DHQ: Digital Humanities Quarterly
2024
Volume 18 Number 3

Deep Learning for Historical Cadastral Maps and Satellite Imagery Analysis: Insights from Styria's Francisean Cadastre

Wolfgang Thomas Göderle <wolfgang_dot_goederle_at_uni-graz_dot_at>, University of Innsbruck; Max Planck Institute of Geoanthropology https://orcid.org/0000-0002-9417-5316
Fabian Rampetsreiter <rampman_at_gmx_dot_net>, University of Graz https://orcid.org/0009-0004-5155-3032
Christian Macher <cmacher_at_know-center_dot_at>, Know Center https://orcid.org/0009-0009-7299-8547
Katrin Mauthner <kmauthner_at_know-center_dot_at>, Know Center https://orcid.org/0009-0003-1227-1190
Oliver Pimas <opimas_at_know-center_dot_at>, Know Center https://orcid.org/0009-0002-5480-1543

Abstract

Cadastres from the 19th century are a complex as well as rich source for historians and archaeologists, the study of which presents great challenges. For archaeological and historical remote sensing, we have trained several Deep Learning models, CNNs, and Vision Transformers to extract large-scale data from this knowledge representation. We present the principle results of our work here and demonstrate our browser-based tool that allows researchers and public stakeholders to quickly identify spots that featured buildings in the 19th century Franciscean cadastre. The tool not only supports scholars and fellow researchers in building a better understanding of the settlement history of the region of Styria; it also helps public administration and fellow citizens to swiftly identify areas of heightened sensibility with regard to the cultural heritage of the region.

Cadastral maps, which were produced with meticulous quality standards from the early 19th century onwards for large parts of Europe, offer a unique and complex representation of knowledge [Göderle 2017] [Göderle 2023] [Wheeler 2023] [Femenia-Ribera et al. 2022]. In the case of Habsburg Central Europe, where cadastral mapping occurred between 1816 and 1861, these maps provide detailed information on individual houses at a scale ranging from 1:720 to 1:5760, with a standard scale of 1:2880. They also document land use, contemporary roads and pathways, land boundaries, and other relevant data [Katastral-Vermessungs Instruktion 1824] [Bundesamt 2017, 44]. Apart from their historical significance, cadastral maps are valuable for contemporary spatial analysis, including environmental and transport history, building history, and economic and social history, as well as questions related to land use and landscape transformation in the Anthropocene era and resource extraction during the time of their creation [Ståhl and Weimann 2022] [Roberts et al. 2024]. 1

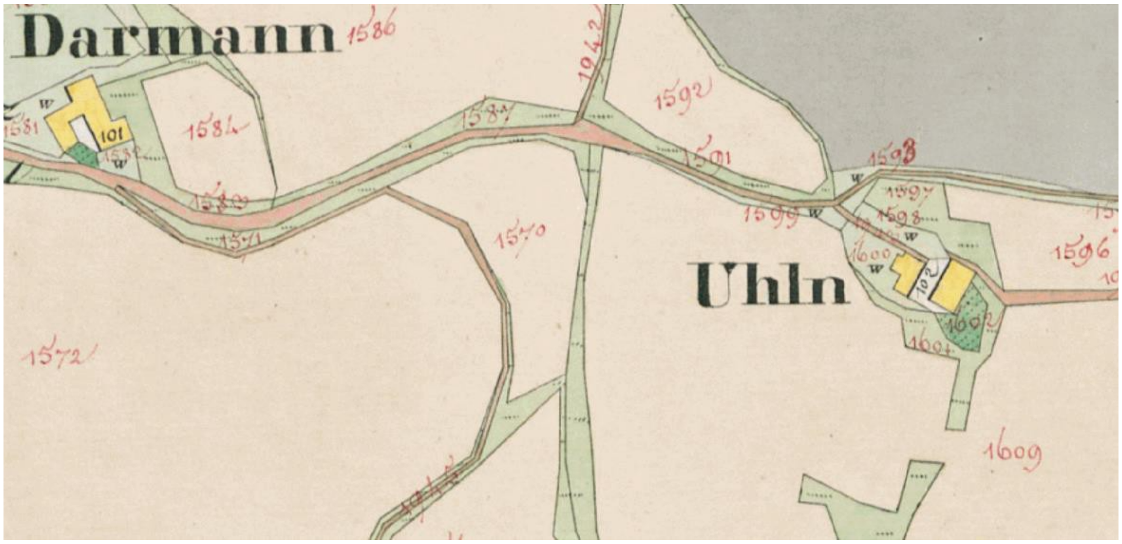

Figure 1. Extract from cadastre from 1828. Wooden houses are visible in yellow, along with paths and agricultural land.

Cadastral maps constitute a key resource for historical research into this region, as they enable the spatialization of 2
historical processes, contexts, and interdependencies. However, accessing and processing cadastral data can be challenging due to their multimodal structure and primary focus on images and geodata. This presents difficulties for digital humanities researchers, whose work is often located in the textual domain [Champion 2017] [Van Noord 2022]. We are convinced that recent advances in the field of deep neural networks enable multimodal analysis of historical sources [Smits and Wevers 2020]. This would, in turn, open up new opportunities to analyze those non-written sources that emerged in Europe (as well as globally) in the long 19th century. Sources such as the cadastre are unique and central representations of knowledge from this period, but have so far been insufficiently analyzed, especially in social and economic history research ([Hosseini et al. 2021].

Digitization has made cadastres more accessible to a wider audience, although accessing the physical volumes stored 3
in archives remains challenging for historians and researchers [Bundesamt 2017] [Femenia-Ribera et al. 2022] [Hagemans et al. 2022] [McDonough 2024]. Despite digitization efforts, the comprehensive evaluation of cadastral data, particularly in Habsburg Central Europe, where 670,000 square kilometers were surveyed between 1816 and 1861, remains labor-intensive due to the sheer volume of individual map sheets documenting every aspect of land use and property [Bundesamt 2017].

The 2022 Austrian Science Fund (FWF) project RePaSE (Reading the Past from the Surface of the Earth) aimed to use 4
AI-assisted extraction to identify historical building outlines and locations from geo-referenced map sheets of the Franziszeische Kataster (Franciscean Cadastre).[1] This involved training an AI model for building extraction from both historical maps and modern, high-resolution aerial and satellite images. By overlaying the extracted historical and present-day building information, the project sought to identify potential archaeological sites, facilitating large-scale archaeological remote sensing. The hypothesis driving RePaSE is that by combining historical cadastral map data with modern satellite imagery, it's possible to identify regions where buildings once stood, aiding in the discovery of vanished structures and potentially significant archaeological sites.

In addition to the cadastre, we worked with current spatial images in our study. We were supported by the relevant A17 5
State and Regional Development – Department of Statistics and Geoinformation of the Province of Styria, which provided us with high-resolution aerial photographs from different flight periods as well as two digital terrain models. It has been shown that even the transfer of information from georeferenced historical map images to different spatial representations of the present has enormous cognitive potential. We were able to draw on high-resolution aerial photographs from 1952; images from the 1970s, 1980s, and 1990s; and current and extremely accurate images from recent years.

The results of RePaSE, and the tools that emerge from it, should create opportunities for digital humanities researchers to analyze the information contained in historical maps. Possibilities for automated segmentation of the information contained therein would enable us to further quantitatively process data contained in maps and to utilize new possibilities for qualitative analysis [Hosseini et al. 2021]. In HGIS systems that contain georeferenced historical maps, it is already possible to calculate areas; thus, automated segmentation would significantly advance historical analysis [Baeten and Lave 2020]. Furthermore, new feature complexes could be defined and searched for automatically, extending the logics of distant reading to map analysis [Arnold and Tilton 2019]. For example, mills not explicitly identified in the French cadastre could be automatically identified, which would allow a profound analysis of hydropower use in Central Europe in the early 19th century and enable the investigation of local or regional differences. 6

We do not consider cadastre, satellite photograph, or aerial photography authentic representations of reality. Rather, we treat them as high-profile sources providing us with several layers of research data [Gil-Fournier and Parikka 2021] [Hosseini et al. 2021]. 7

State of Research

Despite the accessibility challenges mentioned above, there is substantial historical research that deals either with the construction of cadastral maps or with the information stored in them [Dolejš and Forejt 2019]. Prior to their digitization, cadastral maps have already played an important role in historical and particularly archeological research on the micro-level [Petek and Urbanc 2004]. Cadastral maps have long been a research resource with regard to genealogical research. They have been used as sources in agricultural and environmental history, environmental studies, archeological remote sensing, social and economic history, and in the development of a broader understanding of the evolution of landscapes [Drobesch 2013] [Rumpler 2013] [Scharr 2024] [Rumpler, Scharr, and Ungureanu 2015] [Hohensinner et al. 2021]. Their value, however, has frequently been limited by accessibility issues and the difficulties of extending analysis beyond relatively limited borders. 8

Access to historical cadastral maps in Central Europe became publicly available relatively late, with data now accessible through projects like Arcanum Maps (https://maps.arcanum.com/en/).[2] In most cases, the digitization of cadastral maps is accompanied by georeferencing. However, these repositories often lack comprehensive metadata and may require payment for research use. Public authorities have increasingly made parts of the Franziszeische Kataster available online for research purposes, offering data at various quality levels for free [Pivac et al. 2021]. Exploration of historical map data, including cadastral maps, has accelerated due to recent advancements in machine learning [Chiang et al. 2020] [Budig 2017]. Deep learning technologies, in particular, have revolutionized feature extraction [Chen et al. 2016], focusing on streets [Ekim, Sertel, and Kabadayı 2021] [Can, Gerrits, and Kabadayı 2021] [Jiao, Heitzler, and Hurni 2021] [Jiao, Heitzler, and Hurni 2022] [Uhl et al. 2022] and buildings [Heitzler and Hurni 2020] [Uhl et al. 2020], among other applications [Wu et al. 2023] [Petitpierre, Kaplan, and Di Lenardo 2021] [Garcia-Molsosa et al. 2021]. 9

These advancements offer new possibilities for researchers in history, archaeology, and historical geography. Recent research has focused on feature extraction from cadastral maps, particularly in the context of the Venetian cadastre and the 1900 Atlas of Paris [Oliveira et al. 2019] [Petitpierre and Guhennec 2023]. As Petitpierre and Guhennec point out, consistent annotation is the main challenge with regard to automatized vectorization. However, the bulk of current research is directed towards the analysis of aerial and satellite imagery [Jiao, Heitzler, and Hurni 2022]. This research spans a wide range of objectives and interests, with stakeholders including municipal and urban administrations, as well as archaeological remote sensing missions. Significant attention has also been given to LIDAR data and research on urban landscape transformation, in addition to real-time surveillance tasks involving UAVs [Li et al. 2022] [Li et al. 2023] [Ji, Wei, and Lu 2019] [Fiorucci et al. 2020] [Fiorucci et al. 2022] [Bickler 2021] [Ren et al. 2015] [Ding and Zhang 2021] [Luo, Wu, and Wang 2022] [Lee, Wang, and Lee 2023] [Chen, Zou, and Shi 2021] [Uhl et al. 2021]. 10

The Franciscean Cadastre, covering an extensive area of Habsburg Central Europe spanning around 300,000 square kilometers, stands out as an exceptional historical source. Its nearly half-century production period and stringent quality standards provide historians, archaeologists, and environmental scientists with privileged insights into 19th-century land 11 use and spatial organization ([Feucht 2008]. Nevertheless, the cadastre also contains inaccuracies and errors, which are magnified by the size of the operation. It represents a historical primary source and as such requires meticulous and thorough critical reading. Werner Drobesch has recently provided the first comprehensive analysis of the economics of the cadastre. From this and from the savings constraints he has pointed out, there are also indications of the weaknesses of this source [Scharr 2024]. Further, a great deal of the accuracy of the Franciscean Cadastre depends on the quality of the individual's work. The scale of the recording varies considerably, with a high level of detail in urban areas, a medium level of detail in rural areas, and very coarse details in exposed terrain and high mountains. Many of the weaknesses encountered in dealing with the Franciscean Cadastre resemble those that Hosseini et al. describe and scrutinize in detail with regard to the Ordnance Survey [Hosseini et al. 2021].

Objective

The objective of RePaSE was twofold: first, RePaSE should prove that AI-assisted extraction of large amounts of data — in our use case, the location and the form of buildings — from the cadastre is already possible with existing resources and, in the best case, identify suitable models to that purpose.[3] The second objective was to make RePaSE fit for future multimodal applications and to identify and test models that could be used to extract the locations and forms of buildings from current, already available aerial and satellite imagery [Smits and Wevers 2023]. 12

The larger context in which our work is embedded is the question of the possibilities that deep neural networks open up for the work of historians, archaeologists, and other humanities scholars. A central challenge that we are addressing with this project is to make the huge data structures that represent historical cadastres searchable at a higher level of analysis. On the one hand, we see potential for a quantitative economic and social history. Structuring and classifying the information encoded in cadastres would make it possible to analyze land use at regional or provincial level, for example. In the field of archaeology, a cost-effective type of remote sensing in the automated comparison of historical and current visual representations of space would allow faster identification of potentially archaeologically relevant large-scale historical structures. On the other hand, our tool would also make it possible to extend the archaeological view to larger spatial units with relatively little effort and to better visualize translocal or regional contexts for experts. 13

Our research strategy involved two main sub-tasks: Task 1 focused on detecting buildings in scans of historical cadastral map data, while Task 2 aimed to detect buildings in current satellite and aerial imagery. We then superimposed the output layers of both tasks, specifically targeting places where buildings were present in the cadastre but absent in the present. This overlay combined historical cadastral map data with current satellite imagery, allowing us to identify relevant locations for further analysis. These locations were then utilized for analyzing different spatial representations, including satellite and aerial imagery. 14

Approach

As the size of the annotated map patches had a significant impact on the performance of our models, we resized all map patches to the uniform size of 3747x2235, which proved effective. We also employed three different zoom levels: close, medium, and far. We then used the software tool cvat to annotate the map patches [CVAT n.d.]. The first task to be addressed was an image segmentation problem, in order to successfully extract houses from cadastral map material. In our first attempt, we followed a low-key approach and tried clustering algorithms to tackle this challenge, which was unsuccessful. In line with the fail-fast approach of our proof-of-concept study, we subsequently turned to a completely different method. When it comes to deep neural networks in this area of application, convolutional neural networks (CNN) have outperformed other architectures in recent years [Rawat and Wang 2017] [O'Shea and Nash 2015]. A CNN is a type of deep learning algorithm that is well-suited for analyzing visual data. It uses principles from linear algebra, especially convolution operations, to extract features and identify patterns within images. CNNs use a series of layers, each of which detects different features of an input image. Depending on the complexity of its intended purpose, a CNN can contain dozens, hundreds, or even thousands of layers. As building extraction can be considered a computer vision problem, fully convolutional networks are the most widespread state-of-the-art solution to address this kind of problem [Chen, Zou, and Shi 2021]. We therefore chose the Google DeepLabv3 model [Chen et al. 2016], which is very efficient in dealing with the multi-scale problem due to its atrous convolution layer [Chen et al. 2016]. We 15 finetuned DeepLabv3 with 50 annotated example images that we split into 6 squares each to give 300 images in total. The images were split into training, testing, and validation sets according to the following scheme:

Figure 2. Finetuning the DeepLabv3 model. Blue = validation data, red = training data, and green = test data.

The validation set was automatically populated with squares lacking buildings to teach the model the absence of structures, which accounts for its larger size compared to the test and train sets. Each dataset (train, test, validation) contained unique images without duplication across sets. Finetuning utilized a repository designed for transfer learning in semantic segmentation [Singh and Popescu 2021], building upon DeepLabv3 with Resnet 101. Finetuning spanned 20 epochs on various computing platforms until satisfactory results were achieved [PyTorch n.d.]. 16

The second task to be addressed was another image segmentation problem: the extraction of buildings from high-resolution aerial and satellite data. As the quality of freely available satellite imagery has further increased in the two years between the setup of the project idea and the project execution, we could now use higher quality satellite imagery than we originally expected, as well as add high-resolution aerial photography and LiDAR data to the research data under scrutiny. Unlike the 19th century cadastral map data that we were dealing with in Task 1, we were addressing an entirely different problem here, as buildings in ortho- and satellite imagery can take on a wide range of very different appearances. 17

Due to the significantly different challenge, we chose an adapted approach. Vision transformers break down input images into patches, convert each patch into a vector, and then process those vectors using a transformer encoder. This allows the model to understand the content of the image, and this method has been used for various computer vision tasks such as image recognition, image segmentation, and object detection. Vision transformers offer relatively economic computational overhead and memory consumption while still achieving state-of-the-art accuracy [Chen, Zou, and Shi 2021]. We therefore trained a sparse token transformer (STTnet) from scratch, using a GPU. STTnet is a state-of-the-art vision transformer with very promising results in comparable tasks [Chen, Zou, and Shi 2021]. 18

Building on Resnet50 as backbone, we then selected 2,657 labelled images from the AIRS dataset. This was so that we could avoid labelling houses in the satellite images ourselves. The effort for this would have been considerable, because training from scratch requires substantially more training examples than does finetuning. All images resembled the Alpine topography and the landscape characteristics of our target region, Styria. We also used images taken in Austin, TX; Chicago, IL; Kitsap, WA; Tyrol (a historical region in the Alps of northern Italy and western Austria; and Vienna, Austria. The training took place over 98 epochs; once the loss rate and the F1 score did not feature significant changes anymore, the training was considered completed. The model was then tested with 224 labelled images. 19

Subsequently, the extracted layers were superimposed and negative profiles were created. The negative profiles contain potentially interesting locations for archaeological research. While these locations once featured buildings, they are not overbuilty at present. 20

Results

The results for Task 1, the extraction of buildings from cadastral maps via DeepLabv3, were very successful. 21

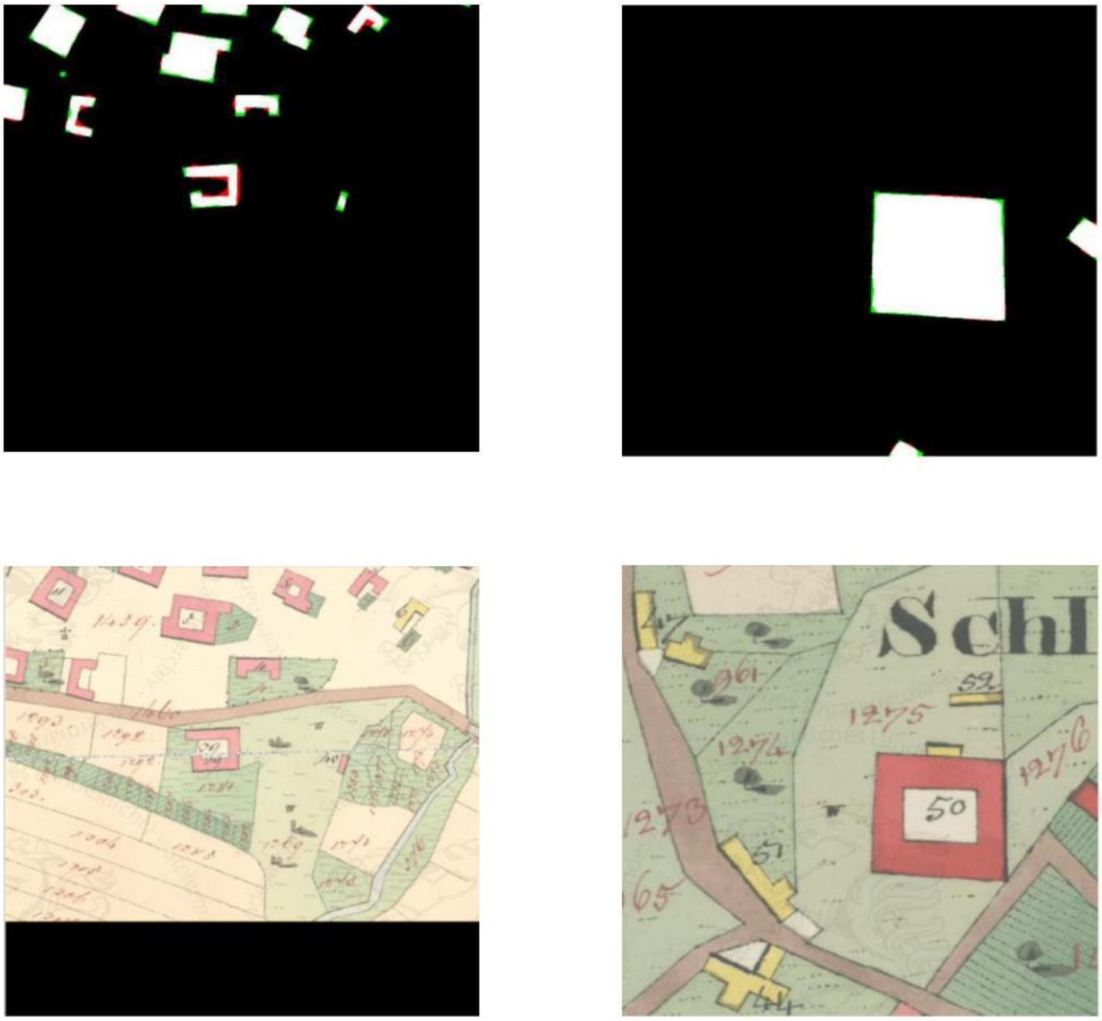

Figure 3. Building extraction from cadastral maps with finetuned DeepLabv3.

As can be seen from the images, the extraction of red houses (stone houses), for which the model was finetuned, 22
worked impeccably, which is reflected in the metrics we obtained. We calculated the parameter intersection over union (IoU), where the labelled mask, or the image that we prepared for training by labelling, is compared with the predicted mask. The IoU scores range from 0 to 1, with a score of 1 indicating a perfect overlap between the predicted and ground truth bounding boxes. A high IoU score indicates that the object detection algorithm is accurately localizing the object in the image. The value was calculated per class, with two defined classes: house (Class 1) and background (Class 2). We calculated two values as follows:

$$\text{Micro average} = \frac{\text{Intersection House} + \text{Intersection Background}}{\text{Union House} + \text{Union Background}}$$

$$\text{Macro average} = \frac{\text{IoU House} + \text{IoU Background}}{2}$$

Example 1.

Our results yielded a mean micro average IoU of 0.990389 and a mean macro average IoU of 0.89516, which represent 23
outstanding values. The accuracy of the model was thus sufficiently high. Among the problems we encountered were some streets that were mistaken for houses, as well as minor differences in the color schemes (shades) in the hand-drawn maps. These were, however, minor issues, which we believe we could resolve by labelling more training examples and retraining the model. Further finetuning could include teaching the model to recognize additional structures, particularly yellow houses (wooden buildings) and streets, which will be our priority in a follow-up project.

The results of Task 2 were slightly different. The achieved IoUs of 0.537755 (macro) and 0.795838 (micro) proved to be satisfactory in this context from a pragmatical perspective, as we were able to work with the obtained results. Although these results are significantly below the values achieved for Task 1, we consider this task successfully resolved. The building outlines contained in the negative layer must be kept more generous, in any case, because it is evident, especially in the terrain model, that the terrain is affected by the development beyond the actual building boundaries, through the levelling of the built-over area. The red markings in the upper layer mark false positives (e.g., a house was detected where no house existed), while the green markings designate false negatives (e.g., the model failed to detect the house), and the white markings indicate correct identifications. The model sometimes displays difficulties correctly identifying houses at the margins of the detection area. Furthermore, it was difficult for the model to achieve good results outside a certain (optimal) zoom range. This problem could be circumvented by properly addressing the issue in the data preprocessing. As a rule, the model tended to minimally enlarge the recognized buildings, which was quite convenient for our work, as already noted. From an academic perspective, we see some possibilities to optimize the model. We are certain that additional training, such as adding more epochs, would substantially improve the model's performance. 24

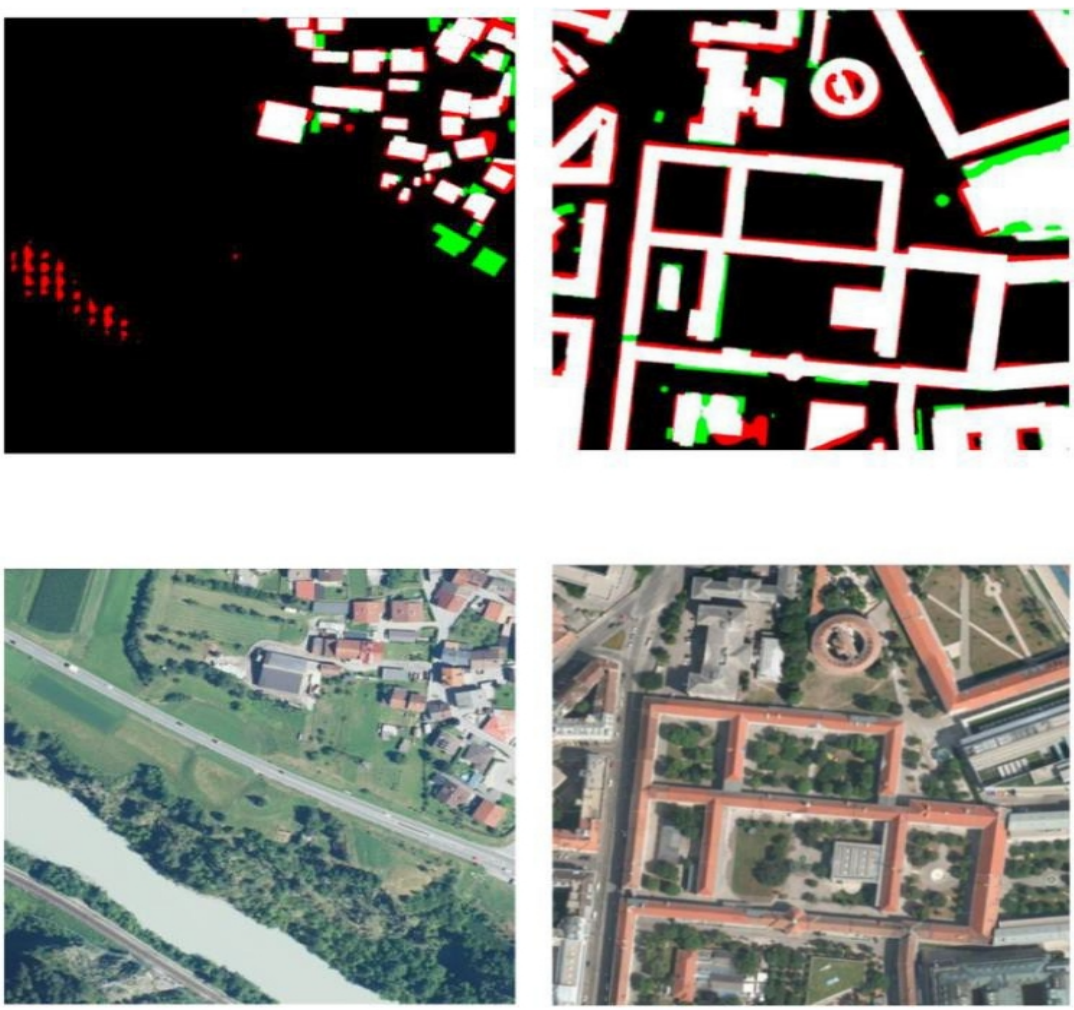

Figure 4. STT results Tyrol (left) and STT results Vienna (right).

Next Steps

RePaSE provided proof-of-concept with regard to AI-driven data extraction from historical maps, and our results represent a valuable way forward. The most important finding is that structures can be extracted from historical cadastral map material accurately and economically in very high quality. This opens up a wide range of new possibilities, and this is a field we plan on engaging further based on the results we gathered in RePaSE and in 25 cooperation with several other groups working in this field of research. In particular, we expect that the findings of our colleagues, Jiao, Heitzler, and Hurni concerning the synthetic production of suitable training data will allow for a significant breakthrough in extracting relevant data from historical maps at scale [Jiao, Heitzler, and Hurni 2022].

In view of the rapid progress in the extraction of graphic structures from high-resolution satellite and aerial photographs, it seems sensible to orient our work with regard to the current state of research rather than investing significant resources in improving the technology. Nevertheless, it is critical to remain up-to-date with advances in the field, as digital humanities and humanities data science often struggle to access state-of-the-art technologies, models, and datasets to process information. We consider the solution we identified, STTnet, viable and expandable, though it will require substantial effort to raise our outputs to a level at which the quality of results will be robust enough to qualify for positive profiles. The STTnet we trained works well in a project context such as RePaSE, but it will require some adjustments to work in a different context where a higher IoU metric might be required. 26

We have identified structures that can be detected very well when using computer vision and which bear important time signatures, promising to render visible to researchers' eyes the past that is hidden in the surface of the Earth. Furthermore, because we have recently found encouraging results using other object detectors for related tasks, we believe we have identified a promising path of development in this field. 27

Our study has shown us that, in principle, our models would also be suitable for recognising and reporting structural changes over time in a more finely graduated manner. With a different model (e.g., experiments with YOLOv5 were promising), building outlines in the cadastral material could be recognised even more finely and, theoretically, buildings that may not have undergone any structural changes could be detected automatically. Although RePaSE was not aimed at determining the structures that replaced historical buildings, the models could also contribute to clarifying such questions with a manageable adaptation effort. 28

A follow-up to RePaSE has recently been granted by the Austrian Science Fund (FWF), which will fully integrate the two models trained and finetuned within RePaSE (DeepLabv3 and STTnet). The objective of this follow-up is to identify and train two further models with a focus on historical road infrastructure. 29

Conclusion

RePaSE provides a wide variety of stakeholders, from municipal authorities and political decision-makers at the local and regional level to historians and archeologists, with a powerful tool to detect sites of potential archeological relevance, as is shown in the following two figures: 30

Figure 5. High-resolution aerial image of a part of the rural municipality Heiligenkreuz in the present day. The red bounding boxes mark the automatically detected locations of buildings that featured in the 1820s cadastre. The yellow shade marks the identified present-day buildings.

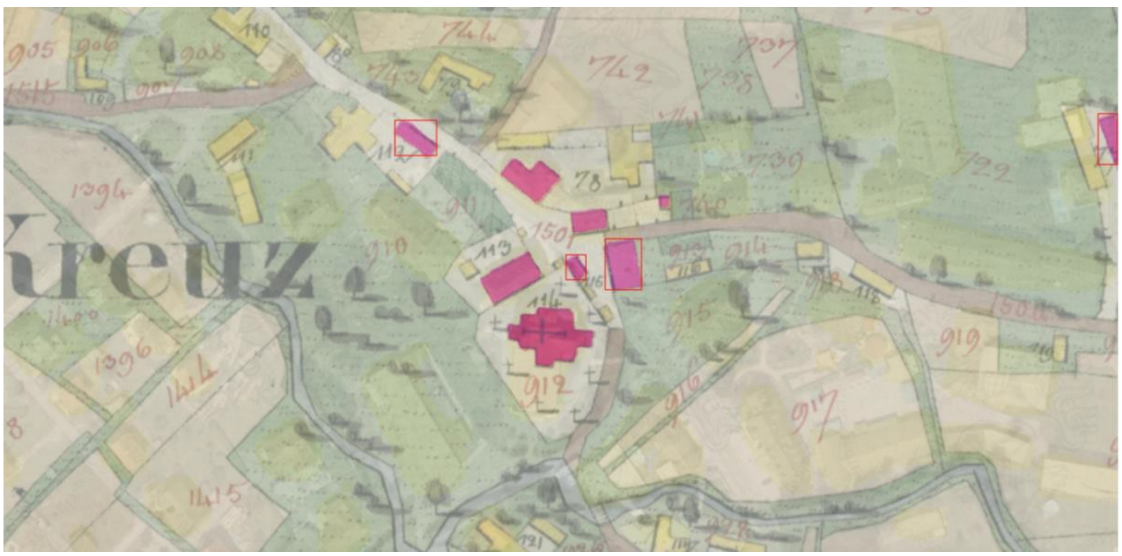

Figure 6. The respective map patch in the cadastre. The bounding boxes mark the buildings that have since disappeared.

Based on our findings, we are convinced that the results of RePaSE will be an important support for various 31
stakeholders in Central Europe. In the field of construction management, our model is likely to be used to identify archaeologically and historically sensitive areas in advance of major construction and infrastructure projects. In the context of RePaSE, we define “archaeologically and historically sensitive” areas as sites that once featured historic buildings and may require archaeological investigation in the course of construction activities or current land use. RePaSE can support the identification of such sites in advance and thus help to prevent so-called emergency excavations (e.g., by changing the route or planned archaeological intervention before construction begins). The model can also help to identify relevant structures in cultural heritage management and make them more comprehensively visible, for example, in existing GIS and HGIS systems. Primarily, however, we see RePaSE as the basis for a research tool designed to support digital humanists in identifying and extracting relevant research data from mapped sources.

The question arises to what extent neural networks can also be trained to detect visually relevant anomalies in different 32
dimensions of surface representations (digital terrain model, different high-resolution aerial photographs) without having

to rely on concrete knowledge from historical knowledge representations. Such detection could be achieved by a neural network by identifying complex feature clusters in multimodal visual input signals, such as the superimposition of DTM and high-resolution aerial images from different aerial surveys.

The RePaSE-team is currently working to set up a server to provide the target group with full access to the entire dataset processed in the project runtime.[4] While RePaSE currently supports research on the region of Styria, we expect that it could easily be expanded for the adjoining regions of Carinthia, Lower Austria, Salzburg, and Slovenia. As Croatia and Burgenland were part of the Transleithanian cadastre, which features slightly different optical features, we expect to encounter some difficulties there, though these issues could likely be resolved through another round of finetuning. 33

RePaSE demonstrates the enormous potential of machine learning, particularly regarding the use of computer vision, image segmentation, and object detection to scale up analysis of historical map material and combine these insights with AI-harvested data gained from current, high-resolution imagery of the surface of the Earth. Models of the latest generation feature outstanding usability and require manageable effort to successfully extract data from large-scale knowledge representations. In the context of this project, however, domain knowledge plays an increasingly important role. 34

Due to the availability, usability, and accessibility of image segmentation, it is increasingly becoming a standard technology. With this in mind, we must consider what tools digital humanities will be using to work with such data in the near future. Based on our results, we believe that object detection bears enormous potential for many crucial tasks, particularly when it comes to the analysis of historical map data at scale. The object detectors we have been working with so far display a wide range of possible applications, which reach far beyond the purposes for which they were developed. The fact that many object identification models combine excellent object recognition with relatively manageable hardware requirements makes them suitable tools for use in DH contexts. That being said, the most relevant potential for these tools is their capacity to bring spatial analysis in a historical context to scale. What is more, a large set of different tools exists to visualize such changes, which will help historians and fellow humanists to develop a deeper understanding of how space transforms over time. 35

Notes

[1] RePaSE was funded by the Austrian Science Fund (FWF) under the reference TAI 591. Further information on RePaSE and access to the demonstrator is available at https://repase.uni-graz.at. The demonstrator can be accessed via https://repase.know-center.at.

[2] Further relevant repositories and projects focusing on the cadastre can be found under https://www.franziszeischerkataster.at. Additionally, the project HiLaK (Historische Landnutzung als Grundlage für Klimaschutzmaßnahmen heute), directed by Kurt Scharr, and its follow-up project HiLuC display state-of-the-art approaches to cadastre-based multi-disciplinary research in the field with regard to Austria.

[3] We were granted access to the highest resolution research data hosted by https://gis.stmk.gv.at/wgportal/atlasmobile. Similar research data are available for all nine provinces of Austria. For example, data for the province of Carinthia is available at https://gis.ktn.gv.at/webgisviewer/atlas-mobile/. Unfortunately, comparable access is not easy in all successor states of the Habsburg monarchy. The Arcanum service offers good usability and orientation possibilities, but no metadata is offered free of charge that would enable scholarly work and the resolution is limited.

[4] We will provide all datasets that we used and produced to the scientific community as soon as the project has been concluded administratively, which will most probably be the case in Q3/2024.

Works Cited

Arnold and Tilton 2019 Arnold, T. and Tilton, L. (2019) “Distant viewing: Analyzing large visual corpora”, *Digital Scholarship in the Humanities*, 34(Supplement 1), pp. i3-i6. https://doi.org/10.1093/llc/fqz013.

Baeten and Lave 2020 Baeten, J. and Lave, R. (2020) “Retracing rivers and drawing swamps: Using a drawing tablet to reconstruct an historical hydroscape from Army corps survey maps”, *Historical Methods*, 53(3), pp. 182–198. https://doi.org/10.1080/01615440.2020.1748151.

Bickler 2021 Bickler, S.H. (2021) “Machine learning arrives in archaeology”, *Advances in Archeological Practice*, 9(2), pp. 186–192. https://doi.org/10.1017/aap.2021.6.

Budig 2017 Budig, B. (2017) *Extracting spatial information from historical maps: Algorithms and interaction*. Würzburg, Germany: Würzburg University Press.

Bundesamt 2017 Bundesamt für Eich- und Vermessungswesen (ed.) (2017) *Österreichisches Kulturgut: 200 Jahre Kataster, 1817–2017*.

CVAT n.d. CVAT (n.d.) *CVAT: Open data annotation platform*. Available at: https://www.cvat.ai/.

Can, Gerrits, and Kabadayı 2021 Can, Y.S., Gerrits, P.J., and Kabadayı, M.E. (2021) “Automatic detection of road types from the third military mapping survey of Austria-Hungary historical map series with deep convolutional neural networks”, *IEEE Access*, 9. https://doi.org/10.1109/ACCESS.2021.3074897.

Champion 2017 Champion, E.M. (2017) “Digital humanities is text heavy, visualization light, and simulation poor”, *Digital Scholarship in the Humanities*, 32(Supplement 1), pp. 25–32. https://doi.org/10.1093/llc/fqw053.

Chen et al. 2016 Chen, L. et al. (2016) “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs”, *arXiv*. Available at: http://arxiv.org/abs/1606.00915.

Chen et al. 2017 Chen, L. et al. (2017) “Rethinking atrous convolution for semantic image segmentation”, *arXiv*. Available at: https://arxiv.org/abs/1706.05587.

Chen, Zou, and Shi 2021 Chen, K., Zou, Z., and Shi, Z. (2021) “Building extraction from remote sensing images with sparse token transformers building extraction from remote sensing images with sparse token transformers”, *Remote Sensing* 13(21). https://doi.org/10.3390/rs13214441.

Chiang et al. 2020 Chang, Y. et al. (2020) *Using historical maps in scientific studies: Applications, challenges, and best practices*. New York: Springer.

Ding and Zhang 2021 Ding, W. and Zhang, L. (2021) “Building detection in remote sensing image based on improved YOLOV5”, *17th Annual International Conference on Computational Intelligence and Security (CIS)*. Chengdu, China, 19-22 November. pp. 133-136. https://doi.org/10.1109/CIS54983.2021.00036.

Dolejš and Forejt 2019 Dolejš, M. and Forejt, M. (2019) ““Franziscean cadastre in landscape structure research: A systematic review”, *Quaestiones Geographicae*, 38(1), pp. 131–144. https://doi.org/10.2478/quageo-2019-0013.

Drobesch 2013 Drobesch, W. (ed.) (2013) *Kärnten am Übergang von der Agrar- zur Industriegesellschaft: Fallstudien zur Lage und Leistung der Landwirtschaft auf der Datengrundlage des Franziszeischen Katasters (1823-1844)*. Klagenfurt, Austria: Verlag des Geschichtsvereines Kärnten.

Ekim, Sertel, and Kabadayı 2021 Ekim, B., Sertel, E., and Kabadayı, M.E. (2021) “Automatic road extraction from historical maps using deep learning techniques: A regional case study of Turkey in a German World War II map”, *ISPRS International Journal of Geo-Information*, 100(8). https://doi.org/10.3390/ijgi10080492.

Femenia-Ribera et al. 2022 Femenia-Ribera, C., Mora-Navarro, G., and Santos Pérez, L.J. (2022) “Evaluating the use of old cadastral maps”, *Land Use Policy*, 114. https://doi.org/10.1016/j.landusepol.2022.105984.

Feucht 2008 Feucht, R. (2008.) *Flächenangaben im Österreichischen Kataster*. Diploma thesis. Vienna University of Technology. Available at: https://repositum.tuwien.at/handle/20.500.12708/11046.

Fiorucci et al. 2020 Fiorucci, M. (2020) “Machine learning for cultural heritage: A survey”, *Pattern Recognition Letters*, 133(May), pp. 102–108. https://doi.org/10.1016/j.patrec.2020.02.017.

Fiorucci et al. 2022 Fiorucci, M. (2022) “Deep learning for archaeological object detection on LiDAR: New evaluation measures and insights”, *Remote Sensing*, 14. https://doi.org/10.3390/rs14071694.

Garcia-Molsosa et al. 2021 Garcia-Molsosa, A. (2021) “Potential of deep learning segmentation for the extraction of archaeological features from historical map series”, *Archaeological Prospection*, 28(2), pp. 187–99. https://doi.org/10.1002/arp.1807.

Gil-Fournier and Parikka 2021 Gil-Fournier, A. and Parikka, J. (2021) “Ground truth to fake geographies: Machine vision and learning in visual practices”, *AI and Society*, 36, pp. 1253–1262. https://doi.org/10.1007/s00146-020-01062-3.

Göderle 2017 Göderle, W.T. (2017) “Modernisierung durch Vermessung? Das Wissen des modernen Staats in Zentraleuropa, circa 1760–1890”, *Archiv Für Sozialgeschichte*, 57, pp. 155–186. Available at: https://library.fes.de/pdf-files/afs/bd57/afs57_10_goederle.pdf.

Göderle 2023 Göderle, W. (2023) “Imperiales Wissen: Zensus und Karte: Landesvermessungen und Volkszählungen im Habsburgerreich in der Sattelzeit (1750-1850)”, in Feichtinger, J. and Uhl, H. (eds.) *Das integrative Empire: Wissensproduktion und kulturelle Praktiken in Habsburg-Zentraleuropa*. Bielefeld, Germany: transcript, pp. 73-96.

Hagemans et al. 2022 Hagemans, E. (2022) “The new, LADM inspired, data model of the Dutch cadastral map”, *Land Use Policy*, 117(June). https://doi.org/10.1016/j.landusepol.2022.106074.

Heitzler and Hurni 2020 Heitzler, M. and Hurni, L. (2020) “Cartographic reconstruction of building footprints from historical maps: A study on the Swiss Siegfried map”, *Transactions in GIS*, 24(2), pp. 442–61. https://doi.org/10.1111/tgis.12610.

Hohensinner et al. 2021 Hohensinner, S. (2021) “Land use and cover change in the industrial era: A spatial analysis of Alpine River catchments and Fluvial corridors”, *Frontiers in Environmental Science*, 9(June). https://doi.org/10.3389/fenvs.2021.647247.

Hosseini et al. 2021 Hosseini, K. et al. “Maps of a nation? The digitized ordnance survey for new historical research”, *Journal of Victorian Culture*, 26(2), pp. 284–299. https://doi.org/10.1093/jvcult/vcab009.

Ji, Wei, and Lu 2019 Ji, S., Wei, S., and Lu, M. (2019) “Fully convolutional networks for multisource building extraction from an open aerial and satellite imagery data set”, *IEEE Transactions on Geoscience and Remote Sensing*, 57(1), pp. 574–586. https://doi.org/10.1109/TGRS.2018.2858817.

Jiao, Heitzler, and Hurni 2021 Jiao, C., Heitzler, M., and Hurni, L. (2021) *A survey of road feature extraction methods from raster maps*, *Transactions in GIS*, 25(6), pp. 2734–2763. https://doi.org/10.1111/tgis.12812.

Jiao, Heitzler, and Hurni 2022 Jiao, C., Heitzler, M., and Hurni, L. (2022) “A fast and effective deep learning approach for road extraction from historical maps by automatically generating training data with symbol reconstruction”, *International Journal of Applied Earth Observation and Geoinformation*, 113(September). https://doi.org/10.1016/j.jag.2022.102980.

Katastral-Vermessungs Instruktion 1824 *Katastral-Vermessungs Instruktion* (1824). Available at: http://www.franziszeischerkataster.at/images/1820_KATASTRALVERMESSUNGSINSTRUKTION.pdf.

Lee, Wang, and Lee 2023 Lee, K., Wang, B., and Lee, S. (2023) “Analysis of YOLOv5 and DeepLabv3+ algorithms for detecting illegal cultivation on public land: A case study of a riverside in Korea”, *International Journal of Environmental Research and Public Health*, 20(3). https://doi.org/10.3390/ijerph20031770.

Li et al. 2022 Li, J. et al. “A review of building detection from very high resolution optical remote sensing images”, *GIScience and Remote Sensing*, 59(1), pp. 1199-1225. https://www.tandfonline.com/doi/full/10.1080/15481603.2022.2101727.

Li et al. 2023 Li, M. et al. “Method of building detection in optical remote sensing images based on SegFormer”, *Sensors*, 23(3). https://doi.org/10.3390/s23031258.

Luo, Wu, and Wang 2022 Luo, X., Wu, Y., and Wang, F. (2022) “Target detection method of UAV aerial imagery based on improved YOLOv5”, *Remote Sensing*, 14(19). https://doi.org/10.3390/rs14195063.

McDonough 2024 McDonough, K. (2024) “Maps as data: Computational spatial history”, in Tilton, L., Mimno, D., and Johnson, J.M. (eds.) *Computational humanities*. Minneapolis, Minnesota: University of Minnesota Press.

O'Shea and Nash 2015 O’Shea, K. and Nash, R. (2015) “An introduction to convolutional neural networks”, *ArXiv*. https://arxiv.org/abs/1511.08458.

Oliveira et al. 2019 Oliveira, A. et al. (2019) “A deep learning approach to cadastral computing”, *Digital humanities conference 2019*, Utrecht, Netherlands, 8-12 July. Available at: https://infoscience.epfl.ch/record/268282?ln=en.

Petek and Urbanc 2004 Petek, F. and Urbanc, M. (2004) “The Franziscean land cadastre as a key to understanding the 19th-century cultural landscape in Slovenia”, *Acta geographica Slovenica*, 44(1), pp. 89-113. http://doi.org/10.3986/AGS44104.

Petitpierre and Guhennec 2023 Petitpierre, R. and Guhennec, P. (2023) “Effective annotation for the automatic vectorization of cadastral maps”, *Digital Scholarship in the Humanities*, 38(3), pp. 1227–1237. https://doi.org/10.1093/llc/fqad006.

Petitpierre, Kaplan, and Di Lenardo 2021 Petitpierre, R., Kaplan, F., and Di Lenardo, I. (2021) “Generic semantic segmentation of historical maps”, *Conference on computation humanities research (CHR2021)*, Amsterdam, Netherlands, 17-19 November. Available at: https://ceur-ws.org/Vol-2989/long_paper27.pdf.

Pivac et al. 2021 Pivac, D. et al. (2021) “Availability of historical cadastral data”, *Land*, 10(9). https://doi.org/10.3390/land10090917.

PyTorch n.d. PyTorch (n.d.) *DeepLabv3*. Available at: https://pytorch.org/hub/pytorch_vision_deeplabv3_resnet101/.

Rawat and Wang 2017 Rawat, W. and Wang, Z. (2017) “Deep convolutional neural networks for image classification: A comprehensive review”, *Neural Computation*, 29(9), pp. 2352–2449. https://doi.org/10.1162/neco_a_00990.

Ren et al. 2015 Ren, S. et al. (2015) “Faster R-CNN: Towards real-time object detection with region proposal networks”, *NeurIPS 2015 conference*, Montréal, Canada, 7-12 December. Available at: https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf.

Roberts et al. 2024 Roberts, P. et al. (2024) “Using urban pasts to speak to urban presents in the Anthropocene”, *Nature Cities*, 1(30-41). https://doi.org/10.1038/s44284-023-00014-4.

Rumpler 2013 Rumpler, H. (ed.) (2013) *Der Franziszeische Kataster im Kronland Kärnten (1823-1844)*. Klagenfurt am Wörthersee, Austria: Geschichtsverein Kärnten.

Rumpler, Scharr, and Ungureanu 2015 Rumpler, H., Scharr, K. and Ungureanu, C. (eds.) (2015) *Der Franziszeische Kataster im Kronland Bukowina/Czernowitzer Kreis (1817-1865): Statistik und Katastralmappen*. Vienna, Austria: Böhlau Verlag.

Scharr 2024 Scharr, K. (ed.) (2024) *Der Franziszeische Kataster im Kronland Österreichisch-Schlesien (1821–1851): The Franciscan cadastre in the crown land Austrian Silesia: Statistik und Katastralmappen*. Vienna, Austria: Böhlau Verlag.

Singh and Popescu 2021 Singh, M. and Popescu, A.C. (2021) *DeepLabv3FineTuning*. Available at: https://Github.Com/Msminhas93/DeepLabv3FineTuning.

Smits and Wevers 2020 Smits, T. and Wevers, M. (2020) “The visual digital turn: Using neural networks to study historical images”, *Digital Scholarship in the Humanities*, 35(1), pp. 194–207. https://doi.org/10.1093/llc/fqy085.

Smits and Wevers 2023 Smits, T. and Wevers, M. (2020) “A multimodal turn in digital humanities: Using contrastive machine learning models to explore, enrich, and analyze digital visual historical collections”, *Digital Scholarship in the Humanities*, 38(3), pp. 1267–1280. https://doi.org/10.1093/llc/fqad008.

Ståhl and Weimann 2022 Ståhl, N. and Weimann, L. (2022) “Identifying wetland areas in historical maps using deep convolutional neural networks”, *Ecological Informatics*, 68(May). https://doi.org/10.1016/J.ECOINF.2022.101557.

Uhl et al. 2020 Uhl, J.H. et al. (2020) “Automated extraction of human settlement patterns from historical topographic map series using weakly supervised convolutional neural networks”, *IEEE Access*, 8. https://doi.org/10.1109/ACCESS.2019.2963213.

Uhl et al. 2021 Uhl, J.H. et al. (2021) “Towards the automated large-scale reconstruction of past road networks from historical maps”, *Computers, Environment and Urban Systems*, 94(June). https://doi.org/10.1016/j.compenvurbsys.2022.101794.

Uhl et al. 2022 Uh, J.H. et al. (2022) “Combining remote-sensing-derived data and historical maps for long-term back-casting of urban extents”, *Remote Sensing*, 13(18). https://doi.org/10.3390/rs13183672.

Van Noord 2022 Van Noord, N. (2022) “A survey of computational methods for iconic image analysis”, *Digital Scholarship in the Humanities*, 37(4), pp. 1316–1338. https://doi.org/10.1093/llc/fqac003.

Wheeler 2023 Wheeler, R. (2023) “European official surveys of the nineteenth century available online”, *Charles Close Society*. Available at: https://www.charlesclosesociety.org/C19EuropeSurveys.

Wu et al. 2023 Wu, S. et al. (2023) “Domain adaptation in segmenting historical maps: A weakly supervised approach through spatial co-occurrence”, *ISPRS Journal of Photogrammetry and Remote Sensing*, 197(March), pp. 199–211. https://doi.org/10.1016/j.isprsjprs.2023.01.021.