Self-Constructing Graph Convolutional Networks for Semantic Segmentation of Historical Maps

Lukas Arzoumanidis [a], Julius Knechtel [a]*, Jan-Henrik Haunert [a], Youness Dehbi [b]

[a] *University of Bonn, Institute of Geodesy and Geoinformation, Bonn, Germany, {arzoumanidis, knechtel, haunert}@igg.uni-bonn.de*
[b] *HafenCity University Hamburg, Computational Methods Lab, Hamburg, Germany, youness.dehbi@hcu-hamburg.de*

* Corresponding author

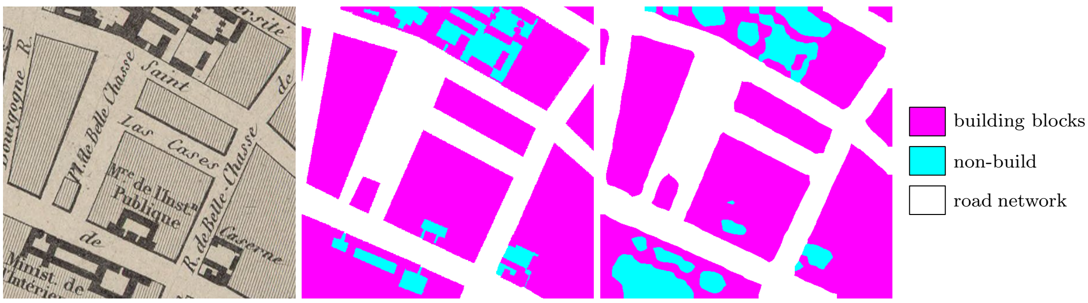
Keywords: Historical map processing, Semantic Segmentation, Graph Convolutional Networks, Self-Constructing Graph

Abstract:

Historical maps precisely capture the spatial reality of the past. Hence, they are a fundamental source of geographic information (Chiang et al., 2020), e.g. as a prerequisite for backtracking environmental changes including land-use change, urbanization, epidemiology and landscape ecology (Schlegel, 2019). To avoid a deterioration of this information and consequently preserve it and make it widely available, often a *digitization* is performed, which mostly corresponds to the process of scanning a historical document and storing the resulting raster graphic in a database. However, in this case the scanned historical documents, i.e. maps, neither have searchable metadata nor are semantically enriched (Chiang et al., 2020). Hence, invaluable spatial information remain unreachable (Schlegel, 2021). Thus, the ultimate goal is to digitize historical maps in a way that their information and content is actually readable, searchable and analyzable by machines. In combination with a database this enables an effective and intuitive analysis and comparison with current map features (Schlegel, 2019, 2021).

As yet, the digitization is often performed in a manual fashion. However, this is laborious and time-consuming and, hence, inefficient for practical use. Additionally, different involved people could interpret features differently resulting in a low level of consistency. Therefore, the implementation of automatic approaches, which are more efficient and yield consistent results, seems desirable. Nevertheless, the complexity as well as the extensive stylistic variety among the historical maps and their individual type faces, i.e. their semiology, pose major challenges when developing automatic approaches. The great success of machine learning-based approaches, especially neural networks, for the semantic segmentation of challenging datasets, for instance, recently motivated Petitpierre et al. (2021) to apply a Convolutional Neural Network (CNN) for the identification and extraction of historical map features. The network is called *dhSegment* and was originally developed for the semantic segmentation of historical documents (Ares Oliveira et al., 2018). It consists of a slightly modified *U-Net* architecture with a *ResNet-50* feature encoder. Petitpierre et al. (2021) applied the pre-trained network in combination with a preprocessing step to identify the frames of the maps on a map corpora that consists of highly heterogeneous images of historical maps from Paris.

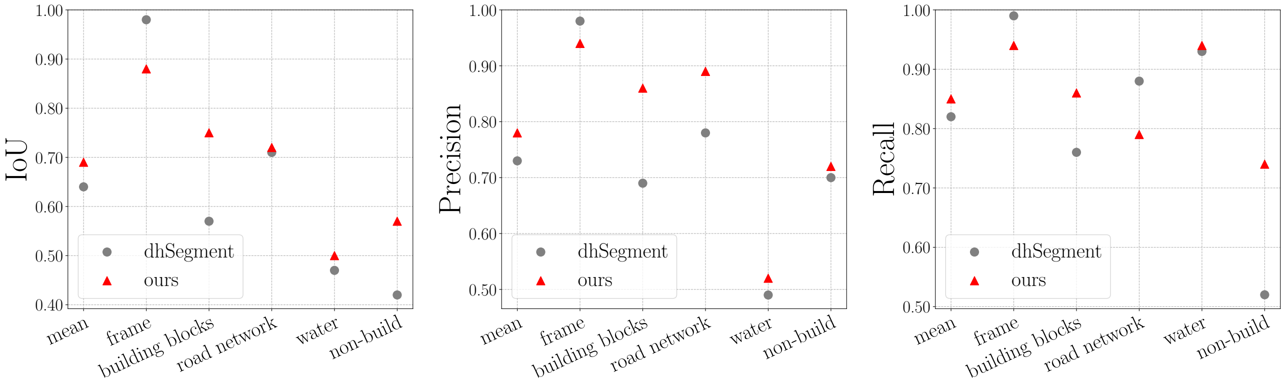
Graph Convolutional Networks (GCNs) are often used for tasks like text classification. Nevertheless, recent publications showed increasing success applying them for the semantic segmentation of images, e.g. aerial farmland images and urban sceneries (Liu et al., 2020). Therefore, we aimed for improving the identification of historical map features by transferring the GCN of Liu et al. (2020). This network consists of three parts: First, a CNN encoder is applied to compute features from the image. Next, a Self-Constructing Graph module has been used to learn a graph structure, with the features representing the nodes and the presence of arcs their respective similarity. This module allows for an automatic induction



(a) image patch (b) ground truth (c) our result

Figure 1. Exemplary qualitative result of our network.

Abstracts of the International Cartographic Association, 6, 11, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-11-2023 |



(a) IoU (b) Precision (c) Recall

Figure 2. Different accuracy measures for our network and dhSegment (Petitpierre et al., 2021).

of the underlying graph and, hence, no manual design of the graph structure is necessary. The third part of the network is a GCN. Herewith, the GCN is able to capture long-range dependencies more efficiently than a CNN which would need a significantly deeper structure and, consequently, an increasing number of parameters.

We tested the network on the above-mentioned Paris dataset (Petitpierre et al., 2021), which consists of maps dating from 1800 to 1950 at a scale between 1:2000 and 1:25000. Stemming from the *Bibliothèque nationale de France* and the *Bibliothèque historique de la Ville de Paris* they are culturally and geographically homogeneous, but figuratively very diverse. In particular, information concerning mobility, such as road or underground network, but also on the water system, the catacombs, or the administrative divisions are present. The dataset provides five different classes, i.e. *non-build*, *building blocks*, *road network*, *water* and *frame*. It originally consists of 300 images for training and 30 for validation and testing with a resolution of 1000×1000 pixels, but for our training we split up the images in patches of 500×500 pixels. We used a *ResNet-101* as encoder, 32×32 nodes for training, 64×64 nodes for the inference, a batch size of 9 and applied a learning rate of 3.47×10^{-4}. Additionally we utilized the *Adaptive Class Weighting Loss* from Liu et al. (2020) which also performs well on highly imbalanced datasets, an issue inherent to the present dataset as well.

Figure 1 depicts an exemplary input image (a), the ground truth (b) and the corresponding result of our network (c). Additionally, Figure 2 provides some quantitative results in comparison with the results from *dhSegment 4+1* (Petitpierre et al., 2021). In this variant, which is the most comprehensive one, they predicted the class *frame* in a preprocessing step and trained their network only for the four remaining classes. In contrast, our network is trained on all five classes, including *frame*, simultaneously. Nevertheless, our network outperforms *dhSegment* considering the mean values of all tested quality measures, i.e. *IntersectionOverUnion* (IoU), precision and recall. Although the pre-trained *frame* prediction performs better, our network shows significant improvement in the other classes considering all three measures, e.g. of 0.18 in IoU and 0.19 in precision of *building blocks* and of 0.15 in IoU and even 0.22 in recall for *non-build* structures. Solely the recall for the *road network* is worse, however, there is a significant improvement of 0.11 in the precision. Precision and recall often form a trade-off, which is also reflected in the comparable values of the IoU for the *road network*.

Based on the comparative study, our approach for the semantic segmentation of historical maps outperforms the CNN-based state-of-the-art *dhSegment* (Petitpierre et al., 2021) and represents a promising approach for the automatic detection and interpretation of historical map features. Additional preprocessing and augmentation of training data will be subject of future work to further improve the performance on heterogeneous datasets.

References

Ares Oliveira, S., Seguin, B. and Kaplan, F., 2018. dhSegment: A generic deep-learning approach for document segmentation. In: *Frontiers in Handwriting Recognition (ICFHR), 2018 16th International Conference on*, IEEE, pp. 7–12.

Chiang, Y.-Y., Duan, W., Leyk, S., Uhl, J. and Knoblock, C., 2020. *Using historical maps in scientific studies: Applications, challenges, and best practices*. Springer.

Liu, Q., Kampffmeyer, M., Jenssen, R. and Salberg, A.-B., 2020. Multi-view self-constructing graph convolutional networks with adaptive class weighting loss for semantic segmentation. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* pp. 199–205.

Petitpierre, R., Kaplan, F. and di Lenardo, I., 2021. Generic semantic segmentation of historical maps. *Computational Humanities Research Conference* pp. 228–248.

Schlegel, I., 2019. Empirical study for a deployment of a methodology for improving the comparability between historical and current maps. *KN J. Cartogr. Geogr. Inf.* 69, pp. 121–130.

Schlegel, I., 2021. Automated extraction of labels from large-scale historical maps. *AGILE GIScience Ser., 2* 2, pp. 1–14.

Abstracts of the International Cartographic Association, 6, 11, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-11-2023