AGILE: GIScience Series, 3, 32, 2022. https://doi.org/10.5194/agile-giss-3-32-2022
Proceedings of the 25th AGILE Conference on Geographic Information Science, 2022.
Editors: E. Parseliunas, A. Mansourian, P. Partsinevelos, and J. Suziedelyte-Visockiene.
This contribution underwent peer review based on a full paper submission.

Representing Vector Geographic Information As a Tensor for Deep Learning Based Map Generalisation

Azelle Courtial[1], Guillaume Touya[1], and Xiang Zhang[2]

[1]LASTIG, Univ Gustave Eiffel, ENSG, IGN, F-94160 Saint-Mande, France
[2]School of Geospatial Engineering and Sciences, Sun Yat-Sen University, Guangzhou 510275, China

Correspondence: Azelle Courtial (azelle.courtial@ign.fr)

Abstract. Recently, many researchers tried to generate (generalised) maps using deep learning, and most of the proposed methods deal with deep neural network architecture choices. Deep learning learns to reproduce examples, so we think that improving the training examples, and especially the representation of the initial geographic information, is the key issue for this problem. Our article extracts some representation issues from a literature review and proposes different ways to represent vector geographic information as a tensor. We propose two kinds of contributions: 1) the representation of information by layers; 2) the representation of additional information. Then, we demonstrate the interest of some of our propositions with experiments that show a visual improvement for the generation of generalised topographic maps in urban areas.

Keywords. Cartography, Deep learning , Map generalisation

Figure 1. Examples that illustrate the main representation issues from (1. Isola et al., 2018)(2. Courtial et al., 2021b)(3. Kang et al., 2019)

1 Introduction

Deep learning is a new way to generate maps without GIS tools, and some recent experiments show it can even include map generalisation transformations. Most of these experiments focus on the network architecture and omit to investigate the training set as a way to improve the map output. The most common way to compile a set of training images is to apply a raw symbolization on the vector data, then to tile the symbolized data, and finally to rasterize the tiles. The way the input data is presented to the neural networks is not always adapted to the target task, i.e. generating a map. In fact, as deep learning relies on knowledge extraction from examples, it is necessary that the input data carries not only the clear location and shape of the geographic objects but also some contextual information and implicit knowledge (e.g. road nature or role in the road network). Especially when the target map includes some abstraction, or generalisation, the notions of importance, patterns and spatial relations are essential, but are badly accessible in the inputs from the literature. Figure 1 summarizes some examples of tiles from previously published experiments that illustrate these representation issues:

1. The images are only small extracts of a map, arbitrarily tiled, and the information just outside the tile is often necessary to properly generate the map. This problem is similar when we want to parallelize the generalization of small adjacent tiles to reduce the amount of processed data (Touya et al., 2017). For example, Figure 1, line 2 includes the collapse of the city centre into a built-up area polygon, but this task requires the whole city to be visible.

2. The resolution of the images is sometimes not sufficient; for example in Figure 1 (lines 1 and 3), the thin roads are not preserved during map generation

because they cannot be really distinguished in the input image.

3. Objects of different types may have a similar colour, and thus may be considered as similar by the neural network. Two examples of this problem can be seen on Figure 1, line 1 with park and roads; and line 2 with buildings and the built-up area that have a close colour. In both these cases, another characteristic (size, arrangement, shape, texture, etc.) allows a human to distinguish them but the network might not be able to make this distinction with a few examples, and does not succeed to generate the rare feature (the park in line 1 and the built-up area in line 2 of Figure 1).

4. Object overlaps or very thin separations involve information loss in the raster input, compared to a vector representation. This problem causes disconnections (Figure 1, line 2).

5. Spatial relations are hardly accessible, as they involve the notion of groups of geographic objects not explicitly encoded in the image. Thus, the preservation of spatial relations is challenging.

We hypothesize that an augmented representation of information with extra layers of information encoded in a tensor would alleviate these issues and allow the network to see the context beyond the tile, understand the object importance, and preserve spatial relations. Tensors are mathematical generic objects that represent an information, images are particular tensor in $\mathbb{R}^3$. Thus, we propose new ways to create input tensors for the neural networks dedicated to map generation, and test their suitability for the use case of urban topographic map generation.

This article is organized as follows: in Section 2 we present a literature review of deep learning techniques employed for map generation. Section 3 proposes an improved representation of geographic information to solve these issues. Then Section 4 presents some experiments and their results.

2 Related Work

2.1 Generating Maps With Deep Learning

The main application of deep learning techniques in geographical information science is remote sensing, e.g. (Zhu et al., 2017), but these techniques also prove successful to recognise features from map images (Chen et al.; Touya and Lokhat, 2018) or to classify map images (Schnürer et al., 2020; Hu et al., 2021). But our focus is on map generalisation (Du et al., 2021; Feng et al., 2019; Courtial et al., 2020a), so we are more interested in deep architectures that can generate a map image. The task of generating an image of a map with deep learning appears to be associated with the generative adversarial networks (GAN) that generate an image from an input image with two adversarial neural networks. In the literature, most experiments of map generation with GAN are about generating an image of a map in the style of GoogleMap from the corresponding aerial photograph, and vice versa. Isola et al. (2018) proposed a generic image-to-image translation model called Pix2Pix, which was later improved by many researchers to better deal with the generation of maps (Ganguli et al., 2019; Chen et al., 2020; Zhang et al., 2020; Li et al., 2020). In their work, the scale of the style-transferred map is similar to that of the aerial photograph.

Some other experiments have explored the generation of other kinds of maps or from stylized geographic information: e.g. Google map from OSM stylized data or artistic map from GoogleMap (Kang et al., 2019), topographic map from detailed national map agency data (Courtial et al., 2021b). These examples involve more complex cartographic processing, especially map generalisation, and are therefore more sensitive to the representation issues highlighted in the introduction of this paper.

2.2 Representing Spatial Data for Deep Learning

Representation is a basic issue in deep learning, where the success of learning depends on how the training examples are a faithful representation of the target task (Bengio et al., 2014). The potential of representing spatial vector data in a graph for deep learning has already been demonstrated (Yan et al., 2020; Iddianozie and McArdle, 2021). On the other hand, because it is also natural to represent map data as raster images for deep learning, the way in which information is encoded is very rarely questioned.

In a map generation task, the representation issues are mostly related to the raster representation of geographic information and are common for all raster approaches for map generalisation (Touya et al., 2019), even without deep learning (Shunbao et al., 2012). Moreover, for a map generalisation task, you often need to look around the feature you generalise to know how to optimally generalise it: for instance, you need to take the whole block into account to generalise a building in a city (Ruas, 1999). In raster mode, some first experiments have been proposed to determine an adapted tiling method that has no border effect. For instance, a tiling method that follows the orientation of the line to improve its generalisation has been proposed (Du et al., 2021), while other proposed a method to produce image tiles entirely included in an urban block in order to improve the land use classification of these blocks (Huang et al., 2018). However, these proposals seem to be insufficient for the above-mentioned issue. Therefore, we propose in the following sections different ways to represent the vector data as an input tensor for GAN models.

3 A Better Representation of Vector Geographic Information

In this section we propose a new representation method to improve the generation of map. This proposal is illustrated in Figure 2 : the generation is based on both a raw input information organized by themes, and additional information. In a usual representation, for a $s-size$ tiles, the input tensor dimension is $s \times s \times 3$ as the image is encoded in RGB. In our proposal, the input tensors dimension is $s \times s \times n$ and $s \times s \times m$, where n is the number of layers in the map, and m is the number of additional information tensor.

3.1 Input Representation for Map Generation

We propose, instead of the symbolized representation of input data, to represent geographic information by layer, called *Layered Representation*, where each layer is a binary mask that describes a single layer of map objects (e.g. buildings), as presented in Figure 3. Such a layered representation should be able to avoid the information loss caused by the overlap of map symbols, limit the confusion between themes with similar colours and simplify the understanding of intra-group spatial relations as objects of the same group are in the same layer.

3.2 Representation of Additional Information

As traditional map generation often requires data enrichment (Mackaness and Edwards, 2002), the deep learning approach also needs additional information. Indeed only showing the location and the shape of an object in a tile (whether in a symbolized or layered representation) does not convey all necessary information for map generalisation. The FuseNet architecture has been proposed to combine two images/tensors (a photograph and a depth map initially) as input (Hazırbaş et al., 2016). We propose to this architecture to combine the additional information to our layered representation. This architecture requires that additional information be structured in a tensor of the same size as the main tensor, where pixels at the same position are associated (describe the same portion of the ground).

The three main missing information are the context of the image, semantic attributes about the presented objects, and information about the relevance or the importance of each map extract.

3.2.1 Adding Semantic Information

First, attribute information provides essential information on objects. For example, in Figure 4 the nature of the road (symbolized on the right of the figure) is needed to understand the situation, i.e. the importance of each road section in the network.

When the semantic information is a categorical variable, we cannot encode it as a unique additional dimension in the tensor; the algorithm considers the values in the tensor as quantities. Consequently, we propose to make a binary layer in the tensor for each category of the variable. For example, if we want to encode the category of each building, we would not make an image with one colour for each category (Figure 5 .a) but several masks (Figure 5 b to d): the additional tensor to be combined with the layered tensor presented in Section 3.1 has a dimension $s \times s \times m$ where m the number of building categories.

3.2.2 Adding Spatial Context

Then, Figure 6 illustrates how the map generation task requires a fine understanding of the spatial context around the tile. On the left image, it is not possible to determine which roads are important and should be kept in case of generalisation, however, the contextual map on the right shows that the T-shape intersection is a crossroad between two important roads, while the others are minor roads.

Thus, we propose to compute contextual global measures, and to encode them on the objects of the tile as additional semantic information. For example, we can calculate the centrality of a road section in a road network, or the density of urban blocks, to give context to an image that does not show a complete block, or calculate for each pixel the slope of surrounding areas.

This contextual measure is often a quantitative variable, so it can fit into a tensor with dimension $s \times s \times 1$ where s is the size of the image in pixels. But the quantitative value has to be ranged through mathematical transformation. Indeed, the additional tensor must have the same range as the main information, i.e. a floating value between 0 and 255. This requirement also enables the visualisation of the information as a grayscale image. This transformation has to be adapted to the sub-type of data (continuous or discrete) and to the distribution of data. For example, the Horton order is a contextual measure calculated on rivers that gives the importance of a river (Horton, 1945; Touya, 2007). The Horton order is an integer value between one and the positive infinite, but the construction method makes that most of the rivers have a small value while very few rivers have a large value (the distribution is a hyperbole). Thus we propose to use the following transformation to convert the Horton order into a tensor value between 0 and 255: $n \rightarrow 255 * (1 - 1/(n+1))$. This example is illustrated in Figure 7.

3.2.3 Adding Focus Areas

Finally, we want to add some information about the importance of a given tile, or of a region within the tile. This addition aims at balancing the effect of under-representation of some situations in the data set (e.g. interchanges, roundabouts, alignments, etc.), and it may indicate the network where the map generalisation does not follow the general

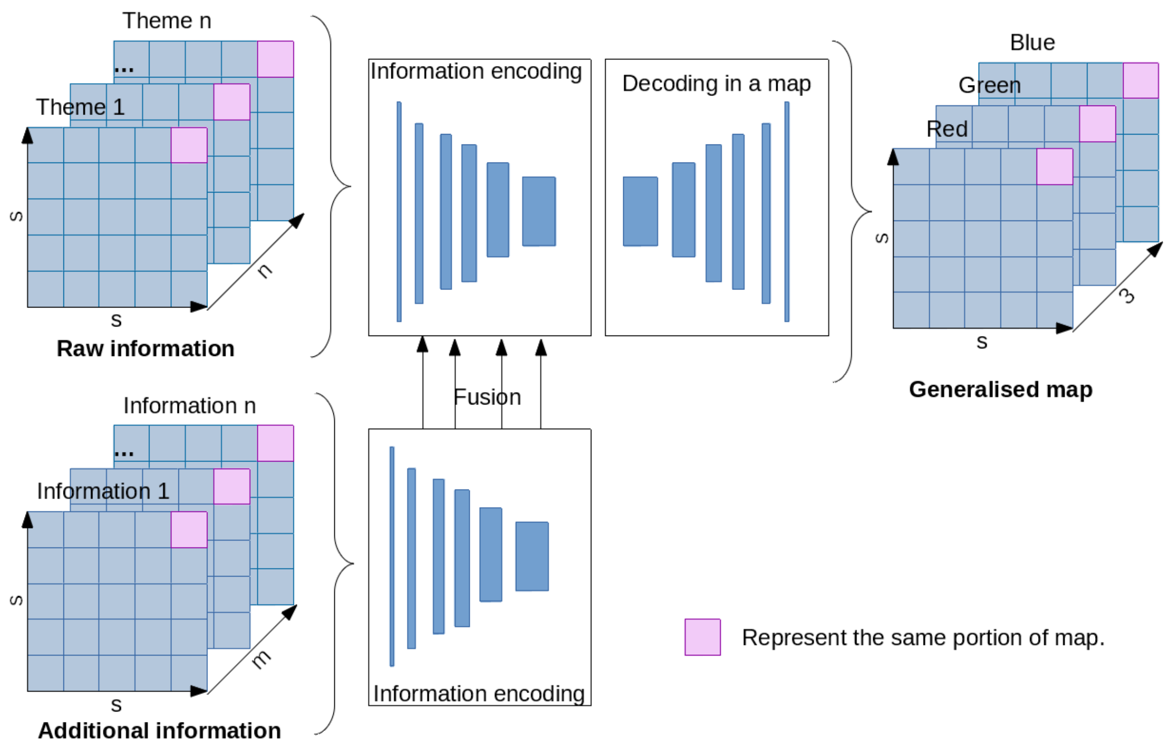

Figure 2. Conceptual diagram of the proposed information representation for map generation.

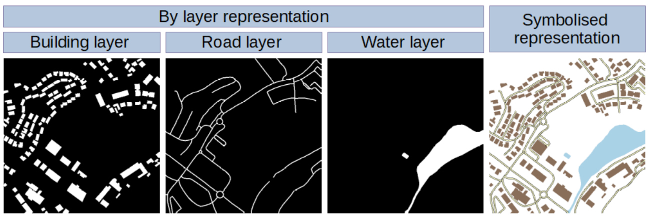

Figure 3. Representation by layers of geographic information.

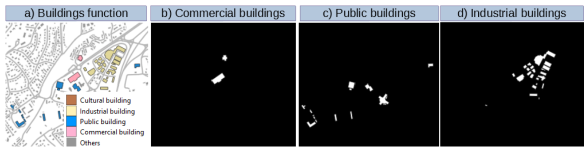

Figure 5. Value and representation of buildings function.

Figure 6. Illustration of the interest of contextual information

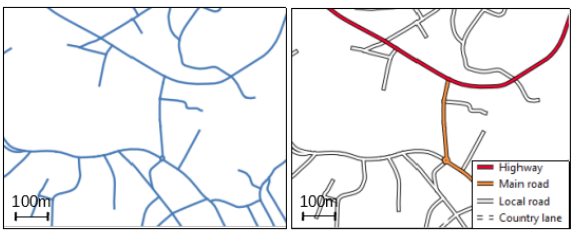

Figure 4. Illustration of the interest of semantic information

rule. For example, Figure 8 illustrates an issue in preserving a building alignment due to the lack of examples of building alignments in the training set. To improve the way the model deals with such situations it can be useful to highlight them with focus areas.

4 Experiments

We carried out some experiments to demonstrate the interest of our proposed representation of vector data for (generalised) map generation.

4.1 Implementation

This section presents the common material for our experiments. To illustrate the interest of representation, the basic and improved training sets are made from the same vec-

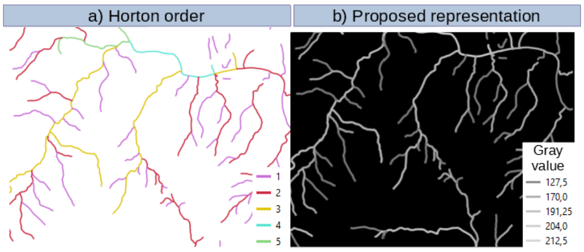

Figure 7. Value and representation of Horton order in an additional tensor.

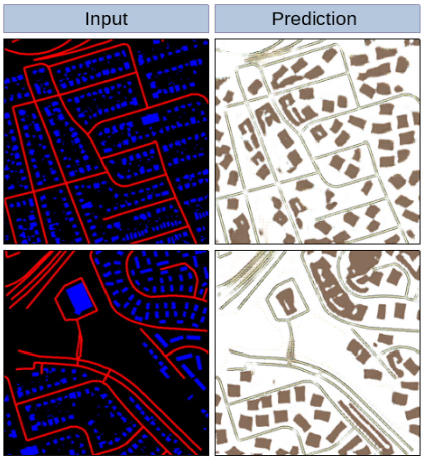

Figure 8. Illustration of the alignment preservation problem: building alignments are not well respected in generated maps.

tor data, that cover an area of 30* 15 kilometres in the southwest of France. A detailed database for maps at the 1:25,000 scale of roads, water features and buildings, and their generalisation for cartography at the 1:50,000 scale. There is also a layer of the city centre that needs to be grayed at the generalised scale (Touya and Dumont, 2017).

To evaluate the success of learning, without real formalisation of the evaluation process for maps generated using deep learning (Courtial et al., 2020b), we decided to only evaluate visually the results for each task. Our visual evaluation is guided by the following constraints for a generalised topographic map (Courtial et al., 2021b) : (C1) Buildings should be bigger than a minimum size; (C2) the smallest edge of the buildings should be greater than a minimum value (granularity constraint); (C3) The buildings should not be too close to the roads symbols; (C4) The buildings should not be too close to each other; (C5) The density of buildings in a block should remain stable; (C6) Building patterns should be preserved; (C7) Topological relations have to be preserved.

This map generation task involves a GAN (Isola et al., 2018) that aims at the generation of a new image that looks like the target images from our training set. This architecture combines a generator that creates the image and a discriminator that evaluates if the generated image looks like the target domain. The generator for map generation without additional information is a U-Net (Ronneberger et al., 2015) while the experiments on additional information use a FuseNet (Hazırbaş et al., 2016).

The code to generate layered images from vector data is published as a plugin of CartAGen open source map generalisation software. Data and code for map generation are openly available here: https://doi.org/10.5281/zenodo.5767663.

4.2 Layered Representation

Figure 9 compares the predictions of models trained with the symbolized and layered representations for the generation of generalised topographic maps at the 1:50 000 scale. We observe that the unexpected deletions and overlaps of roads and buildings are less numerous in the images generated from the layered representation, while the rest of the constraints are similarly satisfied. It shows the usefulness of the layered representation that allows the input image to convey a full and legible shape for each objects.

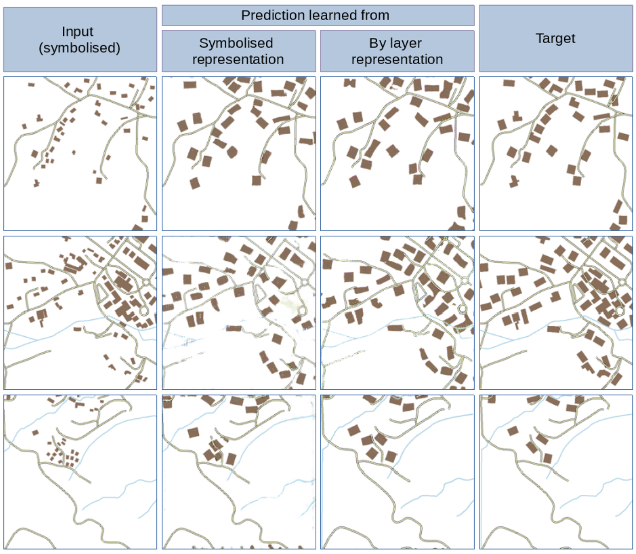

Figure 9. Comparison of images generated with symbolised and layered representations.

4.3 Adding Context for Road Selection

Then, we tested the addition of information for the generation of tiles with road selection, i.e. the deletion of the least important roads. The road network structure and the main roads should be preserved but the road network density should be reduced. This process seems impossible with just an isolated tile where the semantics and context of the road are not accessible (Courtial et al., 2020a). Thus, an additional information is calculated on the whole road network, with a graph convolutional network (Courtial et al., 2021a). Figure 10 illustrates the representation of this probability to be kept during a generalisation at 1:50,000 using the whole road network, to assess the importance of the road in an image. The pixel has the value: $255 * (1 - p)$ with p the probability to be selected.

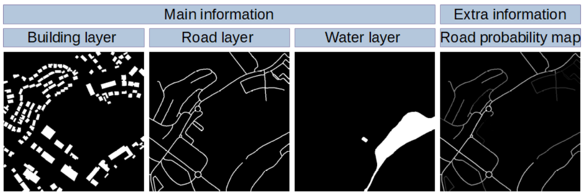

Figure 10. Illustration of the representation of road context as an additional information.

Figure 11 compares images generated with and without this information. It shows that the global structure of the road network is better preserved with extra information, and unexpected disconnections of roads are less numerous with the additional information. This experiment demonstrates that (1) the calculated information is relevant for road selection, and (2) the integration of additional information is really promising when the tile does not convey the complete necessary information.

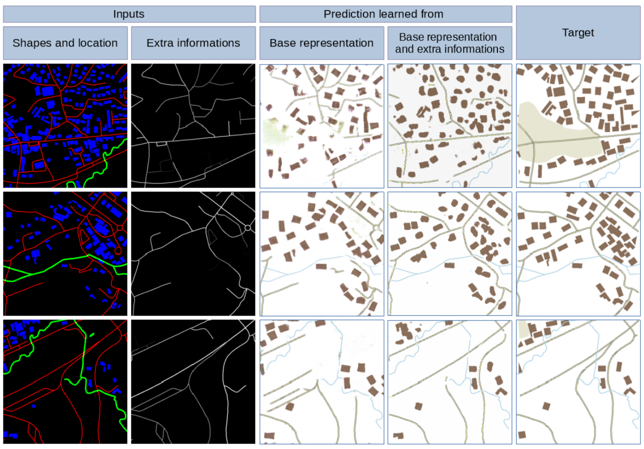

Figure 11. Comparison of images generated with and without contextual information on the road network.

4.4 Focus Areas for Building Alignments

In our last experiment, we tested the interest of integrating information that emphasize spatial relations, and especially building alignments. We hope this representation would encourage their preservation. The alignments are obtained by manual annotation but automated methods exist (Zhang et al., 2013). Figure 12 illustrates several ways to represent this information as a focus mask.

We trained a GAN to generate a map with our data, with and without the focus area, then we tested the model on different area. Figure 13 presents the results of this experiment, we observe that both the prediction are unclear and do not preserve alignments. Images generated with extra information are bad, but this experiment conveys fewer fake structures.

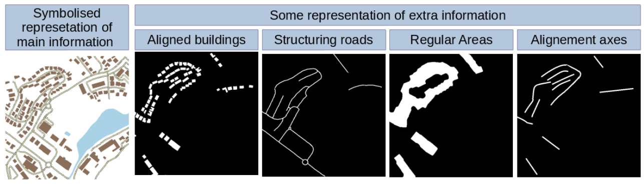

Figure 12. Illustration of different possible representations for alignment focus areas.

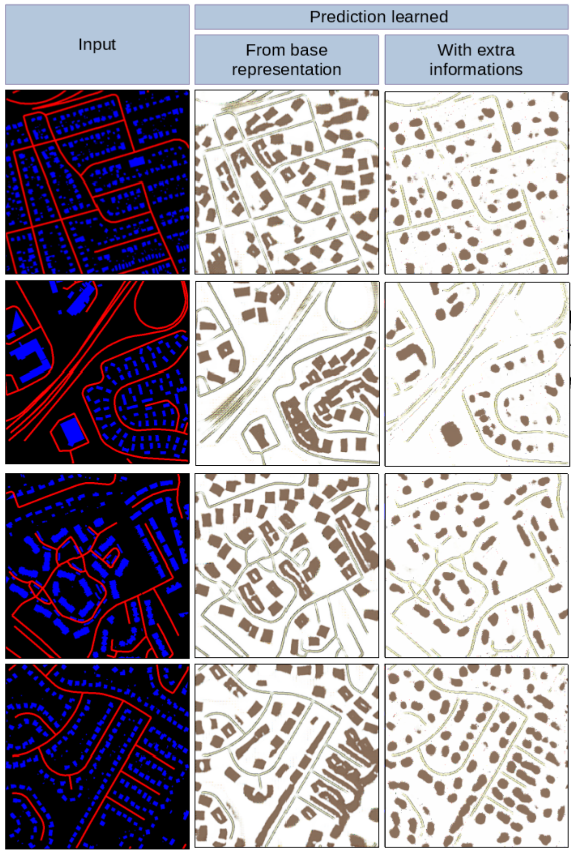

Figure 13. Comparison of prediction realized with and without alignment emphasize.

A similar test with the structuring road focus area show no difference, consequently, the representation using area seems to be better to generate better areas. But using focus areas as additional input data in a FuseNet is not only solution to provide focus areas to the model. We believe that a better solution would be to use the focus areas in the loss function of the network, to highlight on pixels that should be weighted more than the other pixels of the image. This way, the network would mimic more faithfully how the alignments are generalised in the training examples. We plan to test this proposition in further experiments.

5 Conclusion

In conclusion, this paper proposed two ways to better represent vector spatial data for a deep learning based map generation task: a layered representation rather than a simple image, and an additional tensor to convey semantic and contextual information. Our first experiments confirm that the proposed representation improves the quality of generated maps.

The main limitation of this work is about evaluation: a good evaluation method for map generalisation would help us to compare and measure the interest of each proposal, but foremost a measure of training set quality is lacking. It could be seen as an "evaluation for tuning" process (Mackaness and Ruas, 2007), which is an major step in traditional map generalisation. This measure of training set quality should at least include measures of the legibility and completeness of input information and the correctness of the target image.

Finally, the layered organization of data offer the opportunity to apply diverse learning strategy and network architectures to each geographic theme, within a framework for deep learning map generalisation. As map generalisation is a complex problem, and most of the proposed deep neural networks improve only one part of the generated map, we think such an approach is more promising than a single end-to-end model.

References

Bengio, Y., Courville, A., and Vincent, P.: Representation Learning: A Review and New Perspectives, arXiv:1206.5538 [cs], 2014.

Chen, X., Chen, S., Xu, T., Yin, B., Peng, J., Mei, X., and Li, H.: SMAPGAN: Generative Adversarial Network-Based Semisupervised Styled Map Tile Generation Method, IEEE Transactions on Geoscience and Remote Sensing, pp. 1–19, https://doi.org/10.1109/TGRS.2020.3021819, 2020.

Chen, Y., Carlinet, E., Mallet, C., and Dumenieu, B.: Vectorization of Historical Maps Using Deep Edge Filtering and Closed Shape Extraction, p. 16.

Courtial, A., El Ayedi, A., Touya, G., and Zhang, X.: Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation, ISPRS International Journal of Geo-Information, 9, 338, https://doi.org/10.3390/ijgi9050338, number: 5 Publisher: Multidisciplinary Digital Publishing Institute, 2020a.

Courtial, A., Touya, G., and Zhang, X.: Constraint Based Evaluation of Generalized Images Generated by Deep Learning, in: 23rd ICA Workshop on Map Generalisation and Multiple Representation, ICA, Delft, Netherlands, https://hal.archives-ouvertes.fr/hal-02995412, 2020b.

Courtial, A., Touya, G., and Zhang, X.: Can Graph Convolution Networks Learn Spatial Relations?, Abstracts of the ICA, 3, 1–2, https://doi.org/10.5194/ica-abs-3-60-2021, 2021a.

Courtial, A., Touya, G., and Zhang, X.: Generative Adversarial Networks to Generalise Urban Areas in Topographic Maps, in: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XLIII-B4-2021, pp. 15–22, Copernicus GmbH, https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-15-2021, iSSN: 1682-1750, 2021b.

Du, J., Wu, F., Xing, R., Gong, X., and Yu, L.: Segmentation and sampling method for complex polyline generalization based on a generative adversarial network, Geocarto International, 0, 1–23, https://doi.org/10.1080/10106049.2021.1878288, 2021.

Feng, Y., Thiemann, F., and Sester, M.: Learning cartographic building generalization with deep convolutional neural networks., International Journal of Geo-Information, 2019.

Ganguli, S., Garzon, P., and Glaser, N.: GeoGAN: A Conditional GAN with Reconstruction and Style Loss to Generate Standard Layer of Maps from Satellite Images, arXiv:1902.05611 [cs], 2019.

Hazırbaş, C., Ma, L., Domokos, C., and Cremers, D.: FuseNet: Incorporating Depth into Semantic Segmentation via Fusion-Based CNN Architecture, https://doi.org/10.1007/978-3-319-54181-5_14, 2016.

Horton, R. E.: Erosional Development of Streams and their Drainage Basins; Hydrophysical Approach to Quantitative Morphology, Geological Society of America Bulletin, 56, 275–370, 1945.

Hu, Y., Gui, Z., Wang, J., and Li, M.: Enriching the metadata of map images: a deep learning approach with GIS-based data augmentation, International Journal of Geographical Information Science, pp. 1–23, https://doi.org/10.1080/13658816.2021.1968407, 2021.

Huang, B., Zhao, B., and Song, Y.: Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery, Remote Sensing of Environment, 214, 73–86, https://doi.org/10.1016/j.rse.2018.04.050, 2018.

Iddianozie, C. and McArdle, G.: Towards Robust Representations of Spatial Networks Using Graph Neural Networks, Applied Sciences, 11, 6918, https://doi.org/10.3390/app11156918, number: 15 Publisher: Multidisciplinary Digital Publishing Institute, 2021.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A.: Image-to-Image Translation with Conditional Adversarial Networks, arXiv:1611.07004 [cs], http://arxiv.org/abs/1611.07004, 2018.

Kang, Y., Gao, S., and Roth, R. E.: Transferring Multiscale Map Styles Using Generative Adversarial Networks, International Journal of Cartography, 5, 115–141, https://doi.org/10.1080/23729333.2019.1615729, 2019.

Li, J., Chen, Z., Zhao, X., and Shao, L.: MapGAN: An Intelligent Generation Model for Network Tile Maps, Sensors, 20, 3119, https://doi.org/10.3390/s20113119, 2020.

Mackaness, W. A. and Edwards, G.: The Importance of Modelling Pattern and Structure in Automated Map Generalisation, in: Proceedings of the Joint ISPRS/ICA Workshop on Multi-Scale Representations of Spatial Data, pp. 7–8, 2002.

Mackaness, W. A. and Ruas, A.: Chapter 5 - Evaluation in the Map Generalisation Process, in: Generalisation of Geographic Information, edited by Mackaness, W. A., Ruas, A., and Sarjakoski, L. T., International Cartographic Association, pp. 89 – 111, Elsevier Science B.V., Amsterdam, https://doi.org/10.1016/B978-008045374-3/50007-7, 2007.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, vol. abs/1505.04597, pp. 234–241, 2015.

Ruas, A.: Modèle de généralisation de données géographiques à base de contraintes et d'autonomie, phdthesis, Université de Marne la Vallée, 1999.

Schnürer, R., Sieber, R., Schmid-Lanter, J., Öztireli, C., and Hurni, L.: Detection of Pictorial Map Objects with Convolutional Neural Networks, Cartographic Journal The, https://doi.org/10.1080/00087041.2020.1738112, 2020.

Shunbao, L., Zhongqiang, B., and Yan, B.: Errors Prediction for Vector-to-Raster Conversion Based on Map Load and Cell Size, Chinese Geographical Science, 2012.

Touya, G.: River Network Selection based on Structure and Pattern Recognition, in: International Cartographic Conference, International Cartographic Association, Moscow, Russia, 2007.

Touya, G. and Dumont, M.: Progressive Block Graying and Landmarks Enhancing as Intermediate Representations between Buildings and Urban Areas, in: Proceedings of 20th ICA Workshop on Generalisation and Multiple Representation, Washington DC, USA, 2017.

Touya, G. and Lokhat, I.: Deep learning for enrichment of vector spatial databases: application to highway interchange., p. 21, 2018.

Touya, G., Berli, J., Lokhat, I., and Regnauld, N.: Experiments to Distribute and Parallelize Map Generalization Processes, The Cartographic Journal, 54, 322–332, https://doi.org/10.1080/00087041.2017.1413787, 2017.

Touya, G., Zhang, X., and Lokhat, I.: Is deep learning the new agent for map generalization?, International Journal of Cartography, 5, 142–157, https://doi.org/10.1080/23729333.2019.1613071, 2019.

Yan, X., Ai, T., Yang, M., and Tong, X.: Graph convolutional autoencoder model for the shape coding and cognition of buildings in maps, International Journal of Geographical Information Science, 0, 1–23, https://doi.org/10.1080/13658816.2020.1768260, publisher: Taylor & Francis, 2020.

Zhang, X., Ai, T., Stoter, J., Kraak, M.-J., and Molenaar, M.: Building pattern recognition in topographic data: examples on collinear and curvilinear alignments, GeoInformatica, 17, 1–33, https://doi.org/10.1007/s10707-011-0146-3, 2013.

Zhang, Y., Yin, Y., Zimmermann, R., Wang, G., Varadarajan, J., and Ng, S.-K.: An Enhanced GAN Model for Automatic Satellite-to-Map Image Conversion, IEEE Access, 8, 176 704–176 716, https://doi.org/10.1109/ACCESS.2020.3025008, 2020.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., and Fraundorfer, F.: Deep learning in remote sensing: a review, IEEE Geoscience and Remote Sensing Magazine, 5, 8–36, https://doi.org/10.1109/MGRS.2017.2762307, 2017.