Efficient and Accurate Object Extraction from Scanned Maps by Leveraging External Data and Learning Representative Context

Weiwei Duan
Email: weiweidu@usc.edu, Affiliation: University of Southern California, Advisor: Yao-Yi Chiang, SRC: Graduate

Abstract
This paper introduces innovative approaches for extracting geographic objects from scanned map images. Overcoming the challenges associated with labor-intensive data labeling and inaccurate extraction results, we present two data labeling methods and an accurate-enhancing extraction method. The experiment in this paper shows that our approaches outperform the baselines.

CCS Concepts
• Information systems → Geographic information systems.

Keywords
Generative models, Variational Autoencoder, Transformer, Object Detection, Map Processing, Geospatial Information Sciences

ACM Reference Format:

1 Introduction
Scanned historical maps contain valuable, detailed information that often does not exist anywhere else. The information is crucial to many application areas, such as the assessment of critical minerals, which are essential components in many modern technologies vital to modern society and the global economy. Recognizing critical minerals’ significance, the Defense Advanced Research Projects Agency (DARPA) has flagged them as a national security concern. Critical mineral assessment relies on accurate location information about polygonal and linear geographic objects, such as rocks and fault lines. However, extracting objects’ locations in maps involves two main challenges: 1. minimizing human effort for labeling data to train extraction models, as publically available training data are scarce; and 2. learning representative and sufficient context to accurately extract geographic objects.

This paper presents our methods for extracting polygonal and linear geographic objects, which are crucial to critical mineral assessment. Figure 1 shows the extraction workflow for polygonal and linear geographic objects from scanned historical maps. In the workflow, the inputs are scanned topographic or geological maps, and the outputs are extracted vector data representing the desired geographic objects. The extraction process for both polygonal and linear objects includes two components: labeled data generation and extraction model training. Our methods focus on polygonal and linear object labeling and extraction, which are green components in the workflow.

2 Methods
Since our labeling approaches [1, 3] utilize external data to minimize human involvement, here explains external data first. External data refers to data from other sources representing the desired geographic objects depicted in the map images. Due to the different editing years and coordinate systems between the external data and maps, the external data provides incomplete and misaligned labels for the desired objects in maps. Although the labels from the external data are inaccurate, the external data covers or is close to the desired objects in maps. Therefore, our approaches utilize the location information from the external data to generate accurate labels for desired objects with minimal human effort.

Subsequently, this section explains our labeling and extraction methods, as depicted in the green boxes in Figure 1.

2.1 Symbol Labeling for Polygonal Objects
Labeling symbols for polygonal objects is detecting images covering desired symbols within the region-of-interest (RoI) in maps. The external data provides RoI, showing as the blue polygon in Figure 2a. To facilitate the labeling process, the proposed method employs a sliding window approach to generate candidate images within the RoI. Figure 2b illustrates candidate images, including wetland symbols (the blue grass), blue lines, and white backgrounds.

The proposed method groups the candidate images into desired and non-desired clusters. The desired cluster exclusively comprises desired symbols. Figure 2c shows images in the desired cluster for wetland symbols. The advantage of the proposed method is requiring minimal human efforts, while still achieving accurate labeled images. The human efforts required for the proposed clustering method are one or a few manually labeled images covering the desired symbols. In contrast, the state-of-the-art (SOTA) semi-supervised clustering models [8] require a small set of manually labeled desired and non-desired symbols for a desired cluster. While SOTA unsupervised clustering models [4, 7] eliminate manual labels, they may not effectively learn the patterns distinguishing
The proposed Transformer-based extraction model predicts vector lines for desired linear objects by learning representative image contexts and sufficient spatial contexts.

## 3 Experiments and Discussion

For polygonal object labeling and extraction, we test wetlands in six scanned historical topographic maps from the United States Geological Survey (USGS). For linear object labeling, we test railroads and waterlines in three scanned historical USGS topographic maps. For linear object extraction, we test railroads, waterlines, fault lines, and scarps in scanned historical USGS topological and geological maps. We utilize vector data from USGS, published in 2018 as the external data.

The wetland symbol labeling results show that our model achieves superior accuracy compared to unsupervised baselines [4, 7]. Furthermore, the labeling results of our model are comparable to the results from the semi-supervised baseline [8] but with the significant advantage of requiring substantially less manual effort. Subsequently, the extraction model, VGG [6], trained by the labeled data from our model archives 80% average accuracy. Extracted polygons encompass most of the polygonal objects in maps.

The railroad and waterline labeling results show that our algorithm achieves an average 10% higher precision than the SOTA baseline [2]. Consequently, utilizing our algorithm’s labeled data substantially improves the accuracy of extracted linear objects compared to the baseline [2].

The linear object extraction results show that our extraction model achieves an average correctness of 0.87 and significantly improves the connectivity of extracted vector lines by approximately 20% compared to SOTA baselines [5]. Additionally, the proposed model won first place in the DARPA competition 2022 2, significantly outperforming the second place by 184%.

To summarize, this paper presents our methods for extracting polygonal and linear geographic objects from scanned map images, achieving high accuracy with minimal human involvement. The experiment shows that our method [3] extracts most polygonal objects on maps with one manual label for the desired symbol. Additionally, our automatic methods [1] for linear object labeling and extraction achieve more accurate results than baselines [2, 5].

## References


2 https://criticalminerals.darpa.mil/The-Competition

**Figure 2:** The blue area in (a) shows the map area covering the external wetland vector data. (b) shows candidate images generated by the sliding window approach within the blue area in (a). (c) shows the labeled images covering wetland symbols (the blue grass).