Historical Map Toponym Extraction for Efficient Information Retrieval

Ladislav Lenc, Jiří Martínek, Josef Baloun[1], Martin Prantl, Pavel Král

1 Introduction

Information retrieval (IR) in historical documents is a challenging task. Documents occur mostly as digital images with no text layer that makes an efficient information retrieval very difficult and cumbersome. A special group of such documents are historical maps. The maps contain several interesting pictographic elements: boundary stones and milestones, scale, and other symbols determined by the map legend. In addition, they also include so-called toponyms, i.e. local names of towns, municipalities, villages, but also forests, hills, or paths. Such toponyms are crucial for IR and make up the main keywords that are mostly found in user queries in the IR systems.

We focus on toponym detection, classification and recognition in order to annotate each map sheet with a set of keywords to facilitate efficient searching and IR. An important contribution consists in proposition of a novel method for map toponym classification based on KAZE descriptor. We also compare and evaluate several state-of-the-art methods for text and object detection.

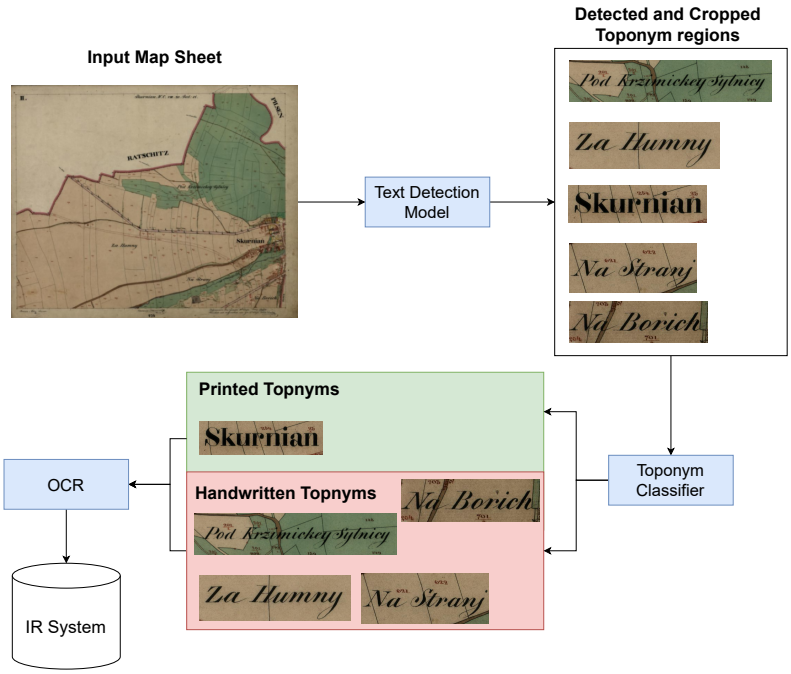

Figure 1: Overall Processing Pipeline

[1]Ph.D student, University of West Bohemia, Faculty of Applied Sciences, Dept. of Computer Science & Engineering, program Computer Science and Engineering, e-mail: balounj@kiv.zcu.cz

2 Method description

The overall pipeline is depicted in Figure 1. Three blue boxes represent models for particular sub-tasks: text detection, toponym classification and OCR.

For the text detection, we examined the baseline based on connected components analysis and more sophisticated approaches are based on neural networks:

- Faster R-CNN model with ResNet-50-FPN backbone (Ren et al., 2015);
- EAST: an efficient and accurate scene text detector (Zhou et al., 2017);
- HP-FCN: High Performance Fully Convolutional Network (Wick & Puppe, 2018);
- YOLO: You Only Look Once (we use YOLOv5 for our experiments) (Redmon et al., 2016).

YOLOv5 and Faster R-CNN are able to realize the detection and classification together. However, the other models predict only a bounding box (region of interest). For classification, we use further a novel algorithm based on KAZE key-point detector (Lenc et al., 2022).

3 Results

Based on presented results (Lenc et al., 2022). EAST is the best text detection model in terms of the average precision. Its drawback, though, is that it is not able to predict a toponym label per se. The proposed toponym classification algorithm reaches excellent classification results even with low amount of training data compared to neural network based approaches.

Although Faster R-CNN and YOLOv5 obtained very good detection and classification results, the best strategy seems to be in separated training: 1) focus on the best possible text detection and 2) use a separate classification algorithm that can be easily adapted for classification into more classes.

Acknowledgement

This work has been partly supported from ERDF "Research and Development of Intelligent Components of Advanced Technologies for the Pilsen Metropolitan Area (InteCom)" (no.: CZ.02.1.01/0.0/0.0/17_048/0007267) and by Grant No. SGS-2022-016 "Advanced methods of data processing and analysis".

References

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.

Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., & Liang, J. (2017). East: an efficient and accurate scene text detector. *In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 5551-5560).

Wick, C., & Puppe, F. (2018, April). Fully convolutional neural networks for page segmentation of historical document images. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)* (pp. 287-292). IEEE.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Lenc, L., Martínek, J., Baloun, J., Prantl, M., & Král, P. (2022). Historical Map Toponym Extraction for Efficient Information Retrieval. In *International Workshop on Document Analysis Systems* (pp. 171-183). Springer, Cham.