Deep Learning for Historical Cadastral Maps Digitization: Overview, Challenges and Potential

Jelena Ignjatić

Urban and Spatial Planning
Institute of Vojvodina
Zeleznicka 6/II
21000, Novi Sad, Serbia
ignjaticjelena@gmail.com

Bojana Nikolić

University of Novi Sad,
Faculty of Technical Sciences
Trg Dositeja Obradovica 6
21000, Novi Sad, Serbia
bojana.nikolic.ftn@gmail.com

Aleksandar Rikalović

University of Novi Sad,
Faculty of Technical Sciences
Trg Dositeja Obradovica 6
21000, Novi Sad, Serbia
a.rikalovic@uns.ac.rs

Dubravko Ćulibrk

University of Novi Sad,
Faculty of Technical Sciences
Trg Dositeja Obradovica 6
21000, Novi Sad, Serbia
dculibrk@uns.ac.rs

ABSTRACT

Cartographic heritage of historical cadastral maps represent remarkable geospatial data. Historical cadastral maps are generally regarded as an essential part of the land management infrastructure (buildings, streets, canals, bridges, etc.). Today these cadastral maps are still in use in a digital raster form (scanned maps). Digitization of cadastral maps is time consuming and it is a challenge for scientists and engineers to find ways to automatically convert raster into vector maps. The process of map digitization typically involves several stages: preprocessing, visual object detection and classification, vector representation postprocessing and extracting information from text. Although neural networks have had a long history of use in the domain, their applications remain limited to extracting the information from text. Recent convergence of advancements in the domains of training deep neural networks (DNN) and GPU hardware allowed DNNs to achieve state-of-the-art results in computer vision applications, beyond hand-written text recognition. This paper provides an overview of different approaches to historical cadastral maps digitization, focusing of the challenges and the potential of using deep neural networks in map digitization.

Keywords

Convolutional Neural Networks, Deep Neural Networks, Map Digitization, Map Vectorization, Pattern recognition

1. INTRODUCTION

Cadastral plans represent a comprehensive register of parcels and facilities of a certain area with related information [Vas16]. Cartographic heritage of historical cadastral maps of the Habsburg Empire (i.e. Austro-Hungarian Monarchy) from XVII and XIX century represent remarkable geospatial data in Europe. The territory of Serbia, which was a part of the Austro-Hungarian Monarchy prior to the First World War, had the geodetic plans made by the graphic method, and cadastral plans were done based on them. In many parts of the territory of Serbia,

Bosnia and Herzegovina and Montenegro old Austro-Hungarian cadastral plans are still in use, divided into sheets according to the regulations that were applied, at that time, in Austria-Hungary [Sap17]. The cadastral maps for parts of Serbia's Northern Province of Vojvodina consist of such analogue maps, which were scanned to obtain raster images and then georeferenced.

While these images are already in digital form, the full process of digitization usually represents the process of conversion of an analogue map into a vector graphics map with all of its associate data [Vas16]. The shift from raster to vector representation is the more challenging part of the process. Therefore, in the study presented here, we focus mainly on this part of the process of map digitization, where the input to the process are digital images representing cadastral maps from XIX century.

ISSN 2464-4617(print)
ISSN 2464-4625(CD)

Computer Science Research Notes
CSRN 2803

Poster's Proceedings
http://www.WSCG.eu

Cadastral maps in vector form have many advantages over the raster ones: faster data processing is achieved, map modification and replenishment is easier, and less storage space is required.

The problem of complete automatization of the conversion of raster maps into vector form has not been adequately solved so far and continues to attract research interest. The literature on the methods for automatic conversion of cadastral plans from raster to vector form has a long history. In the past 30 years, several automatic systems have been developed for digital map processing. Works related to the automatic vectorization of cadastral maps date back to the 1990s [Eji90], [Jan93], [Che96], [Kat99].

Although neural networks have been used in this domain for a long time [Che96], [Kar08], they have recently experienced an expansion in the form of deep learning, achieving significant improvements over the state-of-the-art in a number of applications. Deep Neural Networks (DNNs) have demonstrated prepossessing and even human-competitive results on many object and pattern recognition tasks, especially when it comes to vision classification problems. However, a review of the relevant literature reveals that the applications of deep learning methods to problem of the vectorization of cadastral plans have been very limited [Oli17].

Map digitization is a procedure which requires visual object and pattern recognition, such as the identification of edges (parcel and building lines), symbols, text, and patterns – areas in which DNNs achieve excellent performance.

In the study presented here, we will strive to evaluate the potential benefits that can be achieved through a more extensive use and application of deep learning technology to the problem of vectorization of cadastral maps automatically, or with minimal human intervention. As an aid for researchers the paper presents an overview of the problem, existing state-of-the-art solutions proposed and, more importantly, presents a detailed discussion of state-of-the-art deep learning architectures and approaches that can readily be applied to the domain and or are likely to achieve best results.

2. RELATED WORK

In earlier times, geographic data sets were used only in their analogue form. Today, with the development of technology, geographic information is mainly available in the form of digital data sets.

The problem of digitizing geographic data sets is complex. To illustrate the problem, this paper focuses on digitization of cadastral maps done under the reign of Mary Theresa [McG66]. Fig.1 shows a part of the

original scanned cadastral map, while Fig.2 corresponds to the same map in vector form.
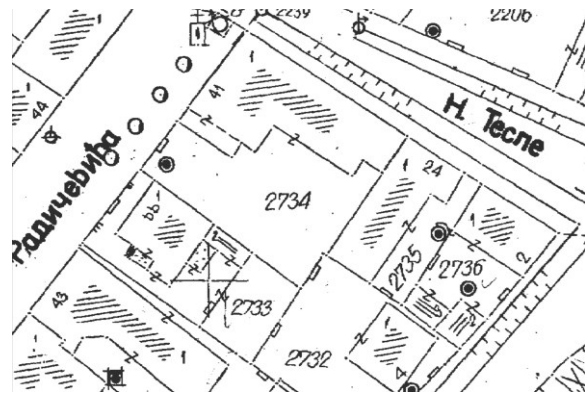

Figure 1. An original scanned cadastral map from the time of Mary Theresa

Figure 2. A sample from a vectorized cadastral map from the time of Mary Theresa (corresponding to the map in Figure 1)

There are many studies that deal with map vectorization [Boa92], [Ill91], [Con97], but the first that attempts to use neural networks to solve the complex problem of the digitization of Chinese cadastral maps and automatic extraction of high-level information was done by Chen et al. in 1996 [Che96]. The system they had developed at the time, consisted of three main parts: the separation of text and graphics, extraction of parcels and recognition of rotated characters. Even though the text and graphic separation and parcels extraction was not done by machine learning methods, they were among the first ones to propose a neural network approach to the recognition of Chinese hand-written and rotated characters. The used neural network was a two-dimensional Hopfield neural network [Hop82] which was able to reflect the degree of similarity of the test characters to previously stored templates in the final states of neurons.

Several years later, Katona and Hudra proposed a slightly different pipeline to deal with the issue of cadastral map vectorization [Kat99]. The authors of the paper developed an application based on the approach, dubbed MAPINT. The main processing phases of their approach are: raster-to-vector conversion, segmentation,

recognition and automatic correction of vectorization anomalies. The raster-to-vector conversion is the phase in which a raw vector image is generated from the scanned map as a skeleton image, through line thinning and line tracing. In the segmentation phase, raw vectors are classified into three basic categories: (text) symbols, dashed lines and continuous lines. The recognition phase is divided into three steps: recognition of symbols, recognition of connection-signs and recognition of buildings and parcels. Only the recognition of symbols was done by a neural network. The full architecture of the net is not specified by the authors, other than that it is trained using back-propagation.

Like Chen et al. [Che96], Katona and Hudra [Kat99] were using the neural network for recognition of rotated hand-written symbols which were pre-segmented as a groups of symbols, i.e. strings. The contribution of their paper is reflected in the distinction between house numbers and parcel numbers, where the distinction is made by size and length of number strings.

In 2008, Karabork et al. [Kar08] used a mulit-layer perceptron neural net to trace lines and detect joints in a skeleton image obtained from scanned maps. The study is remarkable in tha it seems an only relatively early attempt to use neural nets for more than text recognition in the domain of map digitization, but their approach failed to improve on Sparse Pixel Vectorization which was proposed some years earlier [Wen97].

Recently, Oliveira et al. [Oli17] discussed the implementation of a fully automated process based on machine learning algorithms for Napoleonic cadastral map digitalization. They proposed a system for the vectorization of the cadastral map, where their methodology consists of several steps: preprocessing, segmentation, regional classification, parcel extraction, digit extraction and digit recognition.

Once basic preprocessing is done, in order to eliminate variations in hue present in scanned maps, regions are extracted in the form of merged super pixels and a Support Vector Machines (SVM) classifier is used to classify the regions into 3 classes: text, contour and background.

In the final step of their approach, a deep Convolutional Neural Network (CNN), called LeNet [LeC98], is used to recognize digits. The network was used "out-of-the-box", without any additional training or fine tuning, was pre trained on the MNIST dataset [LeC98], and was used only to recognize digits in their pipeline.

This may have contributed to the fact that their method achieved an accuracy of only 10% for identifying hand-written numbers, although the authors propose that this was not due to the recognition algorithm itself, but rather because of an insufficiently good digit segmentation procedure.

3. DEEP LEARNING FOR MAP DIGITIZATION

With the advent of the age of Big Data and the development of new technologies, in recent years deep learning methods have been shown to outperform previous traditional state-of-the-art machine learning techniques in several fields, including computer vision [Vou18].

As we are moving towards more complete computer vision, full historic map understanding becomes achievable, as the crucial steps in this process are visual object and pattern recognition, which is becoming more precise and detailed. Recent years have produced great advances when it comes to training large deep neural networks. However, deep architectures, which are characterized by many building blocks [Yos15], [Pra18], [Cir12], require significant computational resources for training. A fact that was a limiting factor for the advancement of deep learning over several years.

The problem has been remedied by significant progress made in the development of Graphics Processing Units (GPU), which accelerate the whole process and enable the training of large DNN to last for days, instead of months [Cir12]. Today, DNNs represent the state-of-the-art in object and pattern recognition and one can expect the full integration of deep learning technology in the domain of map processing in the near future.

In comparison with more "classical" (shallow) Neural Network (NN) architectures, DNNs are able to learn high-level features, which are more complex and abstract, by integrating feature learning and model construction into a single model. That model is created by selecting different kernels or tuning the parameters by end-to-end optimization with the minimum human inference. where the parameters of DNN model are trained jointly [He15], [Sze17]. Its deep architecture, with many hidden layers, is a multi-level non-linear operation, transferring each layer's features from original input into more abstracted features in the higher layers to find complicated inherent structures. Unlike DNNs, shallow NNs approaches the extraction of the features and the construction of the model are done separately, where each module is formed by step-by-step model training [Wan18].

In the rest of this section, we provide a description of CNN, which is the common DNN method used for computer vision and image recognition.

Map Digitization: CNN approach

To learn about objects and patterns from a cadastral maps, a model with a large learning capacity is needed. The last several years have produced tremendous progress in training powerful DNN models, which are used for challenging computer vision problems. The applications of DNNs to computer vision tasks usually rely on a special form of DNNs, called Convolutional Neural Networks (CNN) [Mur10], which were firstly proposed by LeCun and Bengio [LeC95].

A CNN is a multi-layer feed-forward artificial deep neural network that has fewer connections and parameters than other NN architectures originally intended for two-dimensional image processing [LeC98], [Wan18], [Kri12].
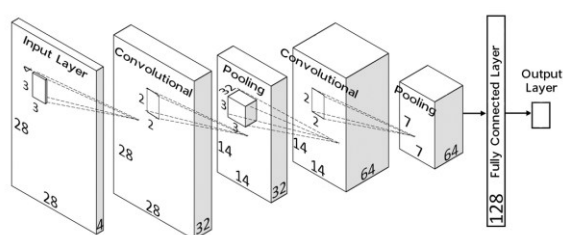

Figure 3. Structure of CNN (adopted from [Che11])

The structure of CNN is illustrated in Figure 3. It consists of three main hidden layer types: convolutional layers, pooling layers and fully connected layers [Vou18]. According to Voulodimos et al. [Vou18] and Cireşan et al. [Cir11], the convolutional layers are parameterized by the size and the number of the feature maps, kernel sizes, skipping factors and the connection table.

In CNN, the pooling layers enable position invariance over larger local regions and summarize the outputs of neighboring groups of neurons in the same kernel map, where the neighborhoods, summarized by adjacent pooling units, do not overlap [Kri12], [Cir11]. The fully connected layers are used for high-level reasoning in the neural network [Vou18] which combine the outputs of the top convolutional layer and convert the 2D feature maps into a 1D feature [Vou18], [Cir11].

The way of achieving the feature learning in CNN is by alternating and stacking convolutional layers and pooling operations. The convolutional layers convolve multiple local kernel filters with input data in order to generate invariant local features [Wan18]. The pooling layers extract the most significant features over sliding windows, which have a fixed-length using different pooling operations. Typical pooling operations are average pooling and max pooling. Average pooling calculates the mean value of the neurons in the region (sliding window) and takes it as the output value for that region [Wan18]. Not surprisingly, max pooling selects the maximum value in the region, so the pooling output is the maximum activation over rectangular regions that do not overlap [Cir12].

Max pooling is well suited to extract sparse features, while other operations might not be optimal on all samples [Wan18]. In comparison with similar models, CNN has several advantages, including easier training, sparse interactions with local connectivity, reduced number of parameter sharing and equivariant representation which is invariant to object locations [Wan18], [Kri12].

4. CHALLENGES AND POTENTIAL OF DNN IN MAP DIGITIZATION

In this section, papers that are surveyed have leveraged deep learning methods to address the main tasks in computer vision, in order to digitized maps, such as object and pattern recognition.

A survey of the literature of done in this study lead to two main findings.

The first is the fact that the different approaches to the problem of extracting vector graphics representations of historical cadastral plans and maps follow a relatively general pipeline, which can be broadly defined to contain the following stages: preprocessing, visual object detection and classification, vector representation postprocessing and extracting information from text. The last two stages are independent and can be done in parallel.

The second finding is that, although neural nets have had a long history of use in these approaches, their applications have mainly been limited to the stage of extracting the information from text and that this has remained so, even with the advancement of deep learning. This, perhaps, is not surprising when one considers that the convergence of advancements in the domain of training deep neural nets and GPU hardware has only recently (as of 2012 [Kri12]) allowed DNNs to achieve state of the art results in computer vision applications, beyond hand-written text recognition.

This state of affairs is likely to end soon. A recent (2017) study by Liu et al. [Liu17] addresses the problem of converting a rasterized two-dimensional image into a vectorized representation using a CNN. Evenhough the floor plan images are used, instead of the maps, their approach is perhaps the one that can most readily be transferred to the domain of vectorizing maps.

The approach of Liu et al. consists of two computational machineries (deep representation learning and integer programming) combined into a

single system that converts a floor plan image through two intermediate representation layers. First, a CNN converts a floor plan image into a junction layer, where data is represented by a set of junctions identified as specific objects by the network. Secondly, integer programming aggregates junctions into a set of primitives, while providing a topologically and geometrically consistent result. After that, a simple post-processing is used for producing the final vector format.

In terms of the general pipeline for cadastral data vectorization, the deep representation learning stage corresponds to the visual object detection and classification and focuses on detecting key points (line joints, types, etc.), while integer programming deals with vector representation postprocessing and is used to construct a topologically and geometrically consistent result. Of the two remaining tasks necessary to achieve fully automated cadastral plan digitization, extraction of the information from hand-written text has always been the domain of (deep) neural networks. Finally, the preprocessing was mainly done to simplify the task of visual object detection and eliminated "noise" present in old maps (such as accidental blotches and variations in color). While some preprocessing might still be meaningful when a full DNN framework for the vectorizatoin is considered, the success of DNNs in the domain of computer vision was in a significan part due to their invariance to such distortions in the input data and it is likely that preprocessign will be eliminated from the pipeline all together.

The main hurdle in terms of using DNNs for historic cadastral plan vectorization is likely to be a limited amount of training data available. E. g. the study of Liu et al. [Liu17] used the 100 870 data set, based on LIFULL HOME'S dataset [Lif18], containing 870 human-annotated floor-plan images and 100,000 unannotated test images to train their DNN. Generating such a dataset for historical cadastral data seems not only prohibitively expensive, but is also likely to be unjustified, when one considers the potential applications of the technology. To make matters worse, the results of Oliviera et al. [Oli17] suggest that using existing pre trained models can lead to a significant degradation of performance (10% vs. over 99% achieved on the MNIST data set the network was trained and tested on [LeC98]).

At the time of writing, the Generative Adversarial Networks approach [Goo14], [Cre16] seems to be the state-of-the-art alternative that could help achieve good performance, even in the absence of labelled training data. In addition, the floor plan dataset by Liu could be used to pre trained the models, and fine-tune them on the cadastral data, as the two domains are different but relatively similar.

Although some challenges remain, a wider adoption of deep learning technology in the process of map and cadastral data vectorization seems immanent and is likely to revolutionize the field in the near future, as it has already done for so many other domains.

5. CONCLUSION AND FUTURE WORK

Map digitization represents a procedure which relies on visual object and pattern recognition, such as the identification of parcel and building lines, line junctions, symbols and text – areas in which DNNs achieve excellent performance. However, over the past three decades, not many studies dealing with the problem of map vectorization problem used NNs outside the scope of written-text recognition. Deep learning, in particular, has barely entered the domain and the survey of literature conducted for the purposes of the study presented here identified only one approach that used a CNN to extract information from text, while relying on more classical computer vision for the rest of the process.

Recognizing that new advances in the domain of deep learning, which have already revolutionizes many areas of computer vision, are likely to be applied to the problem of cadastral map vectorization in the near future, we have attempted to assess the challenges in this process and the potential benefits that can be achieved, as an aid for researchers considering this course of action and in order to stimulate further discussion and a possible faster adoption of the technology. Having identified the challenges, we've also attempted to identify specific approaches in the domain of deep learning that can be used to address them.

Our findings presented in the paper result from the initial stage of the study aimed at achieving beyond-state-of-the-art automatic historical cadastral plan vectorization through the use of deep learning.

REFERENCES

[Boa92] Boatto, L. et al. An interpretation system for cadastral maps. Computer, vol. 25, no. 7, pp. 25–33, 1992.

[Che11] Chen, B. X., Sahdev, R. and Tsotsos, J. K. Integrating Stereo Vision with a CNN Tracker for a Person-Following Robot. Internacional Coference on Computer Vision Systems, pp. 300–313, 2011.

[Che96] Chen, L.-H., Liao, H.-Y., Wang, J.-Y., Fan, K.-C. and Hsieh, C.-C. An interpretation system for cadastral maps. Proceedings of 13th International Conference on Pattern Recognition, pp. 711–715, 1996.

[Cir11] Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M. and Schmidhuber, J. Flexible, high performance convolutional neural networks

for image classification. IJCAI International Joint Conference on Artificial Intelligence, pp. 1237–1242, 2011.

[Cir12] Cireşan, D., Meier, U., Masci, J., and Schmidhuber, J. Multi-column deep neural network for traffic sign classification. Neural Networks, vol. 32, pp. 333–338, 2012.

[Con97] Congalton, R. G. Exploring and Evaluating the Consequences of Vector-to-Raster and Raster-to-Vector Conversion. Photogramm. Eng. Remote Sens., vol. 63, no. April, pp. 425–434, 1997.

[Cre16] Creswell, A. and Bharath, A. A. Adversarial training for sketch retrieval. Computer Vision – ECCV 2016 Workshops, vol. 9913 LNCS, pp. 798–809, 2016.

[Eji90] Ejiri, M., Kakumoto, S., Miyatake, T., Shimada, S. and Iwa-mura, K. Image analysis application, 1st ed. 1990.

[Goo14] Goodfellow, I. et al. Generative Adversarial Nets. Advances in Neural Information Processing Systems 27, pp. 2672–2680, 2014.

[He15] He, K. and Sun, J. Convolutional neural networks at constrained time cost. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5353–5360, 2015.

[Hop82] Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, vol. 79, no. 8, pp. 2554–2558, 1982.

[Ill91] Illert, A. Automatic Digitization of Large Scale Maps. Building, pp. 113–122, 1991.

[Jan93] Janssen, R. D. T., Duin, R. P. W. and Vossepoel, A. M. Evaluation method for an automatic map interpretation system for cadastral maps. Proc. 2nd Int. Conf. Doc. Anal. Recognit. (ICDAR '93), no. November, pp. 125–128, 1993.

[Kar08] Karabork, H., Kocer, B., Bildirici, I. O., Yildiz, F. and Aktas, E. A neural network algorithm for vectorization of 2D maps. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. XXXVII, no. B2, pp. 473–480, 2008.

[Kat99] Katona, E. and Hudra, G. An interpretation system for cadastral maps. Image Analysis and Processing, 1999. Proceedings. International Conference on Image Analysis and Processing, 1999.

[Kri12] Krizhevsky, A., Sutskever, I. and Hinton, G, ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems 25 (NIPS2012), pp. 1–9, 2012.

[LeC95] LeCun, Y. and Bengio, Y. Convolution Networks for Images, Speech, and Time-Series. Handb. brain theory neural networks, no. 1, pp. 1–5, 1995.

[LeC98] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. Gradient-based Learning Applied to Document Recognition. Procedings of the IEEE, pp. 2278–2324, 1998.

[Lif18] Lifull Home's dataset." [Online]. Available: http://www.nii.ac.jp/%0Adsc/idr/next/homes.htm %0A [Accessed: 18.3.2018.]

[Liu17] Liu, C., Wu, J., Kohli, P. and Furukawa, Y. Raster-to-Vector: Revisiting Floorplan Transformation. 2017 IEEE Int. Conf. Comput. Vis., pp. 2214–2222, 2017.

[McG66] McGuigan, D. G. The Habsburgs, First Edit. Doubleday, 1966.

[Mur10] Murray, J. F. Convolutional Networks Can Learn to Generate Affinity Graphs for Image Segmentation. Neural Comput., vol. 22, no. 2, pp. 511–538, 2010.

[Oli17] Oliveira, S. A., Kaplan, F. and Di Lenardo, I. Machine Vision on Cadaster Plans. Premiere Annual Conference of the International Alliance of Digital Humanities Organizations 2017, 2017.

[Pra18] Prakash, C. Ranking with Deep Neural Networks. Fifth International Conference on Emerging Applications of Information, no. January, 2018.

[Sap17] Sapkota, R. K. and Bha, G. P. Technical Aspects of Digitization of Cadastral Maps. Nepal. J. Geoinformatics, vol. 13, pp. 42–50, 2017.

[Sze17] Sze, V., Chen, Y.-H., Yang, T.-J. and Emer, J. Efficient Processing of Deep Neural Networks: A Tutorial and Survey. Proceedings of the IEEE, pp. 1–32, 2017.

[Vas16] Vasiljević, S., Amović, M. and Višnjić, R. Katastarski planovi prema podijeli bečkog vojnogeografskog instituta. Vještak, vol. 2, no. 2, pp. 219–223, 2016.

[Vou18] Voulodimos, A., Doulamis, N., Doulamis, A. and Protopapadakis, E. Deep Learning for Computer Vision : A Brief Review. Comput. Intell. Neurosci., vol. 2018, 2018.

[Wan18] Wang, J., Ma, Y., Zhang, L., Gao, R. X. and Wu, D. Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems, no. January, 2018.

[Wen97] Wenyin, L. and Dori, D. A protocol for performance evaluation of line detection algorithms. Mach. Vis. Appl., vol. 9, no. 5–6, pp. 240–250, 1997.

[Yos15] Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. and Lipson, H. Understanding Neural Networks Through Deep Visualization. Deep Learning Workshop, 31st International Conference on Machine Learning, 2015.