AIM: AI for Map Geolocation and Extraction to Find Critical Minerals

Inferlink STTR Project

Related to CriticalMAAS -- TA1

1 Significance of the Problem

Critical minerals are essential components in many modern technologies used by modern society and the global economy. For example, rare earth elements are used in the production of high-tech electronics such as smartphones, computers, and electric vehicles, while lithium, cobalt, and graphite are critical components in batteries used to power electric vehicles and store renewable energy. The strategic value of critical minerals is reflected in their concentration in a few countries, with China dominating global production and supply chains for several key minerals. This creates potential vulnerabilities and risks for countries that rely heavily on imports of these minerals. As a result, many countries, including the United States, have identified critical minerals as a national security concern and have taken steps to reduce reliance on foreign sources of supply. For instance, the Energy Act of 2020 provides funding for the research and development of critical mineral extraction and processing technologies and a comprehensive program to support the development of a domestic critical minerals supply chain.

Finding new sources of critical materials depends on having accurate data regarding the geology of potential sites. The United States Geological Survey (USGS) has an extensive collection of geological maps containing detailed information about various geological features, such as rock formations, faults, folds, and mineral deposits. Other similar geological maps exist online, such as in the public mining companies' SEC filing or online map repositories. These geological maps are valuable resources for a variety of geological studies, including critical mineral assessment. However, most of these maps exist only as scanned images, requiring significant expert effort to manually convert the information into an analytic-ready format. This manual step is time-consuming, expensive, and does not scale to process large numbers of maps.

The objective of this Phase II effort is to research, develop, and evaluate state-of-the-art machine learning algorithms to help automate critical mineral assessments. Specifically, the algorithms will include automated map georeferencing to accurately geolocate maps of unknown locations, and coordinate system and automated map feature extraction to identify point, line, and polygon features in geological maps.

The proposed work will build on the team's existing machine learning methods and tools, which won first place in the Map Feature Extraction Competition in the DARPA and USGS AI for Critical Mineral Assessment Competition in 2022 (hereafter, the Competition). The new challenges include effectively handling a large variety of cartographic symbols with minimal user-annotated training data, broadening the coverage of map types by transferring learned knowledge from one map type to another, and making the overall machine learning architecture and system able to handle large numbers of scanned maps. To automate the map georeferencing, the proposed work will also build on our past work on accurately geolocating maps using the extracted features from a map.

The expected outcome will be an end-to-end prototype system, called AIM (AI for Maps), for automated georeferencing and feature extraction from scanned geological maps. The AIM system will employ a set of machine learning methods that automate the current manual methods to leverage undigitized map data. The system will be evaluated on two different use cases. First, we will test the system on existing USGS maps with the goal of georeferencing and digitizing the maps in the National Geologic Map Database (NGMDB) catalog. Second, we will test the system using existing reports from mining companies[1] with a rich text and maps regarding potential sites and wider geographic areas. Here the ultimate use case is to georeference and extract features from a figures and maps embedded in the wider literature including mining company reports, academic papers, and other publications (including USGS pubs)

The proposed work will significantly improve today's lengthy process for domestic mineral assessments by automatically converting scanned geological maps into an analytic-ready format. The automation will dramatically reduce the time and effort required to find and extract useful information from scanned geological maps, enabling researchers to focus on their analysis and interpretation. Specifically, the system will benefit USGS by enabling automated and parallel mineral assessments with accurate quantitative results. The resulting data will be useful to any commercial or government agency involved in the assessment of critical minerals and ultimately will help reduce US dependence on foreign resources.

Our team is highly qualified to conduct the research and development of the technology proposed here. The team includes internationally recognized researchers in AI and geospatial sciences from InferLink, the University of Minnesota, and the University of Southern California, all of whom have worked together collaboratively on advanced AI research in the recent past. The PI, Dr. Yao-Yi Chiang, is an Associate Professor in the Department of Computer Science and Engineering at UMN, and a certified geographic information systems (GIS) professional. With a unique background and training in AI and GIS, Dr. Chiang is an international leader in developing machine learning methods and systems for extracting geographic features and text from scanned historical maps. Dr. Craig Knoblock, the Executive Director of the Information Sciences Institute at USC and a Research Professor in both the Department of Computer Science and in the Spatial Sciences Institute, will also have a leadership role in this project. Dr. Knoblock is an expert in geographic information extraction and integration and has been conducting research in geographic information science for over 20 years. He co-founded and sold a startup company, called Geosemble Technologies, that successfully transitioned his research to commercial products.

Our team members at InferLink include Dr, Steven Minton and Dr. Naveen Ashish. Dr. Minton founded one of the leading AI journals (JAIR) and has won international awards for his work on applied AI (including AAAI's Robert S. Englemore Memorial Award and AAAI's Classic Paper Award). Dr. Ashish is InferLink's Chief Scientist and an expert in data science and information integration. His previous work (at NASA Ames, in particular) includes leading projects on geospatial data integration and led to a co-authored book "Geospatial Semantics" (Ashish & Sheth, 2011).

In addition to our research credentials, InferLink has had success in turning SBIR research into commercial technology. In the past six years, InferLink has spun off four vertically focused startup companies based on government-funded technologies developed at InferLink, demonstrating our ability to translate research into deployed applications that support a wide spectrum of users. For example, one of our spinoffs, Evid Science, which was based on DARPA-funded SBIR technology, was recently acquired. Furthermore, the same underlying technology is currently being transitioned to the government in DTRA and NGA phase III contracts. This work was selected as a DARPA SBIR success story (see https://www.darpa.mil/attachments/InferLink-Corporation-Success-

[1] An example of such a report can be find here: https://www.niocorp.com/wp-content/uploads/NioCorp_June-2022_NI_43-101_Technical_Report.pdf

Report.pdf). As described in the Commercialization section, our spinoffs play an important role in marketing and distributing horizontal technologies developed at InferLink.

2 Phase II Technical Objectives

Our work in Phase II will focus on 1) extending the machine learning algorithms that we have developed for the Competition for map feature extraction to improve the accuracy and coverage for accurate feature extraction from maps, 2) building automatic map georeferencing algorithms, and 3) implementing a robust, scalable system to process large numbers of scanned geological map images. These refinements and new capabilities to our existing algorithms are intended to address aspects particularly important for real-world applications dealing with large amounts and varieties of scanned geological map images. Our work will also include developing benchmark datasets, which is critical for evaluating the proposed capabilities, and developing specific applications that interest the government and commercial organizations.

The specific objectives are the following:

- **Objective 1: Accurate Feature Extraction from Maps:** Our first objective is to enhance the accuracy and expand the coverage of our existing geographic feature extraction technologies. Our previous work focused on a limited number of map legends and map types, primarily tested for the Competition. The results of the Competition demonstrated that machine learning models incorporating cartographic design principles can outperform state-of-the-art computer vision approaches for geographic feature extraction from scanned map images. In our proposed work, we will build upon these findings and formalize the incorporation of cartographic design principles into a feature extraction framework that can handle a diverse range of geological map images with various symbologies. Additionally, we will address fundamental machine learning challenges relating to the scarcity and imbalance of available training data and domain shifting, which have practical implications for the generalization of machine learning models to diverse sets of geological maps.

- **Objective 2: Automatic Map Georeferencing:** Our second objective is to develop machine-learning methods for georeferencing scanned geological maps automatically. Leveraging content-based retrieval and representation learning, we aim to eliminate the commonly used requirement of pre-defined matching patterns (e.g., street network patterns or place names) between a geological map and an external georeferenced dataset for georeferencing. Utilizing the extensive archive of already georeferenced USGS historical topographic maps, we plan to develop an approach to identify the base topographic maps of the georeferenced USGS geological maps using map metadata and text information outside the map neat line. Then we will develop a contrastive-learning-based deep learning model to extract and correlate image features from these paired maps, encouraging the extraction of shared image features while ignoring cartographic elements that exist only on geological maps. We expect these shared image features, in the form of visual embeddings, to contain essential information on terrain features and landmarks depicted on topographic maps and geological maps. We will build a content-based retrieval method to compare the visual embeddings from an input geological map and embeddings from all USGS topological maps. The USGS topological map with a visual embedding close to the visual embedding of the input geological map will serve as the base map to enable the localization and, finally, georeferencing of the geological map. Furthermore, we intend to explore additional methods to identify grid lines and text information to aid map georeferencing.

- **Objective 3: Prototype Development and Evaluation:** Our third objective is to design, develop and evaluate the AIM system for automated georeferencing and feature extraction from scanned geological maps. Our work will focus on building an end-to-end system that can rapidly harvest and ingest map data and apply the capabilities described above for 1) automated map georeferencing to accurately geolocate maps of unknown locations and coordinate system and 2) automated map feature extraction to identify point, line, and polygon features in a geological map. The prototype will be designed to address specific uses desired by USGS. One use case focuses on georeferencing and extracting data from USGS maps, including maps in the National Geologic Map Database (NGMDB) catalog. The second use case involves processing documents, such as mining reports, and georeferencing the maps they contain and extracting relevant data about the maps and their context.

3 Phase II Statement of Work

In this section, we describe our plan to achieve the objectives that we have just presented. Our three objectives will each be addressed in a separate task.

3.1 Task 1: Accurate Feature Extraction from Maps

The proposed task 1 includes four sub-tasks: map and legend detection from scanned images, and line, point, and polygon feature extraction from maps.

3.1.1 Map and Legend Detection from Scanned Images

Based on our work on the Competition, the detection of map and legend from scanned images has two challenges: 1) there may be inconsistency in the shape and relative location of the various components on the scanned image, such as the map, the legend, and the auxiliary labels, photos, and text around the scanned images; and 2) because these are scanned images, there may be artifacts in the image from the scanning process, such as creases in the paper that cast a shadow over areas that are not our targets and generate false positives.

To address the first challenge, we will identify background colors to determine the foreground image and then employ a connected-component analysis of the foreground image. In particular, we find the background colors of the image based on the four corners of the image and separate the foreground image into multiple connected components. Next, we calculate the number of distinct colors for each connected component. Accordingly, we select the largest connected component with a decent variety of colors as the map subregion. The reason for limiting the color variety is to prevent selecting the auxiliary texts as our targeted map subregion; while the reason for selecting the largest connected component is to avoid selecting the legend or auxiliary photos. In addition, by limiting the shape, the size, and the variety of colors for the connected components, we can identify the legend from the scanned images. In particular, the geologic maps often present the keys (including lines, points, and polygon features) in a small rectangular region with a limited variety of colors and auxiliary texts nearby.

For the second challenge involving the artifacts from the scanning process, we will perform a series of dilation and erosion of the foreground image. This allows us to better determine the background and foreground of the scanned images. Figure 1 shows a scanned image having a non-rectangular map subregion with auxiliary labels. Figure 2 shows a scanned image with creases and having non-white background color with auxiliary labels and photos.

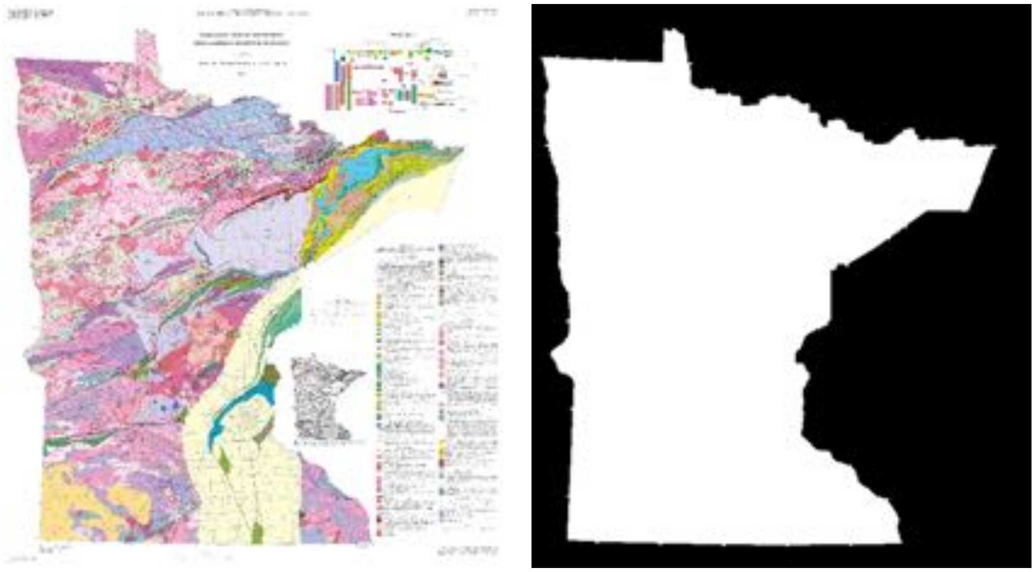


Figure 1. A scanned image having a targeted non-rectangular map subregion to extract features with an auxiliary label showing the taxonomy of keys on the upper right corner and keys from the legend on the bottom right corner (left); with corresponding detected map subregion (right). Note that the keys from the legend come in a specific shape with a limited variety of colors.

Figure 2. A scanned image with a vertical crease in the center of the image having a non-white background color with several auxiliary photos on the bottom part of the image and the keys on the upper right corner (left); with corresponding detected map subregion (right).

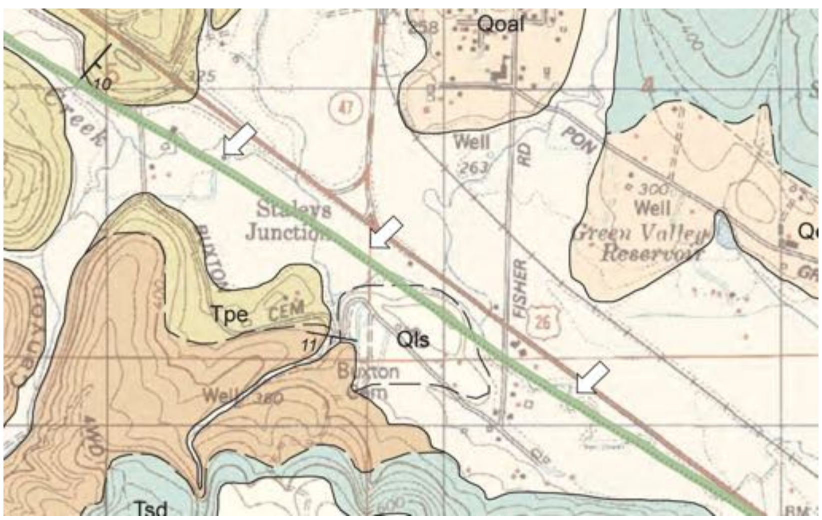
3.1.2 Line Feature Extraction

The extraction of line features from scanned map images presents two primary challenges: 1) the detection of continuous lines while avoiding false negatives and 2) the differentiation of the desired lines from other map symbols to minimize false positives. Conventional image segmentation models struggle to navigate these challenges, as they are typically trained with scenic photos or overhead imagery, thereby neglecting cartographic principles such as the lack of explicitly learning the designs of cartographic symbols, line styles, and lithological ornaments. For instance, topographical and geological maps customarily employ solid or dashed lines illustrated with varying ornaments to signify various features. This subtle difference between features could be difficult for a conventional image segmentation model to learn.

To tackle these challenges, we developed a Transformer-based deep learning model to effectively learn the cartographic context surrounding the target line features in geological and topographic maps. The model contains two stages. The initial stage learns the cartographic context from the ornaments in the legend and their corresponding map content to predict nodes for the desired line vector data. Particularly for a fault line symbol with recurrent ornaments, the detected nodes are typically the location of the ornaments. The subsequent stage learns both local and distant contexts for the detected nodes to determine their connectivity, considering the relative distances and angles between pairs of the detected nodes. The output from this second stage comprises the connected nodes representing the desired line vector.

This model achieved a median F-1 score of 0.576 in the Competition, almost doubling the score of the second-best performer. Figure 3 illustrates two instances of detection results from our model in the Competition, where the green lines represent the detected fault line and thrust fault lines. It is evident that our model effectively detected long and continuous desired lines while successfully eliminating false positives.

However, analysis of the competition's detection results shows three persisting challenges to be addressed in the Phase II project. The first challenge emerges from inaccurately detected joints, a result of our previous method of proposing node locations (Figure 4). To address this challenge, we plan to build on the deformable DETR framework (Zhu et al., 2020) to accurately predict vector nodes. This framework allows for multi-scale, entire-image attention for object detection by dynamically selecting key locations in the image for attention, thereby enabling the proposed model to predict any number of nodes while capturing both local and global context.



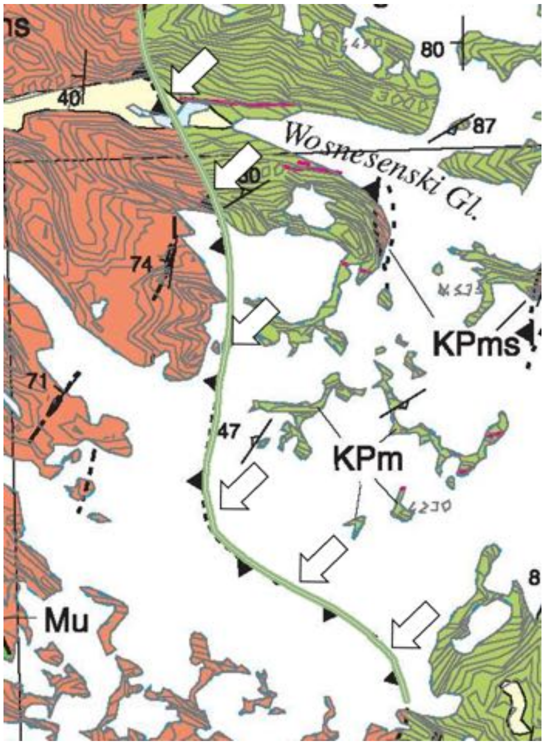


Figure 3. The green lines in both figures are the detected lines using our previous model from the Competition. The desired line features in the top and bottom figures are fault lines and thrust fault lines, respectively.

The second challenge lies in the noisy detection of lines with recurrent ornaments that repeat only a few times for long lines (Figure 5). Because of the rarity of these distant recurrent ornaments, our model could not capture them as representations of the desired line features. To address this challenge, we plan to develop a deep learning module to explicitly enforce the proposed model to learn the connectivity patterns between closeby and distant nodes. This strategy is similar to the Relationformer (Shit et al., 2022) approach, which can also directly generate graphs (vector lines) from images but falls short when dealing with thin lines or multiple lines nearby, such as the line features on geological maps. We also plan to develop automatic post-processing methods that leverage the learned cartographic context to mitigate false detection.

The final challenge is the lack of annotated data to train the detection model. To tackle this issue, we will manually annotate a representative amount of desired line features and apply model distillation, a technique to train a model with limited annotated data. Additionally, we plan to explore active learning, starting with a small set of annotated desired line features and iteratively detecting more desired line features in maps.

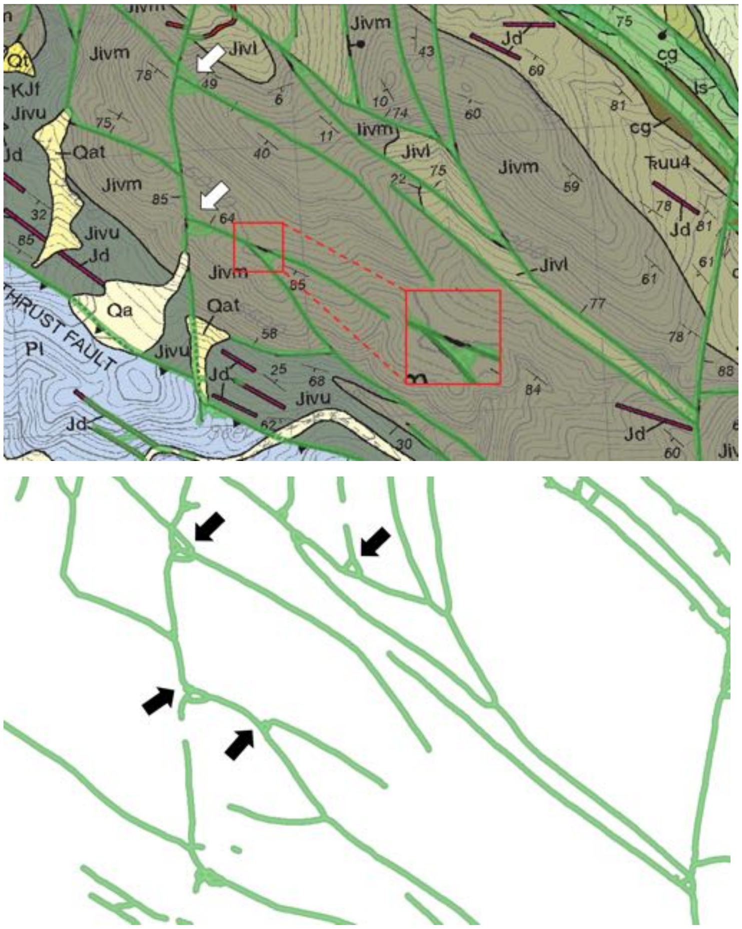


Figure 4. Green lines are the detected fault lines using our previous model from the Competition. The detected joints are inaccurate.

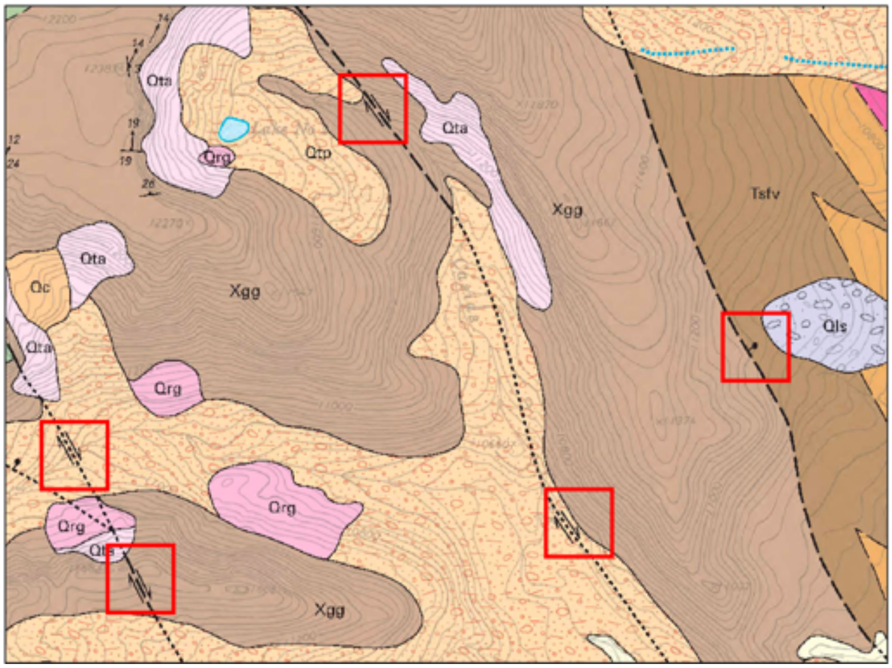


Figure 5. Red boxes show the distant recurrent ornament for a line feature.

3.1.3 Point Feature Extraction

Building an effective machine learning model to extract point features from geographical maps presents several challenges, primarily from the scarcity of available training data. Specifically, scarce training data can significantly impede a machine learning model's result precision (i.e., large numbers of false positives), thus compromising its overall performance. Furthermore, a significant challenge lies in the diverse nature of the symbols used in the geological maps, which vary considerably in shape, color, and size, as well as the wide variety of map content and scan quality, posing difficulties in developing a cohesive model that can handle all point features.

In response to these challenges, our previous machine learning method for point feature extraction comprised a color-centric classical computer vision model, a shape-centric deep learning model (DNN), and a shape-oriented template matching model. The color-centric model identifies symbol locations by scanning for RGB values similar to the target legend. On the other hand, the DNN model, a supervised classifier, examines small image patches to determine the presence of the target symbol. The template matching model, in contrast, searches predefined image patterns that match the legend. The key difference between the DNN and template matching models is the quantity of required training data. The DNN model can automatically learn relevant image features for symbols with multiple training maps for detecting the target symbols. At the same time, template matching relies on human-curated image patterns but can work in cases with limited training data.

The color-based model performs well when no different symbols with similar colors exist on the same map. However, this ideal condition is not always the case. When point features with multiple similar colors exist on the map, the DNN model outperforms the color-based and template-matching models due to its supervised training nature, capable of learning effective image features for detecting the point features. However, our findings from the Competition indicate certain limitations with these methods, particularly with symbols consisting only of simple patterns (e.g., short, inclined lines), which can be challenging for DNN models to learn. Figure 6 shows a number of successful detections of point features, including cross, triangle, and circle symbols. Our DNN models can effectively capture these symbols with distinctive shapes, while symbols that consist of simple patterns are challenging in contrast. Figure 7 shows that there are many false negative detections for the inclined line symbol. The template-based method might be able to detect these simple patterns, but the prediction result is sensitive to the similarity threshold chosen.

In Phase II, we plan to further improve the robustness of point feature extraction and enhance the capability to handle a diverse array of symbols and map types. Specific challenges to overcome include the scarcity of training data, the color mismatch between legend and map symbols, extreme variations in symbol size, and simple patterns (e.g., inclined line symbols) with differing decorators that can confuse the model during training.

As an alternative to employing an ensemble of machine learning models that might not effectively scale up to handle a large number of symbols, our proposed approach for Phase II is to treat this problem as a general object detection/text detection problem. We will build on the deformable DETR framework (Zhu et al., 2020) to develop an object detection model customized for map symbols. Since each map symbol represents some geological phenomenon (e.g., the distribution of different types of rock), we will develop the deep learning module to learn and enforce spatial relationships and patterns within the same symbol type and across different symbol types. A similar approach in our previous work allows the model to learn from limited samples and achieves robust and accurate results in air quality prediction (Lin et al., 2020). We also plan to develop a method to generate synthetic maps with large varieties of point symbol types and conditions (e.g., color and size variations). Similar to our previous work on text spotting from scanned historical map images (Chiang et al., 2023 (in press); Li et al., 2021), we expect these synthetic maps will significantly help improve the coverage of the model.

We will also explore ways to integrate the point feature and line feature extraction models into one framework through meta learning or multitask learning. The idea is that point and line feature extraction tasks can benefit from learning from each other's tasks and their learned cartographic context to effectively solve their own task. For example, Figure 8 shows some of the line intersections incorrectly detected as point symbols using our previous model. This can be solved if the point and line feature extraction model work together sharing the learned knowledge and jointly optimizing for the two tasks.

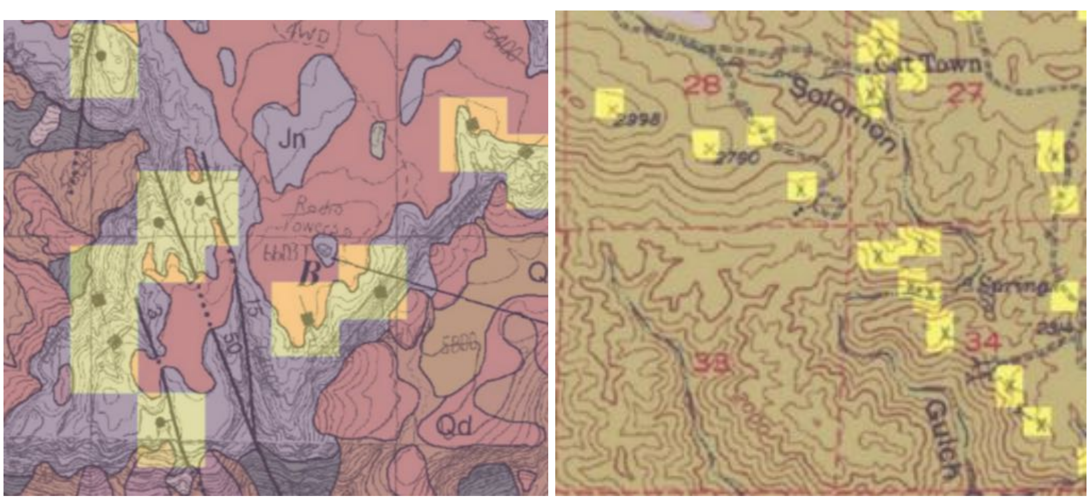


Figure 6. Examples of successful detections of point features using our previous model from the Competition.

Figure 7. Examples of challenging cases for our previous model from the Competition where the target symbol has a simple pattern.

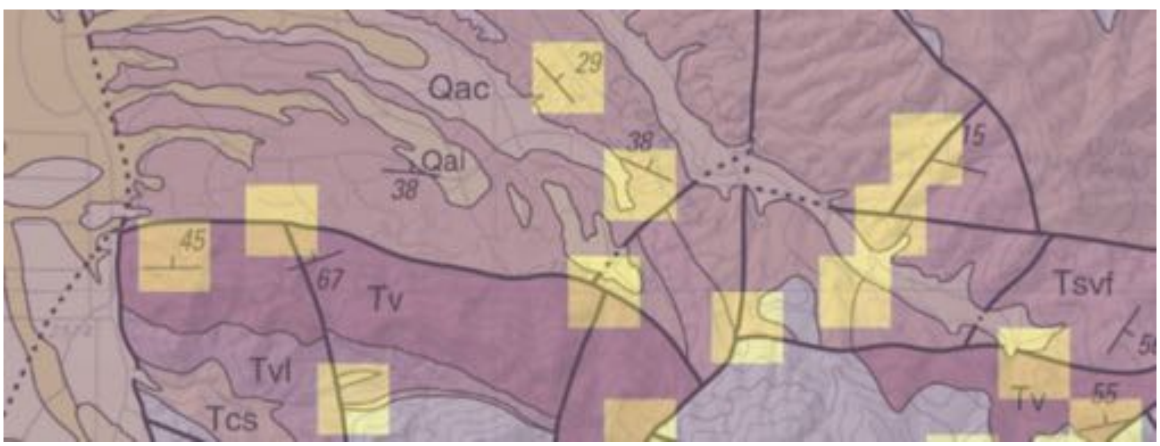


Figure 8. Examples of correctly detected symbols and false positives (road intersections). Combining the point extraction and line extraction tasks in a unified framework can help solve this problem.

3.1.4 Polygon Feature Extraction

The challenge of extracting polygon features from scanned geologic maps is three-fold. First, the keys from the legend in maps come in different colors, texts, or textures. Next, the polygon features in maps often blend in with symbols or textures, making it difficult to identify them from their surroundings. For instance, geological maps often overlap with contours, fault lines, or terrain features. Finally, multiple keys from one map come in the same color, and the only difference among them is the labeled text. This is usually due to the hierarchical relationships in the map legend. All these three factors make it difficult to develop a unified machine learning model that can accurately extract all polygon features.

In the competition, our approach to extracting polygon features consisted of two components: color-based extraction and text-pattern matching. Our approach first uses a set of color-based rules to extract the map pixels with their color similar to the targeted legend. Also, our approach uses pattern matching to rule out the extracted polygon regions labeled with different texts but under similar colors.

Simply applying color thresholding based on either the RGB or HSV space could yield false-negative extraction; therefore, our approach automatically relaxes the tolerance in RGB and HSV values for each key based on other keys with similar colors that we want to extract from the map. Besides, our approach ignores pixels in black or gray and applies the relaxed color on a dilation basis. This allows our approach to overcome noises due to symbols on maps and to achieve high flexibility and performance especially when a map has few keys or keys that have distinct colors among each other.

However, applying color matching without considering the text labels could yield false-positive extraction; therefore, our approach applies text-pattern matching to exclude regions labeled with keys different from the targeted one to retrieve the polygon candidate. Figure 9 is a successful detection of color-based extraction with text-pattern matching. The reason that our approach does not adopt optical character recognition is the inconsistency in text representation between the key and the map subregion. With this polygon candidate, we achieved a 0.629 median F-1 score in the competition.

(a) (b) (c)

Figure 9. An example of successful text-pattern matching that rules out regions labeled with different text. (a) The map subregion from the map and corresponding key of legend to extract; (b) The results obtained based on color thresholding; (c) The polygon candidate after applying text-pattern matching.

Based on the results from the competition, the limitations of the current approach lie in the lack of fault tolerance in color/text representations and the lack of accurate polygon-boundary detection. First, there are inconsistent color and/or text representations between the key and symbols in one map. This makes text-pattern matching yield false-negative extraction. Figure 10 illustrates cases of inconsistent representation. Besides, huge inconsistencies in color can decrease the accuracy of extraction based on color thresholding, clustering, or difference.

(in legend)

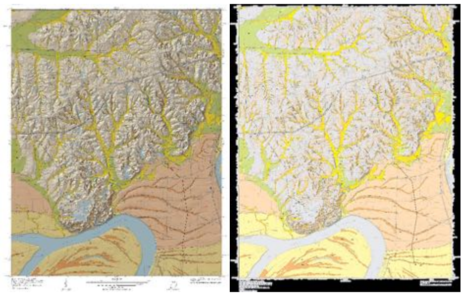
(in map subregion)

(a)

(in legend)

(in map subregion)

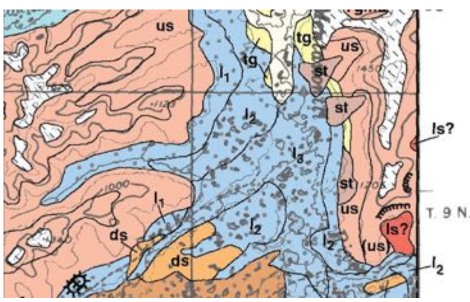
(b)



(c)

Figure 10. Cases of inconsistent representation for a key between legend and the map subregion. (a) Text inconsistency; (b) Color inconsistency; (c) Recoloring the original map subregion (left) based on the colors of keys provided in the legend (right).

Second, the current approach in text-pattern matching can only rule out connected components. Since the boundary that separates polygons comes in different formats, Figure 11 shows that the current approach cannot fully split polygon candidates and therefore limits the accuracy.



(a)

(b)

Figure 11. A case in which the adjacent polygon features are labeled with different text but the same color. (a) Keys ‘l_1’, ‘l_2’, and ‘l_3’ are all in blue but with different texts. (b) The current approach cannot separate them.

In Phase II, to enhance the adaptability of our current approach, we plan to develop a two-stage process for polygon extraction: First, the preprocessing stage applies color thresholding and text-pattern matching to encode each key into a series of bitmaps. Next, we feed this series of bitmaps as multiple channels into a convolutional model (U-Net (Ronneberger et al., 2015)). We then apply shuffle attention (Zhang & Yang, 2021) for the model to learn to put different weights across channels and areas in the map subregion for recognizing the polygon feature for each key from the map subregion. We illustrate the workflow of our approach in Figure 12.

In addition to the polygon candidates that consider both color and text in our approach for the competition, the series of bitmaps that we feed into the convolutional model will include color difference, polygon boundary, and background textures. Figure 13 shows an example of a series of bitmaps with the corresponding output. The adopted convolutional model (U-Net) is a network architecture for biomedical image segmentation (Ronneberger et al., 2015). On the other hand, shuffle attention (Zhang & Yang, 2021) includes channel attention and spatial attention. This will allow our model to learn to adjust the weights between different bitmaps when facing maps with distinct characteristics based on background textures and map information.

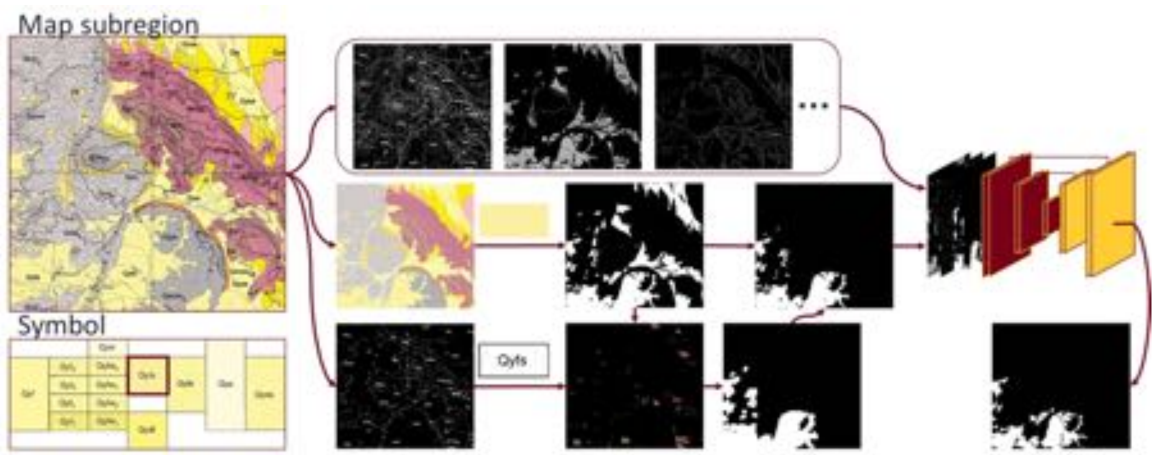
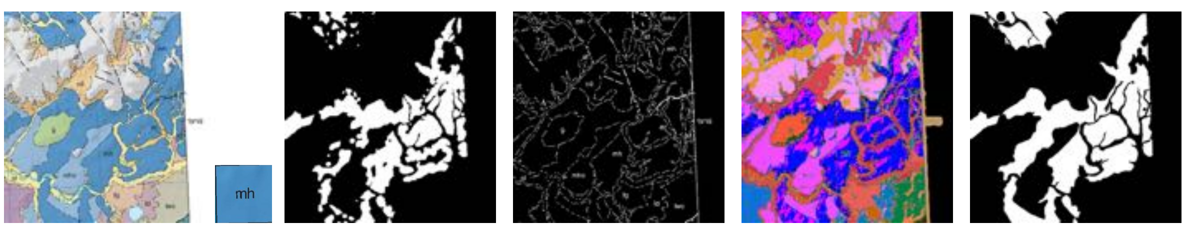


Figure 12. The workflow of polygon feature extraction. We use the color and text from the legend to encode each key into a series of bitmaps and apply a shuffle-attention convolutional model to learn to extract the polygon feature.



(a) (b) (c) (d) (e)

Figure 13. An example for a series of bitmaps. (a) The map subregion and corresponding key of legend to extract; (b) The polygon candidate obtained from preprocessing as a bitmap; (c) The polygon boundary as a bitmap; (d) The color difference as a bitmap; (e) The corresponding ground truth (output).

In the preprocessing stage, we will develop background correction methods for all maps to further address the first limitation where the legend is inconsistent within the map so that the spectrum in the region of interest can match the one from the legend. We will also develop clustering methods for closing boundaries between polygons to address the second limitation of splitting polygon boundaries.

3.2 Task 2: Automatic Map Georeferencing

The process of georeferencing scanned map images typically involves three stages. The initial stage involves detecting potential features from the map and an external georeferenced data source. The subsequent stage searches for matches between the detected features from both sources. The final stage calculates the projection matrix to transform the scanned map images to align with the georeferenced data source using the matched features. While some automatic approaches can manage all or some of these steps when georeferencing overhead imagery, georeferencing historical maps still largely rely on human input. Institutions like the New York Public Library and the Stanford David Rumsey Map Center often hold crowd-sourced georeferencing campaigns (Chiang et al., 2020). Unlike georeferencing overhead imagery—where the georeferencing dataset and the already georeferenced dataset are in the same data domain and share similar features—scanned historical maps and the georeferenced data sources for finding a match are usually in different data domains. For example, a common data source used for georeferencing historical maps is structured map data (e.g., OpenStreetMaps) and satellite imagery. Moreover, scanned maps can be from a variety of sources, each with significantly different content and complexity.

In previous work, we developed a method to georeference and align historical maps with satellite imagery (Chen et al., 2007). Our approach first identifies road intersections from the target maps and satellite imagery. It then searches for matching road intersection patterns between the two datasets. We assume that the map's general geographic area is known (e.g., a city) and that a unique street network pattern exists on the map. This assumption would not work well for many geological maps because many of these maps are in areas with few roads.

In Phase II, we will develop a tiered approach to georeferencing the maps with the most accurate methods applied first to provide the highest quality georeferencing that is possible. In this section, we present three approaches. First, since a large percentage of geological maps are made using a base map from the USGS topographic maps, we will develop a machine learning approach that identifies the precise base map(s) that was/were used. We can then align the geological map to the base map, which gives us the precise geocoordinates as well as the datum and projection used for that particular base map, which then supports accurate transformations on the geological map. Second, since the maps from mining companies may not have used a USGS topographic map as the base map, we plan to develop an approach that exploits the Public Lands Survey System (PLSS) to determine the geocoordinates of a map. Mining companies still use the PLSS system today since they must comply with the Mining Law of 1872 to file claims with the BLM. Third, since the PLSS is only used in 30 states, we will use a fallback of exploiting the text on the map to determine the approximate geocoordinates of a map.

In the subsections below, we describe the work we will do to develop each of these tiered methods.

3.2.1 Learning to Identify the USGS Topographic Base Map(s)

In Phase II, we will develop machine-learning strategies for automatically georeferencing scanned geological maps. We plan to design a content-based retrieval approach to overcome the challenges posed by pre-defined matching patterns (e.g., street networks or place names) between a scanned geological map and a georeferenced external dataset. This strategy will take advantage of the fact that many geological maps (around 9,000) are already georeferenced, and most use the USGS historical topographic maps as base maps. Initially, we will collect the georeferenced USGS historical topographic maps and their metadata from USGS. Then, we will use the geocoordinates of the already georeferenced geological maps and topographic maps to identify base maps for each of the already georeferenced geological maps to generate training data. It is crucial to note that one geological map might be based on multiple topographic maps. Furthermore, multiple topographic maps published at different times may cover similar geographic areas. The challenge lies in using available map metadata and text information outside the map neat line (e.g., detected using our latest text spotting model for map images (Lin & Chiang, n.d.)) to find the best match.

With the georeferenced geological maps and their base maps as training data, we will create a deep-learning approach that learns representative image features from these corresponding map images in a contrastive manner. For example, we will first use convolutional neural networks (CNNs) to extract image features from georeferenced geological maps and their base maps to generate visual embeddings. Then we will train multilayer perceptrons (MLPs) to map the visual embeddings from each georeferenced geological map to the visual embeddings of their base maps. This way, the visual embeddings of their base maps serve as the supervisor to encourage the extraction of novel image features to construct the visual embeddings from the geological maps. For example, since the base maps do not contain geological symbols, their visual embeddings will promote the extraction of shared image features between the geological maps and their base maps and ignore all other features. We will explore methods to provide negative training samples, such as ensuring that image features of non-matched maps in a similar geographical environment are far apart. We will also investigate two additional contrastive learning strategies: 1) training MLPs to map the visual embeddings from the base maps to the visual embeddings of the corresponding geological map and 2) a hybrid approach that maps visual embeddings from both map sources to a new representation space. These contrastive learning strategies have demonstrated powerful capabilities in fusing visual embeddings from different sources for various computer vision applications.

At inference time, the trained model will generate visual embeddings from an input geological map and match them to the visual embeddings from all USGS topological maps to determine the base maps. We will develop automatic methods to align the detected base map (or base maps) to the geological map for localizing the geological map on the base map and use the base map's geocoordinates to georeference the geological map. Once we know the base map that was used, we can then look up the datum and projection used for that base map, which will also apply to the geological map. This information will then allow us to reproject the geological map to a consistent datum and project across all of the geological maps.

3.2.2 Georeferencing using the Public Land Surveying System Grids

The Public Land Survey System (PLSS)[2] is a way of subdividing land in the United States. PLSS surveys were made by the Bureau of Land Management and cover 30 southern and western states. The coverage is shown in Figure 14. The PLSS typically divides land into 6-mile-square townships.

[2] www.usgs.gov/faqs/do-us-topos-and-national-map-have-layer-shows-public-land-survey-system-plss

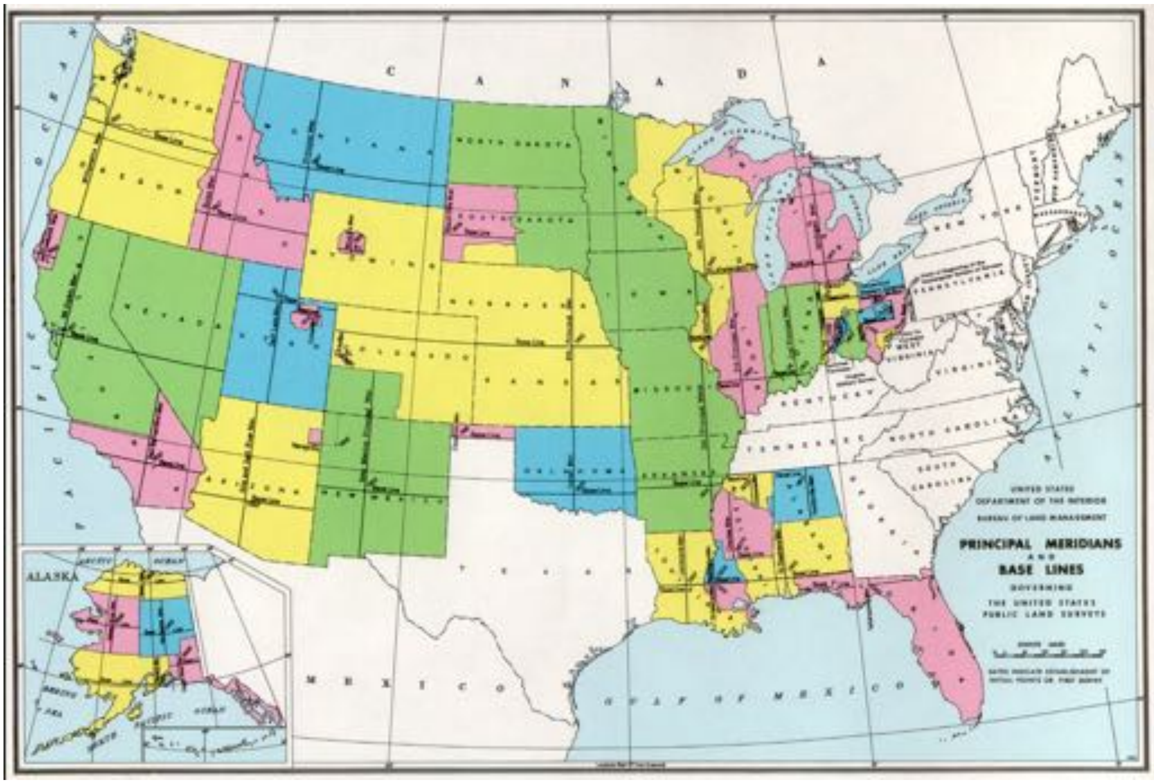


Figure 14: Map showing the coverage of the Public Land Survey System

By law, mining companies are still required to submit claims to the BLM using the PLSS surveys, and as a result, mining maps within the 30 states covered by the system will typically have PLSS survey lines drawn on the maps. The unique pattern of the PLSS surveys, as shown by the grid lines in Figure 15, can be exploited to automatically determine the most likely location of a given mining map. This is similar to our past work on using road lines to georeference satellite imagery. In this case, we can extract the PLSS survey lines from a mining map, determine the intersection of each of those grid lines, and then compare the set of points to a database of the entire set of PLSS grid points. We will create this database from the entire set of PLSS grid lines published by the Bureau of Land Management (BLM).[3] Once we determine the mapping of the grid points, we can determine the geocoordinates of the map.

[3] https://gis.blm.gov/arcgis/rest/services/Cadastral/BLM_Natl_PLSS_CadNSDI/MapServer

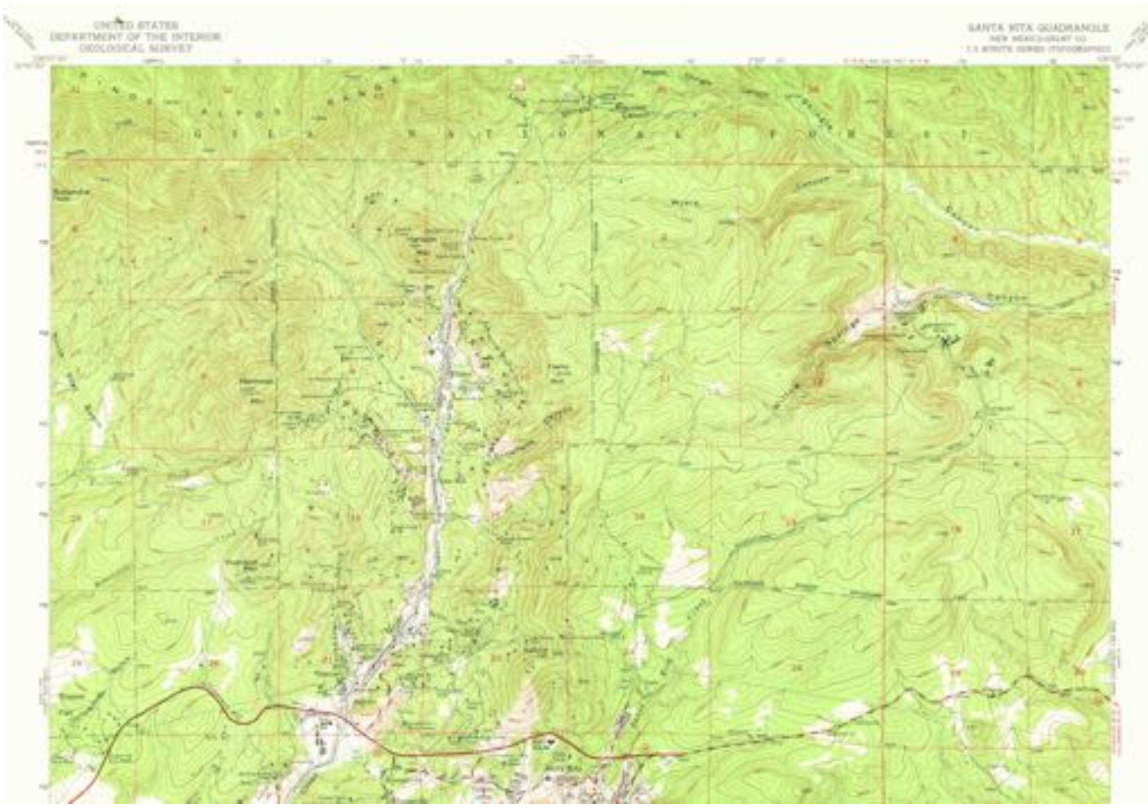


Figure 15: USGS topographic map showing the irregular layout of the PLSS survey lines

3.2.3 Georeferencing using the Text Labels on a Map

In our recent work, we have developed an automated system called mapKurator, which utilizes the text content on historical maps to geolocate them accurately (Li et al., 2020). The mapKurator system employs a two-step process to achieve this goal. First, mapKurator detects individual words using an Optical Character Recognition (OCR) tool and then utilizes deep neural networks to establish connections between these words, linking them to complete location phrases. This linking process considers textual and visual information such as font size, capitalization, nearby background and feature types, and colors. By leveraging these characteristics, mapKurator can generate accurate location phrases, for example, linking "Los" and "Angeles" to form "Los Angeles."

Subsequently, mapKurator utilizes a geocoding service to find a list of candidate geolocations for each detected location phrase. It then analyzes the spatial distribution of these candidate geolocations, identifying spatial clusters and removing unlikely options. mapKurator determines the approximate geolocation of the map by selecting the center of the largest spatial cluster of candidate geolocations. With the approximate geolocation, mapKurator queries LinkedGeoData, a comprehensive geospatial knowledge base, to match the location phrases with entities in the LinkedGeoData. This matching process considers the similarity between text similarity and location proximity between the location phrases and entities in the LinkedGeoData. As a result, mapKurator generates metadata linked to external geospatial knowledge bases in the RDF (Resource Description Framework) format. This linked metadata facilitates complex queries, enabling users to retrieve specific historical maps, such as those covering mountain peaks over 1,000 meters in California.

The effectiveness of mapKurator has been demonstrated through its deployment at Google to support the Kartta Lab project (Tavakkol et al., 2019). Through its end-to-end approach, mapKurator enhances the accessibility and usability of historical maps, opening new avenues for research and exploration.

3.3 Task 3: Prototype Development and Evaluation

In this task, we will focus on implementing a scalable, end-to-end prototype of the AIM system and evaluating the prototype.

Based on our discussions with USGS, we plan to focus primarily on two use cases. The first use case involves georeferencing and extracting data from USGS maps to address the ultimate goal of georeferencing and digitizing all of the maps in the National Geologic Map Database (NGMDB) catalog. The second use case involves processing documents that have a rich text and maps, and georeferencing the maps, and extracting relevant data about the maps and their context. Here the ultimate use case is to georeference and extract features from figures and maps embedded in the wider literature including mining company reports, academic papers, and other publications (including USGS publications).

Initially, we plan to begin this task by collecting and analyzing the system requirements. Given our team's past work in the area, we feel we have a good grasp on the requirements conceptually, but this subtask we will refine the requirements based on discussions with potential consumers of the integrated data that AIM is intended to produce. The work will include meeting with representatives of potential customers, including staff at USGS and other government agencies, as well as DARPA program management, and prioritizing and documenting the requirements. In addition, we plan to develop a set of concrete scenarios that will drive the design process, based on the two use cases above.

We then plan to develop an end-to-end system design, based on the requirements and our designs for the initial algorithms in Tasks 1 and 2. Figure 16 displays an overview of the planned architecture.

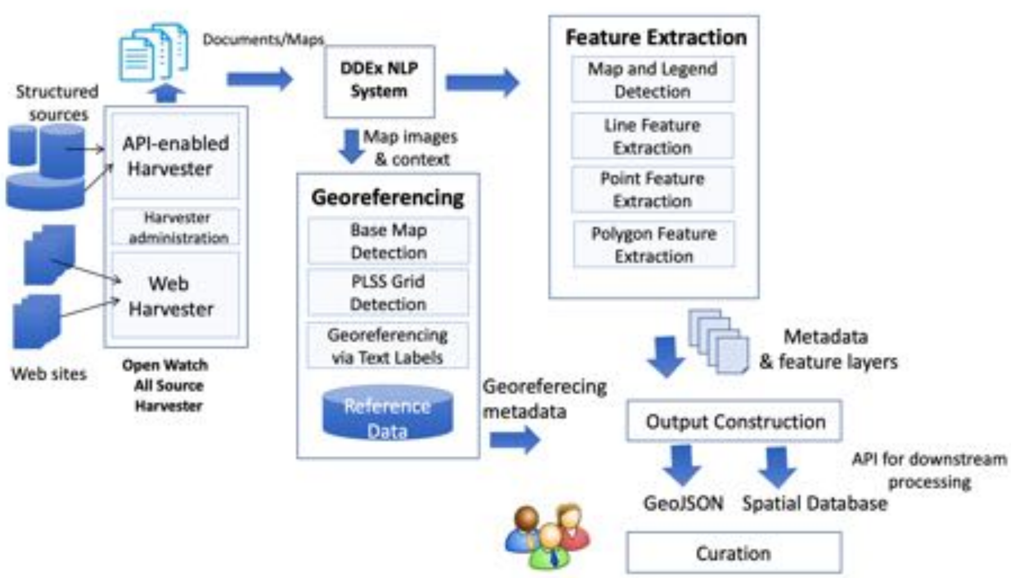


Figure 16: Overview of the AIM architecture

The initial processing in the data ingestion pipeline is provided by two modules that we developed in previous work on the DDEx (Deep Data Extraction) system, that will be customized for AIM. As described in the Related Work system, DDEx is a document processing and NLP system that we originally developed with DARPA funding. DDEx is currently being extended with funding from both DTRA and NGA for text processing applications, and has been used commercially for extracting data from technical documents. The first DDEx module employed for AIM is a harvesting engine that has an API for uploading maps, as well as a web harvester for collecting and ingesting documents from web sites. The second module, a document processing and NLP extractor that we will configure to scan documents to identify embedded images that may contain maps, as well as context (e.g., locations) mentioned in nearby text.

After this initial processing, as shown in Figure 16, AIM will employ the approach developed in Task 1 for extracting data from the maps themselves, and the algorithms developed in Task 2 for georeferencing the maps. AIM will use the determined georeferencing information to convert the extracted data from the image coordinates to an appropriate coordinate reference system (CRS). AIM will output the extracted line, point, and polygon layers in the GeoJSON format, including the “crs” member to store the CRS information (or a similar format, e.g., Shapefile). (The specific output will be determined in consultation with program management and potential users.)

As part of our work in this task, we will also investigate how human curators may be best utilized to check and/or improve the output produced by AIM. We note that we do not plan to design a full-featured user interface. The focus of this aspect of our work will be presenting the data so that users can rapidly validate the results. This will be relevant for helping us to evaluate the system’s performance, as well as potentially relevant for Phase III work where AIM is integrated into a commercial environment.

Once the functional design is fleshed out, we will implement an end-to-end prototype. As noted in the Personnel Section, our team has significant experience developing products for both government and commercial enterprises, using modern software engineering practices.

As the implementation process proceeds, we plan to begin the testing process in parallel, so that components can be tested as they are developed (“test early and often”). We plan to test the system on a range of data, addressing our two target use cases. Specifically, we plan to develop at least two test datasets as part of our requirements analysis, in consultation with potential users. One dataset of maps will be compiled by sampling relevant maps in the National Geologic Map Database (NGMDB) catalog. Another test dataset will be constructed based on documents, such as mining reports, that we will manually compile. We expect to have several hundred maps/documents in each test set. A representative sample of these will be manually labeled to create “ground truth” for evaluation and statistical validation of our results (e.g., accuracy testing). In addition, we will also conduct discussions program management and potential users to identify other sources of test data.

Our evaluation will investigate the key performance characteristics of the system. The anticipated accuracy for Task 1 is projected to reach a median F1 score exceeding 0.8 across all feature extraction tasks. This achievement denotes a substantial improvement compared to our first-place results in the Competition, where we recorded median F1 scores of 0.63, 0.22, and 0.58 for polygon, point, and line feature extraction, respectively.

In terms of expected processing time for Task 1, we project that map and legend detection, alongside point, line, and polygon feature extraction, will take approximately 5 minutes per map with the proposed prototype. This estimate significantly improves our existing, non-optimized system, which required roughly 10 minutes per map during the Competition.

For Task 2, the expected accuracy will fall within 100 meters, assuming a base map can be identified. The expected accuracy will be contingent on the map scale for geological maps lacking explicit geocoordinate information (e.g., maps derived from report figures). In our prior work (Li et al., 2020), we could geolocate historical maps to within less than 1 kilometer of the map center, relying solely on the text information presented on the maps.

Regarding processing time, Task 2 is expected to be completed in under 5 minutes per map. This efficiency is largely because the majority of the feature generation process for the proposed content-based retrieval approach in Task 2 can be pre-generated offline. We will also utilize technologies such as Locality Sensitive Hashing to expedite the retrieval process.

All output files will be hosted in a spatial database (e.g., PostGIS) to support efficient queries by location and feature types and can be easily linked from a geographic information system for visualization. The capability to support efficient queries and visualization of automatically extracted and georeferenced vector features is expected to speed up the expert workflow greatly compared to a manual process.

3.4 Human Subjects Research

No human subjects research will be conducted as part of this project.

3.5 Schedule/Milestones/Deliverables

The project we have described is composed of three tasks. The Phase II deliverables, milestones, and completion schedule are listed below.

- Month 3:
    - Initiate collecting USGS geological maps and historical topographic maps as well as other geological maps (Tasks 1 and 2)
    - Initiate design of the deep learning models for map feature extraction (Task 1)
    - Initiate design of the image retrieval, PLS, and map text approaches for map georeferencing (Task 2)
    - Initiate design of the system pipeline, benchmark datasets, and evaluation methods (Tasks 1, 2, 3)
    - Report on current implementation and plan for first test, including accurate map feature extraction and map georeferencing
- Month 6:
    - Initial prototypes of individual line, point, and area feature extraction models completed (Task 1)
    - Initial prototype of map georeferencing using image retrieval completed (Task 2)
    - Initial prototype of map georeferencing using PLS completed (Task 2)
    - Initial prototype of the system pipeline completed, benchmark datasets and evaluation methods selected (Task 3)
- Month 9:
    - Initiate design of integrated deep learning model for map feature extraction for all features (Task 1)
    - Initial prototype of map georeferencing using map text completed (Task 2)
    - First test evaluation, with domain knowledge provided by our own team and consulted with USGS (Tasks 1, 2, 3)

  - *Interim report on results of first evaluation and any resulting revision in plans, current implementation and plans for second test, including additional benchmark datasets*
- Month 12:
  - Refine individual deep learning models for map feature extraction according to the evaluation results (Task 1)
  - Refine image retrieval and PLS approaches for map georeferencing according to the evaluation results (Task 2)
  - Refine the system pipeline according to the evaluation results (Task 3)
- Month 15:
  - Completion of integrated deep learning model for map feature extraction for all features (Task 1)
  - Second test evaluation, with domain knowledge provided by our own team and consulted with USGS (Tasks 1, 2, 3)
- Month 18:
  - *Final Phase II report documenting final prototype architectures and methods, and results of all tests .*

The deliverables include the AIM system software, a final report detailing the architecture, methods, and evaluation results, and regular progress reports on a schedule determined by the government.

4 Related Work

Our team members have a great deal of published research, as well as commercial and open source products, in the areas related to this proposal, including geographic information sciences, machine learning, computer vision, and document understanding. In this section, we focus on recent and ongoing efforts that relate to the work proposed here, and we refer the reader to our biographies for more detail about our past work.

4.1 BioDDEx and the DDEx Platform (InferLink)

InferLink's BioDDEx (Biological Deep Data Extraction) is an aggregated knowledgebase of global, open source information about institutions/laboratories that work with biological materials (e.g., organisms, biological agents, etc) which is automatically compiled from open source documents..

BioDDEx is based on information extraction technology that was originally funded in an DARPA SBIR project. More recently, BioDDEx development has been funded by two phase III contracts one from DTRA and one from NGA. Both are focused on transitioning the technology for the government's purposes. (We are currently interacting with government analysts who are early adopters. The DTRA TPOC is john.m.ewing6.civ@mail.mil, and our NGA TPOC is Erik.J.Scully@nga.mil).

BioDDEx is automatically compiled through the use of AI and was originally intended to complement the DTRA's BMIP (Biological Material Information Program) database, a government database on dangerous biological materials that is manually compiled and curated by analysts. BMIP provides information related to potential pathogen repositories worldwide to support capability assessments of a country's bio-infrastructure and other bio-related activities that may involve proliferation of knowledge, personnel, and equipment. BMIP helps to better understand where these facilities are located which collect, investigate, and store biological material.

BioDDEx has a similar focus, but contains a vast amount of data that has been automatically extracted, cleaned, and integrated from multiple sources. The system greatly speeds up the manual process of researching and vetting data about facilities, so that data can be rapidly imported into BMIP and other curated databases. In particular, whereas BMIP has manually curated data on approximately 1800 facilities, BioDDEx is completely automated and covers over 170,000 institutions and 460,000 divisions/labs within those institutions. The government is now contemplating an expanded use of BioDDEx, due to the scale it can address.

BioDDEx contains aggregated, entity-resolved information about three types of entities: institutions, people, and biological materials (e.g., pathogens). For each institution, such as a university or an industrial research lab, the system lists locatiion of the facility, and the biological materials that appear to have been possessed by the institution, such as pathogens that are the subject of research articles written by faculty members, equipment used by the laboratory, organisms that the institution has submitted to GenBank, culture collections maintained by the institution (according to the World Federation of Culture Collections), and other information.

In addition to the core work exgtending BioDDEx, with funding from the Army we are currently working on a generic domain-independent platform, called DDEx, which includes all the BioDDEx capabilities for extracting entities, relationships and other technical data, and can be easily targeted to new domains. The DDEx system will also include technology for automatically identifying and extracting tabular data, which we are developing for NOAA. We note that DDEx's capabilities for extracting targeted text and tabular data are relevant to the current proposal, and we plan to rely on these capabilities for processing documents, such as mining reports, in order to identify maps and potentially relevant text.

The DDEx architecture is outlined in the figure below. The system is organized as a pipeline, starting with a highly configurable harvester that can ingest documents from a wide variety of sources. Ingested documents are then processed by an NLP module system that uses a series of ML-based classifiers and extractors to find relevant entities and relations in the content. The data is then entity-resolved to create a knowledge graph that can be queried, and the results can be aggregated, mapped, and displayed to an analyst. We note that the first two modules, the harvester and NLP module, will be utilized in our work in Task3 of this project.

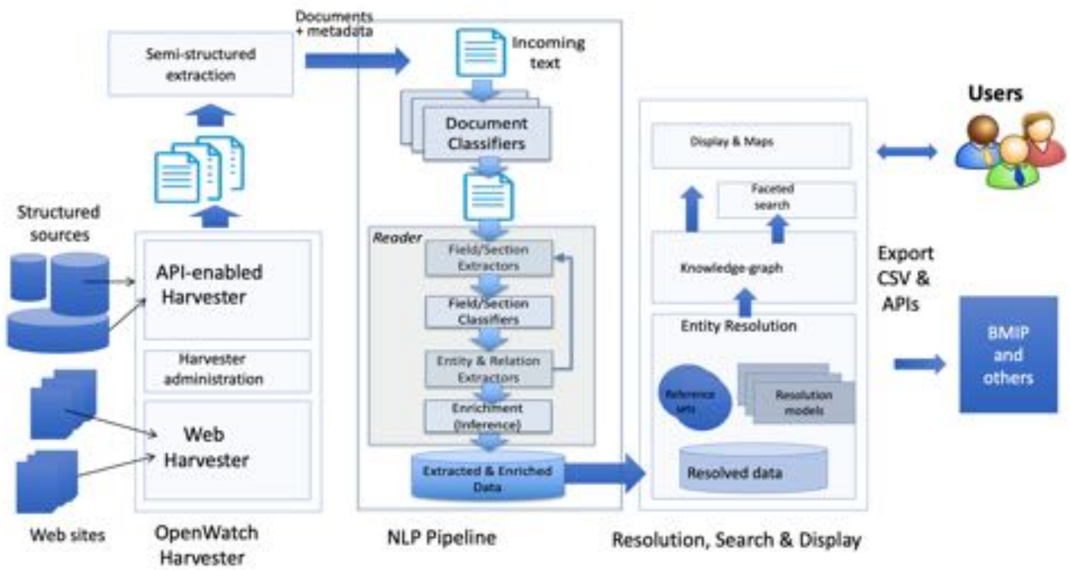


Figure 17: Overview of the DDEx architecture

4.2 Map Georeferencing and Feature Extraction

Our team has worked on automating the extraction and integration of geographic information from historical maps for over 20 years with more than 50 peer reviewed publications. Our work has employed AI technologies to streamline the processing of historical maps, georeference historical maps, extract and link geographic features, spot text content, and analyze large historical map series. We have also built a user-friendly map processing system called mapKurator (Chiang et al., 2023 (in press); Li et al., 2020). Section 3 has described our approach that won the Competition, and this section presents an overview of our recent AI technologies in methods and systems for processing historical maps to generate valuable data, information, and knowledge, further details of which can be found in our prior survey paper (Chiang et al., 2014), book (Chiang et al., 2020), and a recent book chapter (Chiang et al., 2023 (in press)).

One of the major challenges in extracting geographic features from scanned historical maps is the lack of training data. If the map is already georeferenced, using an existing data source overlaying on the map can help provide annotations as training data. For example, Meta's Map with AI uses road networks from Open Street Maps to collect over 100 million road annotations from satellite imagery. However, direct overlaying of external vector data on the map for annotation often led to misalignment and errors due to various data and quality inconsistencies between the two datasets. To address this challenge, we developed the Label Correction Algorithm (LCA) (Duan, W and Chiang, Y.-Y., and Leyk, S. and Uhl, J. and Knoblock, C. A., 2021). LCA optimizes the geometry of external vector data and the map's image features to minimize false annotations. Tested on the extraction of railroads and waterlines from USGS topographic maps, LCA drastically improved the quality of annotations and the results of geographic feature extraction compared to other methodologies.

We also developed approaches to transform the extracted map features into linked spatiotemporal data to support advanced change analysis of map content (Shbita et al., 2023). Our approaches automatically construct a knowledge graph that encapsulates geographic features and their relationships over time, drawing from multiple map editions of the same region. The approaches facilitate the generation of geospatial entities from multi-temporal vector data extracted from historical maps of the same region but at different times, achieve geo-entity linking using OpenStreetMaps, and generate a spatiotemporal knowledge graph supporting the OGC GeoSPARQL standard. The approaches also allow for the incremental addition of new features from subsequent map editions without the need to update existing geographic vector data.

Another line of our work focuses on detecting and recognizing text from scanned historical maps. Historical maps store a wealth of information, not only in geographic features but also in text labels and descriptions. Detecting, recognizing, and linking text in scanned historical maps required a unique approach since map text can have varying orientations, curvatures, spacings, font types, and font sizes. We developed SynthMap, a process that generates unlimited historical-styled map images with text annotations, effectively addressing the lack of training data (Chiang et al., 2023 (in press); Li et al., 2021). Text spotter models trained with SynthMap, such as TESTR (Zhang et al., 2022), demonstrate significantly improved text detection and recognition performance on historical maps. We also developed an end-to-end text spotter for complex documents, focusing on historical scanned maps with wide varieties. A variation of our text spotter integrated with mapKurator has generated 90 million text labels from over 50,000 scanned historical maps to support searching these maps by their text content at the Stanford David Rumsey Map Center.

Simultaneously, we developed mapKurator (Chiang et al., 2023 (in press); Li et al., 2020), an end-to-end pipeline for extracting text content from historical maps and generating metadata linked to external geospatial knowledge bases. The system slices a map image into smaller patches, predicts a bounding polygon, transcribes the text for each map text instance, merges the results, and converts the image coordinates of bounding polygons to geocoordinates, among other functions. To enrich the extracted map text, we developed SpaBERT (Li et al., 2022), a language model that captures the spatially varying semantics of geo-entity names, aiding in geo-entity typing and linking. In our experiment, SpaBERT accurately predicts the semantic types of text labels extracted from historical USGS topographic maps and links them to WikiData entries.

In summary, our past work has not only automated the information extraction process from historical maps but also streamlined the conversion of this data into a structured, standard format readily usable in an analytic environment. Our proposed work aims to build upon these foundations, leveraging advancements in AI and machine learning to enhance the accuracy and efficiency of historical map processing.

5 Relationship with Future R&D

Finding new sources of critical materials depends on having accurate data regarding the geology of potential sites. The objective of this Phase II effort is to research, develop, and evaluate state-of-the-art machine learning algorithms for 1) automated map georeferencing to accurately geolocate maps of unknown locations and coordinate system and 2) automated map feature extraction to identify point, line, and polygon features in a geological map.

If this project is successful, it will provide a foundation for a system that can automate critical mineral assessments. A key to predicting where these materials can be found is having relevant data about the geology of potential sites. The work in this project will provide the technology for harvesting and integrating that data at scale, setting the stage for a Phase III system.

We note that no special clearances, certifications, or approvals are required to conduct the proposed work. We do note, however, that InferLink Corporation and USC both have top-secret facility clearances, which may facilitate the transition of our technology to the DoD in phase III.

6 Key Personnel

Dr. Yao-Yi Chiang, the Principal Investigator, is an Associate Professor in the Computer Science & Engineering Department at the University of Minnesota. Previously, he held the position of Associate Professor (Research) in Spatial Sciences at the University of Southern California. Dr. Chiang is an Action Editor of GeoInformatica (Springer) and an editorial board member for Transactions in GIS (Wiley). He earned his Ph.D. in Computer Science from the University of Southern California and his bachelor's degree in Information Management from the National Taiwan University.

Dr. Chiang is also a certified geographic information systems (GIS) professional. With a unique background and training in AI and GIS, Dr. Chiang is an international leader in developing machine learning methods and systems for extracting geographic features and text from scanned historical maps. His collaborators on this topic include the USGS, the Library of Congress, the British Library, the National Library of Scotland, the Stanford David Rumsey Map Center, Ritsumeikan University (Japan), and the Nanyang Technological University (Singapore). He has developed georeferencing and geographic feature/text extraction algorithms and tools for various map collections, including the USGS historical topographic map series, Sanborn Fire Insurance maps, King's Topographical Collection, the David Rumsey Historical Map Collection, and a variety of historical maps of Asian countries. A recent collaboration with the Stanford David Rumsey Map Center allows Dr. Chiang's team to generate 90 million text labels from over 50,000 scanned historical maps to support searching these maps by their text content at the map center. The search functionality will open to the public in the summer of 2023.

Dr. Chiang also gives regular lectures on AI and historical map processing for the course "Research Topics in Cartography" at ETH Zurich. He has co-authored a survey paper (Chiang et al., 2014), a book (Chiang et al., 2020), and over 50 peer-reviewed conferences and journal publications on this and relevant topics.

In Dr. Chiang's 20-year experience working on AI and historical map processing, he has received funding from government agencies, private companies, and foundations to pursue this line of work. His work on georeferencing historical maps was adopted in a previous Google product called Kartta Labs during his time as a visiting scientist at Google AI (NYC). Kartta Labs was later transferred to Kartta Foundation, where Dr. Chiang is the funder and president. Kartta Foundation is a 501(c)(3) organization providing software and services to distill and assemble geographic knowledge for the public good.

Chiang has over 90 publications in AI. Below is a partial list of sample publications relevant to the proposed work:

Shbita, B., Knoblock A. C., Duan, W., Chiang, Y.-Y., Uhl, J., Leyk, S. (2022) Building Spatio-temporal Knowledge Graphs from Vectorized Topographic Historical Maps, Semantic Web, 1 – 23. IOS Press.

Li, Z., Kim, J., Chiang, Y.-Y., Chen, M. (December 2022). SpaBERT: A Pretrained Language Model from Geographic Data for Geo-Entity Representation. In Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings (accepted), Abu Dhabi

Duan, W., Chiang, Y.-Y., Leyk, S., Uhl, J. H., and Knoblock, C. A. (December 2021). Guided Generative Models using Weak Supervision for Detecting Object Spatial Arrangement in Overhead Images. In Proceedings of the 2021 IEEE International Conference on Big Data, pp. 725-734, online

Chiang, Y.-Y., Duan, W., Leyk, S., Uhl, J. H., and Knoblock, C. A. (2020). Using Historical Maps in Scientific Studies: Applications, Challenges, and Best Practices, Springer (ISBN 978-3-319-66908-3)

Duan, W., Chiang, Y.-Y., Knoblock, C. A., Uhl, J. H., and Leyk, S. (2020) Automatic Alignment of Contemporary Vector Data and Georeferenced Historical Maps Using Reinforcement Learning. International Journal of Geographical Information Science, 34(4): 824-824

Li, Z., Chiang, Y.-Y., Tavakkol, S., Shbita, B., Uhl, J. H., Leyk, S., and Knoblock, C. A. (August 2020). An Automatic Approach for Generating Rich, Linked Geo-Metadata from Historical Map Images, In Proceedings of ACM Knowledge Discovery and Data Mining Conference (KDD), pp. 3290-3298, San Diego, CA, USA

Tavakkol, S., Chiang, Y.-Y., Waters, T., Feng, H., Prasad, K., and Kiveris, R. (November 2019). Kartta Labs: Unrendering Historical Maps. In Proceedings of the Third GeoAI Workshop, pp. 48–51, Chicago, IL, USA

Lin, H., Chiang, Y.-Y. (November 2018). An Uncertainty Aware Method for Geographic Data Conflation. In Proceedings of the 5th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data, pp. 20 – 27, San Francisco, CA, USA

Duan, W., Chiang, Y.-Y., Knoblock, C. A., Vinil, J., Feldman, D., Uhl, J. H., and Leyk, S. (November 2017). Automatic Alignment of Vector Data with Geographic Features for Feature Recognition in Historical Maps. In Proceedings of the First GeoAI Workshop, pp. 45 – 54, Redondo Beach, CA, USA

Uhl, J. H., Leyk, S., Chiang, Y.-Y., Duan, W., and Knoblock, C. A. (July 2017). Extracting Human Settlement Footprint from Historical Topographic Map Series Using Context-Based Machine Learning. In Proceedings of the IAPR 8th International Conference on Pattern Recognition Systems, pp. 15 – 21, Madrid, Spain (best paper award)

Chiang, Y.-Y. (2015) Querying Historical Maps as a Unified, Structured, and Linked Spatiotemporal Source (Vision Paper). In Proceedings of the 23rd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, 16:1–16:4, Seattle, WA, USA (best vision paper award)

Chiang, Y.-Y., Leyk, S., and Knoblock, C. A. (2014). A Survey of Digital Map Processing Techniques. ACM Computing Surveys, 47(1):1–44. doi: 10.1145/2557423

Chiang, Y.-Y. and Knoblock, C. A. (2013). A General Approach for Extracting Road Vector Data from Raster Maps. International Journal of Document Analysis and Recognition, 16(1):55–81. doi:10.1007/s10032-011-0177-1

Chiang, Y.-Y. and Knoblock, C. A. (2009). Classification of Raster Maps for Automatic Feature Extraction. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 138–147, Seattle, WA, USA

Chiang, Y.-Y. and Knoblock, C. A. (2008). Automatic Extraction of Road Intersection Position, Connectivity, and Orientations from Raster Maps. In Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 1–10, Irvine, CA, USA

Chiang, Y.-Y., Knoblock, C. A., and Chen, C.-C. (2005). Automatic Extraction of Road Intersections from Raster Maps. In Proceedings of the 13th ACM International Symposium on Advances in Geographic Information Systems, pp. 267–276, Bremen, Germany

Chen, C.-C., Knoblock, C. A., Shahabi, C., Chiang, Y.-Y., and Thakkar, S. (2004). Automatically and Accurately Conflating Orthoimagery and Street Maps. In Proceedings of the 12th ACM International Symposium on Advances in Geographic Information Systems, pp. 47–56, Washington, DC, USA

Dr. Ju Sun is an assistant professor at the Department of Computer Science & Engineering, the University of Minnesota (UMN). He is directing the Group of Learning, Optimization, Vision, healthcarE, and X (GLOVEX), and serves in the Foundational Data Science Research Committee of the UMN Data Science Initiative (DSI). Prior to this, he worked as a postdoc scholar at Stanford University (2016-2019), obtained his Ph.D. degree from the Department of Electrical Engineering of Columbia University in 2016 (2011-2016), and B.Eng. in Computer Engineering (with a minor in Mathematics) from the National University of Singapore in 2008 (2004-2008).

Dr. Sun's research interests span machine/deep learning, computer vision, machine/deep learning, numerical optimization, data science, computational imaging, and healthcare. His recent efforts are focused on the foundation and computation for deep learning and applying deep learning to tackle challenging science, engineering, and medical problems. Dr. Sun has collaborated with a diverse set of researchers, including, for example, high-energy physicists, materials scientists, fluid mechanics researchers, and medical doctors. He has published over 70 journal articles, book chapters, and conference and workshop papers on these topics and has received 1 best paper award.

Dr. Sun won the best student paper award from SPARS'15, honorable mention of doctoral thesis for the New World Mathematics Awards (NWMA) 2017, and AAAI New Faculty Highlight Programs 2021. He serves as area chair for top AI conferences, such as Artificial Intelligence and Statistics (AISTATS) and Uncertainty in Artificial Intelligence (UAI).

Dr. Craig Knoblock is the Keston Executive Director of the Information Sciences Institute, Research Professor of both Computer Science and Spatial Sciences, and Vice Dean of Engineering at the University of Southern California. He received his Bachelor of Science degree from Syracuse University and his Master's and Ph.D. from Carnegie Mellon University in computer science.

Dr. Knoblock's research focuses on techniques for describing, acquiring, and exploiting the semantics of data. He has worked extensively on source modeling, schema and ontology alignment, entity and record linkage, data cleaning and normalization, extracting data from the web, and combining all of these techniques to process geographic data and build knowledge graphs. He has published more than 400 journal articles, book chapters, and conference and workshop papers on these topics and has received 7 best paper awards on this work.

Dr. Knoblock has also co-founded two startup companies that commercialized his own research. The first, Fetch Technologies, which used machine learning techniques to aggregate data from web sites, was acquired by Connotate Inc. in 2010. The second, Geosemble Technologies, which used data integration techniques to link geospatial data to location, was acquired by TerraGo in 2011.

Dr. Knoblock is a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI), the Association of Computing Machinery (ACM), and the Institute of Electrical and Electronic Engineers (IEEE). He is also past President of the International Joint Conference on Artificial Intelligence (IJCAI) and winner of the Robert S. Engelmore Award.

Dr. Steven Minton is the founder and President of Inferlink. Minton received his Ph.D. in Computer Science from Carnegie Mellon University in 1988. After graduating, Minton worked at NASA's Ames Research Center until 1995. He then joined University of Southern California's Information Sciences Institute, where he served as a Project Leader at ISI.

Minton co-founded Fetch Technologies in 2000. Fetch raised over $14 million in capital and successfully patented and commercialized machine learning technology for data extraction. Prior to its acquisition, the company had more than 50 employees, and its customers included Dow Jones, Lexis-Nexis, AT&T, USIS, ADP, Reuters, and many other enterprises. Fetch was also an InQTel portfolio company, and its government customers included US intelligence agencies. During this period, Minton also helped found GeoSemble (which spun off from Fetch), and served as an officer of the company.

After Fetch, Minton founded Inferlink where has continued to lead research projects, and co-founded all four of InferLink’s commercial spinoffs, as described in the Commercialization Section.

Minton is a fellow of the AAAI and has received international awards for his research in machine learning, planning and AI applications. He received the 2012 Robert S. Englemore Memorial Lecture Award from the Association of the Advancement of Artificial Intelligence (AAAI), which cites his “seminal contributions in scheduling, planning and machine learning” and “advancement of real-world AI systems”. His previous awards include the 2008 AAAI Classic Paper award

In 1993 Minton founded one of the first open access journals on the Web, the Journal of Artificial Intelligence Research (JAIR), now regarded as one of the most highly-ranked AI journals. Minton was JAIR’s first Executive Editor, and he continues to serve as Managing Editor. He is also a member of the advisory board of the Journal of Machine Learning Research (JMLR), as well as the advisory board of the Computing Research Repository (associated with arXiv.org). Minton is a US citizen and holds a TS clearance.

Dr. Naveen Ashish is the Chief Scientist at InferLink. He holds a PhD in Computer Science from the University of Southern California. Ashish started his career as a computer scientist at NASA Ames Research Center, and then held informatics faculty positions at UC Irvine and USC Los Angeles. From 2016-18 he served as founding Director of Data Sciences at the Fred Hutchinson Cancer Research Center. He has managed major data science, health data analytics and NLP platform initiatives and has over 100 publications in top computer science venues, and has authored two books. Ashish was the organizer and program chair of the 1rst Workshop on Semantic Web Technologies in Searching & Retrieving Scientific Data, which became the basis for further collaborations in the Geosciences domain and led to his book titled "Geospatial Semantics and the Semantic Web". He also teaches (advanced) data science courses (part-time) at the University of Washington (Seattle) School of Information. He is a US citizen

.

References

Ashish, N., & Sheth, A. P. (Eds.). (2011). Geospatial Semantics and the Semantic Web: Foundations, Algorithms, and Applications (Vol. 12). Springer Science & Business Media.

Chen, C. C., Knoblock, C., & Shahabi, C. (2007). Automatically and accurately conflating road vector data, street maps, and orthoimagery. In US Patent App. 11. https://patents.google.com/patent/US20070014488A1/en

Chiang, Y.-Y., Chen, M., Duan, W., Kim, J., Knoblock, C. A., Leyk, S., Li, Z., Lin, Y., Namgung, M., Shbita, B., & Uhl, J. H. (2023 (in press)). GeoAI for the Digitization of Historical Maps. In *Handbook of GeoAI*. Springer.

Chiang, Y.-Y., Duan, W., Leyk, S., Uhl, J. H., & Knoblock, C. A. (2020). Using Historical Maps in Scientific Studies: Applications, Challenges, and Best Practices. Springer International

Publishing. https://doi.org/10.1007/978-3-319-66908-3

Chiang, Y.-Y., Leyk, S., & Knoblock, C. A. (2014). A Survey of Digital Map Processing Techniques. *ACM Comput. Surv.*, *47*(1), 1–44. https://doi.org/10.1145/2557423

Duan, W and Chiang, Y.-Y., and Leyk, S. and Uhl, J. and Knoblock, C. A. (2021). A Label Correction Algorithm Using Prior Information for Automatic and Accurate Geospatial Object Recognition. *2021 IEEE International Conference on Big Data (Big Data)*.

Lin, Y., & Chiang, Y. Y. (n.d.). Hyper-Local Deformable Transformers for Text Spotting on Complex Documents. *Submit to the International Conference on Computer Vision (ICCV)*.

Lin, Y., Chiang, Y.-Y., Franklin, M., Eckel, S. P., & Ambite, J. L. (2020). Building Autocorrelation-Aware Representations for Fine-Scale Spatiotemporal Prediction. *2020 IEEE International Conference on Data Mining (ICDM)*, 352–361. https://doi.org/10.1109/ICDM50108.2020.00044

Li, Z., Chiang, Y. Y., Tavakkol, S., Shbita, B., & Uhl, J. H. (2020). An Automatic Approach for Generating Rich, Linked Geo-Metadata from Historical Map Images. *Proceedings of ACM Knowledge Discovery and Data Mining Conference (KDD)*. https://dl.acm.org/doi/abs/10.1145/3394486.3403381

Li, Z., Guan, R., Yu, Q., Chiang, Y.-Y., & Knoblock, C. A. (2021). Synthetic Map Generation to Provide Unlimited Training Data for Historical Map Text Detection. *Proceedings of the 4th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, 17–26. https://doi.org/10.1145/3486635.3491070

Li, Z., Kim, J., Chiang, Y.-Y., & Chen, M. (2022). SpaBERT: A Pretrained Language Model from Geographic Data for Geo-Entity Representation. In *arXiv [cs.CL]*. arXiv. http://arxiv.org/abs/2210.12213

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Shbita, B., Knoblock, C. A., Duan, W., Chiang, Y.-Y., Uhl, J. H., & Leyk, S. (2023). Building spatio-temporal knowledge graphs from vectorized topographic historical maps. *Semantic Web*, *14*(3), 527–549. https://doi.org/10.3233/sw-222918

Shit, S., Koner, R., Wittmann, B., Paetzold, J., Ezhov, I., Li, H., Pan, J., Sharifzadeh, S., Kaissis, G., Tresp, V., & Menze, B. (2022). Relationformer: A Unified Framework for Image-to-Graph Generation. *Computer Vision – ECCV 2022*, 422–439. https://doi.org/10.1007/978-3-031-19836-6_24

Tavakkol, S., Chiang, Y.-Y., Waters, T., Han, F., Prasad, K., & Kiveris, R. (2019). Kartta Labs: Unrendering Historical Maps. *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, 48–51. https://doi.org/10.1145/3356471.3365236

Zhang, X., Su, Y., Tripathi, S., & Tu, Z. (2022). Text Spotting Transformers. In *arXiv [cs.CV]* (pp. 9519–9528). arXiv. http://openaccess.thecvf.com/content/CVPR2022/html/Zhang_Text_Spotting_Transformers_CVPR_2022_paper.html

Zhang, Q. L., & Yang, Y. B. (2021). Sa-net: Shuffle attention for deep convolutional neural networks. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2235–2239. https://doi.org/10.1109/ICASSP39728.2021.9414568

Zhu, X., Su, W., Lu, L., Li, B., Wang, X., & Dai, J. (2020). Deformable DETR: Deformable Transformers for End-to-End Object Detection. In *arXiv* http://arxiv.org/abs/2010.04159