weitere Objekttypen (z.B. Grabhügel, Hohlwege, Wölbäcker, Celtic Fields) sind vorgesehen.



Prinzipskizze maschinelles Lernen mit Convolutional Neural Networks