Historical maps inform landform cognition in machine learning

Samantha T. Arundel [a,*], Gaurav Sinha[b], Wenwen Li[c], David P. Martin[d], Kevin G. McKeehan[a], Philip T. Thiem[a]

[a] *U.S. Geological Survey – sarundel@usgs.gov, kmckeehan@usgs.gov, pthiem@usgs.gov*
[b] *Ohio University – sinhag@ohio.edu*
[c] *Arizona State University – wenwen@asu.edu*
[d] *Duke Energy - david.martin.3234@gmail.com*

* Corresponding author

Keywords: Terrain, geomorphometry, topographic maps, feature extraction, deep learning

Abstract:

Landforms are difficult to delineate in the field or on maps because of inherently indeterminate and fiat boundaries (Smith and Mark, 2003). Research and applications of landform delineation have progressed along two paths in the geosciences. *General geomorphometry* is a continuous field-rooted approach that focuses on computing localized parametric values for mapping land surface shape patterns and deriving land segments or elements that can be considered homogeneous at the chosen scale of analysis for a particular application. *Specific geomorphometry* is an object-oriented approach that applies the geomorphometric parameters to identify, delimit, characterize, and classify individual landform objects. There is no standard list of landform types or methods for mapping individual landforms because landform cognition is influenced substantially by people's cultural, linguistic, and individual backgrounds. Thus, delimiting and classifying individual landforms requires a multidisciplinary approach by incorporating geomorphometry, geomorphology, cognitive science, cartography, remote sensing, and geographic information science knowledge. Cartography plays a unique dual role in named landform representation by providing contextual information as input to the demarcation problem while providing a medium for expressing the human cognization of landforms. Hence, this research aims to improve the automated mapping of named landforms.

The authors have been working on specific geomorphometry projects to advance the theory and best practices for cognition-driven automated mapping of topographically salient landforms (Arundel and Sinha 2018; Sinha and Arundel 2020; Arundel and Sinha 2020; Joly, Sinha, and Hassan 2022). In parallel, they are also studying the strengths and limitations of deep learning-based image analysis for delimiting landforms in the United States (Arundel, Li, and Wang 2020; Li et al. 2022). For example, hillshade, slope values, and natural colour imagery from the National Agricultural Imaging Program (NAIP) were used in a convolutional neural network (FasterRCNN and RetinaNet) to predict landform locations (Figure 1). The chief elements hindering desirable results were the need for more training data and assessment of the quality of extracted landform extents. Both challenges are related to the subjective nature of their delimitation and classification.

In the United States and many other countries, well-recognized (i.e., named) landforms are represented on maps through text labels and point symbols. The U.S. Geological Survey historical topographic map collection (HTMC) illustrates mental representations of officially named U.S. landforms defined through feature capture and text placement specifications developed and refined during its century of production. The extent of a landform was loosely depicted mainly by the placement of its name within the spatial context of other map elements, although terrain characteristics also guide their realization. The HTMC has been scanned into raster graphics, but some of its information has yet to be captured in an easily machine-readable format. Deep learning-based image processing and analysis works well for this task (Arundel and Sinha, 2020). This paper reports two-fold research that (1) incorporates the cognitive element of named landforms for detection and recognition and (2) advances methods to translate information from scanned maps into machine-usable form automatically.

Based on insights from previous research, this work replaced the natural colour imagery with HTMC images to explicitly leverage the cognitive aspects of landform extent detection. This was achieved by creating a custom training dataset by loading images of hillshade and slope surfaces and HTMC images into the RGB bands of single images. This highlights the landform name text on the map, although all loading combinations have been tried with similar results. These images depict the terrain surrounding manually digitized landform boundaries in a GIS file containing almost 90,000 features in 25 classes representing some of the named landform features in the Geographic Names Information System (GNIS;

https://www.usgs.gov/us-board-on-geographic-names/domestic-names), a database of officially recognized topographic features in the United States. Our work with these features began with the GNIS Summit feature class, which contains mountains, some ridges, peaks, and sometimes hills.

Preliminary results indicate that the prediction ability of a convolutional neural network (in this case, a FasterRCNN) is greatly improved by substituting the HTMC for the NAIP. NAIP prediction accuracies for summit features fell below 60%, whereas the HTMC increased the accuracy to over 95% on average. The resulting bounding boxes can be seen to highlight summit features (Figure 1). Although the workflow was also tested with the vector shape segmentation as input, performance was better using bounding boxes around the landform shapes.

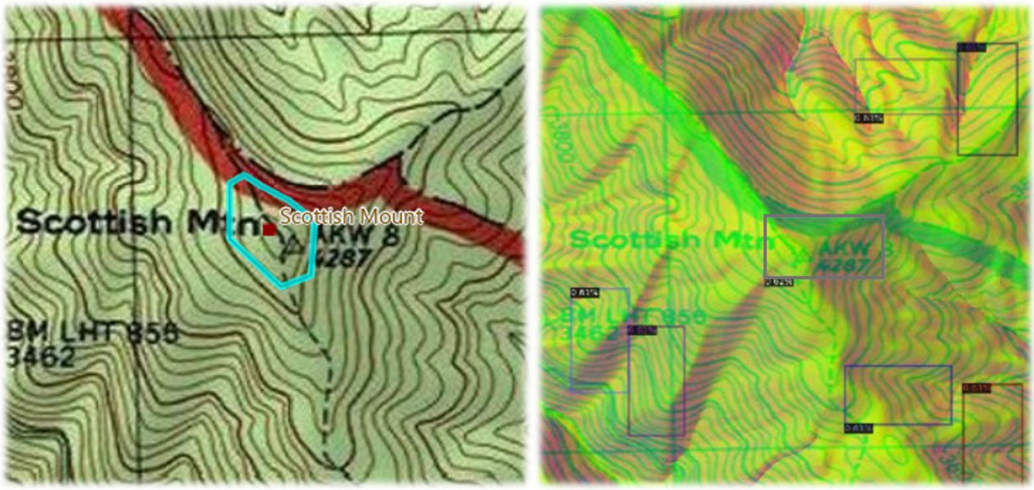

Figure 1. Shape vector representing GNIS Summit feature *Scottish Mount*, North Carolina, USA (left) and bounding box predictions (right) from a FasterRCNN using images with the HTMC loaded into the green channel.

This research indicates that using historical maps to present cognitive information about features with inherently indeterminate boundaries to a learning machine yields better predictions. Although further testing is required to exhaust other reasons for improved performance, results to date are encouraging. Additional enhancements to training data and postprocessing output features may improve their representativeness of ground-truth features.

ACKNOWLEDGEMENTS

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. USGS' work to be published constitutes a 'work of the United States government.' A work of the U.S. government is a contribution by an employee of the U.S. federal government as part of the employee's official duties, and is not subject to copyright protection within the United States.

REFERENCES

Arundel, Samantha T., Wenwen Li, and Sizhe Wang. 2020. “GeoNat v1.0: A Dataset for Natural Feature Mapping with Artificial Intelligence and Supervised Learning.” *Transactions in GIS* 24 (3): 556–72. doi:10.1111/tgis.12633.

Arundel, Samantha T., and Gaurav Sinha. 2018. “Validating the Use of Object-Based Image Analysis to Map Commonly Recognized Landform Features in the United States.” *Cartography and Geographic Information Science* 46 (5). Taylor & Francis: 441–55. doi:10.1080/15230406.2018.1526652.

Arundel, Samantha T, and Gaurav Sinha. 2020. “Automated Location Correction and Spot Height Generation for Named Summits in the Coterminous United States.” *International Journal of Digital Earth* 13 (12). Taylor & Francis: 1570–84. doi:10.1080/17538947.2020.1754936.

Joly, Genevieve, Gaurav Sinha, and Wael Hassan. 2022. “Evaluating Methods for Automated Mapping of Apexes of Non-Linear Eminences.” In *Proceedings of the AutoCarto 2022, November 2-5, 2022*. Redlands, CA.

Li, Wenwen, Sizhe Wang, Samantha T. Arundel, and Chia Yu Hsu. 2022. “GeoImageNet: A Multi-Source Natural Feature Benchmark Dataset for GeoAI and Supervised Machine Learning.” *GeoInformatica*, no. 0123456789. Springer US. doi:10.1007/s10707-022-00476-z.

Sinha, Gaurav, and Samantha T. Arundel. 2020. “Automated Extraction of Areal Extents for GNIS Summit Features Using the Eminence Core Method.” In *Proceedings of the Geomorphometry 2020 Conference*, 38–41. doi:10.30437/GEOMORPHOMETRY2020.

Smith, Barry, and David M. Mark. 2003. “Do Mountains Exist? Towards an Ontology of Landforms.” *Environment and Planning B: Planning and Design* 30 (3): 411–27. doi:10.1068/b12821.