Next Generation Framework for Imagery Recognition and Analysis

Yao-Yi Chiang[1] and Dan Feldman[2]
Spatial Sciences Institute, University of Southern California[1] and Planet Insight[2]
yaoyic@usc.edu; dan@planetinsight.io

Extended Abstract

The volume, velocity, and variety of readily available imagery are growing at a rapid pace. Automatic and accurate object recognition from imagery is critical for understanding the Earth's culture and physical characteristics. Recently, deep learning tools have shown impressive performance in object recognition, detection, and segmentation from imagery (e.g., [1]). However, generating timely insight from imagery is still a time demanding task. **Most of the cost comes from the manual process of teaching/training geospatial software to recognize various types of objects in heterogeneous imagery data.** Commercial solutions such as CrowdFlower provide partial solutions but do not scale well due to the overhead of managing the quality of the crowdsourced training data. **Also, not having control over the quality of individual annotator, requires many redundant annotations.** Moreover, **the manual process currently demands carefully tracing outlines of imagery objects.** Even with access to substantial resources, this process is slow and error-prone.

We envision **a streamlined recognition and analysis framework that will revolutionize how users interact with geospatial software for imagery recognition to enable robust, timely analysis of geo-intelligence.** The framework will include the capability to exploit existing geospatiotemporal data (e.g., vector outlines of agricultural fields in the area-of-interest from seasonal imagery) for automatically generating large numbers and varieties of training data. The idea is that changes on Earth are typically incremental over time, and hence existing data can provide a hint on where to find the desired geographic features in a newly acquired imagery asset. The framework will use these hints to automatically align existing geospatiotemporal data to imagery and then use the aligned locations and geometry to generate diverse samples of the desired objects from imagery from different time periods and scales (e.g., in [2] for recognizing features in scanned maps). Additionally, the framework will include an intelligent user interface in an active learning environment, which will pre-generate object outlines (Figure 1) for semi-automatically generating training data and learn the user intent from each user annotation to improve the quality of the outlines. With more examples from the user, eventually, the number of iterations required for training will be reduced drastically. In sum, manually generating training data for imagery recognition is prohibitively expensive and very often is not able to provide good quality training data. **The ability to quickly teach/train geospatial software for imagery recognition will enable rapid generation of geo-intelligence from imagery.**

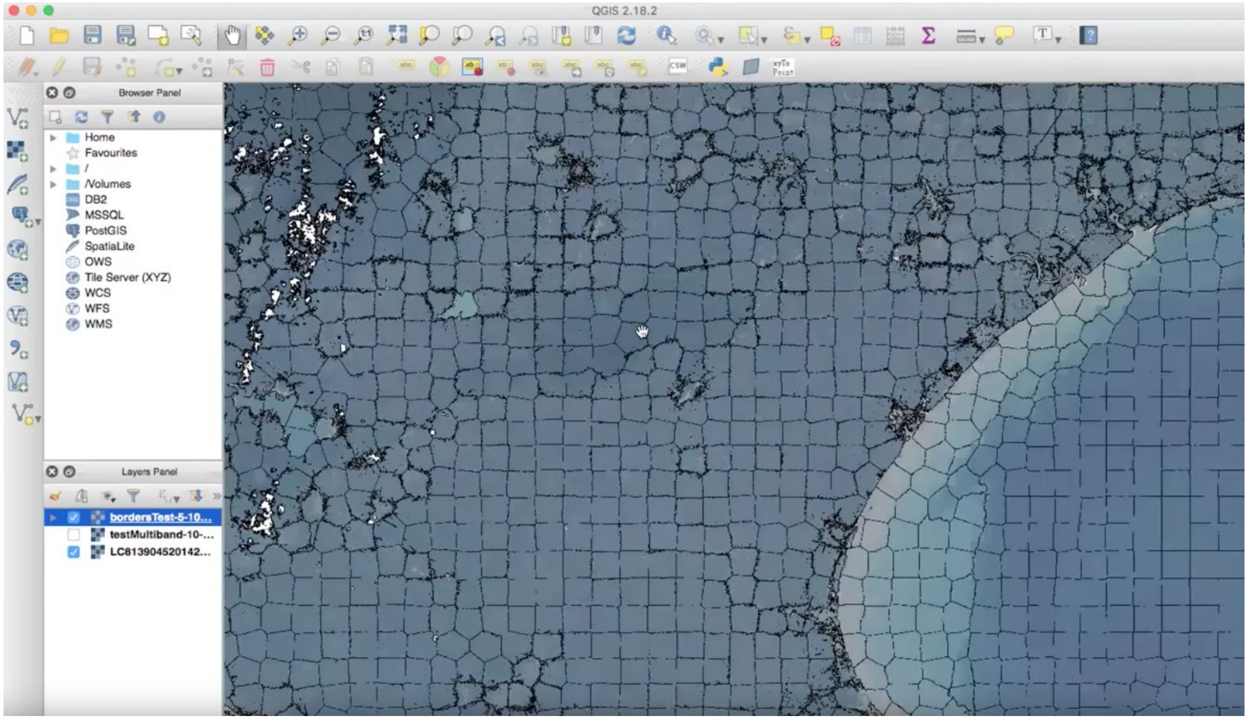

Figure 1. Using superpixels for efficient annotation of training samples (see a demonstration video here: https://youtu.be/uLFnwEfuoAE)

References

[1] Chen, L. C. et al. (2016). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *arXiv preprint arXiv:1606.00915*.

[2] Duan, W. et al. (2017). Automatic Alignment of Vector Data with Geographic Features for Feature Recognition in Historical Maps. In *Proceedings of the GeoAI Workshop*.