École Polytechnique Fédérale de Lausanne

Cartonomics

Semester Project in Digital Humanities

Jiaming Jiang

Supervised by Rémi Petitpierre

8th January 2025

Abstract

Cartography reflects how humans interpret, represent, and navigate their physical and cultural environments. This project explores the application of machine learning to historical cartography, focusing on the detection and classification of cartographic icons. Using a pretrained YOLOv10 model fine-tuned on the Cartonomics dataset, we achieved robust detection performance, particularly with a crop-augmentation strategy that optimized the equilibrium between required visual context and model architecture. Embedding-based clustering provided additional insights into semantic relationships among icons, though classification remains challenging.
The results demonstrate the potential of transfer learning to resolve the problems in cartographic analysis, such as icon variability and stylistic diversity. Despite its successes in detection, the model exhibited limitations in classification. Future work will aim to address these challenges and scale the approach to larger datasets, enabling cross-cultural and temporal studies of cartographic iconography. By bridging the gap between machine learning and digital humanities, this research lays the groundwork for a systematic exploration of the history of cartographic semiotics.

CHAPTER 1

INTRODUCTION

1.1 RELATED WORK

1.1.1 VISUAL SEMIOTICS OF CARTOGRAPHY

Cartography, the science and art of map-making, is central to understanding how humans conceptualize and represent spatial relationships. Beyond their utilitarian role, maps serve as cultural artifacts that reflect not only the geographic landscapes but also the symbolic systems and conventions of their creators. Within this field, historical maps are particularly valuable as they document the geographic, political, and cultural narratives of past societies.

However, interpreting historical maps poses significant challenges due to their physical degradation and diverse design conventions and iconography across region and time. The semiotics of cartography offers a framework to address these challenges by analyzing the visual variables—such as color, texture, and symbols—that maps use for meaning-making (Schlichtmann, 2009). For instance, there is a generally observed transition in cartographic icons over time, shifting from detailed, illustrative depictions to abstract, standardized forms (R. Petitpierre et al., 2024). This progression, driven by the demand for functional clarity and broader accessibility, is also exemplified in a study of Slovenian maps spanning the 16th to 19th centuries, which highlights the evolution of symbols for settlements, vegetation, and relief as reflective of societal and technological advancements (Gašperič and Babič, 2023).

1.1.2 MACHINE-LEARNING APPROACHES

In the past decade, the development of deep learning has transformed the landscape of cartographic analysis. Traditional methods to extract information from maps relied on manual digitization or rule-based algorithms, which were limited in scalability and robustness(Gašperič, 2023). Early work, such as Dhar and Chanda, 2006, focused on extracting geographical features like rivers and roads from scanned maps, laying foundational methodologies for digitizing cartographic information.

Several efforts have aimed to advance automated cartographic analysis. Techniques such as semantic segmentation and object detection have enabled the large-scale analysis of diverse map datasets. For instance, R. G. Petitpierre et al., 2021 introduced a generic semantic segmentation framework capable of processing maps from diverse sources, highlighting the potential for cross-style generalization. Similarly, Chazalon et al., 2021 organized a competition to benchmark the performance of machine learning models on specific tasks, such as map segmentation and building block detection.

Despite these advancements, much of the focus has been on textures, terrains, and transportation networks (Uhl et al., 2022), leaving the detection and classification of cartographic icons underexplored. Recently, Smith et al., 2025 have begun to address this gap by applying object detection techniques to extract symbolic elements like tree icons from historical maps, highlighting growing interest in this area.

1.1.3 CASSINI PROJECT

Another research by our collaborators from EHESS [1] has made significant strides in addressing this gap. Leveraging deep learning techniques, their work focused on detecting and classifying icons on historical maps from the 16th to 18th centuries. Using the Cassini map dataset as a foundation, they developed an object detection framework capable of identifying cartographic pictograms and their associated labels. This approach combined manual annotations with synthetic data augmentation to overcome challenges such as variability in symbol designs and a lack of annotated training data.

Our project builds directly on the achievements of EHESS, extending their pretrained YOLOv10 model (Wang et al., 2024) to a new corpus of historical maps. By fine-tuning the model on a dataset that encompasses a broader stylistic and temporal range, this work aims to generalize the detection and classification capabilities of the original framework.

1.2 MOTIVATION

Understanding the symbolic and representational elements of cartography is essential for answering critical questions about the history of map-making:

- How have conventions for representing elevation, relief, and other natural features evolved over time?
- How do terrain types, such as urban environments or water bodies, influence the choice of cartographic symbols?
- What role do historical survey techniques and advancements in geodesy play in shaping these conventions?

Analyzing the complex iconography of historical maps requires scalable solutions beyond manual methods. Deep learning models enable automated extraction and classification of cartographic symbols, bridging the gap between theoretical semiotics and large-scale analysis. This approach reveals the evolution of cartographic conventions and their cultural and technological influences.

1.3 GOALS

The project's primary goals can be summarized as follows:

1. **Developing a Robust Detection Model**: Fine-tune a pretrained YOLOv10 model on a diverse dataset of historical maps to improve its ability to detect and classify cartographic icons.
2. **Testing and Optimization**: Evaluate the model's performance on unseen datasets, identify limitations, and refine its capabilities through iterative testing and enhancement.
3. **Laying the Foundation for Future Research**: Generate high-quality, labeled outputs (icon detections and classifications) that can serve as the basis for subsequent analysis of cartographic semantics and symbol evolution.

[1] https://github.com/CRH-EHESS/cocass-benchmarking

Chapter 2

Data and Methods

2.1 Dataset

2.1.1 Data collection

Our dataset was collected from a diverse array of reputable sources, ensuring a comprehensive representation of historical maps. These sources include the Bibliothèque nationale de France (BNF), the Leventhal Map Center, the David Rumsey Map Collection, among others. The dataset comprises 794 high-quality images, representing maps from various regions across the world. Spanning the 16th to the 20th century, this collection captures a wide temporal range, offering insights into cartographic practices and geographical perceptions over four centuries. The diversity of sources and time periods ensures that our dataset is both rich and varied, suitable for historical and geographical analyses.

2.1.2 Data Annotation

We manually annotated the maps using the following labeling system. Each cartographic icon was categorized into one of 13 classes, as detailed in the table below:

Table 2.1
Cartographic Icon Labeling System

Class ID	Description	Class ID	Description
1	Trees	2	Miscellaneous (Misc)
3	Mill	4	Bridge
5	Building (City)	6	Blank
7	Grave	8	Bushes
9	Marsh	10	Grass
11	Vine	12	Religious Edifice
13	Mountain		

Each label corresponds to a specific type of cartographic icon. In particular, class 2 (Misc) is used for icons that do not fit into any other category, while class 6 (Blank) serves as a mask to obscure specific areas in the images.

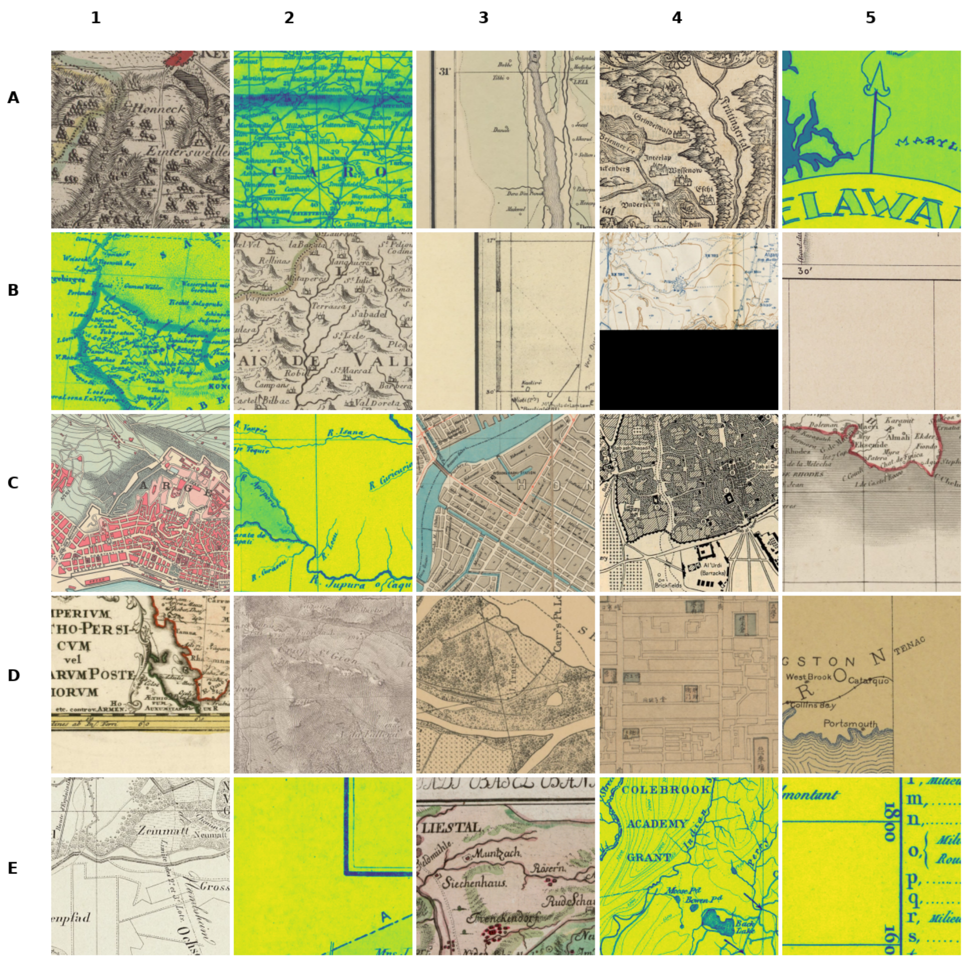


FIGURE 2.1
Samples from the corpus

2.1.3 DATA SPLITTING

We split our dataset into three subsets: train, validation, and test. The dataset was divided using the following proportions:

- **Train**: 70% of the dataset, used for model training.
- **Validation**: 20% of the dataset, used for hyperparameter tuning and model validation.
- **Test**: 10% of the dataset, used for evaluating the final model performance.

This splitting strategy ensures that the training set is sufficiently large for learning, while the validation and test sets provide robust benchmarks for model tuning and evaluation, respectively.

2.2 METHODOLOGY

This project builds on the foudation of the Cassini project, which benchmarked various object detection models—YOLO (Redmon, 2016), DETR (Carion et al., 2020), and Faster R-CNN (Ren et al., 2016)—on the Cassini dataset. Based on their evaluation, YOLO demonstrated the best performance for detecting and classifying cartographic symbols. Fig. 2.2 outlines the methodology.

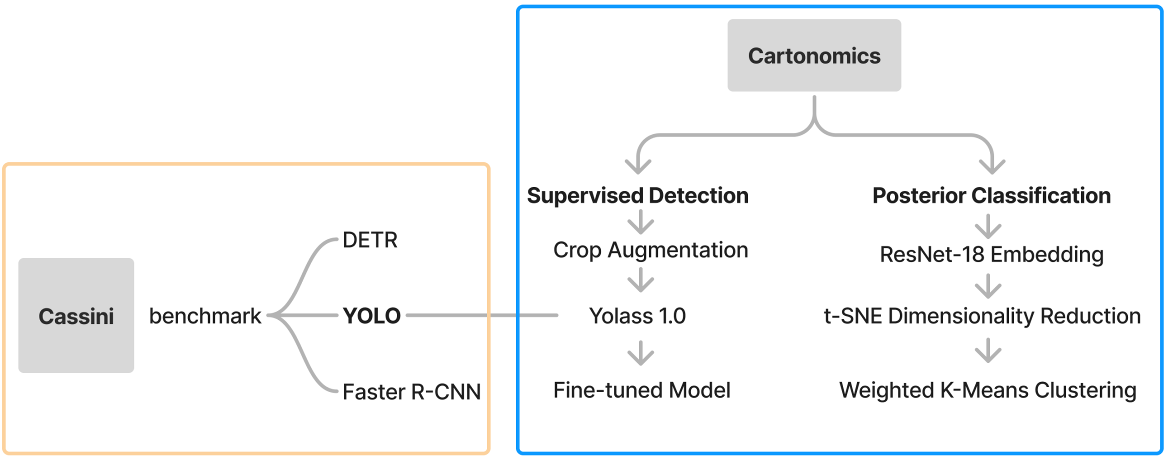


FIGURE 2.2
Methodology Flowchart
Yellow box is the Cassini project and blue box is Cartonomics project, with grey area being their dataset.

2.2.1 TRANSFER LEARNING ON DETECTION

Using the YOLOv10 model (named *yolass-1.0*) [1] pretrained on the Cassini dataset as a foundation, we applied transfer learning to adapt it to the Cartonomics dataset. This leverages the pretrained model's feature extraction capabilities while fine-tuning it to address the characteristics of the Cartonomics dataset, such as stylistic variations and icon diversity, in a computationally efficient manner.

CROP AUGMENTATION

One of the key contributions of our work is the implementation of crop augmentation to enhance the model's detection performance. In our dataset, the images samples are rather large-scale: either 768×768 or 1000×1000 pixels. To address this, we preprocess the images by cropping them into smaller patches of varying sizes—specifically, 256×256, 512×512, 640×640, and 768×768. This augmentation strategy expands the training set, ensuring better generalization to small, intricate map symbols, highlighting overlooked smaller symbols, and enhancing robustness through variations in scale and context.

MODEL CONFIGURATION

We set the learning rate to **0.002** for the default configuration, with a batch size of **8**. For smaller crop sizes of **256×256** and **512×512**, we increased the batch size to **16** to speed up training, adjusting the learning rate to **0.003** accordingly. The **AdamW** optimizer was employed for smooth and stable training.

Standard YOLO augmentations, such as mosaic, HSV variations (hue, saturation, brightness), random scaling, translation, and erasing, were disabled. Cartonomics dataset already contains adequate variability, and these augmentations could introduce noise, hindering the model's learning ability.

To account for the expected fluctuations in loss during training, the patience parameter for early stopping was set to **100**, effectively disabling early stopping. This ensures the model has ample time to converge, as the learning process for detecting small icons requires extended training due to the complexity of the task.

[1] https://huggingface.co/crh-ehess/yolass-1.0

The YOLOv10 model is fine-tuned on the crop-augmented dataset using the train and validation splits. During this step, all annotation labels were converted to a single generic label to focus solely on evaluating the detection performance, rather than classification accuracy. This simplified labeling approach provides a clear assessment of the model's ability to locate cartographic symbols within the images.

2.2.2 POSTERIOR CLASSIFICATION

While the detection framework effectively located cartographic icons, direct classification into predefined categories proved challenging due to high variability in icon styles and the incompatibility between the classification systems of Cassini project and our project. To address this limitation, we explored an embedding-based approach, which allowed us to group icons into clusters that capture latent semantic relationships, facilitating posterior classification.

FEATURE EMBEDDING

To extract robust and meaningful features from the detected icons, we employed a ResNet-18 model (He et al., 2016) pretrained on ImageNet (Deng et al., 2009) to project the extracted features into a 128-dimensional embedding space, considering the balance between representational richness and computational efficiency. The model processed all cropped and normalized icon images, generating embeddings that encapsulate their semantic content.

DIMENSIONALITY REDUCTION

To enable visualization and uncover latent structures within the embedding space, we applied t-distributed Stochastic Neighbor Embedding (t-SNE). This nonlinear technique reduced the 128-dimensional embeddings to two dimensions, while preserving their local structural relationships. The t-SNE parameters were set with **2 components**, a **perplexity of 30**, and **1000 iterations**. The reduced embeddings were then used to explore patterns and visualize semantic similarities among the icons.

WEIGHTED CLUSTERING

To group semantically similar icons, we applied K-means clustering to the two-dimensional embeddings. We set the number of clusters to 13, corresponding to the predefined classes in the dataset, and assigned weights to each sample based on the class distribution. The sample weights were computed using a balanced weighting scheme to mitigate the effects of class imbalances in the dataset.

CHAPTER 3

RESULTS

3.1 RESULTS AND ANALYSIS

3.1.1 DETECTION

The detection performance of the YOLO model was evaluated across different crop sizes, as summarized in Table 3.1. It should be noted that images without annotations are excluded by default during the conversion from COCO to YOLO format. And the validation sets in Table 3.1 don't contain those images without annotations.

The best results were achieved with the **train_640** dataset, which demonstrated the highest precision (0.958), recall (0.724), and mAP@50 (0.832). This suggests that a crop size of 640 strikes the optimal balance between context and resolution, allowing the model to effectively detect cartographic icons.

TABLE 3.1
Performance Metrics for Datasets of Different Crop Sizes (Excluding Unannotated Images)

Train set	Val set	Epochs	Batch Size	Learning Rate	Precision	Recall	mAP@50
train_768	val_768	50	8	0.002	0.924	0.625	0.724
train_640	val_640	50	8	0.002	**0.958**	**0.724**	**0.832**
train_512	val_512	50	16	0.003	0.935	0.724	0.812
train_256	val_256	50	16	0.003	0.809	0.614	0.696

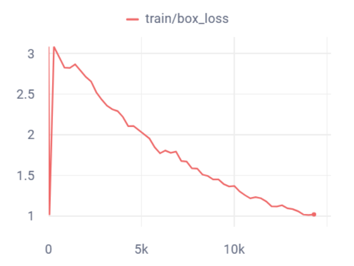
However, for a more comprehensive evaluation, we also examined the model's performance on the validation sets that included these unannotated images. The results, shown in Table 3.2, represent a more challenging and realistic scenario where the model must process images even when no objects are present. Under these conditions, performance metrics are understandably lower, as the inclusion of unannotated images reduces the overall precision, recall, and mAP@50.

The training and validation loss curves for the best-performing **train_640** model are shown in Figure 3.1. The loss curves demonstrate stable convergence over 50 epochs, indicating that the model effectively learns the task without overfitting. The consistency between the training and validation loss further confirms the model's generalizability.

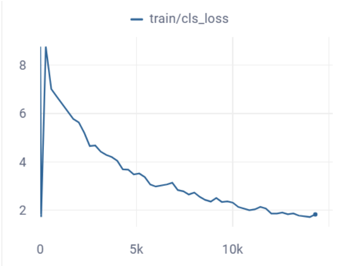
To further evaluate the model's predictions, we present sample batches of validation data along with their corresponding predictions in Figure 3.2. The labeled examples on the left show the ground truth annotations, while the predicted examples on the right show the icons the model detected.

TABLE 3.2
Performance Metrics for Datasets of Different Crop Sizes **(Including Unannotated Images)**

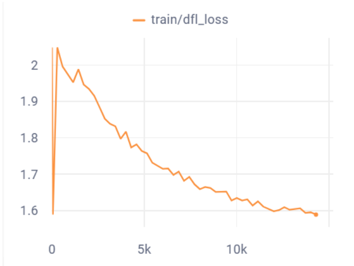
Train set	Val set	Epochs	Batch Size	Learning Rate	Precision	Recall	mAP@50
train_768	val_768	50	8	0.002	0.346	0.179	0.160
train_640	val_640	50	8	0.002	0.449	0.243	0.276
train_512	val_512	50	16	0.003	0.464	0.262	0.276
train_256	val_256	50	16	0.003	0.358	0.277	0.240



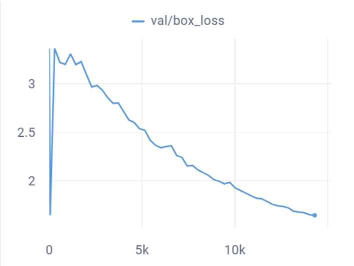
(A) train/box_loss VS step



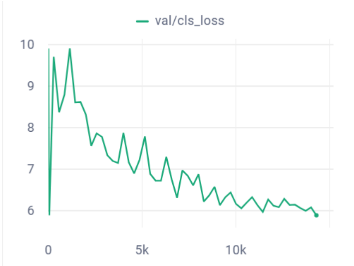
(B) train/cls_loss VS step



(C) train/dfl_loss VS step



(D) val/box_loss VS step



(E) val/cls_loss VS step

(F) val/dfl_loss VS step

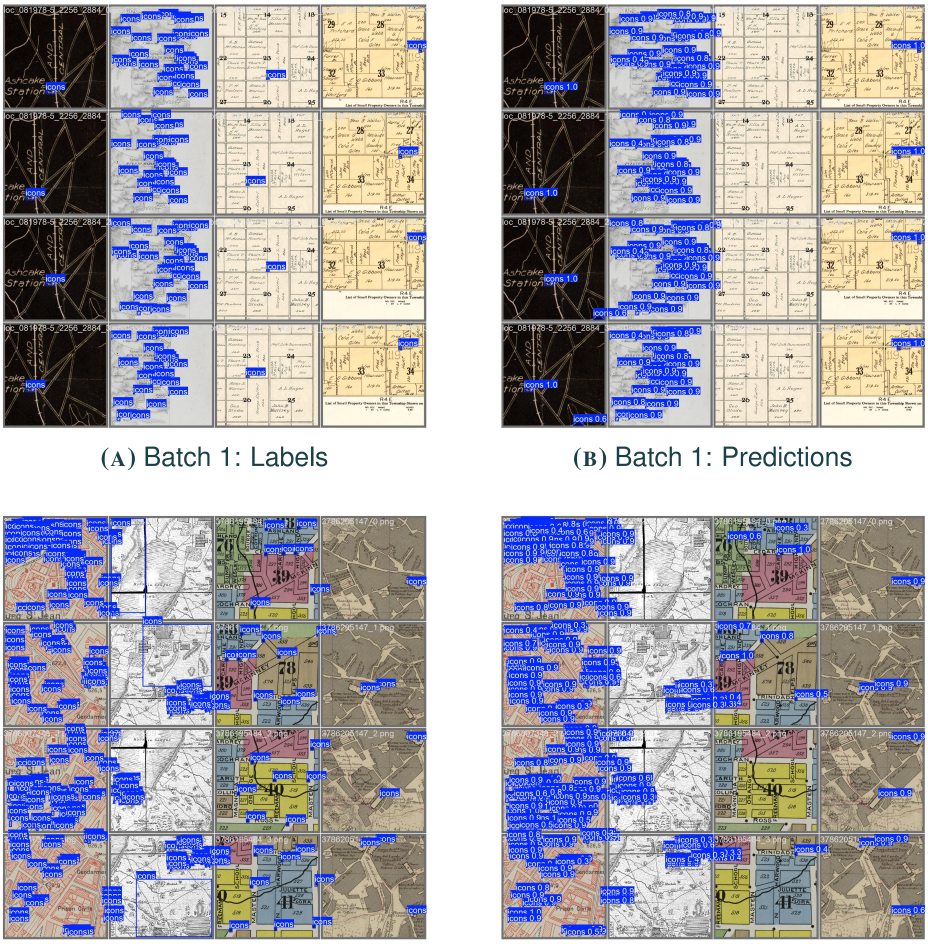
FIGURE 3.1
The training and validation loss curves for the best-performing **train_640** model

In Batch 1, the model detects most cartographic symbols with a reasonable level of accuracy, as seen in the alignment between predicted bounding boxes and the labeled ground truth. The predictions appear well-placed in regions with moderate icon density. However, some false negatives are found in the third column.

In Batch 2, the model also performs quite well, particularly in less crowded areas where individual icons are visually distinct. In more densely populated regions, missed detections are observed, with some symbols either undetected or ambiguously grouped. Despite these inconsistencies, the model demonstrates a solid ability to capture a significant portion of the labeled icons.

3.1.2 POSTERIOR CLASSIFICATION

To evaluate the semantic grouping of cartographic icons, we applied the posterior classification pipeline to both the combined training (70%) and validation (20%) sets, as well as the independent test set (10%). This process utilized the original annotations rather than model predictions so that the results reflect the embedding model's capacity to capture semantic relationships. The results are analyzed using t-SNE and weighted K-Means clustering.



(A) Batch 1: Labels

(B) Batch 1: Predictions

(C) Batch 2: Labels

(D) Batch 2: Predictions

FIGURE 3.2
Validation batches: labels and predictions for the **train_640** model.

t-SNE VISUALIZATION

Figures 3.3 and 3.4 illustrate the t-SNE projections of the embedding space for the train (including validation) and test sets, respectively. These visualizations provide insights into the latent structures of the embedding space:

- Train Set (Figure 3.3): The left subfigure shows the distribution of original category labels, while the right subfigure depicts the clustering results. Certain categories are notably dominant and dispersed as shown on the left, while the clustering algorithm inherently assigns clusters of more clear-cut boundaries and similar size.
- Test Set (Figure 3.4): Similar patterns and issues are observed in the test set, even though both subfigures look more sparse with fewer datapoints.

LABEL-CLUSTER MATRIX

As shown in Table 3.3, there exists a significant imbalance among the categories: class 1 (trees) and class 5 (building & city) are apparently more dominant. To further quantify the alignment between the clustering results and original labels, we generated a label-cluster matrix for the train (including validation) set (Figure 3.5) and the test set (Figure 3.6).

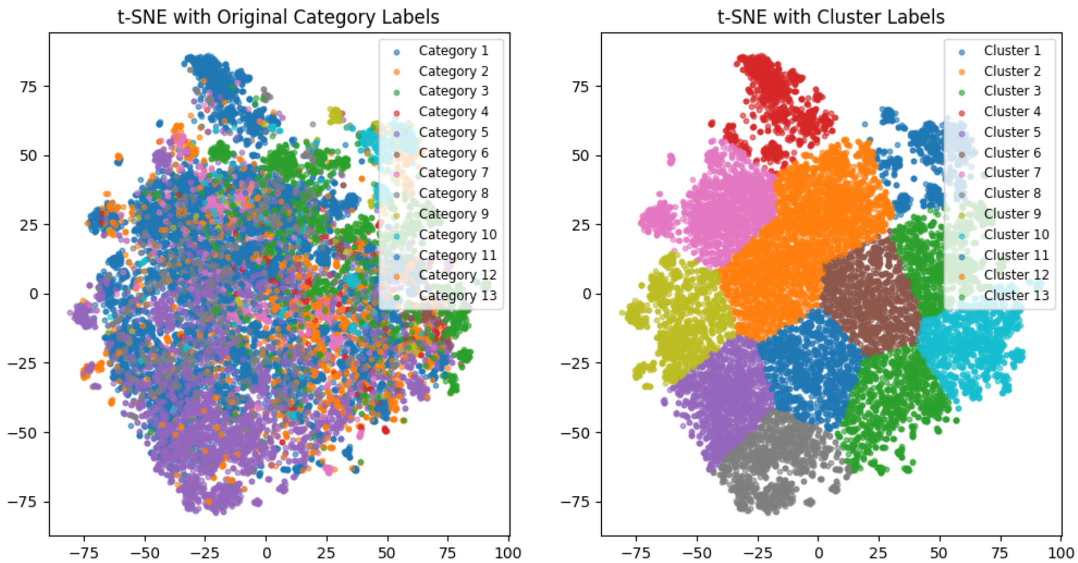


FIGURE 3.3
t-SNE cluster results compared to human labels in the train set (including val).

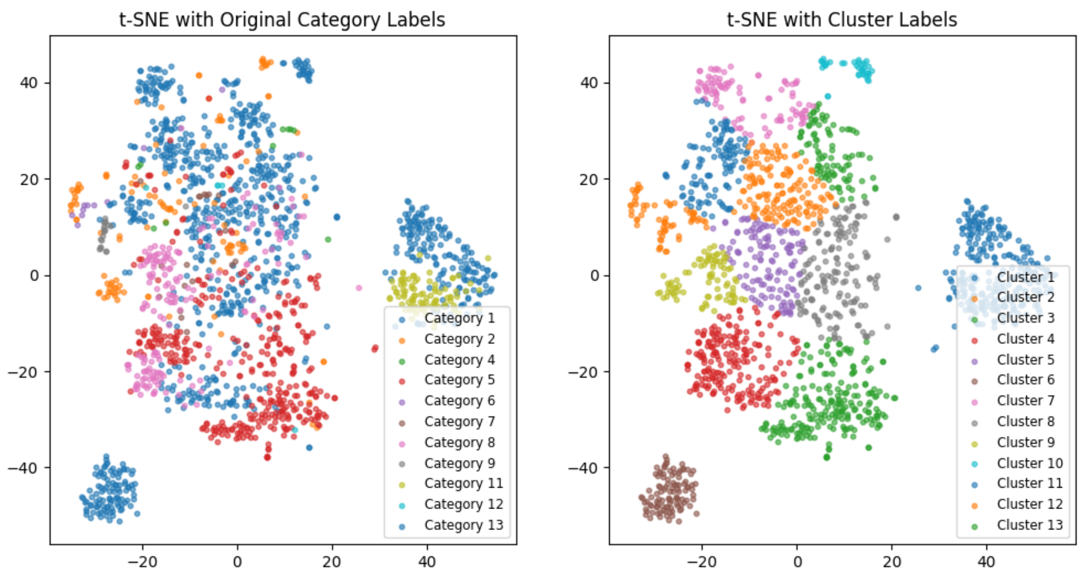


FIGURE 3.4
t-SNE cluster results compared to human labels in the test set.

TABLE 3.3
Category distribution for test and train+val splits.

Category ID	1	2	3	4	5	6	7	8	9	10	11	12	13
Test (10%)	858	119	–	10	356	10	18	166	19	–	92	5	49
Train+Val (90%)	7564	1223	73	182	3010	122	455	655	157	502	156	239	1201

CLUSTER MOSAICS

To qualitatively assess clustering performance, mosaics of images within each cluster were created to visualize the semantic grouping of cartographic icons.

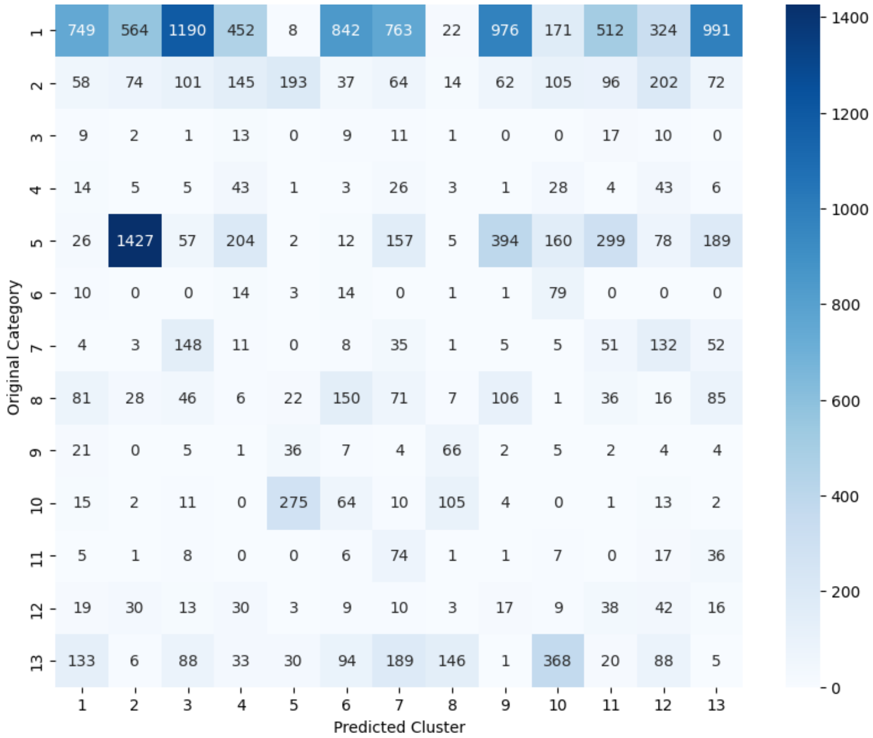


FIGURE 3.5
Label-Cluster Matrix: clustering results compared to original labels in the train set (including val).

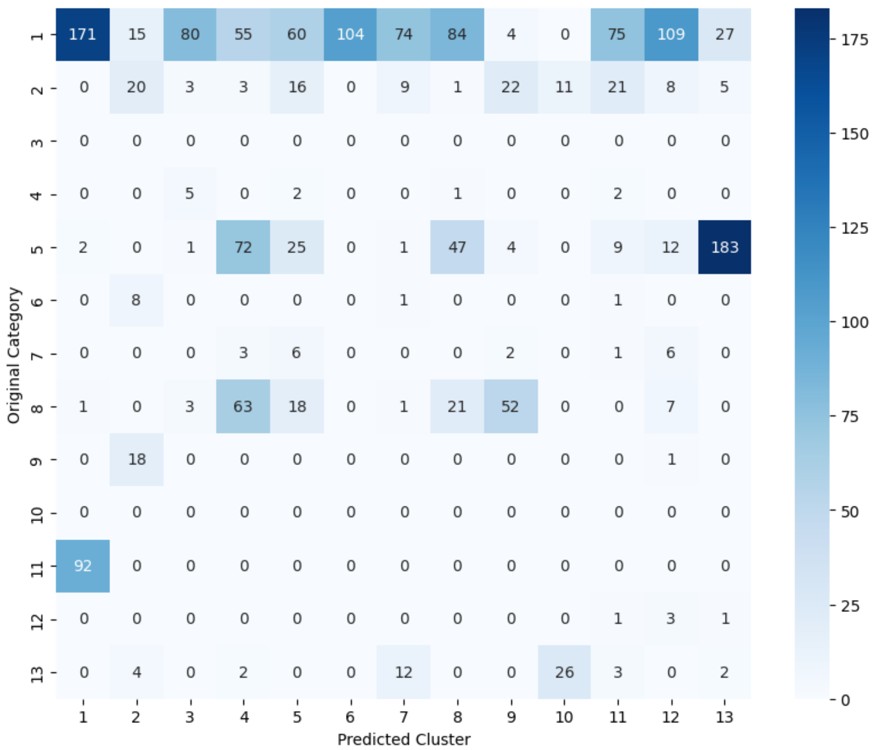


FIGURE 3.6
Label-Cluster Matrix: clustering results compared to original labels in the test set.

Figure 3.7 depicts **Cluster 5**, which predominantly contains icons from class 2 (misc) and class 10 (grass), as indicated in the label-cluster matrix. The inclusion of Grass icons suggests that these symbols may share visual attributes, such as texture or simplicity, with certain elements in the Misc category.

Figure 3.8 illustrates **Cluster 10**, where the majority of icons belong to class 13 (mountains). The mosaic reveals a cohesive grouping of symbols characterized by triangular or jagged shapes typically used to represent mountains on historical maps. This cluster demonstrates the embedding model's ability to handle distinct geometric forms effectively, resulting in well-defined semantic groups.

Figure 3.9 showcases **Cluster 2**, which is primarily composed of icons from class 5 (building and cities) and class 1 (trees). The mosaic highlights a degree of overlap between Building and Tree symbols. This is likely due to the way these features are depicted on historical maps. Cities are often represented by small circles, symbolizing settlements, while trees are sometimes depicted by circular shapes representing tree canopies. This similarity in iconography, particularly when circles are used for both, might have contributed to the embedding model grouping these classes together.

FIGURE 3.7
Mosaics of cluster 5 in the train (including val) set.

FIGURE 3.8
Mosaics of cluster 10 in the train (including val) set.

FIGURE 3.9
Mosaics of cluster 2 in the train (including val) set.

Chapter 4

Discussion

4.1 Key Achievements and Findings

The project successfully demonstrated the capability of machine learning, particularly the fine-tuned YOLO model, to detect cartographic icons with remarkable accuracy across an exceptionally heterogeneous dataset. This dataset spans multiple centuries (16th to 20th), diverse cartographic styles, and varying levels of iconographic detail, showcasing the model's robustness in adapting to stylistic and temporal variability. Despite the challenges posed by such diversity, the model achieved high precision (0.958) and mAP@50 (0.832) on iconic maps, especially with a 640x640 crop size, which optimized the balance between context and resolution.

A significant achievement lies in the model's scalability and potential for large-scale applications. The detection pipeline excels when applied to maps known to be icon-rich ("iconic"), irrespective of their stylistic or geographic origin. This opens pathways for automating iconographic analysis in extensive historical map archives, provided that pre-segmentation is used to focus on regions with icons. By leveraging this approach, researchers can streamline workflows and extract meaningful data from maps that were previously too complex to analyze manually.

The model's success in such a heterogeneous dataset demonstrates its potential to generalize across varied historical contexts, a critical prerequisite for cross-regional and temporal studies of cartographic semiotics. The ability to handle stylistic variability while maintaining high accuracy highlights the feasibility of applying this approach to larger, more complex datasets, provided the maps meet the conditions of iconicity.

4.2 Challenges and Insights

Unannotated Images in Evaluation. The conversion of the dataset from COCO format to YOLO format inherently excludes images without annotations by default. However, as shown in Table 3.2, retaining these unannotated images during evaluation greatly impacted the performance metrics. For example, the train_640 model, which achieved a precision of 0.958 and mAP@50 of 0.832 in the absence of empty images, saw these metrics drop to 0.449 and 0.276, respectively, when evaluated on the validation set that included empty images. This decline underscores the model's limited ability to handle true negatives and distinguish them from missed detections. The inclusion of unannotated images presents a more realistic and challenging scenario, revealing gaps in the model's generalizability. Future work should address these gaps by explicitly training the model on empty images to enhance its capacity for detecting true negatives, thereby improving overall robustness and reliability.

GENERALIZABILITY IN CLASSIFICATION. While the model demonstrated robust icon detection performance following fine-tuning, its classification capabilities significantly declined after transitioning from a single-label to a multi-label system encompassing the original 13 classes. Training was conducted using the YOLO model fine-tuned on the Cartonomics dataset with the best crop size of 640x640. This experiment yielded suboptimal results: precision dropped to 0.411, recall to 0.0959, and mAP@50 to 0.0965. The stricter mAP@50-95 metric further highlighted performance gaps with a value of 0.0481.

These challenges are attributed to differences in the classification systems of the Cassini and Cartonomics datasets, as well as significant stylistic variability among cartographic icons. These findings underscore the need for advanced domain adaptation techniques, multi-task learning, or hybrid architectures that integrate icon detection with context-aware classification strategies. Future work should also explore training approaches explicitly tailored for imbalanced and stylistically diverse datasets to improve classification performance.

REFINING THE LABEL SYSTEM The current label system 2.1, while effective for initial analysis, offers room for refinement to enhance classification and detection:

1. **Subdividing Broad Categories:** The “Misc” category contains diverse icons that differ significantly in appearance and function. Breaking it into more specific subcategories can reduce ambiguity and improve the model’s ability to learn finer distinctions.
2. **Grouping Related Classes:** Certain classes like “Trees,” “Vines,” “Grass,” and “Bush” share commonalities as plant symbols. Consolidating them into a broader “Vegetation” category could simplify the label structure and improve classification accuracy by reducing class overlap.
3. **Addressing Class Imbalance:** Imbalanced datasets, where dominant classes overshadow rare ones, can hinder model training. Techniques like downsampling frequent classes, upsampling rare ones, or using synthetic data generation could mitigate this issue and improve the model’s performance across all categories.

4.3 FUTURE DIRECTIONS

SCALING TO LARGER DATASETS. With further improvements, the model can be applied to larger and more diverse datasets, enabling cross-regional and temporal analyses of historical cartographic icons. This would facilitate systematic studies of patterns and variations in iconography, reflecting technological and cultural changes over time.

ANALYZING THE HISTORY OF CARTOGRAPHIC ICONS. Using outputs from the improved model, researchers can address key questions posed in the introduction:

- How have cartographic conventions evolved across periods and regions?
- What role have technological advancements and measurement skills played in shaping icon designs?
- How do different terrain types influence the choice and style of symbols?

These insights could uncover historical trends in cartographic semiotics and inform broader interdisciplinary studies.

INTERDISCIPLINARY INTEGRATION. The labeled datasets and insights generated by the model could support digital humanities research, offering historians and geographers tools to explore annotated maps, visualize icon evolution, and conduct custom analyses.

Chapter 5

Conclusion

This project demonstrates the successful application of machine learning techniques to the study of historical cartography, with significant advancements in cartographic icon detection. By fine-tuning a pretrained YOLOv10 model on the Cartonomics dataset despite its inherent heterogeneity of the dataset, which spans centuries, styles, and iconographic conventions, we achieved robust detection performance, particularly with crop-augmented images. The results highlight the scalability and adaptability of the detection framework to diverse datasets, provided they meet the condition of iconicity. The use of embedding-based clustering also provided insights into the semantic relationships among cartographic icons, though challenges in classification and handling class imbalance persist.

Key findings include the optimal performance of the model on a 640x640 crop size, the effectiveness of transfer learning for icon detection, and the potential for embedding techniques to uncover latent patterns in iconography. However, limitations such as poor generalizability in classification, reliance on annotated datasets, and the impact of class imbalance were identified as areas for further improvement.

Looking forward, future work will focus on refining the model's classification capabilities, addressing class imbalance, and enhancing feature extraction methods. Once improved, the model can be applied to larger datasets, enabling systematic analysis of cartographic icon evolution across regions and historical periods. This research contributes not only to digital humanities but also to the broader understanding of how humans have represented space through cartographic semiotics.

Bibliography

Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020). End-to-end object detection with transformers. *European conference on computer vision*, 213–229.

Chazalon, J., Carlinet, E., Chen, Y., Perret, J., Duménieu, B., Mallet, C., Géraud, T., Nguyen, V., Nguyen, N., Baloun, J., et al. (2021). Icdar 2021 competition on historical map segmentation. *International Conference on Document Analysis and Recognition*, 693–707.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. *2009 IEEE conference on computer vision and pattern recognition*, 248–255.

Dhar, D. B., & Chanda, B. (2006). Extraction and recognition of geographical features from paper maps. *International Journal of Document Analysis and Recognition (IJDAR)*, *8*, 232–245.

Gašperič, P. (2023). A new standardized methodology for analyzing cartographic information on old maps. *Acta geographica Slovenica*, *63*(2), 23–49.

Gašperič, P., & Babič, S. (2023). The semiotics of cartographic symbols on old maps. *Acta geographica Slovenica*, *63*(2), 51–71.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Petitpierre, R., Uhl, J. H., Di Lenardo, I., & Kaplan, F. (2024). A fragment-based approach for computing the long-term visual evolution of historical maps. *Humanities and Social Sciences Communications*, *11*(1), 1–18.

Petitpierre, R. G., Kaplan, F., & Di Lenardo, I. (2021). Generic semantic segmentation of historical maps. *CEUR Workshop Proceedings*, *2989*(27), 228–248.

Redmon, J. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(6), 1137–1149.

Schlichtmann, H. (2009). Overview of the semiotics of maps. *Proceedings, 24th International Cartographic Conference, Santiago, Chile*, 15–21.

Smith, E. S., Fleet, C., King, S., Mackaness, W., Walker, H., & Scott, C. E. (2025). Estimating the density of urban trees in 1890s leeds and edinburgh using object detection on historical maps. *Computers, Environment and Urban Systems*, *115*, 102219.

Uhl, J. H., Leyk, S., Chiang, Y.-Y., & Knoblock, C. A. (2022). Towards the automated large-scale reconstruction of past road networks from historical maps. *Computers, environment and urban systems*, *94*, 101794.

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). Yolov10: Real-time end-to-end object detection. *arXiv preprint arXiv:2405.14458*.