



HAL
open science

Exploring the potential of deep learning for map generalization

Azelle Courtial

► **To cite this version:**

Azelle Courtial. Exploring the potential of deep learning for map generalization. Geography. Université Gustave Eiffel, 2023. English. NNT : 2023UEFL2001 . tel-04089883

HAL Id: tel-04089883

<https://theses.hal.science/tel-04089883>

Submitted on 5 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Doctoral thesis from Université Gustave Eiffel
Doctoral school MSTIC
Specialty "Geographic information sciences"

Exploring the potential of deep learning for map generalization

Azelle Courtial

Publicly defended on January 18th, 2023

Jury

Mr Cyril De Runz,	Université de Tours	Rapporteur
Ms Anne Ruas,	Université Gustave Eiffel	Rapporteur
Mr Mathieu Aubry,	École des Ponts ParisTech	Examiner
Ms Monika Sester,	Leibniz Universität Hannover	Examiner
Mr Guillaume Touya,	Université Gustave Eiffel	Supervisor
Mr Xiang Zhang,	Sun Yat-Sen University	Supervisor

Abstract

Map generalization is a process that aims to adapt the level of detail of geographic information for cartography at a small scale. Automating the process is complex but essential in map production. We think this research field could benefit from the recent advances in deep learning that make it possible to solve more and more complex tasks, using numerous training examples.

This thesis proposes exploring the potential of deep learning for map generalization. This exploration is built upon three map generalization use cases: recognition of spatial relations, graphic generalization of mountain roads, and generalization of topographic maps at medium scales. These three use cases enable us to address research questions relative to the concrete implementation of deep learning models for map generalization (including dataset creation and architecture), the evaluation of such models and their integration in existing generalization processes.

In addition to the models and training set adapted for each of our case studies already mentioned, we propose evaluation methods adapted to the challenges of cartographic generalization by deep learning. Finally, we propose a partitioning of the cartographic generalization into sub-problems facilitating the resolution by learning and allowing the generation of generalized map images.

Keywords : Map generalization, deep learning, generative adversarial networks.

Résumé

La généralisation cartographique est un processus qui vise à adapter le niveau de détail de l'information géographique présente sur une carte pour la rendre lisible à une échelle plus petite. L'automatisation de ce processus est complexe mais essentielle à la production de cartes. Nous pensons que la recherche dans ce domaine pourrait bénéficier des avancées récentes en apprentissage profond qui rendent possible la résolution de tâches de plus en plus complexes, grâce à l'utilisation de nombreux exemples déjà bien généralisés.

Cette thèse propose d'explorer le potentiel de l'apprentissage profond pour la recherche en généralisation cartographique. L'exploration est construite autour de trois problèmes de généralisation cartographique : la recherche de relations spatiales, la généralisation graphique des routes de montagne, et la généralisation des cartes topographiques complètes à des échelles intermédiaires. Via ces trois cas d'études, nous abordons les questions de recherches liées à la mise en œuvre concrète des modèles d'apprentissage profond pour la généralisation cartographique (jeu de données et architecture), l'évaluation de tels modèles, et leur intégration dans des processus cartographiques préexistants.

Outre les modèles et jeux d'entraînements adaptés pour chacun de nos cas d'études, déjà mentionnés, nous proposons des méthodes d'évaluations adaptées aux défis de la généralisation cartographique par apprentissage profond. Finalement, nous proposons un découpage de la généralisation cartographique en sous-problèmes facilitant la résolution par apprentissage, et permettant la génération d'images de cartes généralisées.

Mots clefs : Généralisation cartographique, apprentissage profond, réseaux génératifs antagonistes.

Contents

Introduction	12
Part I A new paradigm for map generalization	19
Chapter A. Literature review	20
A.1 Map generalization	20
A.1.1 Definition	20
A.1.2 Map generalization automation	22
A.1.3 Map generalization is still a challenging question	28
A.2 Deep learning	29
A.2.1 Deep learning principles	29
A.2.2 A short history of deep learning	29
A.2.3 Deep learning usage towards map generalization	30
A.3 Thesis objectives	35
A.3.1 Needs in map generalization automation	35
A.3.2 Summary of the theoretical motivation	36
A.3.3 Approach	37
Chapter B. Formulating map generalization as a deep learning task	39
B.1 Geographic information representation for deep learning	39
B.1.1 Map generalization as an image problem	39
B.1.2 Map generalization as a graph problem	41
B.1.3 Other approaches for deep learning-based map generalization	44
B.2 Deep learning tasks and map generalization	45
B.2.1 Classification and map generalization	45
B.2.2 Segmentation and map generalization	46
B.2.3 Regression and map generalization	47
B.2.4 Image generation and map generation	48
B.2.5 Approach combination	48
B.2.6 Other tasks	49
Chapter C. Designing a framework for deep learning based map generalization	51
C.1 The characteristics of an adapted deep neural network	51
C.1.1 What is a deep neural network?	51
C.1.2 Some deep neural architectures for map generalization	54
C.1.3 Choosing or designing deep neural network for map generalization?	60

C.1.4	How to choose a relevant learning objective for map generalization	64
C.2	Designing a good training dataset	68
C.2.1	What is an efficient dataset?	68
C.2.2	An example of an image dataset for shape generalization	71
C.2.3	An example of an image dataset for map generation	74
C.2.4	A proposal for a better representation of geo-information in images	76
C.2.5	Two example graph datasets for learning spatial relations	79
 Part II Exploration of deep learning for map generalization		85
 Chapter D. Can graph neural networks model spatial relations?		86
D.1	Motivations	86
D.2	The use case of pattern recognition in a map	87
D.2.1	Presentation	87
D.2.2	Method	88
D.2.3	Results	89
D.2.4	Discussion	89
D.3	The use case of the road network selection	92
D.3.1	Presentation	92
D.3.2	Method	93
D.3.3	Results	94
D.3.4	Discussion	96
 Chapter E. CNN for the generalization of roads		98
E.1	Materials and methods	98
E.1.1	Use case	98
E.1.2	Approach	99
E.1.3	Experiment setting	100
E.2	Results	101
E.2.1	Segmentation approach	101
E.2.2	Generation approach	105
E.3	Discussion	109
E.3.1	Learning strategy	109
E.3.2	What can we learn?	110
E.3.3	Remaining limitations	112
 Chapter F. The generation of topographic map with several themes		114
F.1	Materials and methods	114
F.1.1	Approach	114
F.1.2	Implementation	115
F.2	Results	117
F.2.1	General results	117
F.2.2	Improving the default models	121
F.3	Discussion	128
F.3.1	Are fusion models useful?	128

F.3.2	How to learn the preservation of spatial relations?	130
F.3.3	Is it possible to generate a complete generalized map?	130
Part III	The future of map generalization with deep learning	133
Chapter G.	Usages of deep learning models for map generalization	134
G.1	Transfer learning	134
G.1.1	Location transfer	134
G.1.2	Transfer to geographic information from a new data source	135
G.2	Usage as a tool in a mixed process	136
G.2.1	Deep enrichment	136
G.2.2	Deep generalization	137
G.3	Usage as an end-to-end process	139
G.3.1	DeepMapScaler: a workflow of deep neural networks for the generation of generalized maps	139
G.3.2	Experiment	141
G.3.3	Discussion	146
Chapter H.	Evaluation of deep learning predictions	149
H.1	The specificities of evaluation of deep learning-based generalization	149
H.1.1	Why evaluating deep learning predictions?	149
H.1.2	Evaluation in map generalization	151
H.1.3	Evaluation in deep learning research	152
H.1.4	Proposed evaluation methods	154
H.2	User test	155
H.2.1	Design of the user test	155
H.2.2	Results	156
H.2.3	Limitations	158
H.3	Constraints for raster based evaluation	159
H.3.1	Method	159
H.3.2	Proposed constraints	160
H.3.3	Interpretation and usage of constraint values	167
H.3.4	Discussion	169
H.4	Vector based evaluation	172
H.4.1	Vector conversion for tiles of generalized buildings	172
Conclusion		175
Bibliography		186
Publications		205
Appendices		206
A	Comparison of models for road shape generalization	206
B	Estimation of constraint adaptability for raster calculation	209
Résumé substantiel en français		211

List of Figures

1	Indicators of the evolution of interest in map generalization.	13
A.1	A ground and a raster and vector representation of this ground.	21
A.2	Main steps of map production.	21
A.3	Example of global and iterative approaches.	24
A.4	Superpixel and map generalization, from (Shen, Ai, and Li 2019).	26
A.5	Action diagram of machine learning the global principles.	30
A.6	Example of use-cases of the generic image to image networks, from (Isola, Zhu, Zhou, and Efros 2017).	34
A.7	Scope of this thesis.	36
B.1	Law of Gestalt perception: similar objects, close objects, connected objects, etc. are perceived as more related than others.	40
B.2	Some elements of a geographic database and an image representation of these data.	41
B.3	The topological relations according to the 4-intersections model (Egenhofer and Herring 1990).	42
B.4	Example of a graph of spatial relations.	42
B.5	Example of graph representation of building shape (from Yan, Ai, Yang, and Tong 2020).	43
B.6	Graph object representations.	43
B.7	Graph representations: list of edges, graphical representation, and adjacency matrix.	44
B.8	Illustration of the segmentation definition: a) in common meaning, b) in image processing, c) in a graph.	46
B.9	An example of segmentation for map generalization conflicts detection.	47
B.10	Learning shape generalization with shape-graph node regression.	48
B.11	Learning scale transitions with a video approach.	50
B.12	Main possible formulation of map generalization as a deep learning problem.	50
C.1	Neurons and deep neural network.	52
C.2	Comparison of the sketch-to-drawing conversion and mountain road shape generalization.	53
C.3	Image processing in CNN. Convolution with a kernel of size $(3 * 3)$. Max-pooling of size $(2 * 2)$	55
C.4	The functioning of a CNN.	55
C.5	A U-Net architecture and a residual block.	56
C.6	An example where pixel-level loss is not adapted for map generalization.	57
C.7	Conditional GAN principles.	58

C.8	Example of tiles from paired (a) and unpaired (b) datasets for roads generalisation.	59
C.9	Illustration of graph convolution with a simple example.	60
C.10	Comparison of models accuracy over time for node classification on Cora and Pubmed dataset.	61
C.11	Comparison of max pooling and G*pooling create different given a feature map as an input (from Deng, Tian, and Newsam 2021).	63
C.12	Example of alteration of the connectivity of the road network and number of set of background and roads.	66
C.13	A 256×256 Google map tile and its 255×255 gradient map.	67
C.14	The partition of the dataset for deep learning.	68
C.15	Illustration of the most common image augmentation in deep learning applied to an image of a mountain road.	69
C.16	The two proposed tiling methods. Left: the sliding windows. Right: the object's shell.	72
C.17	(Du, Wu, Xing, Gong, and Yu 2021) tiling methods.	72
C.18	Example of two tiles with different cartographic resolution.	72
C.19	Comparison of prediction a: with and b: without selection.	74
C.20	Some examples of naive geographic information representation for cartography learning.	77
C.21	Organization of the geographic in tensor for map generalization.	78
C.22	Example of symbolized and layered representation of geographic information.	78
C.23	Example of representation of the binary information location of city center with a black and white mask.	79
C.24	Example of representation for an ordered categorical variable.	79
C.25	Example of representation for a un-ordered categorical variable.	80
C.26	Different proximity graphs in a simple situation of two parallels alignment perturbed by a building without relation.	81
C.27	Extract of the annotated graph for aligement detection.	81
C.28	The characteristics and context of proximity links.	81
C.29	Attributes that describe the difference or ratio of related buildings.	82
D.1	Different generalization solutions when contextual relations are ignored and observed.	87
D.2	The input, the prediction, and the goal of the alignments detection experiment.	88
D.3	Process for the prediction of alignments.	88
D.4	Prediction of the model trained for alignment detection.	90
D.5	Illustration of the limitations of neighboring definition.	91
D.6	The generalization of patterns in a road network is not a subset of initial sections.	93
D.7	The generalization of the road in source databases and the annotations that the neural network aims to learn.	93
D.8	A road network and the associated learning graph.	94
D.9	Annotation and prediction of selection in a small town.	95
D.10	Annotation and prediction of selection in a rural area.	96
D.11	Annotation and prediction of selection in the city center.	96

E.1	Some examples of operations during mountain road generalization.	99
E.2	Simple illustration of our approach.	99
E.3	Evolution of the evaluation value across epochs.	102
E.4	Some results of the segmentation approach (input, prediction and target from the testing set).	102
E.5	Comparison of the predictions for different tiling methods.	103
E.6	Comparison of methods to avoid coalescence in the input image.	104
E.7	Effect of windows size on generalization.	104
E.8	Illustration of the main defaults of generation models with two predictions example.	105
E.9	Comparison of CycleGAN and pix2pix prediction in situations including an important bend series.	106
E.10	Comparison of CycleGAN and pix2pix prediction in situations including an intersection.	107
E.11	Comparison of CycleGAN and pix2pix prediction in situations including a main road.	107
E.12	Comparison of the model trained without and with the connectivity loss.	108
E.13	Comparison of CycleGAN trained with different sizes of the training set.	109
E.14	Comparison of prediction when the model is trained (a) without the selection preprocess. (b) with the selection preprocess.	110
E.15	Example of CycleGAN predictions for images with roads and rivers.	111
F.1	Simple illustration of our approach for generalized map generation.	114
F.2	Extract of Plan IGN, inspiration for our target representation.	117
F.3	The fusion GAN architecture.	118
F.4	Pix2pix prediction for generalized map generation.	118
F.5	Some zoomed image on prediction and target, the predicted building shape may be blurred, irregulare or not credible.	119
F.6	Pix2pix prediction for generalized map generation in city center areas.	120
F.7	Examples of results from pix2pix (paired) and CycleGAN (unpaired) to generate generalized maps.	121
F.8	Comparison of images generated by pix2pix trained with agent-based generalization, and typification-based generalization.	122
F.9	Comparison of images generated by pix2pix trained with layered and symbolized representation as input.	123
F.10	Prediction of a model trained to generate a generalized map that includes road selection.	124
F.11	Comparison of prediction of base model and fusion model with road selection probability information to generate a generalized map that includes road selection.	124
F.12	Prediction on test area with many strong alignments.	125
F.13	Some proposals to represent the alignment on an additional image.	126
F.14	Prediction of the two proposed alignment improvement methods.	126
F.15	Prediction made with the geometry consistency loss.	127
F.16	Results of the model trains with the structure loss.	128
F.17	Prediction of the model trained with the stack of the main and additional information for map generation that includes road selection.	129

F.18	The recommended organization of input data to learn generalized map generation.	132
G.1	Situation map of the training and tested transfer areas.	135
G.2	Prediction of the model trained in the north Alps for tiles from other regions in the same country.	135
G.3	Input tile a) in the USA; b) in France.	136
G.4	Illustration of the main steps of DeepMapScaler workflow.	140
G.5	Illustration of an application for our workflow.	141
G.6	Proposed steps for scale adaptation of road network at 1: 50 000 scale.	142
G.7	Prediction of the scale adaptation for the road at 1:50 000.	143
G.8	Prediction for building shape adaptation.	144
G.9	Prediction made by the model trained for graying segmentation.	145
G.10	Input and target of map generation step.	145
G.11	Predictions of the map generation steps; good predictions are presented on the left and more controversial ones on the right.	146
G.12	Illustration of the alternative transition between steps based on embedding.	148
H.1	Main steps of evaluation when designing a deep learning model.	149
H.2	Illustration of the main specifications of the buildings on a topographic map.	151
H.3	Illustration of the limits of the intersection over union.	153
H.4	Evaluation of the different methods according to user marks.	156
H.5	Examples of images badly evaluated from the validation set.	158
H.6	Distribution and examples of the value for edge-density clutter measure.	161
H.7	Effect of the smoothing operation on a road and its generalization.	161
H.8	Distribution and examples of values for the smoothness measure.	162
H.9	Steps for detecting coalescent pixels.	163
H.10	Effect of coalescence detection on a road and its generalization.	163
H.11	Distribution and examples of values for the coalescence reduction measure.	164
H.12	Intersection and union of roads and their generalization.	165
H.13	Distribution and examples of values for the position measurement.	165
H.14	Distribution and examples of values for the connectivity alteration measure.	166
H.15	Distribution and examples of values for the color measure.	167
H.16	Distribution and examples of values for the noise measure.	168
H.17	Comparison of models strengths and weakness, with a visual representation of constraint satisfaction.	170
H.18	Illustration of the patches to detect border effects: green patches would have a greater value than purple ones.	171
H.19	Vectorization of the building, and a Douglas Peucker simplification.	173
2	Multi-scale map generation as a smooth interpolation problem: comparison between example of smooth interpolations and geoportals.	185

List of Tables

C.1	Comparison of several methods and parameters for the creation of a mountain road shape generalization dataset.	73
C.2	Conversion table road nature and width on map-like image.	73
C.3	Comparison of the quality of several datasets for the generalization of the shape of mountain roads.	74
D.1	Confusion matrix for alignments detection using a GCN.	89
D.2	Confusion matrix for the selection of road sections using a GCN.	95
E.1	Comparison of approach for image-based graphic generalization.	113
F.1	Table of constraint threshold in pixel for a display at 1:50 000 scale.	116
H.1	Confusion matrix for the selection of road section at 1:50 000 using a GCN.	153
H.2	Possible and tested evaluation methods.	154
H.3	List of the models compared during the user test.	155
H.4	Table of possible levels of analysis for vector and raster data.	159
H.5	Alterations are applied to create fake test images.	160
H.6	Interpretation of constraints in a level of satisfaction (at the image level all values are in percent except noise that is in pixel).	168
H.7	Table of statistiques on regularity measures for prediction of model with and without geometry consistency (GC).	173
H.8	Applicability of proposed evaluation methods to main evaluation objectives.	174
1	List of the experiments compared in this appendix.	206

Introduction

Context

Geographic information sciences have evolved widely since the 1960s with the birth of digital cartography made possible by the development of computers. However, it has only been in the last 20 years that online maps have been accessible to the general public. At that time, online map consultation was rare, and paper maps remained the main means of conveying geographic information. Since the launch of Google Maps in 2004, the use of online maps has reached a wide audience. Quickly, the usage of online maps has become democratized, the list of such geoportals has expanded, and the development of the web allows the integration of cartographic views in numerous web applications. At the same time, the field of geographic information has been subject to the multiplication of heterogeneous geographic data sets. Indeed, since the emergence of Web 2.0 in the 2000s, the internet allows not only to find and access information but also to share and collaborate. Then, the success of geographic information initiatives (VGI) as well as the dynamics of open data in the National Mapping Agencies (NMA) marks the beginning of the era of big data. On the one hand, this explosion of data is associated with an increase in the need for synthesis, harmonization, selection, and abstraction of geographic information. On the other hand, it also is conducive to the development of automated learning strategies that require a large amount of data.

In parallel, the internet induces a change in the use of maps. Today, the use of online, multi-representational, and multi-scale maps is increasingly common, and the number of cartographic representations at disposal increases. With mobile devices, maps are accessible anywhere at any time; however, the time dedicated to map reading seems to decrease. The general public, as well as experts, expect adapted content and readable information, with a smooth transition towards a more direct transmission of information. The production of maps adapted to this new usage (simpler, more legible, etc.) is required.

While Web 4.0 (or intelligent) is currently being announced by the specialized media, mapping is struggling to reach Web 3.0 (semantic). Indeed, Web semantics promises to organize the volume of information according to context to meet specific user needs and highly portable and interactive content, while on-demand mapping is still a myth. On-demand mapping is ‘the creation of a cartographic product upon a user request appropriate to its scale and purpose’ (Cecconi 2003, p17). It includes the following issues: understanding the user’s needs, dynamically selecting the most relevant data source for the user’s needs (and eventually including the user’s own data), and generating a product on-the-fly. Today, a few models designed for on-demand mapping exist, but they have no generic map specification (Balley and Regnauld 2011). Whether to understand a geographic phenomenon or for more

common uses (e.g., building an itinerary or finding a point of interest), completeness, and especially abstraction are required, and automatic generation of a readable map for these purposes from a geographic database remains an issue. Thus, this objective cannot be achieved without an effective automated map generalization. Currently, both printed maps and map panels on computers and mobile devices are neither specific nor easy to navigate. For instance, the gap in representation between scales in a multi-scales map produces confusion during zooming in and out (sometimes called desert fog effect, Touya 2019). Map production and cartography research currently require developing techniques to reduce this confusion (e.g., smooth transition, anchored zoom, and creation of intermediate generalization) (Touya, Lobo, Mackaness, and Muehlenhaus 2021).

Map generalization is the adaptation of a dataset for a representation in a map at a smaller scale. This step allows the map content to be legible and is among other things essential in previously mentioned technologies (on-demand mapping, and smooth multi-scale maps). Generalization automation is the foundation of the automation of map creation. Traditional automatic map generalization is often processed by a combination of operators, each resolving a specific conflict in the map (e.g. enlargement resolves building legibility conflict, displacement resolves overlapping). However, the interest in map generalization has recently decreased. As shown in Figure 1, despite the general tendency to increase research publication, the publication related to map generalization is slowing and the proposed presentation at the annual workshop organized by the ICA Commission on Generalization and Multiple Representation is decreasing.

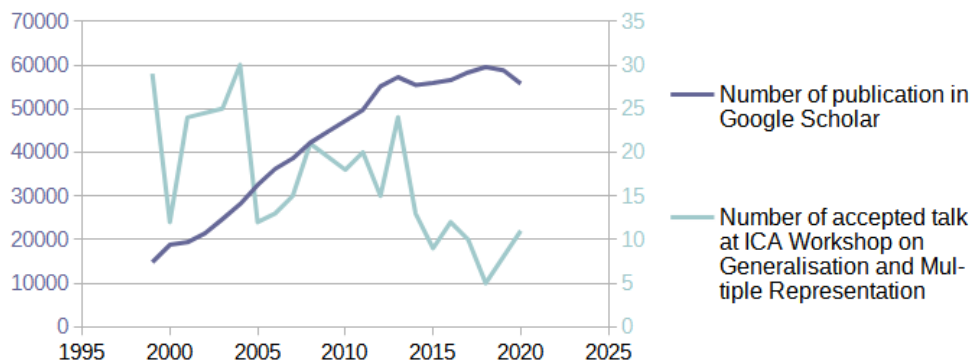


Figure 1: Indicators of the evolution of interest in map generalization.

This tendency can be explained by two main elements. On one hand, the NMA and industries have chosen a compromise between map quality and production time and cost and reduced their investment in map generalization. Indeed, currently, the algorithms used are semi-automatic or automatic generalization algorithms that are mostly acceptable but include weaknesses: they can produce unreadable parts, unnecessary information loss, inconsistencies between scales, etc. On the other hand, map generalization is facing a technical gap: the choice and combination of operators require implicit knowledge acquisition and a difficult orchestration.

In recent years, the emergence of big data and the diffusion of powerful machines with important computing capacities have allowed for a large development of deep learning. Indeed, these artificial intelligence techniques are based on the extraction of implicit information from a set of important examples to learn to mimic human

behavior in the resolution of a data management task. It is succeeding in solving more and more tasks in the field of geographic information management, and GeoAI projects appear in NMA and research institutes (Usery, Arundel, Shavers, Stanislawski, Thiem, and Varanka 2021). However, it is currently used for geoinformation acquisition tasks and is underemployed in cartography or map generalization tasks. We believe that the numerous existing generalized map and dataset openly available since the Web revolution could be used as training examples for deep learning models and allow for automatically extracting the implicit knowledge necessary for map generalization. Thus, research on deep learning for map generalization could unlock research on automating map generalization and accelerate the path to Map 3.0 (i.e., effective map in the Web 3.0 era). Indeed, this generalization would be integrated into an on-demand process and contribute to the creation of more smooth transitions in multi-scale maps (continuous generalization intermediate scale or smooth transition).

Research context

The research presented in this thesis is carried out in the research laboratory (Lastig) on geographic information from IGN the French national mapping agency. This agency has provided printed and online topographic maps for years, and we benefit from archive maps at diverse scales to construct our training sample.

Moreover, this laboratory (and its predecessor COGIT) has an important history in map generalization automation, and we can reuse previously produced automatic algorithms for map generalization for training set creation, complementary approaches, model comparison, and evaluation.

Currently, research on map generalization automation in the laboratory is no longer a central topic, but many projects still require the automation of map generalization processes. Especially, the question of multi-representation map studied in the project MapMuxing (ANR 2014-2018), the navigation in multi-scale map studied in the project LostInZoom (ERC 2021-2026), or tactile map creation with the project ACTIVmap (ANR 2020-2024). This thesis is not founded on any of these projects, but each has a specific need in map generalization automation, and they illustrate how research on automatic map generalization is still of interest.

Finally, the IGN has a growing interest in artificial intelligence techniques, as shown by the creation in December 2021 of an agora on artificial intelligence, and the common meeting with production members, researchers, and teachers about artificial intelligence panorama for cartography in June 2022. Furthermore, since February 1, 2022 IGN is collaborating in the SNSF project (Swiss National Sciences Foundation) “DeepGeneralization: Utilizing Deep Learning in Map Generalization”.

Problem

The use of deep learning for map generalization is a recent approach to solving a old problem. At first sight, this approach is the opposite of the evolution of map generalization, where one first formalizes map rules and then performs a generalization operation to fit those rules. Indeed, with a deep learning approach, rule identification and generalization are learned together from a set of examples. The main interest of this approach seems to be the deep learning ability to learn to produce

generalized maps without explicitly defining map rules but by inferring them from examples. One of the goals of this thesis would be to ensure that it is actually possible. Then, we propose an exploratory approach that aims to identify and describe the potential and challenges of the deep learning approach for map generalization. The objective of this thesis is to:

Explore the potential of deep learning to contribute to map generalization research.

This exploration is done through several experiments that describe a specific generalization problem. The first three proposed experiments correspond to individual map generalization operators, while the last one is more global and the target tasks aim at combining several generalization processes during learning:

- **Exploring the potential to learn shape generalization: the example of mountain road shape prediction at 1:250 000 scale.** Mountain road is a classical sub-problem of road generalization because the sinuous shape of these entities requires an important generalization. Indeed, in a bend series, roads suffer from symbol coalescence. Depending on the shape and context of the road, enlargement, displacement, and schematization of the bends must be applied in some parts of the road. The goal of this experiment is to show that deep learning can learn to reproduce map generalization operators for shape simplification. Especially, the particularity of this use case is the generation of linear shape that is challenging in image approach, and the target is not the output of generalization algorithms but the element of a cartographic database constructed including human corrections.
- **Exploring the potential to learn patterns: the example alignment detection in an urban area.** An alignment is characterized by a set of similar building disposals along an axis. Alignments are salient structures in a map that are expected to be preserved during map generalization. However, they are exclusively implicit in the map and never explicitly encoded in the database. Detecting such structures using deep learning would show that deep learning can effectively learn structure and relation in spatial data that are essential for map generalization.
- **Exploring the potential to learn network structures: the example of road network selection for cartography at 1:50 000 scale.** The road network is a complex structure and during map generalization, all individual road sections cannot be displayed (e.g. pedestrian roads and dead-ends are only relevant in detailed hiking or local map). At 1:50 000 the selection is important: main and regional ways should be kept, but local ways have to be erased when they reduce the legibility or if they do not belong to an important itinerary, the choice of selecting or not a road is depending on multiple factors (e.g. all possible important paths must remain possible even after generalization). Learning this selection would show the deep learning ability to learn from both complex features and multi-scale relation networks.
- **Exploring the potential to learn map generation: the example of a topographic map at 1:50 000 scale.** Finally, this experiment investigates

the deep learning ability to generate a generalized map. This generation involves generalization and the combination of several cartographic themes into a legible map at the desired level scale.

For each experiment, we try to demonstrate the capability and limitations of deep learning, we identify the issues of such an approach, and we proposed and tested improvements. So, our contributions are the following :

- We participate in the creation of benchmarks for the development of deep learning-based map generalization (accessible on zenodo).
- We provide one or several architectures and some loss function.
- We propose some metrics for evaluation and one user test.
- We provide a pre-trained model that is robust for specific input and target scale and tested for transfer in other regions.

This thesis then aims to collect and organize the lessons learned from these experiments. Especially, our contribution also includes some demonstrations of the ability of deep learning for a specific aspect of map generalization (shape simplification, pattern recognition, object selection, and complete map generation), and clues and discussions on deep learning potential, usages, and limits.

Structure of the manuscript

This thesis is structured in three parts: first, we present the main issues related to the use of deep learning for map generalization (Part I), then we present a comparison of deep learning approaches for a common map generalization task (Part II), and we discuss the future of map generalization with deep learning (Part III).

The first part of the thesis aims to explore the main issues related to the use of deep learning for map generalization. It consists of a review of the literature (Chapter A) that first explains the current need for new approaches in map generalization research. The purpose of this chapter is also to explain why deep learning would theoretically be beneficial to map generalization, and this chapter also helps to state the motivation of this thesis and to clarify its objectives. Next, Chapter B presents how to formulate map generalization as a deep learning problem. This chapter describes the need for map generalization and explores what is the most appropriate approach for these needs. As a preliminary experimentation step, it is necessary to delineate the capabilities of each approach. Finally, Chapter C is a theoretical review of the adaptation methods required to make deep learning models match the generalization problem. In this chapter, we present the two main issues of such an approach: designing an architecture and creating the training set, adapted to the use case of map generalization.

In the second part, we study deep learning approaches for solving three map generalization problems: spatial relation preservation (Chapter D), graphic generalization (Chapter E), and the generation of complete generalized maps (Chapter F). All these chapters are based on one or several experiments and are organized in a similar way: first, the problem, materials and methods are presented, then our results are described, and finally, a discussion allows us to conclude on the potential of deep learning for this use case.

The third part presents experiments and discussions on the future of map generalization with deep learning. Specifically, Chapter G explores the possibility of using the models trained during this thesis (on French 1:50,000 data) in other contexts. Then, two possible uses of deep learning models in map generalization: as an end-to-end process and as a mixed process. This chapter includes a proposal to integrate the individual models designed during this thesis to the existing map generalization model and an initial proposal for an end-to-end process to generalize the 1:50,000 map. Finally, Chapter H deals with the evaluation of the generalization of the map processed by deep learning. This is one of the main issues in the development of deep learning for map generalization, because without relevant evaluation methods, the prediction made by deep learning will not be credible and will never be used. Evaluation is an important issue in map generalization in general but the deep learning approach also implies new problems.

Part I

A new paradigm for map generalization

Chapter A

Literature review

This thesis proposes to explore the combination of two research fields: map generalization and deep learning. This chapter aims to present these two fields and to expose interest in a deep learning approach for map generalization.

A.1 Map generalization

A.1.1 Definition

Map and scale

A common definition of a map could be the following: a map is an abstracted representation of a geographic area. Lapaine, Midtbø, Gartner, Bandrova, Wang, and Shen (2021) proposes a more accurate definition that corresponds to contemporary cartographic development: "A map is a medium designed for communication of generalized spatial information and relationships." Both these definitions contain the notion of representation (convey by medium designed for communication in the second definition) of spatial information, and this representation is generalized or abstracted. This means that the map must not only represent spatial data, but also mostly preserve spatial arrangements. Fairbairn, Gartner, and Peterson (2021) add that another important aspect of a map is its function: this representation is designed for a specific usage (i.e., constructing an itinerary, visualizing and analyzing a localized phenomenon, etc.), and propose a definition of map based on this functioning aspect.

The two main questions during map production are "What to represent on the map?" and "How to represent information?". Indeed, the map content and representation depend not only on the map objective but also on its scale. The notion of scale is central in cartography and have two dimensions (Mackness 2007). First, it basically is the ratio between object size on the map and on the ground. Then, the scale has a reasoning meaning: it serves to determine a level of analysis (Ruas 2002). For example, a map at the scale of the continent is completely useless for understanding the distribution of income in a city, even if the map represents the average income per area and contains the considered city. These two definitions of scale constrain the way in which we can represent geographic information on a map.

Maps are made from a geographic database (databases that contain localized information). Whether encoded as a raster or as a vector, the information in geographic databases is neither exhaustive nor exact. Indeed, it is already a simplified

representation of the location of objects in a continuous space described by a reference system (Figure A.1). The objects represented in vector form are located using a point or a set of points, while in raster format the continuous space is discretized using a partition in a finite number of units (pixel), each representing a regular portion of the ground of fixed size.

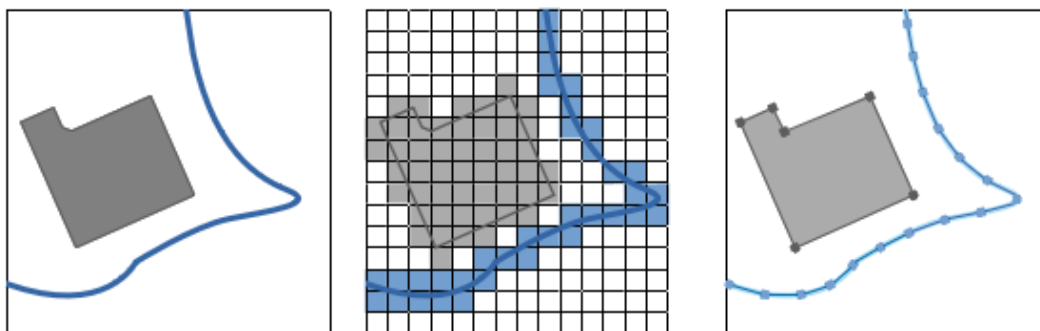


Figure A.1: A ground and a raster and vector representation of this ground.

By definition, a database does not have a scale. However, the geographic database has a geometric level of detail (i.e. resolution of the raster, geometric accuracy of points in vector mode, the granularity of the point set in vector mode) and a semantic level of detail (selected entities, the precision of categories, etc.) (Ruas 2002). Both levels make a database more or less adapted for cartography at a scale. Nevertheless, the map cannot be a simple symbolization of a geographic database, and scale adaptations are necessary. Figure A.2 explains the main steps in creating a map and illustrates the difference between a symbolized database and a map.

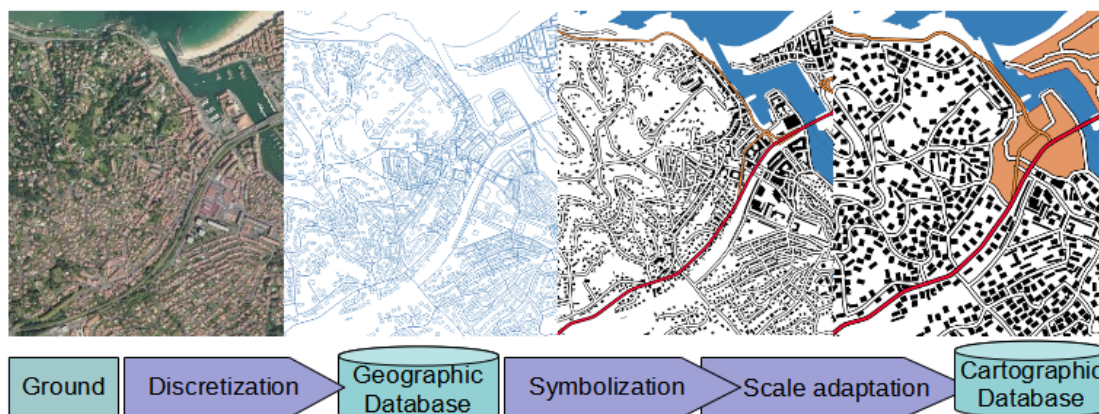


Figure A.2: Main steps of map production.

Scale adaptation

The map creation process is summarized in Figure A.2. The scaling transformation has the effect of reducing the size of represented geographical entities, which can lead to ambiguity of interpretation. For instance, elements are no longer visible or differentiable and, therefore, the map is no longer usable. To avoid this problem, cartographic rules include thresholds of perception and dissociability. For instance, the human eye cannot distinguish too small symbols, too small detail in a shape, too

close objects, etc. To make the representation of geographic objects fits these rules, cartographers proceed to the dilation of elements that do not respect the specified thresholds. However, this simple dilation is not enough; indeed, after dilation, the size of the object no longer corresponds to the space occupied on the ground and this causes clutter on the map. This scale adaptation is called map generalization. Map generalization seeks to adapt a data set for cartography in a map at a lower smaller scale. The goal is to improve the legibility of the map while minimizing information loss. Information selection, simplification, and caricature are required. To maintain the usefulness of a map at the desired scale, generalization has to keep the "main information" such as structures or patterns. This process allows for the construction of multiple maps at multiple scales from one detailed dataset, and its automation would encourage the reduction of data collection and on-demand mapping.

A.1.2 Map generalization automation

Map generalization was first performed manually by cartographers; the question of map generalization automation appears with the use of computers and geographic databases. The following section describes the research advancement in map generalization automation.

Constraints

For the automation of map generalization, one of the most important steps was the formulation of map requirements. A constraint formalizes map specifications (Beard 1991). For example, "a building must have a size greater than 2 mm²" is a usual constraint for automated map generalization.

Constraints first aim at formalizing precisely map generalization objectives, but they also have guided research in map generalization automation. From our perspective of learning map generalization, constraints have several roles: first, they are the most important part of the knowledge the network must acquire. For instance, to learn map generalization, deep neural networks have to understand from images that the overlapping has to be avoided, the element enlarged, the structures emphasize, etc. Moreover, these constraints can guide the map generalization, and their satisfaction is a kind of assessment of map generalization quality. Consequently, in a learning approach, they can be used to evaluate or guide the learning process.

The first attempt to classify constraint separate constraint on legibility called graphical from other (topological, semantics, structural, of Gestalt etc.) (Ruas and Plazanet 1996; Weibel and Dutton 1999). However, these first classifications are misleading, as most constraints are also graphical (e.g. a topological relation, a structure, or a perceived Gestalt are graphically salient in the input map and must remain in the target map). Currently, one of the most common constraint classifications is from (Burghardt, Schmid, and Stoter 2007), it typifies constraints according to their nature (legibility or preservation), the type of geometry (point, line, or area), the number of objects involved (individual, pair of objects, or groups) and the thematic of objects involved. Further, Touya (2011) proposed a model of geographic constraints, which first distinguishes the level of the objects they involve and then their types. Both classifications better emphasize that most map generalization constraints are graphical constraints (and thus can be learned using deep learning). Moreover, these questions on constraint classification expose problems on

the level of analysis, object cardinality, and type; all these problems may also be a center point during training set creation for a deep learning approach.

Finally, this formalization has also contributed to the research on collaborative generalization, where common and shared objectives are a prerequisite for comparing and making work together diverse algorithms and operators. Despite the deep learning approach for map generalization being a shift in map generalization automation, its development requires (as any innovation in map generalization automation) to be inserted and compared with traditional generalization.

Operators

To resolve the conflict on the map and perform map generalization, many operators have been proposed (Stanislawski, Buitenfield, Bereuter, Savino, and Brewer 2014). The following paragraphs present the research on some of the most important operators. Not only do these operators contribute to resolving previously announced constraints, but they mostly provide an answer to the main questions of map generalization (What and How). Indeed, elimination operators aim to indicate which geographic entity or structure must be displayed on the map at the target scale. Then, shape simplification and displacement allow exploring how they must be presented to avoid overlapping.

Shape simplification First operators proposed are for line simplification and are based on the reduction of the number of vertices of the polyline (Douglas and Peucker 1973). However, these operators are not very efficient at dealing with the specificity of geographic data. Then, specific methods for geographic objects have been developed (e.g. Visvalingam and Williamon 1995). When the line to simplify is a road, we distinguish two approaches (illustrated in Figure A.3): (1) a global approach that tries to find the optimal shape and (2) iterative processes that separate roads into homogeneous sections, apply several algorithms to each of them, and merge them together (Lecordix, Plazanet, and Lagrange 1997; Mustiere 1998; Duchêne 2014a). The algorithms of the first approach are often based on the optimization of a set of constraints on the position of the points of the line (Sester 2005), or on a physical model (Bader and Barrault 2001). This is also the approach of the processes based on data filtering and image compression. In the second approach, the process has the following steps: 1) separating the road into homogeneous parts; 2) characterizing each part; 3) choosing and applying the relevant operator; 4) reconstructing the objects.

Similar processes have been developed for other kinds of geographic objects including building (Haunert and Wolff 2010; Ai, Zhou, Zhang, Huang, and Zhou 2014; Cheng, Liu, Li, and Wang 2013) and other more complex shape simplification (Haunert and Sester 2008).

Displacement Then, operators for placement or displacement of objects are popular as they avoid overlaps between objects. The limited available space and the large number of constraints for this displacement (position, structure, and shape preservation, etc.) make the operation complex. For this task as for simplification, global and iterative approaches are possible: the global one defines a displacement for all objects in one step, often inspired by physics models (Liu, Guo, Sun, and Ma 2014; Li, Ai, Shen, Yang, and Wang 2020; Ai, Zhang, Zhou, and Yang 2015;

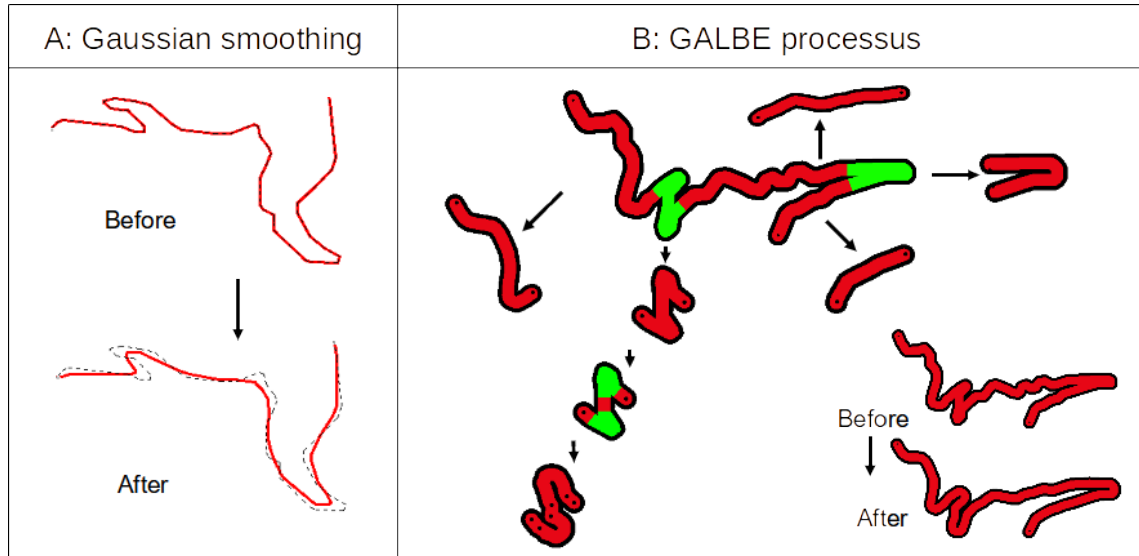


Figure A.3: Example of global and iterative approaches. A: Gaussian smoothing is a technique from image processing that is sometimes applied for smoothing of road shapes, in a global approach. B: GALBE is an iterative process proposed by (Mustiere 1998) that proposes to divide the road into homogeneous parts and characterize each part to choose the most appropriate operator.

Bader, Barrault, and Weibel 2005), and the iterative approach resolves conflicts one by one and iterates until achieving a satisfying result (Aslan, Bildirici, Simav, and Cetinkaya 2012; Ruas 1998; Mackaness and Purves 2001).

Elimination The elimination operators aim to determine which objects should be kept on the generalized map. One of the main issues in this process is maintaining the structure of the information while reducing the number of entities. For instance, Gulgen and Gokgoz (2008) analyze the relations between the scale and the number of roads kept .

Selection inside a road network is based on shape characteristics (Thomson and Richardson 1999), semantics features (importance, status, width, etc.), and the role played in the road network (Garcia-Balboa and Ariza-López 2008; Weiss and Weibel 2014; Jiang and Claramunt 2004). Methods succeed in combining these factors and properly selecting roads (Touya 2010; Benz and Weibel 2014; Stanislawski, Briat, Punt, Howard, Brewer, and Buttenfield 2012). Similarly, rivers selection operators have been developed (e.g. Touya 2007; Benz and Weibel 2014; Ai, Shuai, and Li 2008).

For some other themes, such as buildings, elimination is not the main approach to reducing the number of objects represented. The aggregation (or amalgamation) (Allouche and Moulin 2005; Regnaud and Revell 2007; Ai and Zhang 2007) and the typification (Burghardt and Cecconi 2007; Regnaud 2001; Wang and Burghardt 2019) are preferred. Instead of generalizing a group of buildings with a subset of buildings from this group, in amalgamation, buildings are progressively fused and their shape is generalized into an amalgam built-up area; and in typification, the building group is replaced by another group that is representative of the initial group. Thus, the authors focus on the operators' ability to recognise (Yan, Weibel, and Yang 2008; Yan, Ai, Yang, and Tong 2020; Liqiang, Hao, Dong, and Zhen

2013; Zhang, Ai, Stoter, Kraak, and Molenaar 2013; Christophe and Ruas 2002) and characterize building groups (Steiniger, Lange, Burghardt, and Weibel 2008; Yan and Ai 2018).

All these operators are common and challenging; they require understanding both the graphical aspect and semantics of geographic information. One of the first tasks of our thesis would be to verify that deep learning is able to learn to reproduce them. Verifying the learnability of these operators has later guided the choice of use case in our exploration of deep learning potential for map generalization.

Orchestration

Then, the combination of individual operators in a process able to achieve automatic generalization is still an important problem. Indeed, a fully automatic process should include decisions about the level of analysis (which object to process), the order and parameters of the operations, and the evaluation (Ruas and Plazanet 1996). Brassel and Weibel (1988) state that orchestration must be more than just a sequence of operators and identify the following steps: structure recognition, process recognition, process modeling, process execution, and display. There are several approaches to combining generalization operators: constraint-based optimization, agent-based generalization, or machine learning (discussed in a later paragraph), among others.

First, optimization-based orchestration aims at defining and optimizing an objective function to lead a generalization process. Harrie (2003) try to parameterize simultaneous graphic generalization by providing weight on a set of constraints. Then, optimization algorithms such as gradient descent, least-squares adjustment, or simulated annealing can be applied to find a sequence of operators (or transformations) that minimize constraints violations (i.e., the objective function) (Edwardes and Mackaness 2000; Neun, Burghardt, and Weibel 2009; Sester 2000a).

Then, multi-agent systems (MAS) have been applied to map generalization orchestration in order to contribute to process modeling. In MAS, each agent (object, part of object, or group of objects) tries to resolve its own objective (reach its better-generalized form in our case) by interacting with other agents in these systems and applying operators to itself. Many multi-agent generalization systems exist (Baeijs, Demazeau, and Alvares 1996; Lamy, Ruas, Demazeau, Jackson, and Mackaness 1999; Galanda and Weibel 2002; Duchêne 2003; Gaffuri 2007; Zhang and Guilbert 2011; Maudet, Touya, Duchêne, and Picault 2014; Yan, Guilbert, and Saux 2017). Duchêne, Touya, Taillandier, Gaffuri, Ruas, and Renard (2018) describe in a technical report the potential, advancement and limitations of this approach. This paper explains how MAS distributes complexity among a set of agents. This decentralization is the main advantage of this system.

If the hope of end-to-end map generalization using deep learning would avoid this step of orchestration, it is even useful to question how it would be possible to orchestrate traditional map generalization operators and deep learning-based ones.

Raster-based map generalization

As deep learning methods are mainly image-based, for now, we propose a quick overview of image-based experiments for map generalization.

In the early years of research on map generalization, raster-based and vector-based were developed in parallel as both approaches were equally promising (Peu-

quet 1979; Monmonier 1983; McMaster and Monmonier 1989; Li 1994; Li and Su 1995). In the 1990s, vector-based generalization finally out-performed raster-based generalization (Peter and Weibel 1999; Daley, Goodenough, Bhogal, Bradley, and Yin 1997). As generalization was increasingly monitored by constraints on objects of the map, the vector-based methods appeared to be more practical to use. Besides terrain generalization where raster-based methods are prominent (e.g. Raposo and Samsonov 2014; Guilbert 2013; Li, Wilkinson, and Khaddaj 2001), new raster-based generalization algorithms were very rare. For instance, Pantazis, Karathanasis, Kasoli, Koukofikis, and Stratakis (2009) discuss the potential of morphing in the generalization of raster data, the main limit noticed is the integration of multiple themes in an entire generalized map, and Cheng, Liu, Li, and Wang (2013) design raster-based detector that combines backpropagation neural networks and cartographer’s expertise to learn the relation between the generalized and the detailed shape of the building.

We can also mention a series of articles that propose generalization operators based on the superpixel. Superpixel (Ren and Malik 2003) is an image segmentation method based on features derived from the classical Gestalt cues, including contour, texture, brightness, and good continuation. This segmentation is then used as a basis for the generalization of lines and surfaces (Figure A.4). In this manner, authors propose waterline simplification (Shen, Ai, and He 2018), building simplification (Shen, Ai, and Li 2019), aggregation (Shen, Ai, Li, Yang, and Feng 2019), building selection (Shen, Ai, and Zhao 2021), and building typification (Shen, Li, Wang, Zhao, and Wang 2022). These approaches give promising results for individual object generalization and generalization of a group of objects. However, like most raster-based approaches, they are not able to deal with other spatial relations.

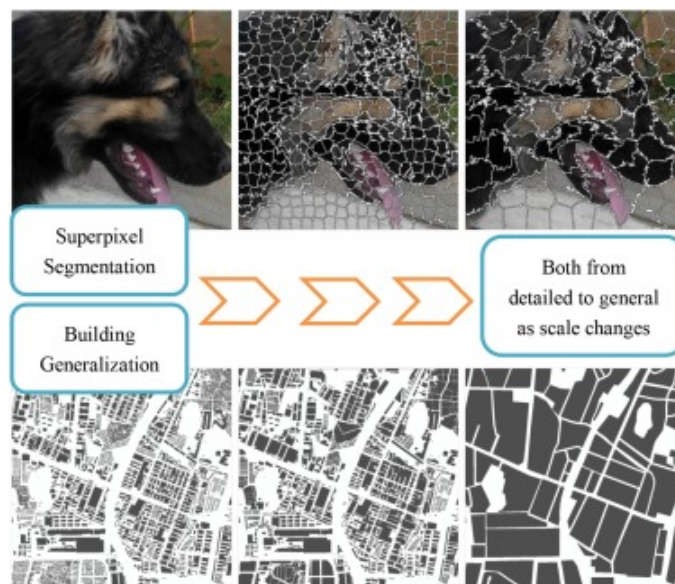


Figure A.4: Superpixel and map generalization, from (Shen, Ai, and Li 2019).

Learning map generalization

In this section, we focus on map generalization methods that involve machine learning techniques. Research in this domain can be organized according to learning

objectives: learning cartographic rules, interpreting spatial data, predicting a set of operations, and evaluating.

The idea of using artificial intelligence to overcome the knowledge acquisition bottleneck appears in the 90s (Weibel, Keller, and Reichenbacher 1995). They identified the human expert as the main source of cartographic knowledge and proposed to construct an interactive system that would allow cartography experts to realize maps and the system to learn cartographic rules from their actions.

The approach that tries to infer cartographic rules is quite rare in the literature, because the diversity of cartographic rules and the lack of metrics limit formalization and usage of acquired knowledge (Kilpelainen 2000). Karsznia and Weibel (2018) is a recent example of such an approach, that tries to improve the selection of settlement points in low-scale generalization maps, by inferring cartographic rules from existing maps.

On the contrary, the approach that aims at the interpretation of spatial data allows making explicit the knowledge necessary in map generalization among other (image interpretation, map-matching, etc.) (Sester 2000b). In particular, she proposes an object-oriented structure where objects are described using a set of attributes and semantic relations. The model learns information on objects by recursively subdividing the training set using an entropy criterion to construct a decision tree that links the input to the target. Thus, machine learning can be used to enrich the initial data and identify patterns to generalize or preserve. For instance, Steiniger, Lange, Burghardt, and Weibel (2008) works on learning to characterize buildings in urban areas. Buildings are characterized according to their geometrical properties and neighborhood, then they are classified into five categories of urban patterns (inner city, urban, suburban, industrial, and rural), and this characterization can guide the generalization of buildings. Similarly, learning strategy can provide knowledge for road generalization as the classification of mountain roads (Garcia-Balboa and Ariza-López 2008) or their selective omission (Zhou and Li 2017).

Then, strategies for learning orchestrating automatic map generalization using deep learning via the prediction of a sequence of operations emerge. Indeed, the choice of the relevant algorithm or sequence of algorithms for the generalization of a set of an object requires an important understanding of implicit spatial knowledge about objects. Especially for road lines, generalization of bend series, rough parts, individual bends, etc. require different generalization algorithms, and the choice of the appropriate algorithm for each part of the entity is an issue (Plazanet, Bigolin, and Ruas 1998). For instance, Mustière, Zucker, and Saitta (2000) describes a process in several steps for the generalization of a mountain road: first, the road is split into homogeneous parts, then, the characteristics of objects are learned, and finally, these characteristics are used to learn which generalization operator is relevant to apply. Another example is from (Burghardt and Neun 2006), who proposes a reinforcement approach for learning a sequence of operations for the generalization of a set of buildings. However, this kind of approach has progressively been abandoned for the agent approach, which allows for more efficient orchestration of map generalization operators.

Finally, machine learning was also experimented with to learn to evaluate generalized maps. For instance, Allouche and Moulin (2005) use a Kohonen self-organization map in order to detect anomalies in patterns and identify regions that need to be generalized; and Harrie, Stigmar, and Djordjevic (2015) develops a mea-

sure that learns to fit the visual evaluation of a generalized map.

If machine learning usage in map generalization research appears as spares and almost anecdotal in this review of the literature, it is mainly due to the following limitations: the acquisition of data, the extraction of characteristics, and the annotation to create an example set was time-consuming and still challenging. Furthermore, the computational resources in the cartography laboratory were limited. Today, the context is different: expert cartographers are no longer the main source of cartographic knowledge, and today we hope to extract this information directly from the existing maps. The open-data dynamic allows for the acquisition of an important example set of maps at several scales. The computational ability has increased and the learning models have progressed. In particular, deep learning architectures promise to skip feature extraction and learn from raw data.

A.1.3 Map generalization is still a challenging question

The literature review also reveals that research on map generalization is still challenging and useful, even if this question has been explored for years. In one hand, the generalization of topographic maps has been widely explored in NMA, and many operators and orchestration methods have been developed. Thus, it is possible to automate or semi-automate the creation of a satisfactory generalized topographic map in a predefined scale range (Duchêne, Baella, et al. 2014). However, some questions about the parameterization, integration, conservation, and sharing of existing code remain important. For instance, (Burghardt, Duchêne, and Mackaness 2014, in Chap. 12, p. 399) formulate the problem this way "How do we standardize generalization components and make them available as interoperable generalization services that could then be combined ? ". On the other side, thematic maps have an increasing usage, and their generalization is rarely explored (Raposo, Touya, and Bereuter 2020). The diversity of content and objectives makes the creation of generic generalization rules and operators more difficult, and the generalization of the base map and thematic information have to fit together (Duchêne 2014b).

Then, big data and open data streams also involve the multiplication of large heterogeneous data sources, and new generalization techniques are required for the level of detail harmonization and to create multi-representation databases (Touya 2017). The application of algorithms to big data and large areas is also an issue. Indeed, data partitioning, processes parallelization, and calculation cost for on-the-fly cartography are modern issues of generalization directly involved by data multiplications.

Touya (2017) states that another challenge is the inability to collaborate the styling processes and map generalization processes. Indeed, both are often applied in chains; however, they are interdependent: the generalization depends on the chosen symbol, and the chosen symbol can be more or less adapted with the generalization. He said that a better collaboration of the two-step would make more efficient maps.

Finally, the evolution of map usage is also a shift in map generalization needs: maps are more and more multi-scale, mobile, and interactive. 1) The optimization of navigation in multi-scale maps requires smooth and coherent generalization through the scales. Indeed, several promising approaches to deal with this problem are investigated: intermediary scale creation (Touya and Dumont 2017), aware generalisation (Girres and Touya 2014), continuous map generalization, and interactive derivation from a unique detailed database (Oosterom, Meijers, Stoter, and Šuba

2014). 2) Map supports widely evolved first from paper to online maps and then from computer to mobile devices (mobile phones and smartwatches). The cartography on such a device requires responsive and even mobile-first map designing with a generalization adapted to the reduced space dedicated to the map (Ricker and Roth 2018). 3) Interactive cartography as its own challenges: (Burghardt, Duchêne, and Mackaness 2014, Chap. 12, p. 398) explains that "building a system that is able to understand a user's needs, and respond with appropriate generalization techniques to produce a tailored map remains a considerable challenge".

A.2 Deep learning

A.2.1 Deep learning principles

Artificial intelligence is the field of computer science that aims to make a machine capable of reproducing human reasoning. Deep learning belongs to the subfield of artificial intelligence that aims at reproducing human learning mechanisms (machine learning). The principle of these mechanisms is to learn a task from many examples rather than explicitly programming the algorithm to resolve it. The particularity of deep learning compared to machine learning algorithms is the following: it does not require feature extraction, it can solve complex tasks, and it requires a very important amount of training examples. Indeed, deep learning includes representation learning, and models can be fed with raw data to automatically discover knowledge from the data itself (LeCun, Bengio, and Hinton 2015). Moreover, the deeply layered structure of the network allows the representation of an increasing level of abstraction from the input, and consequently, deep implicit knowledge extraction from the data.

The learning mechanism is organized into two main phases, the training phase, which aims to define the good weight for the model using the example set, and the calculation phase: when the model is applied to an unknown input and a prediction is made. The results obtained when applying the model to the input are called predictions. During the training phase, each prediction is compared with the expected output of the network (called target). The loss function make this comparison. The goal of the training is to optimize the loss function. As illustrated in Figure A.5 during this phase, the weights are first initialized, then the training examples are used to make the first prediction, and finally the weights are adjusted to produce a better result (backpropagation). These steps (prediction, comparison, and backpropagation) are performed until the prediction is satisfactory.

A.2.2 A short history of deep learning

The premises of deep learning appeared with the idea of a neural network (McCulloch and Pitts 1943), and then in the 1960s with the backpropagation concepts that allow the neuron to learn. In 1979, the first convolution neural network (CNN) finally made it possible to learn visual patterns within images (Fukushima 1980).

The GPU (graphics processing units) development since 1999 allows computers to become faster at processing data and makes the deep learning processes evolve significantly. Moreover, the increasing use of the internet has favored the creation and sharing of large training sets. Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, and Bengio (2014) proposed a new architecture that is

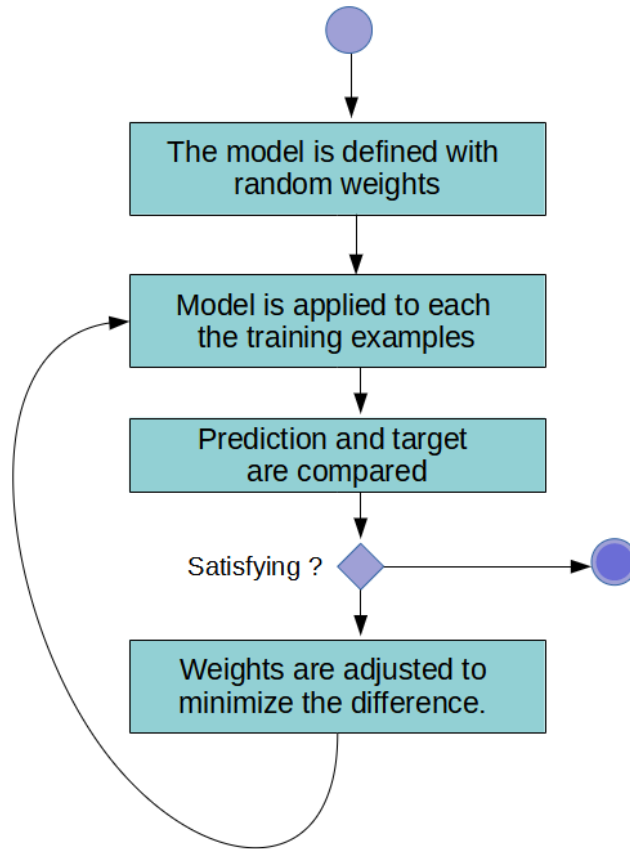


Figure A.5: Action diagram of machine learning the global principles.

able not only to recognize visual patterns, but also to produce images with these patterns. Indeed, generative adversarial networks (GAN) oppose two deep neural networks: a classifier trains to distinguish a real image (from the training set) from a false image (predicted by another learning network), and a generator trains to predict an image that fools the discriminator. This opposition allows the model to obtain more and more realistic images.

All these evolutions have conducted to a multitude of models to solve a large and growing number of diverse applications. This evolution explains why the subject of machine learning for map generalization can today be treated in a completely different way than it was in the 2000s. Deep neural networks erased the common limitation of feature extraction. Then, the image convolution models show the potential to encode spatial information represented in images and overcome the knowledge acquisition bottleneck. In parallel, the graph convolution network aims at resolving the spatial relation limitations of image-based approaches. Finally, GANs are promising for generating parameters images and raising the hope of end-to-end generalized map generation.

A.2.3 Deep learning usage towards map generalization

The goal of this section is to provide a quick overview of the current state-of-the-art in deep learning and expose elements that lead us to consider this technique for cartographic generalization.

Some interesting analogies

In this paragraph, we present some successful applications of deep learning that are not directly related to map generalization or even geographic information, but where an appropriate analogy could suggest the interest of deep learning for map generalization.

First, map generalization is often compared with **text summarizing** (Ruas 2002) and it is assumed that researchers in map generalization automation could benefit from the advancement of these domains (Touya 2015). Indeed, **video, text or music summarizing** aims to extract the main information and structures from a video, text, or music to produce a less detailed but comprehensible new video, text, or music. These tasks, as map generalization are complex cognitive tasks. They require high-level context understanding, and there is no perfect solution or reference solution: different humans can make different acceptable summarization or generalization. Finally, the evaluation of summarized media and generalized maps is a difficult task. Thus, map generalization and media summarization have similar issues but different objects to process. Recently, the emergence of deep neural network mechanisms capable of understanding the context in a video (such as the recurrent neural network (Cho, Merriënboer, Gulcehre, Bahdanau, Bougares, Schwenk, and Bengio 2014) and attention network (Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gmoez, and Kaiser 2017)) has encouraged the use of deep learning for video summarization (Zhao, Li, and Lu 2019; He, Hua, Song, Zhang, Xue, Ma, Robertson, and Guan 2019). Despite these elements, deep learning has been underexperimented for map generalization compared to research on summarization.

On the other hand, map generalization is a graphical problem, and there is a recent trend to use deep learning to solve graphical problems with similarities to map generalization. For example, **image compression** aims to reduce redundancy in an image and reduce its weight in memory without degrading the information. As for map generalization, redundancy reduction, and information preservation are the main goals. Yasin and Abdulazeez (2021) propose a review and comparison of the numerous deep learning approaches for this task. Map generalization can also be compared with **style transfer**, commonly style transfer is the transformation of an image to correspond to a target style extracted from one or several other images (i.e., photography to painting, manga to photograph, day photograph to night photograph, etc.). Indeed, the goal of a style transfer is to apply another style to an image while preserving the main information of this one. The preservation of information in this task and in generalization is relatively close. For style transfer, the preservation of information means preserving the structure of objects that are salient in the image, while in generalization, it means preserving the shapes, relationships, and organization (structure) of the important geographical objects on the map (and often salient in the map). Convolutional networks (Gatys, Ecker, and Bethge 2016) and GAN (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville, and Bengio 2014) have permitted more and more style transfer applications. However, most of the style transfers mainly imply modification in colors and texture, which is not really the goal in map generalization; we rather expect selection or extraction of synthetic information. A more similar example is the style transfer from **sketch to clean drawing** (Simo-Serra, Iizuka, Sasaki, and Ishikawa 2016) that involves the selection, simplification and smoothing of drawing lines, which are quite similar to line selection in a road network at some scales.

Deep learning and geographic information

This paragraph is a brief examination of the state of the art of deep learning techniques for the acquisition and enrichment of geographic information.

Most research projects deal with the segmentation of geographic objects from an image acquired using remote sensing. The network classifies the pixels of the image according to their nature. Therefore, the network can find objects (e.g., building, village) or make a map of land-use/land-cover (Zhu, Tuia, Mou, Xia, Zhang, Xu, and Fraundorfer 2017). However, some use-cases that involve other data sources are possible and promising. For instance, Xu, Chen, Xie, and Wu (2017) and Yuan, Chowdhury, McKee, Yang, Weaver, and Bhaduri (2018) includes volunteer geographic information to improve the mapping of buildings. Méneroux, Dizier, Margollé, Van Damme, Kanasugi, Le Guilcher, Saint Pierre, and Kato (2018) use GPS path for the detection of traffic signals that are not visible in aerial images. Kazimi, Thiemann, and Sester (2019) and Satari, Kazimi, and Sester (2021) use the digital elevation model to detect the structure in the landscape. Finally, some attend to learn structure and object recognition with paper map exist (Schnürer, Sieber, Schmid-Lanter, Öztireli, and Hurni 2020; Chen, Carlinet, Chazalon, Mallet, Dumenieu, and Perret 2021).

Even if these works aim mainly at the identification and segmentation of a particular geographic object or structures and, therefore, are located in the database creation step prior to cartography in Figure A.2, some common elements with map generalization problems exist. Indeed, the use of geographic information as input for deep models is common to map generalization and acquisition or enrichment of geographic information. Consequently, the following two issues arise from this analysis:

- Geographic information is spatially correlated, so nearby features are more affected than distant ones (Miller 2004). The design of a network that respects this principle is the first common issue (Mai, Janowicz, Hu, Gao, Yan, Zhu, Cai, and Lao 2022). For instance, Deng, Tian, and Newsam (2021) propose to replace the max-pooling principle from CNN with a Gi*Pooling that better takes into account spatial correlation between close pixels. This pooling allows an improvement of satellite scene segmentation, especially to reduce the noise at the border of a geographic object and the detection of small objects, and we believe that this principle could be beneficial for most of the learning algorithms that deal with spatial data.
- The constitution of the training set is also affected by the characteristics of geographic information. In fact, the detection of graphic objects as well as their generalization require a certain context around each object, and the constitution of the relevant training set is not as obvious as other computer vision problems (Iddianozie and McArdle 2021; Huang, Zhao, and Song 2018) (later detailed in Section C.2)

Among the diversity of possible geographic information that deep learning can help to acquire, some not only feed the geographic database, but are also often crucial for cartographic (and even generalization) purposes. For instance, the following spatial patterns with specific cartographic needs can be detected using deep learning: urban morphology (Chen, Wu, and Biljecki 2021; Yan, Ai, Yang, Tong, and Liu 2020), interchanges (Touya and Lokhat 2020), road type (Iddianozie and McArdle

2021), building shape (Yan, Ai, Yang, and Tong 2020), etc. This detection is a prerequisite for their correct generalization and representation in a map (Mackness and Edwards 2002). Thus, deep learning has shown its potential to solve enrichment tasks, preliminary to cartographic generalization.

On one hand, it has shown a capacity to produce clues for mapping. For instance, Lan, Li, Wang, Gong, and Ti (2022) employ an artificial neural network to predict the directions and positions of labels in a schematic map and Jenny, Heitzler, Singh, Farmakis-Serebryakova, Liu, and Hurni (2021) use a CNN to predict the shading location on a map. CNN has shown potential for predicting the shape that objects should have in a specific generalized map from their detailed shape building (Feng, Thiemann, and Sester 2019). Then, Kang, Rao, Wang, Peng, Gao, and Zhang (2020) investigate the potential of GAN for this same task on building and Du, Wu, Xing, Gong, and Yu (2021) on the coastline. The experiment on the building did not succeed in improving the quality of the predicted images compared to the segmentation approach: some fuzzy/unrealistic shapes appeared and several buildings were not simplified enough, but for coastline generalization, GANs seem to be efficient. Furthermore, in the coastline experiment, the authors propose to segment the study area into samples that maximize the portion of an object represented in each tile without introducing any distortion. All these attempts have shown the potential of deep learning to produce abstract representations from cartographic data and for graphical generalization, but they have not been tested for the generalization of multiple objects.

On the other hand, deep learning has shown the ability to understand the content of a map and its organization (cartographic rules) from an image. Indeed, Touya, Brisebard, Quinton, and Courtial (2020) succeeded to infer scale and content from images of the thematic maps; Hu, Gui, Wang, and Li (2021) propose a deep learning approach to enrich the metadata of map images by predicting spatial extents and place names, Zhou, Li, Arundel, and Liu (2018) learn to classify six styles of maps and thus show the ability to learn map characteristics. Thus, deep learning models are able to learn the content of a map from an image of the map, Wu and Biljecki (2021) try to go further and experiment how a GAN can fill an incomplete map, using the context of the missing area.

Deep learning and cartography

In this section, we explore how deep learning can be used to generate maps. The task of generating an image of a map with deep learning appears with generative adversarial networks (GAN), which aims to create an image from another. In the literature, most of the experiments of map generation with GAN are about generating an image of a map in the style of GoogleMap from the corresponding aerial photograph. The first models employed for this purpose are designed for a generic image-to-image conversion from paired (Isola, Zhu, Zhou, and Efros 2017) and unpaired (Zhu, Park, Isola, and Efros 2017) dataset. These models are not designed specifically for cartography, but generic models that can learn also map conversion, day-to-night conversion, or label-to-scene conversion (Figure A.6).

These models demonstrate GAN's ability to reproduce main characteristics of the map and preserve salient geographic information. However, preservation and representation challenges proper to cartography remain. Thus, several variations of this architecture have been proposed to improve this specific use case and deal

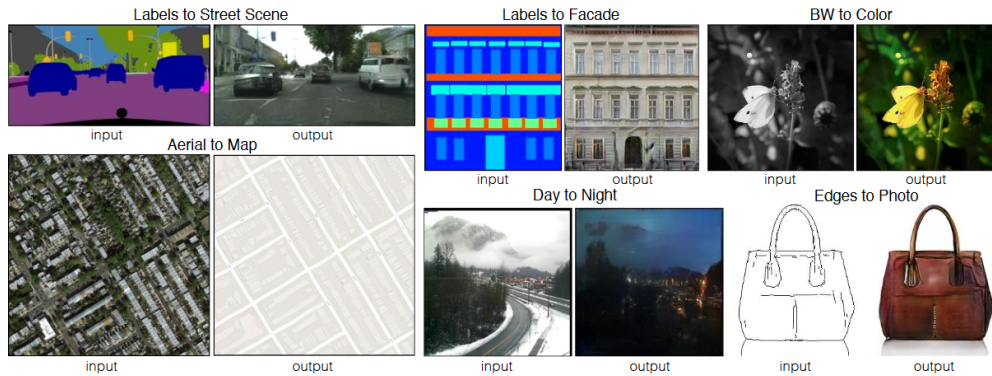


Figure A.6: Example of use-cases of the generic image to image networks, from (Isola, Zhu, Zhou, and Efros 2017).

with requirements specific to map generation. Ganguli, Garzon, and Glaser (2019) compare three different GAN architectures for paired training sets and experiment with the addition of a style loss (from the style-transfer network). Chen, Chen, Xu, Yin, Peng, Mei, and Li (2020) propose to rather use a semi-supervised learning strategy to pre-train the model on rich unpaired samples and fine-tune it on limited paired samples. They also propose an improved loss function based on the image gradient that aims to guarantee the preservation of topological relationships. Zhang, Yin, Zimmermann, Wang, Varadarajan, and Ng (2020) experiment the interest of external geographic data as implicit guidance, for example, by adding an image of GPS traces to the input to improve the shape of generated roads. Fu, Gong, Wang, Batmanghelich, Zhang, and Tao (2019) propose to resolve the problem of geographic object distortion in GAN predictions with a geometry consistency loss (eq. A.1). Indeed, they state that the model G must not be sensitive to a geometric transformation: the prediction of the same transformation t of any image i must look like the prediction transformation.

$$G(t(i)) \approx t(G(i)) \quad (\text{A.1})$$

These publications experiment with end-to-end map generation for a defined map function and scale, but the output is not a generalized map, as the target and input have a similar level of detail. Very few other experiments have explored other map generation tasks. Li, Chen, Zhao, and Shao (2020) have shown that it is possible to similarly convert aerial images into other styles without scale conversion. They used an additional classifier to learn different styles of maps. In a similar way Chen, Yin, Chen, Li, and Xu (2022) tries to learn this conversion (aerial to google map) at several scales at once in order to benefit from the similarity between scales to learn how a map looks, and Henry, Storie, Alagappan, Alhassan, Swamy, Aleshinloye, Curtis, and Kim (2019) train a deep neural network to generate land use map. Finally, Kang, Gao, and Roth (2019) propose to generate maps from raw symbolized OSM data instead of aerial images. The difference in information between input and target allows the authors to note some generalization transformations to avoid symbol coalescence: selection, enhancement, and typification. All these examples include more complex cartography mechanisms and are closer to map generalization.

A.3 Thesis objectives

A.3.1 Needs in map generalization automation

This literature review has shown that map generalization remains an interesting question with unsolved issues. However, interest in this field seems to decrease. In its history, research on automating map generalization has often followed a dynamic driven by innovation in computer sciences and artificial intelligence (e.g., with MAS for orchestration, computational geometry for shape simplification, etc.). We believe that map generalization can benefit from the recent deep learning stream. In particular, the deep learning approach can bring solutions to the following issues:

- **Map generalization must deal with various heterogeneous data.** Deep learning is adapted for big data management, and learning is capable of extracting common information from several data sources. Moreover, both thematic map generalization and multi-source data integration are problems of data source combination, and deep learning is able to easily deal with such problems.
- **On-the-fly and parametrization of map generalization are challenging.** In deep network setting, several parameters are required; then during the calculation, the parameters are no longer required, and the models can be adaptive. Thus, applying a deep learning-based generalized topographic map generation trained model would no longer raise any parametrization challenge. Moreover, training such a model requires important computation and time resources, but then it is easy and quick to apply to new data even on-the-fly.
- **Multi-scale map creation is a challenging task.** Most of the proposed approach for this task would benefit from a deep learning approach: basically, improvement in map generalization in general would simplify the generation of map at intermediate scale; moreover, attention networks could improve context-aware map generalization, and the technique for spatio-temporal data and video generation (such as LSTM Mahasseni, Lam, and Todorovic 2017) are promising for continuous generalization (with the time replaced by scale).
- **Map styling and map generalization are performed separately.** With a deep learning approach, style and generalization can be learned together from examples.
- **User requirements for map generalization are difficult to handle.** As deep learning is able to understand implicit knowledge in images provided by users with specific requirements and generate images with similar characteristics, an approach to on-demand mapping would be to predict the desired map style from user interaction with examples and then learn to fit the user data to the desired style and generalization.

This thesis is one of the first studies on deep learning-based map generalization. As summarized in Figure A.7 map generalization can involve a multitude of methods, including those of artificial intelligence; however, the scope of this thesis is restricted to the exploration of deep learning-based methods.

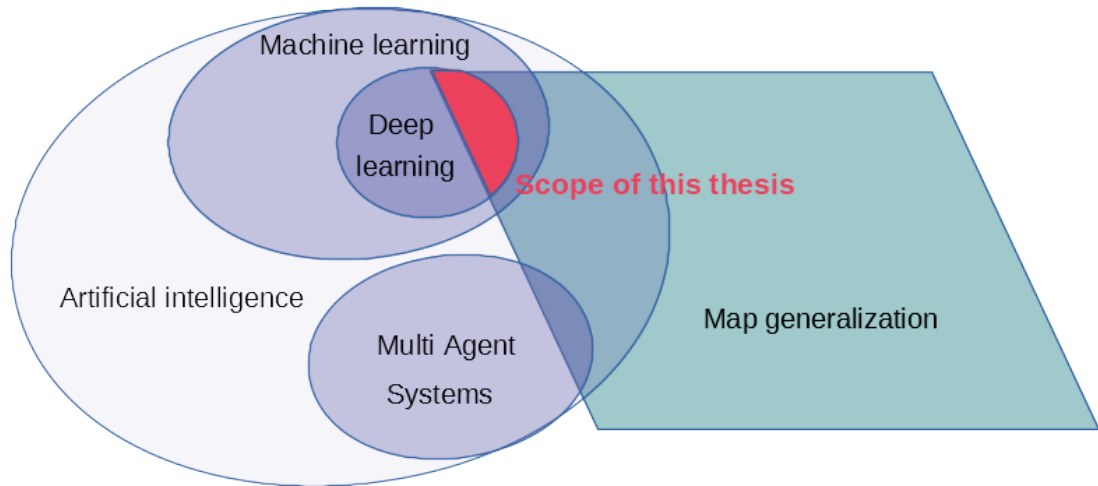


Figure A.7: Scope of this thesis.

A.3.2 Summary of the theoretical motivation

Thus, this thesis is based on the assumption that deep learning is a powerful tool that has the potential to solve problems of map generalization. The main clues to this postulate are the following:

- Deep learning succeeds in solving similar tasks or with similar objectives.
- Deep learning is adapted for graphical problems where the knowledge is implicit in the data.
- Numerous examples of generalized and detailed maps are available for learning.

However, the use of deep learning for map generalization presents at least the following limitations.

- The image representation seems to be the main approach to learn about spatial data in deep learning. However, the use of images to represent spatial data is known to be less efficient in map generalization and may involve loss of information, as shown by the current inferiority of raster-based generalization methods.
- Deep learning models are complex to parameterize. Moreover, they may need adaptation to deal efficiently with spatial data.
- Deep learning intern mechanisms are hidden, and unexplained results can be considered unreliable or difficult to integrate into semi-automatic processes.
- Training in the deep learning model requires powerful computer materials and high computational time. Such resources can be difficult to obtain for researchers in cartography.
- Deep learning has computational limits and can especially be economically, technically, and environmentally unsuitable.

Moreover, it is part of the following moonshot motivation of artificial intelligence: "Design[ing] a software agent that takes a user's GIS-related domain question, understands how to gather the required data, how to analyze them and how to present the results in a suitable form" (Janowicz, Gao, McKenzie, Hu, and Bhaduri 2020, p. 631). Indeed, cartography (and consequently map generalization) is still the most suitable solution for presenting to the human user geographic information or phenomena.

A.3.3 Approach

This thesis aims to contribute to the research on map generalization automation. The use of deep learning for map generalization is a new research topic and the potential of this approach remains to be defined. Thus, this thesis has two consecutive objectives: first, to demonstrate the potential of this technique by reaching the level of existing methods from the literature review, then going further toward fully automatic generalized map generation. This thesis is positioned as an exploration to better understand the issues in deep learning map generalization. This exploration is permitted by the exploitation of IGN archives that provide a large example of a generalized map for multiple objectives and at multiple scales with manual correction produced for years. Moreover, it deals with the following research questions:

1. How to automatically construct an adapted training set from IGN data?
2. How to conceive an adapted deep neural network?
3. How to deal with the specifics aspect linked to geographic information and map generalization (e.g. cartographic context and spatial relation representation in an image.)
4. How to evaluate the result of deep learning-based map generalization?
5. How could the deep neural network collaborate with previous techniques for automatic map generalization (e.g., in a collaborative process Touya, Duchêne, and Ruas 2010)?
6. To what extent is it possible to transfer or apply the models learned with deep learning to large areas?

Thus, our work is organized into two experiments of map generalization: mountain road generalization at the 1:250 000 scales, and urban area generalization at the 1:50 000 scales. These two use cases cover most generalization operators, including simplification, enlargement, deformation, displacement, selection, and typification. These use cases allow us to test graph and image-based approaches and were useful in this thesis for three main purposes:

- To explore the main issues in the deep learning model setting for map generalization:
 - The formulation of an adapted learning task,
 - The practical model implementation (including the architecture of the neural network, the training set, and the design of the learning objective).

- To explore the potential and compare deep learning approaches for common map generalization tasks:
 - Preserving spatial relations while simplifying,
 - Learning complex geometric transformations,
 - Avoiding symbol overlap in a map.
- To bring out important discussions on the future of map generalization with deep learning, especially on the:
 - Contribution of deep learning in map generalization research,
 - Evaluation of deep learning prediction.

Chapter's conclusions

In conclusion, map generalization is an essential and complex process that requires understanding the implicit knowledge in the data. Thus, research for map generalization automation involves more and more advanced information and computer sciences techniques, including artificial intelligence. Current research on map generalization has slowed down, while research on deep learning is expanding to more and more use cases. This lets us think that deep learning can contribute in map generalization research.

Consequently, our thesis is an exploration that aims to:

- Identify the potential and issues of this approach (Part I).
- Explore concretely the ability of deep learning for the most important generalization tasks (Part II).
- Gives directions for the future usage of deep learning in map generalization (Part III).

Chapter B

Formulating map generalization as a deep learning task

Chapter A has shown that deep learning has potential in contributing to map generalization. Some tasks of map generalization can be clearly seen as a classification problem (e.g. object selection); but, most map generalization operators are less simple to formulate as a deep learning task. The goal of this chapter is to describe how to formulate a map generalization task so that a deep learning algorithm can solve it. This formulation is mainly composed of geographic information representation (i.e. as image, graph, etc.) and task definition (i.e. regression, classification, etc.).

B.1 Geographic information representation for deep learning

Map generalization is traditionally mainly formulated as a vector problem: cartographic objects are points, lines, or polygons, encoded using the coordinate of one or a set of points and the interpolation function (see Figure A.1). However, most deep learning algorithms are not adapted to vector data. A tensor is a generic data structure that is used as input for deep learning. Mostly, deep learning can deal with images, graphs, text, and sequences. In the following section, we explain how map generalization can be formulated as an image or a graph problem that can be resolved using deep learning.

B.1.1 Map generalization as an image problem

Even if a raster-based approach for automated map generalization is rarely explored and is said to be limited to dealing with spatial relations, all the information for map generalization are implicitly present in a map image and generalization is mainly a graphical problem.

A map is a representation or medium of communication (Lapaine, Midtbø, Gartner, Bandrova, Wang, and Shen 2021); this representation is perceived by the user as an image printed on a computer, a mobile device, or a paper map. A human can extract most of the knowledge required for map generalization from this image, we can see crowded and free areas, spatial relation and structures:

- **Crowded and free areas**, these areas are essential in map generalization, crowded areas need generalization, and may contain many conflicts. Free

spaces are also important for hosting displaced spatial information that cannot be represented in its place.

- **Spatial relations**, ask a human to describe a map: the description may look the following: "the church is in the park at the north of the river, etc.". Spatial relations are the main elements of this description.
- **Main structures**, the psychology research around Gestalt Theory has shown that humans perceive structure (especially using the principle illustrated in Figure B.1). These principles are involved when using a map and allow users to perceive spatial relations. Moreover, these principles have been widely used in cartography and map generalization to define a relevant group of objects (Thomson and Richardson 1999; Steiniger, Burghardt, and Weibel 2006).

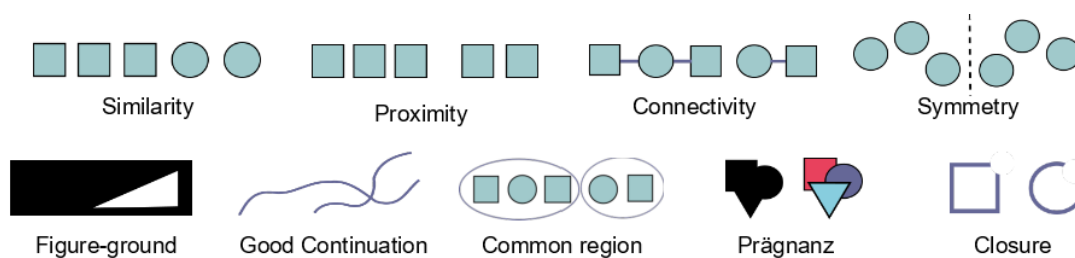


Figure B.1: Law of Gestalt perception: similar objects, close objects, connected objects, etc. are perceived as more related than others.

The usability of map images despite the raster representation shows that the map information is implicit in the image representation and that vector encoding seems to be not a necessity for the human perception of the shape, nature, or relationships between geographic objects. Moreover, in cartography history, cartographers were able to draw a generalized map from a detailed one only. So, why would not an advanced artificial intelligence mechanism be able to deal with implicit data in an image, such as deep learning?

In computer science, images are encoded using a matrix of pixel values of size $height \times width \times channel$. A classical colored image has three channels for the Red, Blue and Green values, each color displayed on the computer is a combination of different levels of these three colors. Other image formats are possible using a different range of pixel values and number of channels. In an image approach of deep learning for map generalization, two strategies are possible: 1) directly using maps that are images (e.g. raster maps) and 2) constructing map-like images from the vector geographic database. We believe that the second strategy has to be preferred as it allows one to set up exactly the desired information, generalization, and style in the training example, and making generalization more adaptive and flexible.

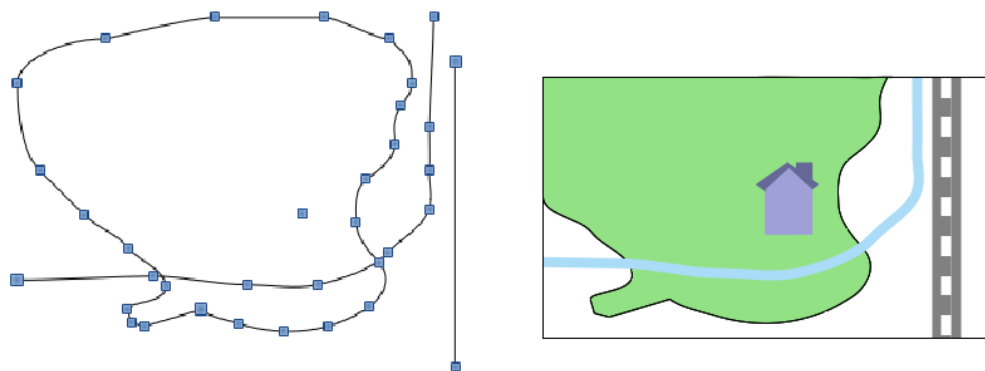


Figure B.2: Some elements of a geographic database and an image representation of these data.

However, the representation as an image requires rasterization as illustrated in Figure B.2, and especially choosing symbolization, order, and rules for affecting a value to pixels (e.g. "if the pixel crosses the shape of the forest its value is light green, e.g., RGB(146,226,133), this rule is not prioritized in case of conflicts with river roads or building representation"). Some examples of this construction for the creation of a deep learning dataset are detailed in Chapter C.

B.1.2 Map generalization as a graph problem

The graph representation is particularly appreciated for map generalization, as this structure allows the generalizers to model the spatial relations (Regnauld 2005), and especially, how concepts such as isolation, connectivity, adjacency, and neighborhood influence map generalization (Mackaness and Beard 1993). The following paragraphs present the three most common graph representations.

Network-graph In geographic information, graphs are often used to encode physical networks, such as hydrographic or transportation networks. They allow the calculation of paths, flows, measures of centrality, etc. For instance, Thomson and Richardson (1995) use graph techniques, such as the shortest path between network nodes and spanning trees, to provide a solution for deriving measures of the functional relevance of network road segments. Variations of such graphs for a road network representation are possible: the primal graph (Porta, Crucitti, and Latora 2006b) where each vertex represents a segment and connected segments are linked, and the dual graph (Porta, Crucitti, and Latora 2006a) where each vertex represents a named street and a link an intersection.

Relation-graph Spatial relations are primordial in understanding maps and their identification is central in map generalization. Touya, Bucher, Falquet, Jaara, and Steiniger (2014) state that: as a prerequisite to developing relation-driven generalization, we need to formalize spatial relations. The nature of relations can be varied and includes topological relations (containing, intersection, adjacency, etc.), proximity relations, and relative position and relative orientation. Egenhofer and

Herring (1990) propose one of the first models to describe the topological relations between two objects using the existence of the interior and frontier intersections (see Figure B.3).




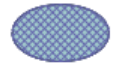




 $\begin{pmatrix} \emptyset & \emptyset \\ \emptyset & \emptyset \end{pmatrix}$ Disjoint	 $\begin{pmatrix} \neg\emptyset & \neg\emptyset \\ \emptyset & \emptyset \end{pmatrix}$ Contains	 $\begin{pmatrix} \neg\emptyset & \emptyset \\ \neg\emptyset & \emptyset \end{pmatrix}$ Inside	 $\begin{pmatrix} \neg\emptyset & \emptyset \\ \emptyset & \neg\emptyset \end{pmatrix}$ Equal
 $\begin{pmatrix} \emptyset & \emptyset \\ \emptyset & \neg\emptyset \end{pmatrix}$ Meet	 $\begin{pmatrix} \neg\emptyset & \neg\emptyset \\ \emptyset & \neg\emptyset \end{pmatrix}$ Covers	 $\begin{pmatrix} \neg\emptyset & \emptyset \\ \neg\emptyset & \neg\emptyset \end{pmatrix}$ CoveredBy	 $\begin{pmatrix} \neg\emptyset & \neg\emptyset \\ \neg\emptyset & \neg\emptyset \end{pmatrix}$ Overlap

Figure B.3: The topological relations according to the 4-intersections model (Egenhofer and Herring 1990).

Spatial relations are one of the most critical information in a map and have to be preserved during map generalization: When a human describes a map, he mainly describes the relation in the map¹. Thus, it is necessary for map generalization to provide a structure that describes relations. For that purpose, a graph representation is often involved. As illustrated in Figure B.4, the relations between the objects on the map are multiple and can be represented in a graph that links each related object.

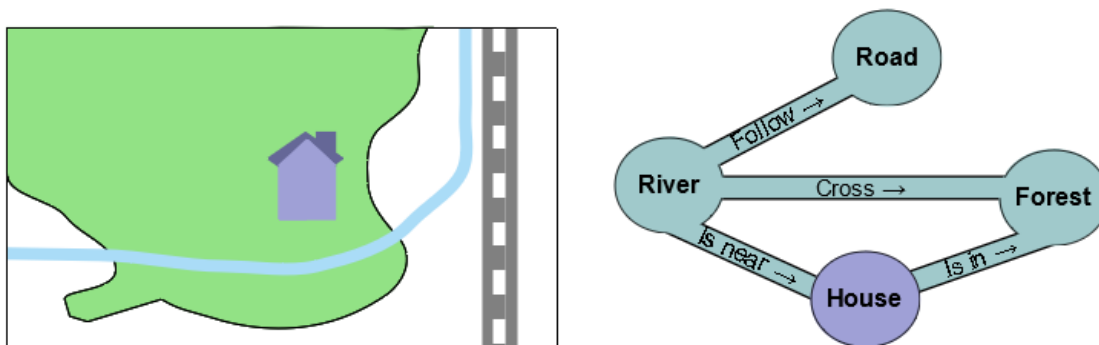


Figure B.4: Example of a graph of spatial relations.

In the literature, the construction of structures like the Voronoï diagram is the basis to identify neighboring (Hangouet 1998). Then, it is used to construct triangulation that reflects the proximity, to identifying conflicts and process displacement (Ruas 1998; Sester 2000a), or typification (Regnauld 2001). Thus, in map generalization graphs are often used to describe proximity (Zhang, Ai, Stoter, Kraak, and Molenaar 2013) or relative position (Bader, Barrault, and Weibel 2005) of buildings during or after their generalization. However, the nature of possible relations

¹Readers can make the test and describe the map in the Figure B.4, they may say that the house is in the forest (topological relation), at the north of the river (relative position), near the road (proximity relation), etc.

in a map is heterogeneous and is not limited to proximity, Iddianozie and McArdle (2021) experiment how a graph neural network can deal with both homogeneous (which describe only one kind of relation) and heterogeneous (which describe various relations between diverse objects) graphs.

Shape-graph Finally, graphs can be used to describe and encode the structure of vector objects. For instance, Harrie (2000) uses a graph issued from the triangulation between points of geometry to describe and generalize the object shape with an optimization approach, Sester (2000b) proposes an object-oriented structure with the relation between pair of objects, for the description of arbitrary geometric 2D-objects in a learning approach. And Gaffuri (2008) defines submicro object for geographic object shape generalization. It is composed of points in an object, linked by segments organized in triangles, each of them has constraints and is used in an agent system. However, such a formulation is rare in the literature, as vector-based approaches are often efficient enough to resolve object-shape problems. A such graph is interesting in our case where the vector representation is not yet compatible with deep learning, while the graph representation is. Yan, Ai, Yang, and Tong (2020) propose a first simple graph representation for deep encoding of the building shape, each point of the shape outline is a vertex of the graph, and the connected points by the outline are linked (as illustrated in Figure B.5).

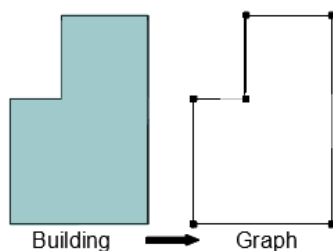


Figure B.5: Example of graph representation of building shape (from Yan, Ai, Yang, and Tong 2020).

Once the geographic information is encoded in a graph, a data model is required to store it on a computer. Graphs are mathematical structures of objects composed of objects (nodes) and relations between pairs of objects (edges). Traditionally, object-representation is privileged due to its completeness and efficiency in terms of storage space (Regnauld 2005). This representation stores for each node its coordinates, its feature, and the connected edge, for each edge a feature list, a starting and ending node. Finally, a graph is an edge and node list (Figure B.6).

Graph	Node	Edge
EdgeList NodeList	EdgeList FeatureList X Y	StartNode EndNode FeatureList

Figure B.6: Graph object representations.

However, some other models can be involved; Figure B.7 illustrates some of them: the list of edges and the adjacency matrix for a graph. Graphical representation is only convenient for human reasoning. The list of edges is a convenient transition

between human and computer representations, as it is easily constructed by listing the relation visible in the graphic representation and easily convertible in an adjacency matrix. The adjacency matrix is more practical for path computation and graph analysis.

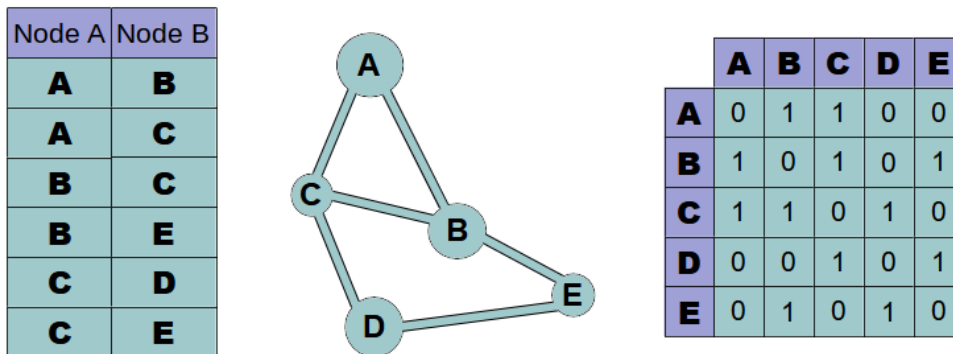


Figure B.7: Graph representations: list of edges, graphical representation, and adjacency matrix.

B.1.3 Other approaches for deep learning-based map generalization

Map generalization can be a sequence problem. On one hand, shapes of a geographic object can be encoded in a sequence of elements with a fixed structure (coordinates, attribute, etc.). This representation is done the following way: an origin point and an order are defined, then each point of the shape is stored in a sequence. This representation is more natural to encode the shape of the linear individual object that has a natural direction (waterline), but can also be applied to other objects with arbitrary choices. On the other hand, predicting a sequence of operations (or objects to deal with) for map generalization can also be an object of interest in deep learning. The interest in such representation is the large quantity of deep neural networks designed for sequence from research on text and video learning.

Map generalization can be a video problem A video is the ordered compilation of several images over time. A multiscale map can be seen as a similar structure: it is the compilation of images over the scale. This representation is based on the same principles as the image, since it is a succession of images. Moreover, it allows us to represent and learn from multi-scale phenomena.

Map generalization with an ad hoc tensor. Finally, we could imagine using none of the approaches presented above and defining an ad hoc tensor for map generalization. Indeed, raster, sequence, graph, etc. are particular examples of tensors. For example, a point cloud can directly be encoded as the tensor of point coordinates and attribute and would be the simpler representation for POI simplification. This representation is easy in point, as the points have a fixed size representation, contrary to other geometries (Knura 2021).

B.2 Deep learning tasks and map generalization

Deep learning allows various tasks, from classification to generation, that produce a new object with the same or even a different structure as the input object (e.g. predicting a text that describes the content of an image). In the following sections, we review some common deep learning tasks that can be used for map generalization.

B.2.1 Classification and map generalization

Classification is already a classic task in machine learning for data enrichment prior to map generalization. Similarly, deep learning classification can be used for this task. Indeed, data enrichment is essential in map generalization automation, as the information on object structure and hierarchy has to be explicitly known to apply some generalization algorithms or orchestrate them.

Classification and data enrichment First, image classification could contribute to data enrichment prior to map generalization. For example, a classifier can detect images containing objects that need a specific generalization (e.g. highway interchanges (Touya and Lokhat 2020)). It can also be used to infer the content and scale of a map (Touya, Brisebard, Quinton, and Courtial 2020; Zhou, Li, Arundel, and Liu 2018) and guide its generalization. However, the image representation is often not the most relevant for such a problem, as it classifies an image that contains objects instead of the object (e.g. for interchange detection we not only need to know if the interchange is in the map, but also where it is).

Complementary, graph classification can also be used on shape-graphs, network-graphs, and relation-graphs for data enrichment. The classification on a shape-graph (or network-graph) would provide information on the shape of the object itself (or on the structure of the network itself). For example, Yan, Ai, Yang, and Tong (2020) use the graph representation of buildings to encode their shape and learn a shape label (L-shape, T-shape, etc.), this label is useful to determine the most adapted shape generalization algorithm for each building. The classification of relation-graphs can be used to obtain information on a group of objects. The same author then uses a Delauney triangulation to describe relations between buildings in an urban block, then a classification model allows to determine if the block has a regular pattern or not (Yan, Ai, Yang, Tong, and Liu 2020). This grouping is essential for generalizing groups. In a similar way, the classification of other objects like sequence, etc. could also be used for data enrichment (predicting features on an object represented by a sequence) or to guide map generalization (predicting if an operator sequence is adapted for the generalization of an object).

In general, classification can be used to obtain global information on the content of the cartographic database. Even if a result at the level of graph or image is often not sufficient, this demonstration is another clue to the feasibility of our thesis. In fact, the experiment's success with learning to classify images and graphs for data enrichment has shown that deep learning can understand some underlying knowledge for map generalization (e.g., there is a complex situation, the image is in urban area, the graph represent a complex building, etc.). With the hypothesis that deep neural networks (and especially convolution networks) are able to understand and encode this underlying knowledge, we expect to design models that learn map generalization without explicitly enriching the data before generalization. Finally,

the classification for data enrichment is not a new possibility offered by deep learning (it was already possible with feature-based machine learning). That is why in this thesis we do not explore this approach much in detail and focus on the next tasks.

Classification and evaluation Classification can also be used for the evaluation of map images, as it could classify map images between legible ones and those that require map generalization (not yet studied in the literature). During this thesis, we do not experiment with this approach alone, but the map generation experiment includes a classifier that learns if the predicted map looks like a generalized map during the map generation approach (see Chapter F).

B.2.2 Segmentation and map generalization

Segmentation in general is the fact of dividing a structure into a subset consistent with common characteristics. For example, in image processing, segmentation allows one to construct pixel groups. So the segmentation output is another image of the same size as the input, where the pixel value describes the group to which it belongs. In graph learning, the term segmentation is not common; however, node classification is a sublevel segmentation that allows us to group nodes in subsets (see Figure B.8), and its usage for map generalization is analogous to image segmentation.

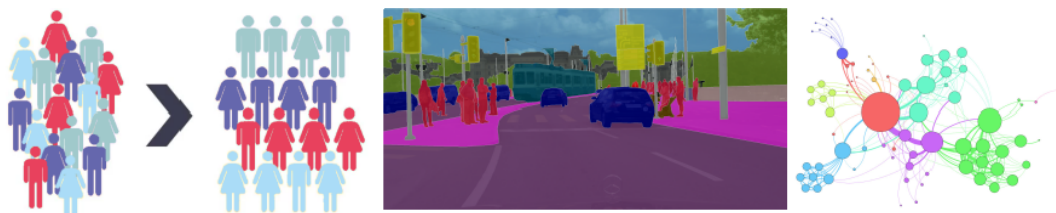


Figure B.8: Illustration of the segmentation definition: a) in common meaning, b) in image processing, c) in a graph.

Segmentation is common in image-based deep learning to detect objects in a scene photograph. This section describes how it can be used for map generalization.

Segmentation and data enrichment Segmentation models are often used in remote sensing images to extract the location of an element. Similarly, segmentation can be used to detect a particular element on an image of the map and contribute to data enrichment. This is a more precise way for data enrichment than image classification (Touya and Lokhat 2020) because it is an enrichment at the pixel level rather than the image level. Especially, segmentation of conflict areas that need correction or generalization (rather than object) in the image is interesting, as it allows evaluating, guiding, or post-processing of map generalization (as illustrated in Figure B.9).

Similarly, graph segmentation allows for class prediction at the node level. Graph segmentation (or node classification) can be applied to relation-graphs to predict information on an object regarding its characteristics, its relation with other geographic objects, and even the characteristic of its neighbor. This task is primarily used for data enrichment: for example, to classify buildings that are part of a structure or group (as we do for alignment in Chapter D). Then, it allows for map

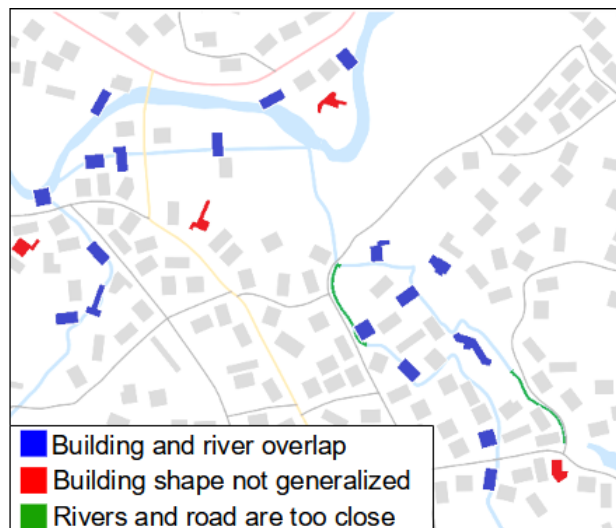


Figure B.9: An example of segmentation for map generalization conflicts detection.

rendering or operation prediction. For instance, it allows for the prediction of object selection with labels "kept" and "deleted" (as we did for the road network in Chapter E). However, a class for other operators is possible, and we could predict if an object needs smoothing, simplification, schematizing, typification, etc.

Segmentation and shape prediction Furthermore, the segmentation models can be used to predict a shape that is not in the input image when all the clues for this segmentation are. In particular, an approach is to predict the shape of an object on a target generalized map from an input detailed map. During our thesis, we experimented with this technique for mountain road shape generalization (more details in Chapter E) while other authors experimented with it for buildings (Feng, Thiemann, and Sester 2019) and coastline (Du, Wu, Xing, Gong, and Yu 2021). The simplification of isolated polygons, such as buildings, seems more efficient than line simplification, which can suffer connectivity alterations due to resolution issues.

Segmentation on a shape-graph can also allow managing the shape of an object by predicting if a point has to be kept or not, but line simplification only based on point selection is often not the most appropriate for map generalization.

B.2.3 Regression and map generalization

Regression models are often similar to classification or segmentation models, but the output is not a class in the set of possible classes but a real value. At the object level, they allow models to rate rather than label objects (images, graphs, pixels, nodes, etc.). Continuous values are more practical for some map generalization steps, especially evaluation and shape generalization.

Regression and evaluation Due to the continuous characteristic of prediction, regression models are more efficient for generalization evaluation and can be used to predict at which level the image is not legible, contains overlap, preserves information, looks like a map, etc. These techniques can also be used to predict information on object generalization, such as clues of the desired level of detail or scaling.

Regression of displacement and deformation Moreover, regression on the nodes of a graph can be used to predict generalization of an object. The model predicts a new position (X, Y) or a displacement (distance and direction). Once these values are predicted, the reconstruction of the vector with the new position for each point may be the generalized shape of the object (Figure B.10). This approach as well applies to the individual building deformation (using a shape-graph), as to the object displacement to avoid coalescence (using a relation-graph). For instance, the building displacement inside a block to avoid coalescence could be resolved by a model trained on a relation-graph that describes the relative position of roads, buildings and rivers.

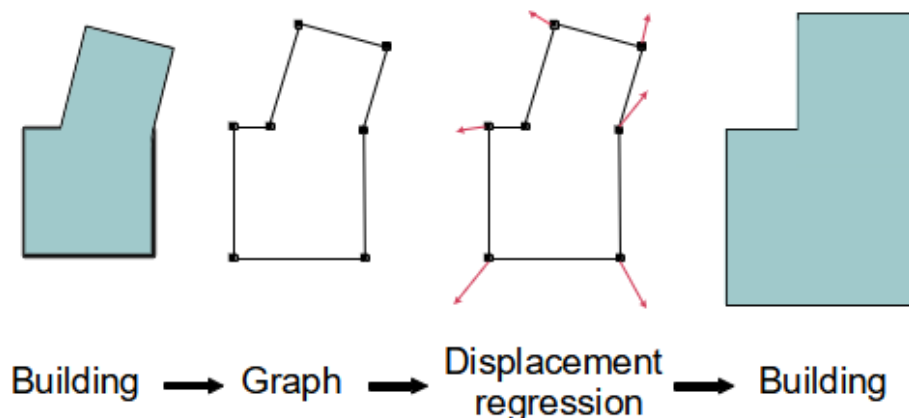


Figure B.10: Learning shape generalization with shape-graph node regression.

B.2.4 Image generation and map generation

Finally, one of the most promising deep learning tasks is its ability to generate images. Indeed, the map is an image; thus, we hope that deep learning will be able to generate an image of the generalized map. Various image generation networks have been proposed to produce images of a map from raw data, such as aerial images (Isola, Zhu, Zhou, and Efros 2017) or symbolized OpenStreetMap (OSM) data (Kang, Gao, and Roth 2019). In this thesis, we test the naive approach to learning complete map generalization using a generative network in Chapter F.

B.2.5 Approach combination

All these formulation proposals have limitations, and we believe that a combination of approaches instead of a unique approach could be beneficial. For example, a model able to combine graph and image based learning (Yang, Lu, Lee, Batra, and Parikh 2018) enables to deep scene description and could be adapted for a deeper analysis of relation in a map.

Other approaches that combine tasks within a single representation are also promising. For instance, the shape prediction using the combination of shape-graph segmentation and regression could achieve better results than individual formulations. Indeed, segmentation is used for predicting if a node has to be kept and structure simplification, then regression predicts the position of the node in the gen-

eralized shape. Thus, the combination faces both the limitations of the structure of regression and the limitation of the position of the segmentation.

Similarly, in Chapter G, we propose an approach for generalized map generation that combines the segmentation of the generalized shape of the object and learning about map rendering through a generation approach.

B.2.6 Other tasks

Finally, our review of tasks that deep learning can resolve is not exhaustive and we mostly mention the most common and promising formulations. However, a deeper exploration of the deep learning capacity could provide highly innovative perspectives.

- Text generation from an image (Kinghorn, Zhang, and Shao 2017) could be used for map description and even map requirements description.
- Text-to-image models such as Dalle² are for now very bad at predicting images of a map but could be reworked to make possible the generation of maps from a description and then for on-demand mapping.
- Video generation (Rathore, Nagar, Arora, and Jawahar 2019) is promising for the generation of smooth transitions between scales. Indeed video summarization and generation often learn together and enable to create a summary of a few images from a video, on the contrary, generating a video from a few images (see Figure B.11) applied to map images at several scales such approach would allow generating smooth transition between the image of the map at different scales.
- Scene-graph generation (Yang, Lu, Lee, Batra, and Parikh 2018) extracts relation between object in an image (e.g. in this image a man is on a bike, on the street), and could be used to create schematic maps from map images or be combined with map generation to enrich the learning of generalized map generation.
- The prediction of a link in a graph is often used by recommendation algorithms. For a node, it predicts which other nodes in the graph it is most likely to be linked to. This approach can be used for generalization recommendations, for example, if there is a change in the level of detail of this element and which other elements may be affected.

²<https://huggingface.co/spaces/dalle-mini/dalle-mini>

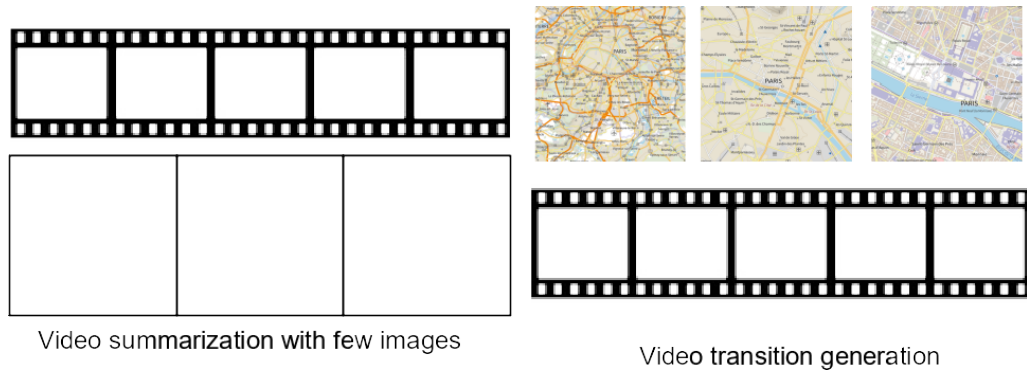


Figure B.11: Learning scale transitions with a video approach.

Chapter's conclusions

As summarized in Figure B.12, deep learning can solve various data management problems, and most map generalization tasks can be formulated as a deep learning problem. However, several formulations are often possible and may lead to a different result and a different level of information in the results. Finally, combinations of several approaches and more innovative approaches are possible and promising.

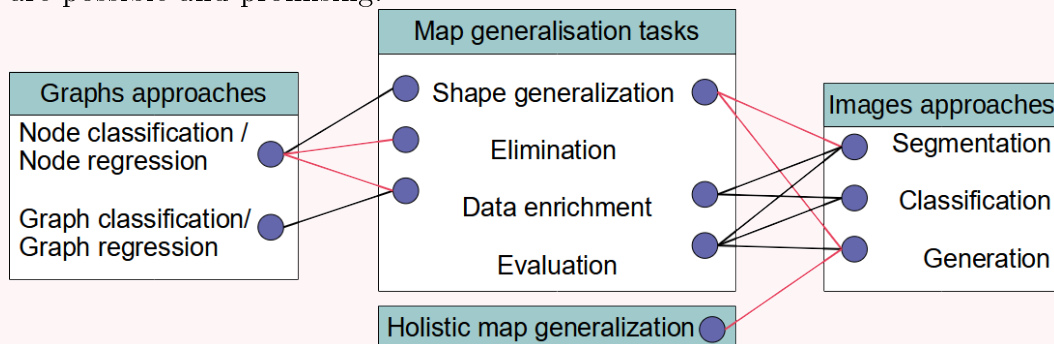


Figure B.12: Main possible formulation of map generalization as a deep learning problem (the approaches that are experimented in this thesis are highlighted in pink).

Chapter C

Designing a framework for deep learning based map generalization

In the previous chapter, we presented various way to formulate the desired map generalization task into a deep learning problem. In this chapter, we propose to go deeper into the definition of a learning strategy and explore how deep learning can be implemented in real-world map generalization applications. This chapter presents some concrete insights about deep neural network design and training set creation.

C.1 The characteristics of an adapted deep neural network

The creation of new architectures that converge faster, encode thinner information, need less data, etc. is one of the most important points of interest in deep learning research. This section aims to provide keys for using an adapted deep neural network for map generalization among the numerous existing deep neural networks.

C.1.1 What is a deep neural network?

This first section explains more precisely what a deep neural network is. The different components of the deep neural network are reviewed to allow the reader to understand how a deep neural network is designed.

Architecture A deep neural network is a succession of neurons, as illustrated in Figure C.1. Each neuron summaries a set of inputs in one output value. The organization of these neurons is called architecture.

The model architecture is responsible for the encoding and decoding of the information to produce the target. The structure of this model (number of layers, number of neurons per layer, and neurons activation function) can vary depending on the target task. For instance, a deeper model (with more layers) will encode more information levels and will be able to perform more complex tasks. The structure of the network also has an impact on the robustness of the model. For example, dropout layer randomly omits a part of the weight at each iteration of the training, this induces randomness in the information encoding and avoids the model overfitting. In deep learning, as in life sciences, the function of a neuron is to transmit

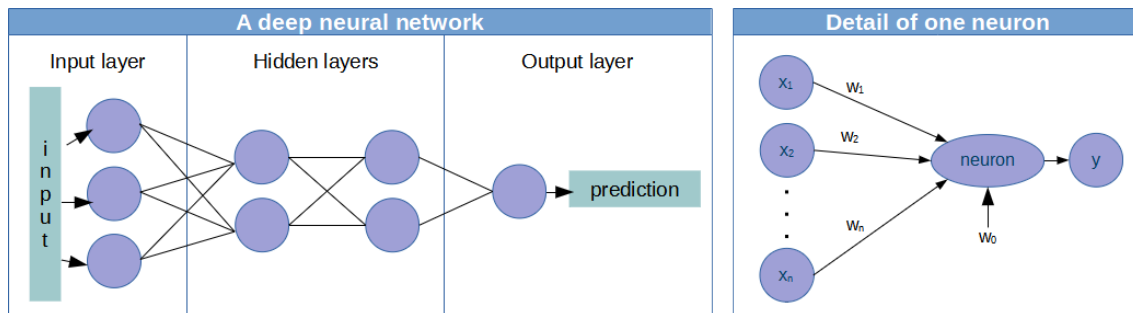


Figure C.1: Neurons and deep neural network.

information. In fact, a neuron is a bounded parametric function of the following form (eq. C.1) with the input x a tensor with n elements, w_i the weight, w_0 the bias, and f the activation function.

$$y = f(w_0 + \sum_{i=1}^n w_i \times x_i) \quad (\text{C.1})$$

Initialization The model has initial weights that are often randomly defined. This initialization is one of the reasons for the non-determinism of deep neural networks¹. The random choice is often performed using a Gaussian or Uniform distribution between small/near zeros values, (Goodfellow, Bengio, and Courville 2016):

The choice of Gaussian or uniform distribution does not seem to matter very much but has not been exhaustively studied. The scale of the initial distribution, however, does have a large effect on both the outcome of the optimization procedure and the ability of the network to generalize.

However, choosing a non-random point of departure can be beneficial when the model has already been trained in other use cases. In fact, using the weights from one step of a previous training may reduce training time. If this choice is relevant in some common use case from computer vision (photograph of an indoor scene, outdoor scene, animals photograph, etc.), is it true of map generalization? The idea is to skip some of the initial learning steps that consist of learning the encoding of the global property of objects, because they may be shared. Is there sufficient common knowledge between map examples you give to your network and the one the initialization network has already seen? For example, we think it may be relevant to choose the weight of a style transfer network for regularization of hand drawings to initialize a road generalization model because they are both composed of lines on a background (Figure C.2).

Iterations Learning is processed by iteration. Each iteration shows the model a certain number of samples (grouped in a batch) and complete the optimization for this batch. The use of batch allows computer to parallelize calculation as optimization is performed for several image at once, its size affects the calculation time and memory required. An epoch is reached when all batches have been seen once. The learning often stops when a certain number of epochs is reached. However, this number

¹The model will be different each time it is trained.

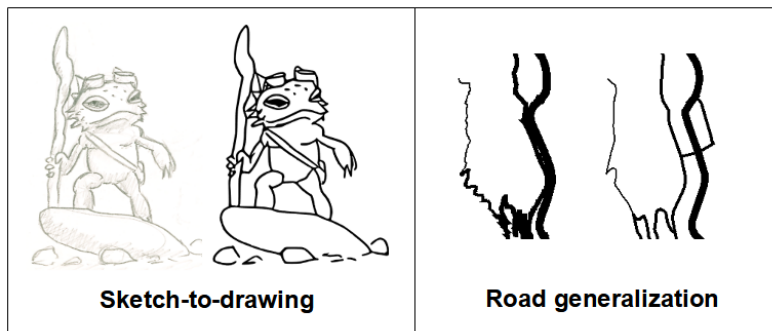


Figure C.2: Comparison of the sketch-to-drawing conversion (from Simo-Serra, Iizuka, and Ishikawa 2017) and mountain road shape generalization.

has to be adapted. If it is too small, learning may not be finished and the results may be poor. On the contrary, when it is too large, the model may overfit the exact distribution, but is unable to predict a relevant result for any unseen sample. To avoid over-fitting we have to stop the training at the right time (when the loss converges?); this moment is usually determined empirically.

Learning objective The learning objective (loss function) is the function that measures the deviation between model output and expected result. It is calculated for each prediction and optimized at the end of each iteration during training. An adapted loss function can improve learning. The objective function is basically based on a distance between the prediction \hat{Y} and the target Y (both composed of n elements). For example, to compare two images, a measure derived from the least absolute error (L1) and the least square error (L2) (eq. C.2) can be used as a loss. L2-based losses are often preferred, as they generally perform better; however, it is more sensitive to outliers and can produce less satisfying results on some datasets. The choice of an adapted learning objective for map generalization is important and detailed detailed in Section C.1.4.

$$\begin{aligned} \text{L1} &= \sum_{i=1}^n |Y_i - \hat{Y}_i| \\ \text{L2} &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \end{aligned} \tag{C.2}$$

Optimization Several optimization methods (optimizers) exist that define how the weights are adjusted. At each iteration the optimizer is responsible for changing the weight values (w_i) to minimize the loss function. The choice of an optimizer affects the speed and accuracy of the training. The learning rate provides a scale of how much model weights should be updated at each iteration. The first common optimizer is gradient descent. It is defined as follows: let g the objective function be some continuously differentiable real function. In this method a fixed step size, α is chosen and used with the derivation of the loss function (that gives the direction of the functions slot) to determine the new value of weight (eq. C.3).

$$w_i = w_i - \text{alpha} * \frac{\partial g(w)}{\partial w_i} \quad (\text{C.3})$$

This strategy is long and not very adapted for not convex functions. So, several adaptations from this method have been proposed. For instance, the Stochastic Gradient Descent (SDG) uses randomness to reduce computation time. At each iteration the optimization is made for only a small portion of the data, thus, the path for reaching a minimum is noisier, and more iterations are needed, however the time for computation each iteration is reduced. Today, the most popular optimizer is the adaptive moment optimizer (ADAM); unlike the SDG, which maintains a fixed learning rate during training, it updates for each weight individually. It estimate the first and second moment of the gradient using the exponentially moving averages of the gradient m_t and of the square gradient v_t at iteration t (see eq. C.4, β_1 and β_2 are hyper-parametres and g_t is the gradient from eq. C.3, v_0 and m_0 both are initialized at 0).

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \end{aligned} \quad (\text{C.4})$$

Then it introduces a bias correction that consists in dividing m_t (resp. v_t) by $1 - \beta_1$ (resp. $1 - \beta_2$), and use the corrected moving average to updates weights (see eq. C.5, with ϵ a fixed small value to avoid division by zero).

$$w_t = w_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (\text{C.5})$$

This optimizer is often preferred as it is faster to compute, requires low memory, and less tuning. However, it is most likely to produce a model that does not fit the data.

C.1.2 Some deep neural architectures for map generalization

The goal of this section is to give some examples of deep neural networks and to identify why they are suitable or not for map generalization. Considering the diversity of deep learning models and applications, we choose to focus on the models able to resolve some formulations expressed in Chapter B.

Convolutions neural networks for image based data enrichment Convolution neural network are models able to learn image based tasks. We hope to use the convolution neural network to learn data enrichment and shape simplification for map generalization (see Chapter B). A convolution is an operation on matrix (and images represented as a matrix) that uses a kernel (small matrix with a predefined value) to reorganize the patterns in an image (such as edges, textures, object parts, etc.). This kernel is sliding over the input image (as illustrated in Figure C.3). The value of an output pixel is the sum of a element-wise product of the kernel and the input. In CNNs the values of kernel are learned from data rather than predefined manually, so they are intialized according the choosen initialization method and then optimized during training. CNNs also include pooling (also illustrated on figure C.3), it allows the model to reduce the dimension of information; it summarizes (via maximum, average, or other methods) the information in a neighborhood.

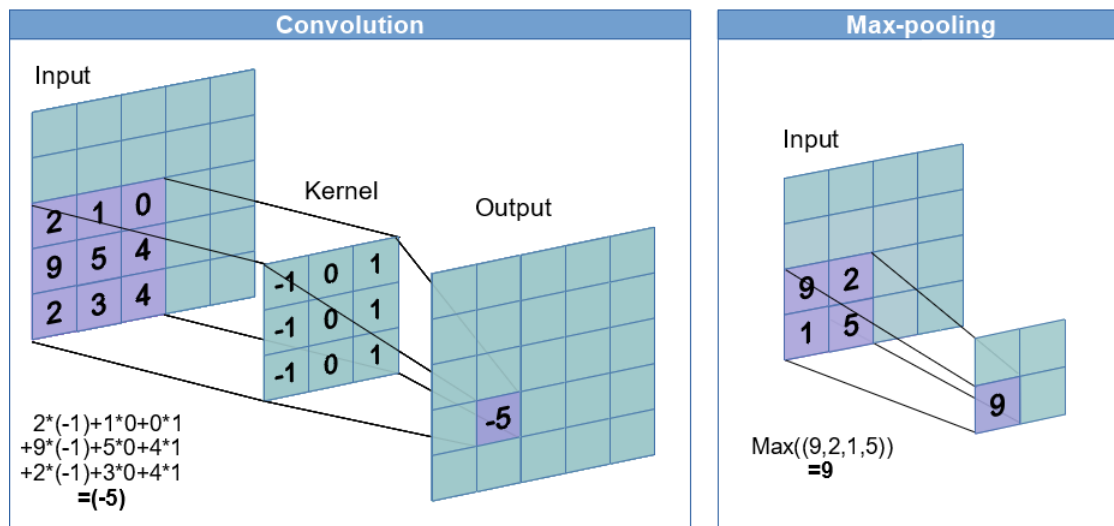


Figure C.3: Image processing in CNN. Convolution with a kernel of size $(3 * 3)$. Max-pooling of size $(2 * 2)$.

Most deep learning models based on images are derived from convolution neural networks (CNN). CNNs are organized in two steps (Figure C.4): 1) a series of convolutions that filter each image and extract features from an image, and 2) a layer connects each input to all neurones (fully connected), it decodes the feature to achieve the target task (e.g. predict a label by image for classification). Most convolution layer consists of convolution, pooling, and activation.

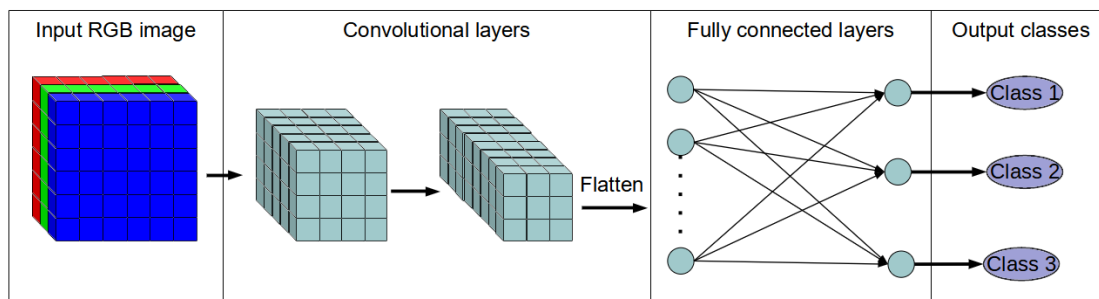


Figure C.4: The functioning of a CNN.

Since the first functioning implementation of a CNN, named LeNet (Lecun, Bottou, Bengio, and Haffner 1998), CNN architectures have evolved widely. They tend to increase the number of layers (depth). Furthermore, the layers structure management involved a large improvement in performance: the average pooling was replaced by max-pooling, and padding was introduced. The order of layer component also changes, in particular the systematic usage of max-pooling on top of each convolution layer is questioned and abandoned (Krizhevsky, Sutskever, and Hinton 2012). The Rectified Linear Unit (ReLU) is today one of the most popular activation functions, it replaces the negative value with zero (eq. C.6).

$$y = \max(0, x) \quad (\text{C.6})$$

Finally, a smaller filter size (combined with increasing model depth) allows the model to encode a different level of relation in the image (Simonyan and Zisserman

2015). All these improvements have contributed to make CNN able to extract various implicit information from images in general. We think such model would also be able to understand implicit knowledge from images that represent geographical information and particularly maps that are full of relations between pixels.

In parallel, fully convolution networks were developed for image segmentation (Long, Shelhamer, and Darrell 2015), they allow to predict a classification of pixels in the image. Feng, Thiemann, and Sester (2019) compare two of them for building generalization: a U-net (Ronneberger, Fischer, and Brox 2015) and a residual U-net (Zhang, Liu, and Wang 2018). Both models have shown promising results and seem to be able to encode the different levels of relation in a map image. The U-net (illustrated in Figure C.5) is structured in a series of convolutional layers into a contracting path with the typical architecture of a convolution network and an expansive path that reconstructs the spatial size of the input. Both parts are connected through skip connections that they stack the embedding of each layer with the input of each decoding step information, this allows for managing border effects and introduces several levels of information encoding in the prediction. The residual U-net is a U-net where the convolutional layers are replaced by the residual unit from (He, Zhang, Ren, and Sun 2016) (i.e. batch normalization BN, ReLU activation, and convolution, illustrated on Figure C.5). Furthermore, the input of the residual unit is skip-connected after the next convolution layers. This model improved the generalization of the building shape without increasing the depth of the network (Feng, Thiemann, and Sester 2019).

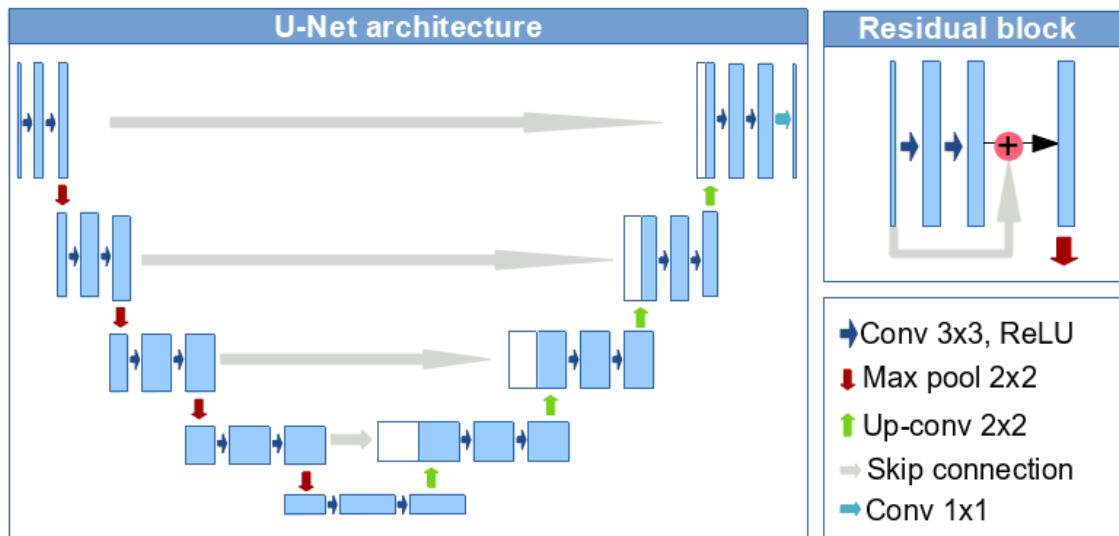


Figure C.5: A U-Net architecture and a residual block.

CNNs produce categorical predictions (classification/segmentation) and thus involve specific learning objectives. The Cross-Entropy also called Log Loss and the Dice coefficient are the most common learning objectives for such models. Cross-Entropy (eq. C.7) measures the difference between the distribution Y for target and \hat{Y} for prediction and x a possible output class, $Y(x)$ is the probability of the event "the target is x " and $\hat{Y}(x)$ the probability of the event "the prediction is x ". This measure can be biased in case of imbalanced classes whereas Dice D is preferred. The Dice coefficient is based on the intersection of the set of predicted elements \hat{Y} in one class and the set of elements Y in this class in the target (eq. C.8). The objective

is to maximize this coefficient.

$$H = - \sum_x Y(x) \log \hat{Y}(x) \quad (\text{C.7})$$

$$D = \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \quad (\text{C.8})$$

These measures compare values element by element and summarizes pixel-level differences; it can be misleading in several cases including map generalization. Figure C.6 gives an example of this problem for an image of a road: the target generalization avoids the proximity conflict between the two roads by displacing the road to the right and enlarging the turn of the road at the left. In contrast, the 'good' prediction does not displace the right road and attenuates the bend, even if this generalization is relevant and correct, the measure would really badly evaluate it. On the contrary, the "bad prediction" has many road pixels common with the target but is not good at all. The problem is the following: several different generalizations can be good, errors in some pixels are more important than others, and the overall aspect of the prediction is more important than the local value of the image.

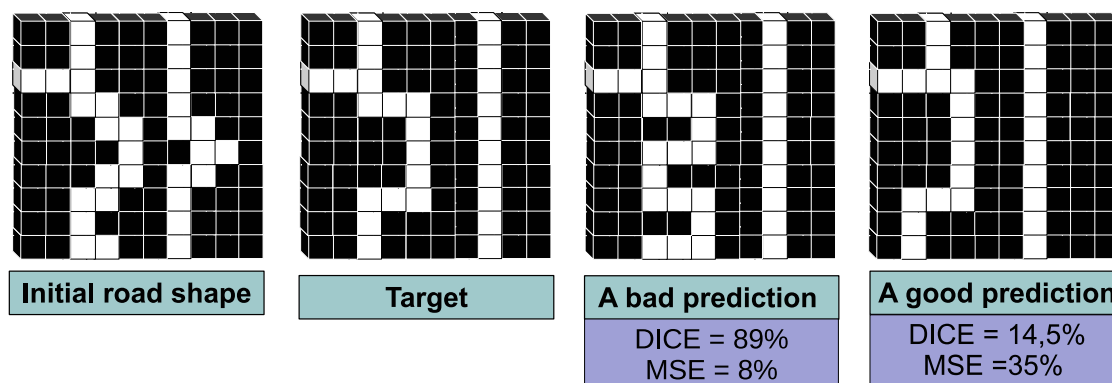


Figure C.6: An example where pixel-level loss is not adapted for map generalization.

Thus, there is a need for an appropriate measure of prediction quality, that is not based on pixel level or that seeks to measure the specifics of map generalization for objective function definition (see section C.1.4).

Models for map generation

Several strategies for image generation are possible. The two most common are style transfer and generative adversarial network (GAN), both based on the combination of several CNNs to allow the generation of the desired image.

On one hand, style transfer networks are composed of two CNNs: one that extracts the content from the source images and one that extracts the style features from the source images (Gatys, Ecker, and Bethge 2016). Then it applies the style of an image to the content of another image. Style transfer approaches are not really adapted for map generalization, as it is not only a style issue but also a content one. The content has to be simplified to make the style applicable (see Figure A.2).

On the other hand, GANs can perform various transformations, especially those that contain changes in content and level of detail. They are based on the combination a generator and a discriminator (see Figure C.7):

- A generator is a CNN that derives an output from the the input image using a random noise z .
- A discriminator is a classifier that assesses whether an image is false or a real target image.

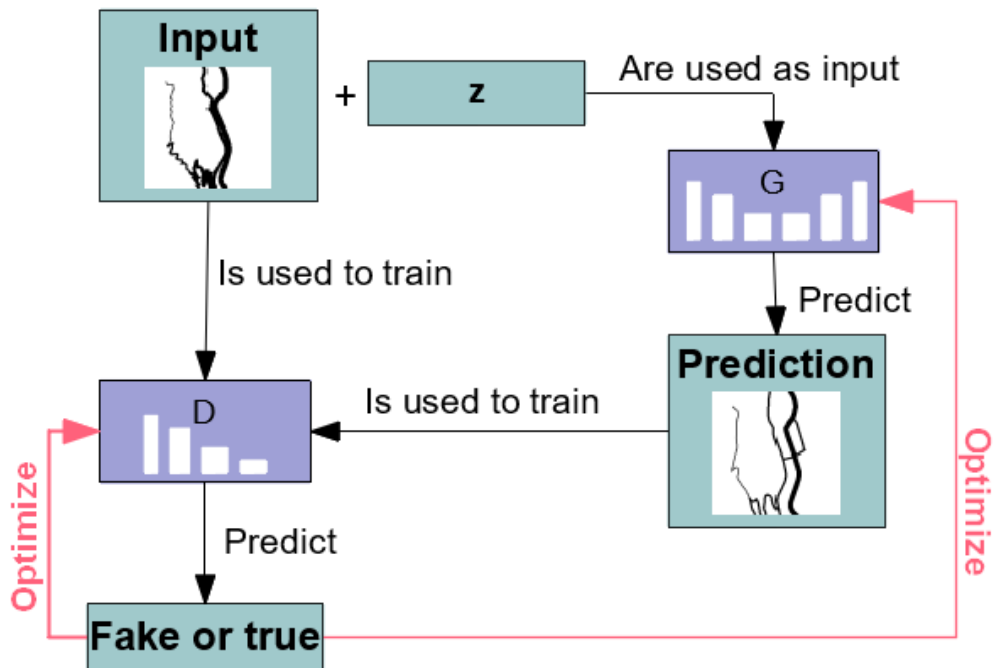


Figure C.7: Conditional GAN principles.

The loss of a GAN includes a similarity measure that compares the prediction and the target using the loss L1 or L2 (or the input and reconstruction in the case of an unpaired circular model like CycleGAN or DualGAN) and an adversarial loss that investigates the overall credibility of the prediction, compared to that of the target. For x and y the associated target, $G(x, z)$ is the generator prediction, the adversarial loss is the sum of the probability of fooling the discriminator $\mathbb{E}[\log(1 - D(x, G(x, z)))]$ and its probability of correctly classifying the target $\mathbb{E}[\log D(x, y)]$ (eq. C.9).

$$L_{cGAN} = \mathbb{E}[\log D(x, y)] + \mathbb{E}[\log(1 - D(x, G(x, z)))] \quad (\text{C.9})$$

There are mainly two types of GAN: those that can deal with paired images (Isola, Zhu, Zhou, and Efros 2017) and those that deals with unpaired images² (Zhu, Park, Isola, and Efros 2017; Yi, Zhang, Tan, and Gong 2017; Gan, Chen, Wang, Pu, Zhang, Liu, Li, and Carin 2017). Paired images mean that one example is made up of an input image with the corresponding output while using unpaired images means that two independent example sets are used: a set of good input images and

²GANs designed for paired images are often called supervised and those for unpaired unsupervised, even if it is not properly unsupervised learning as the output is described.

a set of good output images without matching both (paired and unpaired image are illustrated in Figure C.8). Paired examples often allow learning some features that unpaired example do not when input and target are similar. In an unpaired approach, the learning objective is less to reproduce the operation of automatic map generalization but more on producing image with the aspect of a the generalized map. This approach is preferred when input and output are very different or when it is difficult to find paired data. For the use case of map generalization, paired learning will learn to link the detailed and the generalized map, while unpaired learning would rather learn what a detailed and generalized image look like and then how to transform the detailed one so that it looks like a generalized map. Finally, semi-supervised learning represents a key issue for GAN improvement (Simo-Serra, Iizuka, and Ishikawa 2017). It consists in combining learning on a few paired and many unpaired examples, and this process is expected to benefit from the advantages of both methods.

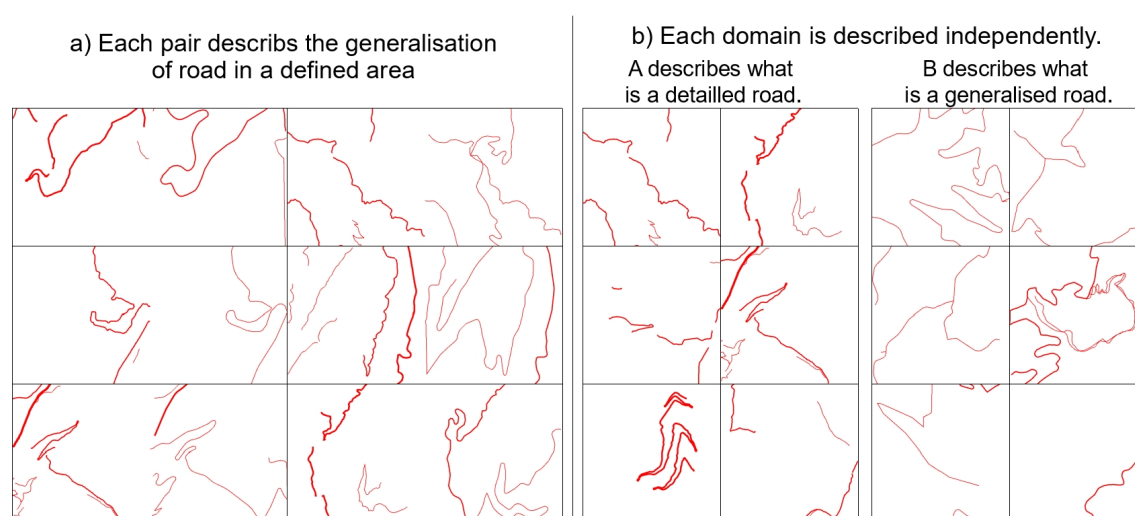


Figure C.8: Example of tiles from paired (a) and unpaired (b) datasets for roads generalisation.

Models for graph problems.

Graph neural networks (GNNs) are deep neural networks that are capable of dealing with graphs as input. Diverse approaches have been explored to extend deep learning techniques to graphs. However, the graph structure is less regular than an image, and the interdependence between objects is complex.

(Wu and Biljecki 2021) proposes a review of GNN and proposes the following classification:

- Recurrent graph neural networks (RecGNNs) learn a target node's representation by propagating neighbor information in an iterative way until a stable fixed point is reached.
- Convolutional graph neural networks (ConvGNNs) try to generalize the notion of convolution to a graph. Research in this field has two main approaches, spectral-based which proposes a matrix decomposition of the graph, and spatial-based which works on a local neighbor of the node.

- Graph autoencoders (GAEs) encode nodes/graphs into a latent vector space and reconstruct graph data from the encoded information.

For geographic information, the graph convolution network (GCN) is mainly used, it is a convolution graph neural network that defines a graph convolution using the node attribute matrix X and the adjacency matrix A (these notions have been presented earlier in this thesis, see Figure B.7). For a layer l , the convolution result is the matrix product of the adjacency matrix A with the previous layer and a weight matrix W (eq. C.10).

$$\begin{cases} H(0) = X \\ H(l+1) = \sigma(A.H(l).W(l)) \end{cases} \quad (\text{C.10})$$

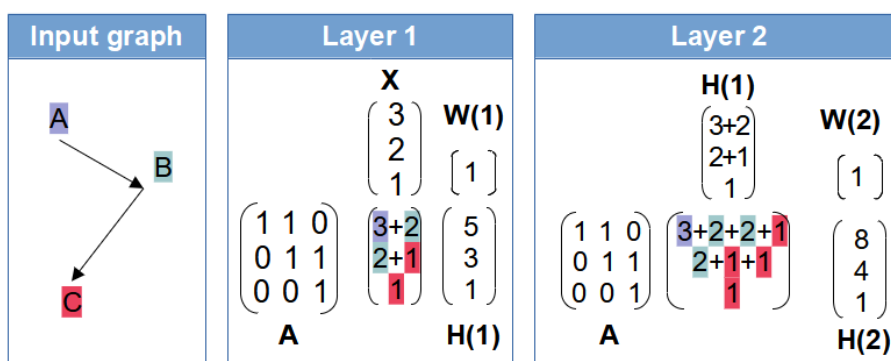


Figure C.9: Illustration of graph convolution with a simple example (for simplification the activation function is identity and not represented).

This definition allows the content of the graph to be encoded while propagating the value of the node through the edge of the graph (see Figure C.9). Indeed, the matrix product sums the attributes of neighboring nodes. Thus, this model learns the structure in neighboring, and the deeper the model is, the more neighbors are considered for each decision prediction on a node.

C.1.3 Choosing or designing deep neural network for map generalization?

The previous section has shown that there is a large diversity of models, and, at first sight, very little information allows us to determine whether a model is adapted for a defined use case. This section tries to provide guidelines for the choice of models. We explore the following approach: 1) choosing an existing model, and, eventually, 2) adapting it.

Re-using a model from the multitude of existing models

Once the geographic information representation and task have been chose (see Chapter B), it is possible to identify model for this task with this kind of data. Moreover, many generic models exist (i.e. designed to resolve the same task in diverse use cases). For instance, pix2pix (Isola, Zhu, Zhou, and Efros 2017) is designed for a generic image-to-image task and can be trained for many tasks. However, choosing

the best model for **our** task is not easy. The latest published model is not always the best. A model can be the best on one dataset and not adapted at all for another, even if the new task looks similar. For example, Pubmed and Cora are both benchmarks for graph node classification that include a graph of scientific publication linked by citation relations, and Figure C.10 shows that at one time, the best models differ for Pubmed and Cora benchmarks.

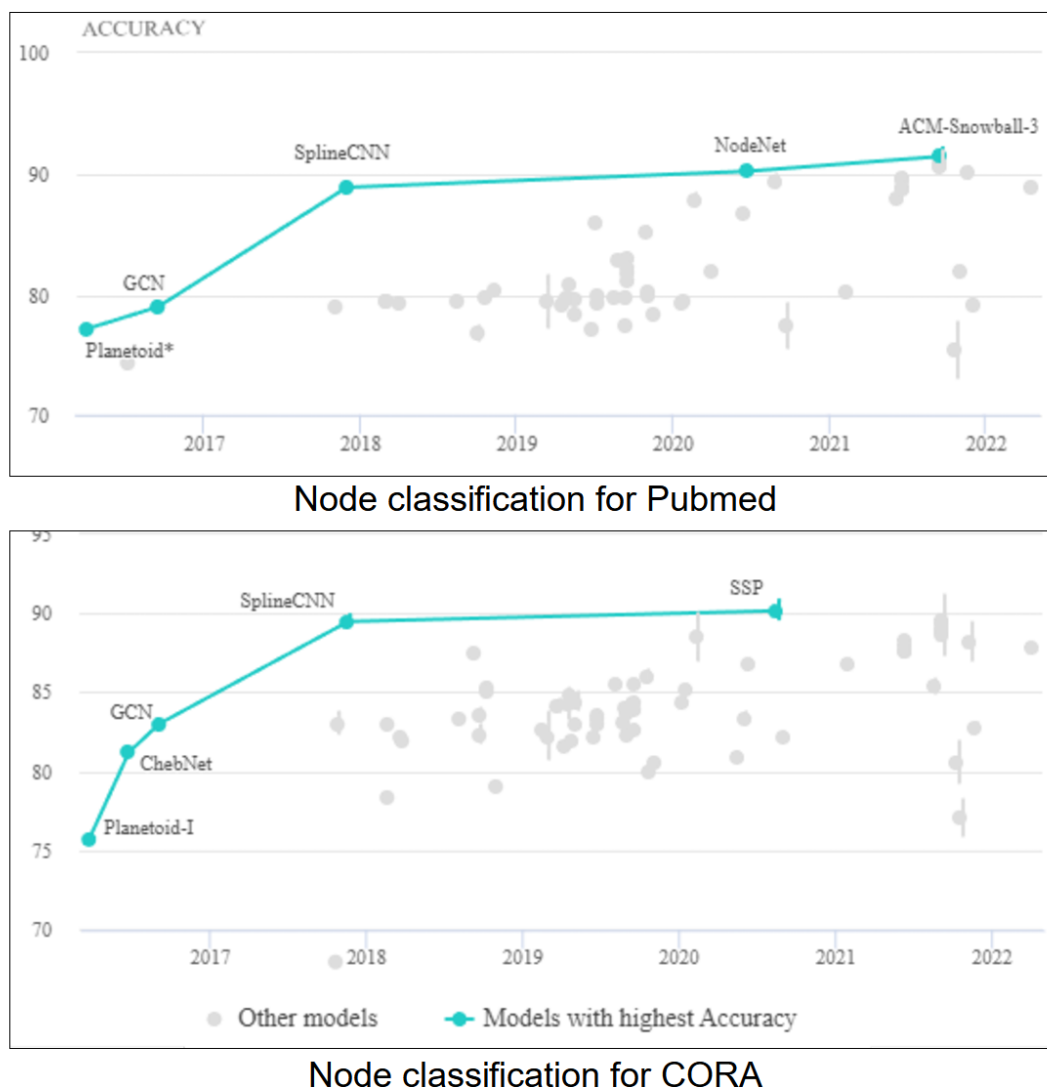


Figure C.10: Comparison of model accuracy over time for node classification on Cora and Pubmed datasets (from paperswithcode.com visited on 05/07/2022).

Finally, for this example, the models are compared using accuracy; however, evaluating the task is not always as simple as classification. For instance, the assessment of the image generation network is an issue, since it is difficult to define a distance to a good solution. Thus, the comparison of architecture is tricky, and it is impossible to determine a priori if an architecture will work for our use case (e.g. can an architecture capable of digit recognition learn scale recognition?).

Very few models have yet been applied to use cases similar to map generalization, so how do we make sure that we choose an adapted deep neural network? Our strategy during our exploration was based on two ideas:

1. **Analogies and common characteristics:** If a network has been successfully

applied to a use case with similar or common characteristics, it is most likely to be adapted. For instance, (Simo-Serra, Iizuka, and Ishikawa 2017) designed a network capable of learning the image-to-image conversion from sketch to drawing. This use case is very similar to mountain road generalization, as line smoothing, simplification, and typification are required. Furthermore, this model is robust and has good transferability due to the semi-supervised structure, so we think that this model could be adapted for the generalization of the mountain road shape (see Figure C.2).

2. **Ability criteria:** we determine a list of criteria to determine if a model is likely to be adapted. Our list of criteria includes abstraction ability, interpolation ability, preservation ability, transferability, spatial ability, variability, and gluttony, and is discussed in greater detail in the following paragraphs.

Abstraction, or interpolation ability Usually two main types of networks are adaptable for map generalization: those able to interpolate and abstract. On one hand, models for interpolation (e.g. colorization, super-resolution, terrain generation (Guérin, Digne, Galin, Peytavie, Wolf, Benes, and Martinez 2017), etc.) convey the idea that deep neural network can do data enrichment. On the other hand, map generalization is an abstraction task, and a deep learning experiment that makes data abstraction may be adaptable to the generation of a generalized map. For instance, Kang, Gao, and Roth (2019) mentioned during their experiments on map-style transfer, that some abstractions are learned by GAN models, and with a specific dataset, it would be able to achieve map generalization.

Preservation Another required aspect of a network is the ability to preserve input information. For example, in image generation, a model with the output images really different from the input ones would probably be un-adaptable for map generalization, where maximum information preservation is required. In map generalization, the stringency of desired preservation is varying according to the use case and the scale gap. Therefore, a network with an adaptable level of preservation could also be beneficial.

Transferability Deep learning is trained on a set of examples, its ability to be applied to other situations is called transferability³. This ability is really important in the case of geographic data, where we want to apply a model train in a geographic area to a large region sometime in another country. Moreover, application to other map scales and styles can be hopeful if the model has good transferability. This problem traditionnaly exists in other map generalization process (Touya 2011).

Spatial-ability Most models are not designed especially for spatial data. An encouraging characteristic would be that the model already solved spatial problems i.e., problem where the localization and geometry of the data have an important role in the problem solving. In the absence of such a test, it does not mean that the model cannot deal with spatial data, but the spatial ability cannot be evaluated. Then, some networks are designed especially for geographic data and include elements

³Generalization-ability can also be used to describe this notion but in this thesis, we prefer transfer-ability to avoid the confusion with map generalization.

to improve this ability. For example, Deng, Tian, and Newsam (2021) propose pooling methods that take into account the first law of geography to aggregate data from an image during convolution. They propose to replace the max-pooling in the convolution layer with G_i^* pooling adapted to geographic data. This pooling selection takes the interpolated value at the center location of the windows and thus depends on the spatial cluster (Figure C.11).

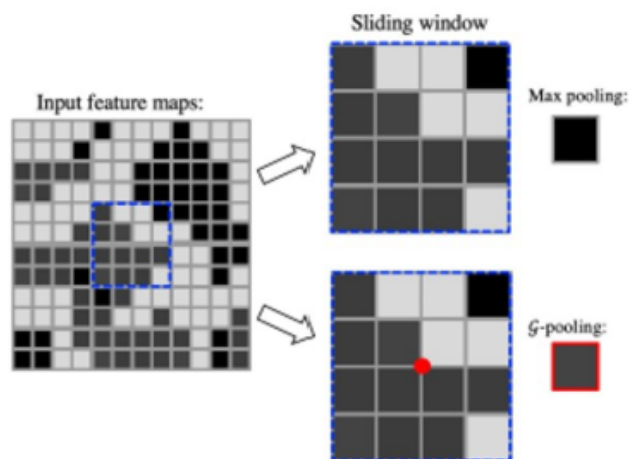


Figure C.11: Comparison of max pooling and G_i^* pooling create different given a feature map as an input (from Deng, Tian, and Newsam 2021).

Variability Some models are more or less deterministic (especially image generation models). An important variability is particularly adapted for the use case where several different predictions can be good and, in the end, the user choice is suitable. This can be the case for map generalization, e.g., in a perspective of generalization for on-demand mapping, users can prefer one variant of a map. However, this variability is also a limitation for model evaluation, and if the reconstruction/assembly of consecutive image tiles is required, the homogeneity of predictions can be required.

Gluttony On one side, models have different learning data requirements; some need thousands of examples to learn, while others need only hundreds. On the other hand, map generalization tasks also have different numbers of possible examples. For instance, the detection of ring roads, for simplifying their generalization cannot be performed using a model that requires a large amount of examples as such structures are rare in maps, and few-shot models may be most adapted (Potié, Mackaness, and Touya 2022). However, for map generation, numerous maps exist, and even a gluttony model can be adapted. Therefore, the gluttony of the chosen model must be adapted to the generalization task.

Adapting a model

In the previous section, we describe methods to ensure that a chosen network is adapted for map generalization. However, in reality, it is not probable to find a model that meets all our criteria and has no specific limitation. Thus, the development of an improved or adapted network may be required. In this section, we explain how

we can adapt a model by successive modifications of an existing model. In particular we can propose adaptation of the following elements:

- **Settings** The first adaptation for a model is to manage its settings (choosing an initialization, optimization methods, and a number of epochs). These adaptations may affect the convergence of the model.
- **Neurons** The type of the neuron have an important role in the efficiency of the network, indeed it is the manner the input is transformed into output in each layer. For instance, the introduction of convolutions and the reduction of kernel size have greatly improved the deep learning potential to deal with images. We can apply some improvements, such as the pooling proposed by (Deng, Tian, and Newsam 2021) to the chosen model. Finally, a different choice of activation function can be made and allow a more efficient passing of information.
- **Structure** The depth (number of layers) and the number of neurons in each layer must be adapted to the desired use case: the more abstraction required, the more layers would be needed. The denser the information to encode, the more neurons would be needed.
- **Combine models** The combination of several elements from different models is the simplest way to take advantage of models with different abilities. For example with an adversarial approach or by including other network components or modules such as fusion (Hazırbaşı, Ma, Domokos, and Cremers 2016) that encode and fuse separated information, skip connections (Ronneberger, Fischer, and Brox 2015) that link the encoding and decoding path, etc., can be used to create links between layers and create a more adapted model.
- **Loss** The change of the objective function for a more adapted loss function could finally improve the model (see C.1.4)

In Part II we describe experiments that include the choice, comparison, and some adaptation of deep neural networks for map generalization. In particular, in Chapters E and F, we provide a comparison of some models for generalized shape prediction and generalized map generation. Chapter E includes a study of the most adapted setting. Then, Chapter G proposes an adapted model for generalized map generalization, constructed by the combination and adaptation of neurons from existing models. However, the goal of these experiments is not to design the most adapted model, but rather to show the potential of an approach, and the time dedicated to each experiment was limited; so we did not experiment with all adaptations mentioned here.

C.1.4 How to choose a relevant learning objective for map generalization

The loss function is a function that summarizes a set of variables (here the prediction) into a real value that represents the error of this set, it aims to guide learning and is the function optimized during the training. The choice of a loss function depends mainly on the architecture and the dataset. It can be a combination of

several measures (e.g., GAN loss is the sum of the adversarial loss and the distance L_1). To demonstrate the interest of such custom learning objectives for a deep neural network, an ablation study, that compare model with and without the term is necessary. In this section, we present some loss terms for image based approaches that can be helpful to map generalization (formulation from B.1.1).

Adapted distances

First, we have seen that the comparison of images at pixel level (like in L_1 , L_2 and Dice measure, see Figure C.6) is often not adapted. This paragraph presents some losses that are not based on the pixel value, but rather compare the target and the prediction globally. A such comparison would encourage the credibility of the prediction.

Style losses In general, style losses from style transfer experiments are designed to deal with high-frequency similarity and are more appropriate for map generalization than local losses. For instance, (Ganguli, Garzon, and Glaser 2019) propose a style loss for map images that compares the gram matrix of the prediction and the target. The gram matrix is obtained by multiplying a matrix by its transpose, this matrix encodes the aspect of the image and "hide" its content. The distance is a mean square difference, but since it is computed on the gram matrix and not on the image itself, it allows us to compare the aspect of the map independently from the content concordance. This style loss has shown the potential to understand and apply a map style to geographic information. As map generalization is not only a map style problem, some may state that the interest of the loss is limited. However, it does not measure the difference in the map style, but rather the difference in the image style. And this measure includes some characteristics of scale adaptation (e.g. the level of granularity or smoothing of the content is a part of the style loss and more a scale adaptation than map style elements).

Patch-based loss Moreover, some losses can be based on a distance calculation at the patch level. Patches are small portions of pixels from an image (e.g. a windows of 24 pixels). Isola, Zhu, Zhou, and Efros (2017) propose a discriminator for their GAN (patchGAN) that is computed at the patch level instead of predicting similarity at the pixel level. A second example is the spatial correlative loss (Zheng, Cham, and Cai 2021), based on patches. It computes for each patch the relation of each pixel with the center point of the patch and compares these values for each center point of the prediction and target. This comparison allows us to compare the position of object on map and thus spatial structure of the image. In the case of image generation, it leads to better preservation of the global shape of the presented object. In the case of a map, it can be efficient to evaluate if the empty and filled spaces are similar. It may be adapted for map images with a main structure that does not varies much, such as thematic maps or large-scale maps. However, it is not relevant when the contour of the object in target and prediction vary a lot, as in our map generalization experiments during our exploration.

Distance loss Another proposed approach to encourage GANs convergence is to compare the distance between two images in the training set (Benaim and Wolf 2017). This distance must be similar before and after the transformation. It has

the advantage of giving an indication of the credibility of the prediction in an unsupervised way and without relying on reconstruction. However, map generalization is not a homogeneous transformation, the generalization of some part of a map may involve a big difference while in other part the input and generalization are really close. Thus, this loss does not appear to be adapted for most map generalization tasks.

Preservation measures

The loss function can also be used to measure information preservation. The principle is to compare the input X and the prediction \hat{Y} ; the more similar they are, the more the input information is preserved. In traditional map generalization, information preservation is an important constraint and we expect this constraint to be met even with deep learning based map generalization.

Content loss In style transfer, the content loss aims at preserving the image content, and in GANs, the identity loss computes a distance between the input and prediction. In the case of map generalization, such a measure is important as the main information has to be preserved. The choice of the similarity measure for this term has similar issues with the comparison between prediction and target, as we do not want to learn a pixel similarity, but rather a global similarity: the main geographic information has to be preserved. Therefore, a patch-level or structure-level measure would be preferred to pixel comparison.

Connectivity loss We proposed a connectivity loss in our experiment on mountain road segmentation. It compares the connectivity of the road network in the input and the predicted image. This connectivity is approximated by the number of pixels on the road and in the background. In fact, these quantities are affected by alterations in the structure of the road network (Figure C.12).

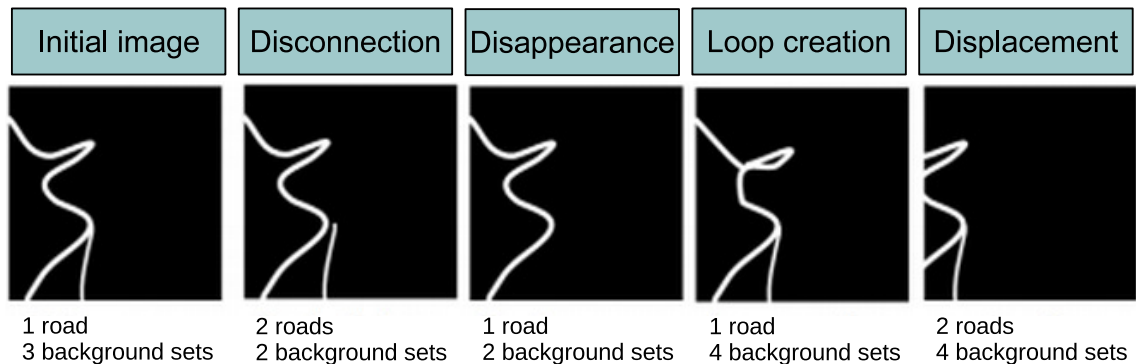


Figure C.12: Example of alteration of the connectivity of the road network and number of set of background and roads.

This loss aims to penalize changes in the connectivity of an image. The formula is given in (eq. C.11): it sums up the absolute difference between the number of sets of pixels of objects of interest (road) and the number of sets of background (not

road). A detailed comparison of the interest in this loss is given in Chapter E.

$$L_{connectivity}(X, \hat{Y}) = E_X \frac{|nback(X) - nback(\hat{Y})| + |nroad(X) - nroad(\hat{Y})|}{nback(X) + nroad(X)} \quad (\text{C.11})$$

Topology loss The topology loss from (Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020) is based on the gradient of the image (Figure C.13). Instead of calculating the distance, the Pearson correlation is computed between the gradient image, this correlation measures to what extent the value of the gradient in the input determines that in the output. Thus, this measure is interesting for comparison of salient areas in the image, such as the urban block that must remain the same before and after generalization. We tested this measure for the generalization of building shape but it is not relevant, as the gradient image of two generalized building groups is too different due to displacement and simplification.

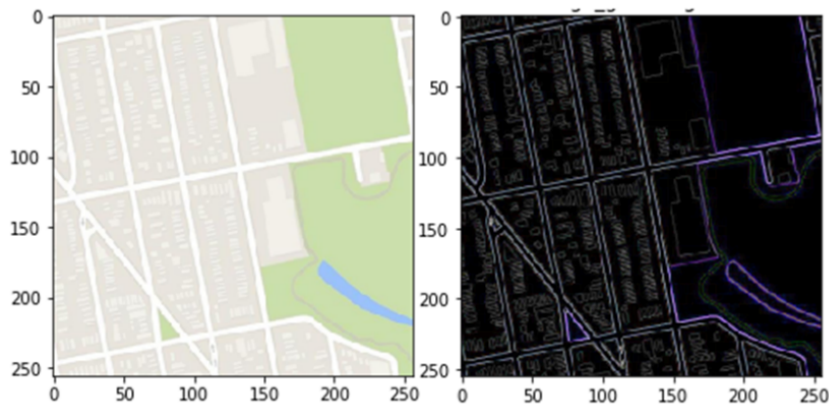


Figure C.13: A 256×256 Google map tile and its 255×255 gradient map (from Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020).

Map characteristics-based losses

We can imagine an objective that evaluates whether the image is a map (and even a good map) or not. The goal is not to provide a unique function for map generalization evaluation (if we could, deep learning would not be useful, and map generalization automation no longer remains a challenge). The goal is to encourage the prediction to follow of desired cartographic rules.

For example, since cartography is a transformation insensitive to rotation, (Fu, Gong, Wang, Batmanghelich, Zhang, and Tao 2019) proposes a loss that encourages this characteristic by comparing the prediction of rotation of an image with the rotation of the prediction. This loss has shown interest in the generation of maps in the style of Google Maps. Interest in this measure for the generation of generalized topographic maps is tested in Chapter F.

Moreover, in map generalization, it is quite easy to identify places that require a fixed generalization and places where the generalization varies more. Thus, we could use a spatially weighted loss. Indeed, in Chapter F we propose employing such a mask for building generalization: an error in this area is more important and "costs more" than in other areas.

C.2 Designing a good training dataset

Deep learning models learn from examples. Examples are organized in a dataset that contains at least examples of the input and target data. The creation of an efficient training set is a critical point in learning success. Even the most powerful model could be totally useless if it is trained with non-relevant data. In this section, we present what an adapted training set is and how to create it.

C.2.1 What is an efficient dataset?

During the training phase, several batches of examples are used to optimize the weight of the models. During this phase, some unseen examples can be required to verify the model progresses and is not over-fitting. The validation set is a subset of the dataset not used to train the model; however, the model is applied to each example in this set after each epoch and the evaluation is calculated. It is used to determine the number of epochs required during training. Finally, during the testing phase, another set of examples (test set) is used to fairly evaluate and compare models. Thus, the dataset is generally divided into three image sets: the training, validation and testing sets (Figure C.14). This separation is not strictly regulated and must be adapted to the complexity of the model (the more parameters, the more validations are needed) and the total size of the dataset. Finally, each set must at least cover most of the situation expected. For example, in the case of map generation, the train, validation and test sets must contain images of maps with different land uses (urban **and** suburban, **and** rural map image, etc.) and different landscapes (mountain, forest area, coastal area, etc.) since all these situations involve different cartographic knowledge.

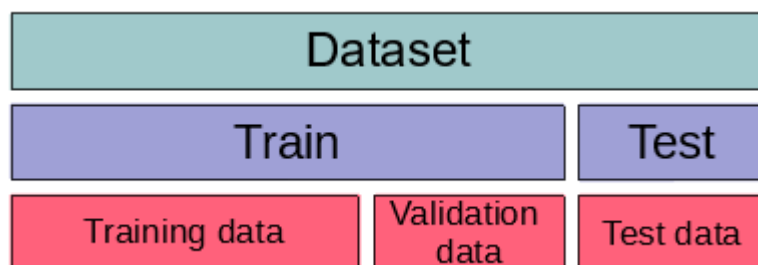


Figure C.14: The partition of the dataset for deep learning.

Characteristics of an efficient dataset

In machine learning, the relevance of a dataset is based on the quality, quantity, and diversity of the examples.

Quantity The number of examples in deep learning is often a critical point. There are several methods to artificially increase the number of examples (e.g., image augmentation see Figure C.15). Image augmentation consists in adding new examples derivated from real ones by simple geometric or color transformation. However, this augmentation is not always relevant for map generalization. For instance, the cropping and scaling changes the scale of some of the images, and in map generalization,

the objective is to learn scale relativ characteristics. Similarly, all color augmentations are only relevant for natural images like photographs, but not for maps, where each color have a particulare meaning.

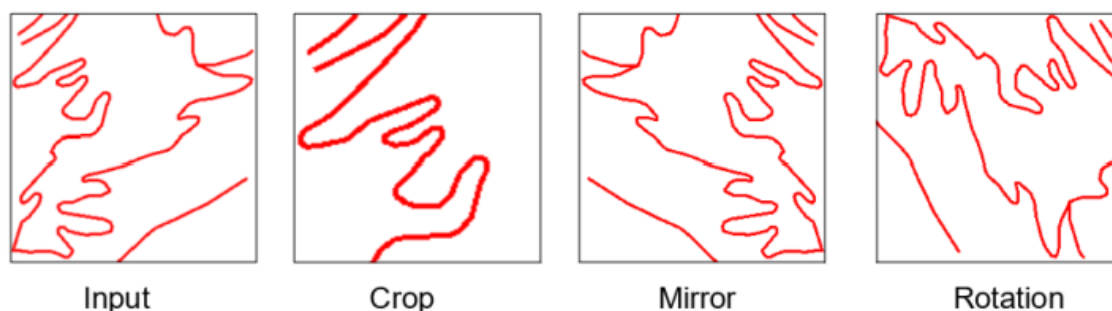


Figure C.15: Illustration of the most common image augmentation in deep learning applied to an image of a mountain road.

Diversity If a training set includes only similar examples (e.g., only a sleeping cat or a map of an urban area), the model will not be able to be applied to other situations (e.g., a walking cat or a rural map). However, it is not possible and useless to constitute a dataset that includes all possible individual situations⁴. Furthermore, the more similar the situations in the training set are, the simpler the model would be and the faster it would converge. Thus, it is necessary to find a balance in which the dataset is representative of most of the situations we expect the model to be able to deal with and gives enough examples of each situation to allow learning. Finally, the dataset should avoid being really unbalanced, unless the model may apply the knowledge learned in a common situation to other situations. For example, if our dataset includes mostly smooth roads and few sinuous roads, the model may not be able to generalize a series of bends. To measure the diversity of images, we can compare statistics on images (e.g., clutter); however, it is often not sufficient to describe the diversity of represented situations. This diversity depends on the object and the task. Indeed, evaluating the diversity of an image set of cats is really different from evaluating the diversity of map images. For cat diversity, we can study the number of different scales, cat positions, or cat breeds. For a map, we consider rather a diversity of landscape, landuse, map style, map scale, conflict and map content.

Quality If an example contains errors, the model could learn to reproduce these errors. Moreover, a pattern must exist to be learned: If there is no consistency between the input attributes and the output in the example, learning is impossible⁵. This means that to be qualitative, the input must contain the necessary information to learn the task (map generalization). For example, the nature of the road (highway, national, regional or local) is required for road selection. Many useless or not relevant information may confuse the learning and hide the real target knowledge.

⁴If all individual situations are in the training dataset, the model has no more situation to be applied on.

⁵In the case of image-based deep learning the attribute are pixel values in Red, Blue, and Green (RGB) channel.

For instance, if the image includes rivers but they are not used for road generalization and not generalized themselves, this may make the learning of road shape generalization complicated. Finally, the information must be comprehensive. For instance, images with too small resolution are not comprehensive, and the shape of the un-generalized road is coalescent, so the model cannot grasp it and generalize it. Thus, the quality evaluation of a dataset cannot be summarized as the evaluation of the quality of the target generalization; it must evaluate the comprehensiveness, relevance, and consistency of the dataset. For example, a dataset composed of generalization example performed with different constraint and operators may miss of consistency and the model may be unable to learn any rules.

Dataset creation

A deep learning dataset must be an example set that represents the target task (Bengio, Courville, and Vincent 2013). The main steps for the creation of the training set are the following.

1. **Location** We need to choose a study area (where the model takes examples) and an area of interest (where the model should work). This choice is important because it defines the scope of the model. However, this is often not a real choice, as the area in which we have various, qualitative, and annotated data can be limited.
2. **Content** We have to determine the content necessary for the generalization of the map. The deep learning approach can allow us to understand the implicit knowledge in the data; however, we have to make sure that this information is in the data (and if not, perform the necessary data enrichment). On the contrary, unnecessary content must be avoided as much as possible to simplify learning. For example, for road shape generalization at small scales the model must contain the shape, location and importance of roads but not the shape of rivers (in case where they are generalized independantly).
3. **Representation** We need to choose a representation for input information that can be used by a deep learning model (see Chapter B). The correct representation of vector information in other forms is a challenge. For example, objects in the vector database have relations, and the image representation must show this relation. However, the raster cannot encode relations, especially when objects overlap. Humans are able to see the whole even if a part is hidden due to Gestalt principles of visual perception, which may not be captured by current deep learning models. Thus, if a road crosses a river through a bridge, a human will see the river as a line, while the pixel at the cross is not a river at all, and the image represents two disjoint parts of rivers. This loss of information can be misleading, and the model will learn that a river can be discontinuous.
4. **Splitting** We have to separates the dataset into examples (e.g., image tile, sub-graph, etc.). The study area is divided into several examples. This separation or tiling is an issue. In fact, each tile must have the relevant level of detail to allow for generalization. The question of the relevant level of analysis and context for a generalization task is even common with vector approaches (Ruas 1999). How to separate the study area into relevant tiles? For example, to

generalize a road, it is more practical to see the entire stroke (Thomson and Richardson 1999).

5. **Refinement** Refining the training set to improve quality, quantity, or diversity. This refinement step can include filtering to remove some un-relevant examples or correct an unbalanced dataset, and, on the contrary, data augmentation can be applied to increase the number of examples in some situations.

In the next section, we present and discuss some examples that apply these steps to create learning dataset.

C.2.2 An example of an image dataset for shape generalization

In this section, we will detail an experiment on the creation of an image set for the use case of mountain road shape segmentation at 1:250,000 scale.

Area of interest

For this first experiment, we mainly used data covering a small portion of the French Alps (2155 km^2). The input is an extract from the database used to create topographic maps at 1:25,000 scale at IGN. The generalized dataset is an extract of the database used to create a map at 1:250,000 scale at IGN. It only represents important roads, and their shapes are simplified, smoothed, and often schematized for better visualization of the sinuous bends. It was derived from the detailed one many years ago using generalization algorithms with the GALBE process (Mustiere 1998), and then manual updated.

We also include examples from Corsica, where mountain roads are more uniform and thinner (there are fewer highways and national roads and more local roads). These examples were used to test the impact of increasing the number of training data, and the effect of the chosen study area. Our first test shows that the dataset in the Alps is more adapted than the one in Corsica due to the less diverse situations and the thinner roads that make learning difficult⁶. Finally, we constructed a similar small dataset in another important mountain area in France to test the transferability of our model (this experiment is detailed in Chapter H). It includes roads in an other part of the Alps, in Corsica, in Jura, and in the Pyrenees.

Tiling

The study area is separate in tiles. We propose two tiling methods (presented in Figure C.16): tiling using a regular window that slides into the study area or using a window centered on each object. To complete the overview of possible tiling methods we can mention Du, Wu, Xing, Gong, and Yu (2021), who proposes a third tiling method illustrated in Figure C.17.

For the window approach, we choose an image resolution (number of pixels), a cartographic resolution (ground size represented by each pixel), and a step for window displacement between two tiles. The image resolution influences the image legibility and the calculation time, as more pixel have to be processed. The cartographic resolution also impacts the legibility but also the size of the ground in the

⁶We compared datasets of the same size and created with similar processes.



Figure C.16: The two proposed tiling methods. Left: the sliding windows. Right: the object's shell.

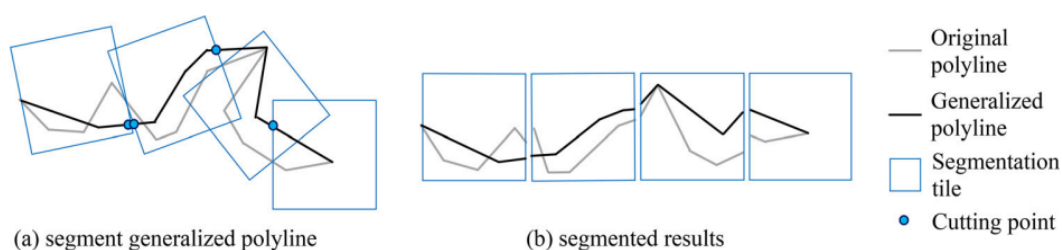


Figure C.17: (Du, Wu, Xing, Gong, and Yu 2021) tiling methods.

image, and so the level of analysis: with small pixels, the image represents a small portion of terrain really precisely, so it is able to learn a thin local transformation but not a global generalization (see Figure C.18). Finally, the step mainly changes the number and diversity of the examples.

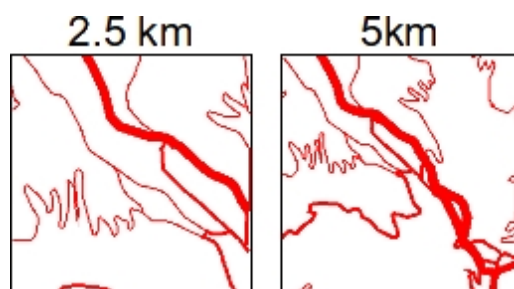


Figure C.18: Example of two tiles with different cartographic resolution.

The object method centers the tile on one road object, the maximum image size is defined, and if the object is larger than this size, it is resized to be completely encased in one tile. This method ensures that the road is completely visible during its shape generalization. However, the rescaling is confusing because we want to learn tasks that are scale-related. This method requires us to define an image resolution and an object size at which a re-scaling is needed.

For the last method, the tiles follow the axis of the line. This method allow simpler tiles, however, it cannot be applied to our use case where the road line is part of a complex network and the junction of the objects is as important as the central shape.

To choose the relevant tiling methods and parameters, we compare several configurations, presented in Table C.1. This quick test allows us to choose tiles created using the windows method with a cartographic resolution of 2.5 km.

Table C.1: Comparison of several methods and parameters for the creation of a mountain road shape generalization dataset (+: good, -: bad, -: very bad).

Method	Quantity	Dice	IOU	Visu
Window 2.5 km	560	0,2	0,3	+
Object 2.5 km	687	0,05	0,1	-
Window 1 km	2128	0.05	0.1	-
Windows 5 km	182	0,3	0,4	-

Content and representation

We choose to include information on the position, shape, and nature of the roads. The nature of the road indicates the future symbolization on the map and impacts the level of generalization required: the more important road will have a larger symbol and require more important generalization. Table C.2 defines the number of pixels in the image for each type of road. The proposed width of the road allows the shape of the road to be large enough to learn the continuity and to be informative about the importance of the road; however, it is not small enough to avoid all the coalescence of the shapes, and the shape of the input road is sometimes not legible. Nevertheless, our first experiment with several scales of symbol (from 1 to 1, from 1 to 3 pixel, and from 1 to 5 pixel) shows this symbolization to be the best compromise.

Table C.2: Conversion table road nature and width on map-like image.

Width in pixel	Symbolization at 1:25 000	Attribute values translation at 1:250 000
1	5	Irrelevant, forbidden, local narrow roads
2	4	Regional roads and narrow regional roads
3	3	Regional roads with bike path
4	∅	Major roads
5	2	Highway

We first tested learning with a dataset that includes all roads in both the input and output; however, as the scale gap is important, the selection in output is important, and our dataset does not include the necessary information for this selection. Moreover, our goal in this experiment was to learn shape generalization (displacement, bend simplification, schematizing, enlargement, etc.). So, such a dataset is irrelevant, the input and output are really different, which is confusing and unnecessary. Then we pre-process the input data to include only the road kept after generalization in the input image. This pre-process is done using a multi-criteria data matching algorithm (Olteanu-Raimond, Mustière, and Ruas 2015) and refined manually to match both scales. Figure C.19 illustrates the interest in the pre-process.

Our idea was to first compare the training with images of all roads in our area and only the sinuous road, to show the interest of a more restricted dataset for a

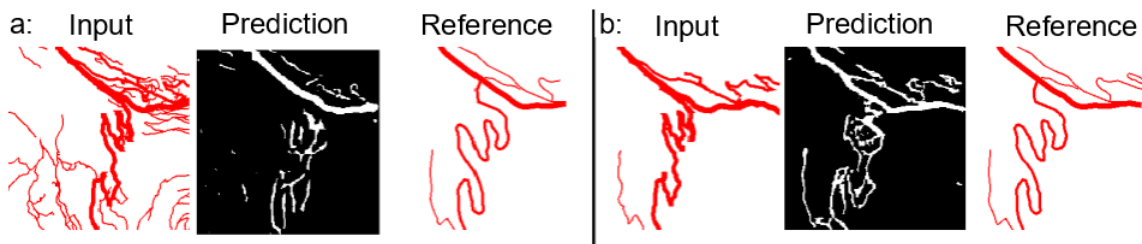


Figure C.19: Comparison of prediction a: with and b: without selection.

really specific use case. However, our dataset does not include enough roads for this experiment. Finally, the shape of the road also depends on its environment, especially topography and hydrography. We also trained a model with a variation of this training set that includes the rivers' shapes. This experiment is discussed in Chapter E.

Refinement

We tested several refinement methods, especially the filtering of training set to remove irrelevant examples, and several data augmentation methods. Table C.3 compares the different proposed datasets. The quantity column gives the number of examples in the dataset. The efficiency of a dataset is described via a visual evaluation and quality measures including the intersection over union and Dice (for both these measures a greater value indicates a greater quality of prediction).

Table C.3: Comparison of the quality of several datasets for the generalization of the shape of mountain roads (+: good, -: bad, -: very bad).

Method	Quantity	Dice	IOU	Visu
Win 2.5 km	560	0,2	0,3	+
Win 2.5km filtered	284	0,1	0,2	-
Win 2.5 km rotation	1120	0.2	0.4	+
Win 2.5 km crop	1120	0.3	0.4	-
Win 2.5 km without selection	560	0,1	0,2	-

We observe that our dataset is not large enough to filter out some of the examples. Moreover, increasing the data does not seem to improve the quality of the prediction. Indeed, such increases do not improve the diversity of examples in the dataset, which is the main lack of our dataset.

C.2.3 An example of an image dataset for map generation

In this section, we will detail some options tested for the creation of an image set for our use case of urban map generation at the 1:50k scale.

Area of interest

In the case of the map generation experiment, we used an area of 30 * 15 km east of Saint-Jean-de-Luz in the south-west of France. We have an intermediate

generalization in this area from the work of (Touya and Dumont 2017). This area includes a few urban areas and mainly peri-urban and rural situations, on the land and on the coast.

Tiling

We split our area using a sliding window. Each tile represents $500 \times 500 \text{m}^2$ of ground and has a size of 512×512 pixels. This resolution allows one to avoid the most important object overlapping in the input image while showing a large portion of the terrain of a few small urban blocks, which is the minimum necessary context to perform building density reduction, in our opinion. This tiling is not related to the data; thus, there is no guarantee that it is relevant, especially, since each tile can suffer from the border effect. For instance, the limits of the tile can hide important information (the end of a dead end, the dense part of a dense block, etc.).

For this use case, we did not experiment with different tiling methods. However, a more flexible method could be possible, for example, based on Huang, Zhao, and Song (2018) methods to produce tiles completely included in an urban block. This tiling method use the point of the skeleton of map block as center for processing unit (images), it was proposed to improve the land use classification of these blocks. To apply such a method, regular blocks are needed, which is not the case for our dataset.

Content and representation

We choose to represent water, roads, and buildings, as they are the main themes in topographic maps at this scale. The input image is made from data from the BD TOPO®, the vectorial database of french spatial infrastructure produced by IGN and use for topographic map with scales between 1:2 000 and 1:50 000. The target is a generalized map obtained using semi-automatic generalization on the input dataset for cartography at a 1:50 000 scale (Touya and Dumont 2017) completed with som manual corrections. Both vector datasets are symbolized in a topographic map style and rasterized.

We represent roads with an outline in input and target representations. The outline enlarges the road symbol and makes learning simpler (because thin linear object are traditionnally hard to preserved). Moreover, this representation allows us to verify that our model is capable of understanding and producing complex symbolizations. During generalization, the road network is selected. Small and local roads are removed. The selection requires a deep understanding of the context of the map image and object attributes. To test whether just the image is enough to learn this change, we constructed a version of the dataset with and without this selection.

The buildings are symbolized in brown. We test both with and without outline; the outline allows us to show the limits and real shape of the building and so avoid losing this information in case of overlapping; however, it enlarges the shape of the objects and so increases the risk of overlapping; moreover, it adds information in an already dense area and so brings confusion. A quick test shows that the symbolization of the building outline is not relevant. We tested and compared the dataset constructed from data generalized using both AGENT method (Barrault et al. 2001; Touya and Dumont 2017) and Typify method (Burghardt and Cecconi

2007). This comparison has shown that the model can also learn to reproduce each generalization method (a comparison of the prediction for both methods is presented in Chapter F).

The output vector dataset also includes graying: some dense areas of the city center are generalized using a gray area instead of buildings. This symbolization has some defaults: the color used in the topographic map for the center line of the road and the graying is close; however, for the model, close color means close value, which means related objects. This proximity of colors leads to confusion during learning and a more distinct symbolization may be required.

The hydrography includes surface polygons of water and river polylines symbolized in blue with fixed width and without an outline. The generalization of this theme is less important: some lines are smoothed, and small water surfaces are removed.

C.2.4 A proposal for a better representation of geo-information in images

Both this use case and the literature review have shown that the creation of an image dataset is a problem, even if the representation of geographic information on a map-like image seems natural. In this section, we present our proposal for the representation of geographic information in tensors to form an adapted training set. This model is applied to the use case of topographic map generation, and its interest has been demonstrated for this use case. However, we made it generic and easy to apply, and we think it could be as beneficial for other use cases.

Problems of naive solutions

The naive solution to create a map generalization image dataset is to represent in the input an image of detailed geographic information and in the output an image representation of the generalized geographic information, with the same input and output size (as we do in C.2.3).

Figure C.20 illustrates some naive solutions for learning cartography tasks based on images. The examples in line 1 are created by tiling a map-like image (Google map for the target and Google aerial image for the input). However, the chosen resolution is not sufficient to make the example set comprehensive, so the lines and thin objects are not legible and imprecise in the prediction. those at line 2. are our naive dataset for map generalization (presented earlier). This dataset reveals another drawback on quality: the target generalization contains many errors, the representation is also in-adapted as much overlapping produces a loss of information and essential information is missing (e.g., the road importance for road selection, or the city center limits for block graying). In line 3, the input is OSM data presented in a raw style (minimal symbolization); the output is a GoogleMap tiling. The representation includes comprehensiveness problems. The resolution is not sufficient; different objects are too close and too similar, and the model does not understand how to differentiate them and eliminate them. In fact, this lack of comprehensiveness is due to **symbolization problem** (building outline and minor road have the same symbol and are confused), **consistency problems** (some important roads are symbolized with two lines and the other with one, the labels are not in the input,

so cannot be produced in the output), and **relevance problem** (the input contains un-relevant confusing objects).

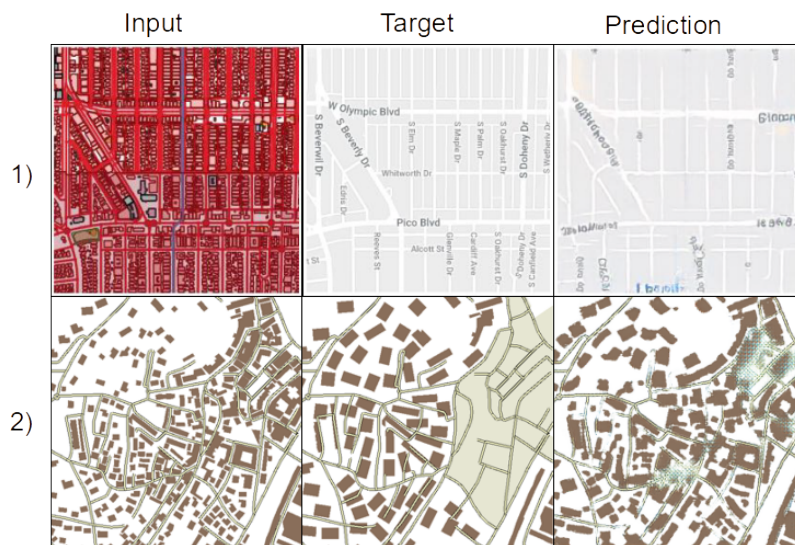


Figure C.20: Some examples of naive geographic information representation for cartography learning, from 1) Kang, Gao, and Roth 2019, 2) our experiment.

A hierarchy in information representation

We propose a representation of vector geographic information for image-based map generalization based on two tensors: one representing the main geographic information (i.e., object of interest shape and position) and one representing all additional information necessary to map generalization. This hierarchy is an original proposal for map generation, inspired by the work of (Hazırbaş, Ma, Domokos, and Cremers 2016) about combining scene photograph and depth map to learn indoor scene segmentation. In this study, the photograph is the main information; the segmentation is constructed from this information, it contains information about object position and shape; however, the depth map gives additional information that improves the segmentation. The first tensor, rather than a simple rasterization of the symbolized data, is the stack of masks of position and shape of geographical. Then an adapted model allows us to efficiently combine both information. This separation is also relevant for the perspective of map generalization, where the object shape and location are the base information for map generalization (the main goal is to avoid conflicts and preserve content), and then additional information serves to precise the generalization process. The two tensors are organized so that units in the same position in each layer describe the same part of the terrain (Figure C.21).

Main information representation

We propose to represent the main information provided by layer rather than symbolized in RGB. A binary mask of the shape and location of the object is provided for each layer of the input map object. The masks are stacked in the input tensor. This representation of the main information no longer requires symbolization and is free of symbolization bias. Moreover, the loss of information in the case of an overlap there is limited, and similar objects are in the same layer, so groups are

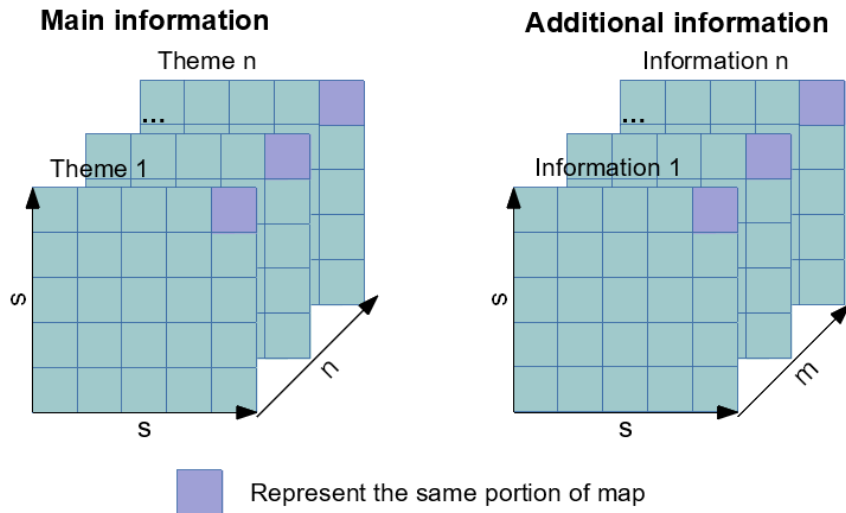


Figure C.21: Organization of the geographic in tensor for map generalization.

easier to identify. Figure C.22 illustrates this representation for our use case of map generation. However, the number of themes stacked is not limited to three. One of the inconvenient features of this representation is that the visualization for humans is less convenient than the symbolized representation.

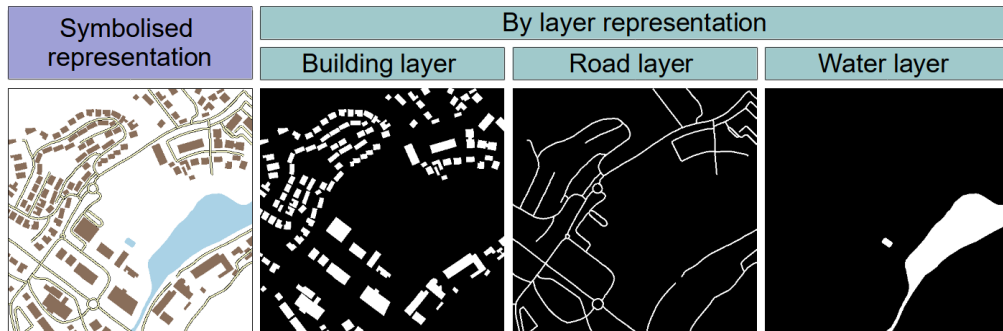


Figure C.22: Example of symbolized and layered representation of geographic information.

Enrichment tensor representation

The main information alone is often not sufficient to generalize the map. Map generalization, often require data enrichment prior to the generalization, we think such information is also needed in deep learning based map generalization. Indeed, even if deep learning models are able to use implicte knowledge in the data, the data representation (especially as image) often hide some information. Mostly, semantic attributes and context are needed and may be missing. We propose to create another image stack that contains the information for data enrichment. The representation of additional information varies depending on the nature of the variable we want to represent.

Binary variable For a binary variable (e.g., the block is in the city center or not, the road section belong to a round about or not, the road is a dual carriage way

or not, etc.), a simple mask with the presence (true) value is a good representation (see Figure C.23).



Figure C.23: Example of representation of the binary information location of city center with a black and white mask.

Quantity variable or ordered category For a quantity variable (e.g., sinuosity measure, density measure, etc.) or ordered categories (e.g., Horton order, ranking of city population, etc.), a simple pixel mask in the entity is a good representation (see Figure C.24). We normalize the value to be between 0 and 255, then the normalized value is the level of gray for representation of the enrichment tensor.

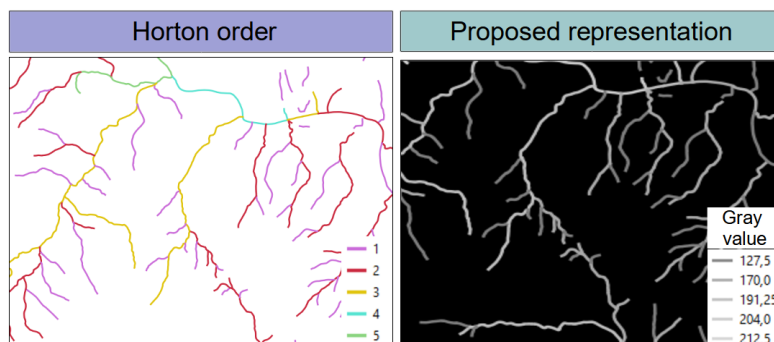


Figure C.24: Example of representation for an ordered categorical variable.

Un-ordered categories For unordered categories (e.g. building type), one mask per category is required; this solution avoids the symbolization problem of the colored image representation and the false order problem of the grayscale representation (see Figure C.25).

The improvement brought about by this proposal and several additional information is detailed in Chapter F.

C.2.5 Two example graph datasets for learning spatial relations

The interest in choosing the relevant representation of spatial information within a graph for deep learning has already been demonstrated (Iddianozie and McArdle 2021). In this section, we present how we design a graph dataset to test the ability of

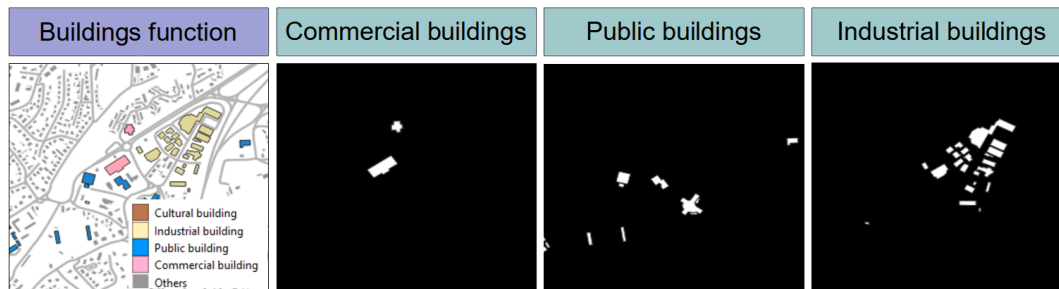


Figure C.25: Example of representation for a un-ordered categorical variable.

a graph neural network to deal with spatial relations, and we especially constructed a dataset for road selection and one for alignment detection.

Alignment detection

For the alignment detection dataset, we extracted almost 10,000 buildings from OpenStreetMap in urban and suburban residential areas of the United States of America, where buildings are organized in a regular way⁷. We constructed a relation graph that links the building in each urban block. We manually annotated the links as part of a curvy-linear alignment, part of a straight alignment, or not part of any alignment (Zhang, Ai, Stoter, Kraak, and Molenaar 2013).

Representation as relation-graph In map generalization, proximity graphs are often used to analyze urban structures (Regnauld 2005). The structure can be constructed from the center point of the objects or from another set of points (e.g., vertex). We choose to use center points because they involve a smaller amount of information redundancy and simpler structures. It also simplifies the reconstruction of alignment after learning.

The chosen graph structure must contain all necessary relations for alignment detection but exclude inconsistent links that confuse the learning. Moreover, the graph must be easy to split into examples, so a big graph of all relations is not relevant) and we chose to make one graph by city block. A such separation is relevant in urban areas but not in suburban and rural areas, where block can be really large. The construction of a proximity graph constrained by the road network is also a solution to the more relevant relations in the examples.

We examine variations of Delaunay triangulation (DT) (Figure C.26). DT is a triangulation that links objects in adjacent cells of a Voronoi diagram; it gives a more complete view of the proximity between buildings, but not all edges of this graph are necessary to identify alignment structures. Relative neighborhood graphs (RNG) connect the nodes that are at least as close to each other as they are to any other points (Toussaint 1980). However, some edges that are part of the alignment are often missing from the RNG. We consequently propose a proximity graph that fixes this problem based on RNG and refined by adding the edges of DT that are short enough, following the criterion proposed by (Bader, Barrault, and Weibel 2005) to enrich the graph by important edges. This criterion is based on the comparison of distance between building and cost to travel between them through the graph.

⁷We first experimented with datasets in France but the very low occurrence of clear alignments in the landscape makes the dataset too unbalanced and learning impossible.

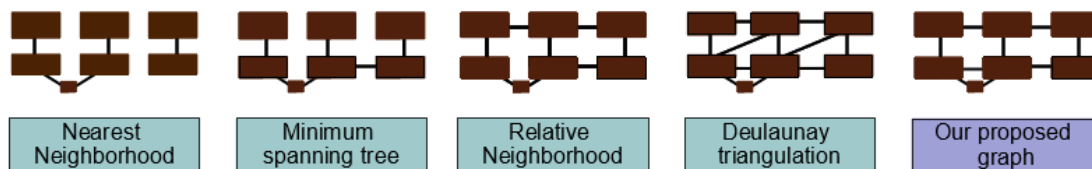


Figure C.26: Different proximity graphs in a simple situation of two parallels alignment perturbed by a building without relation.

The graph obtained consists of 13993 edges labeled *no alignment*, 2976 edges labeled *straight alignments*, and 2010 edges labeled *curvilinear alignments*. An extract of the input dataset is presented in Figure C.27.

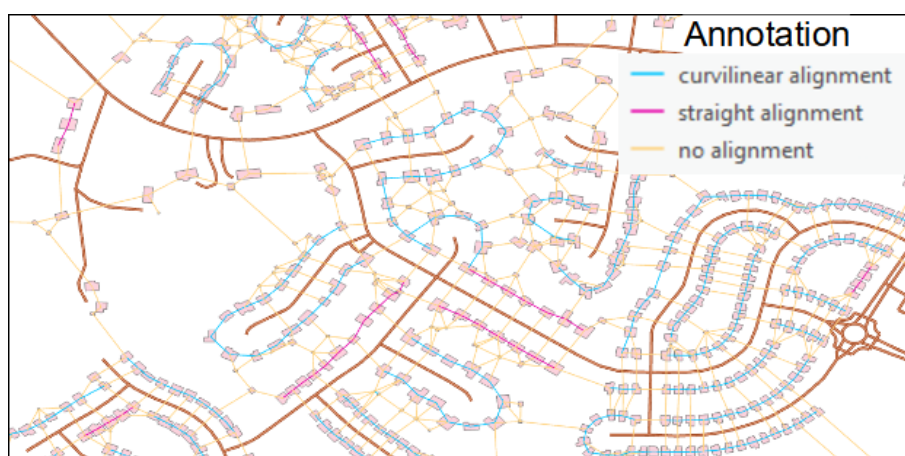


Figure C.27: Extract of the annotated graph for alignment detection.

Content enrichment The relation graph of the center point is not enough; the relations and the building must be characterized with attributes. Thus, we first characterize the relations between the building (see Figure C.28 and the explanations below).

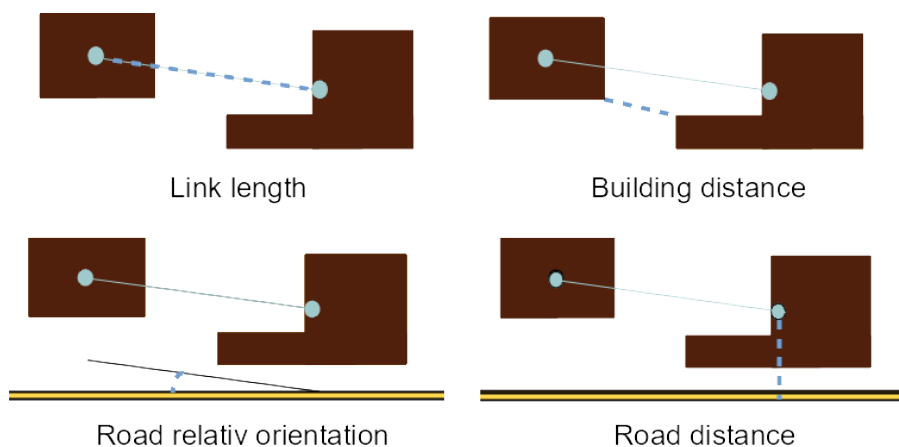


Figure C.28: The characteristics and context of proximity links.

The **length of the link** represents the distance between the center points of

the buildings; it reflects the strength of the proximity relation, however, it should be considered according to the size of the buildings (a large building would have a longer link). The **distance** between buildings does not have this bias, as it computes the minimum distance between the two buildings; however, it can be inconsistent in the case of a complex building with a long outgoing indentation. Finally, the combination of both attributes is the best description of the proximity strength between buildings and is essential for alignment detection. Both these measures are expressed in meters. The context information is related to the closest road because buildings are generally constructed along the roads (or served by a road depending on the point of view) and an alignment thus often is along a road. The **closest road relative orientation** is the angle formed by the segment that connects the centers of the two buildings and the tangent of the road at its closest point to the segment. This angle is given between 0 and 90° because the side of the angle is meaningless. The **closest road distance** is calculated between the closest point of the road and the segment that represents the link. It represents the relevance of the angle (the more roads are far, the less relevant it is to use the angle to determine whether it is an alignment).

Then, we introduce some attributes about the similarity relation of the two linked buildings for each possible alignment (see Figure C.29 and the explanations below).

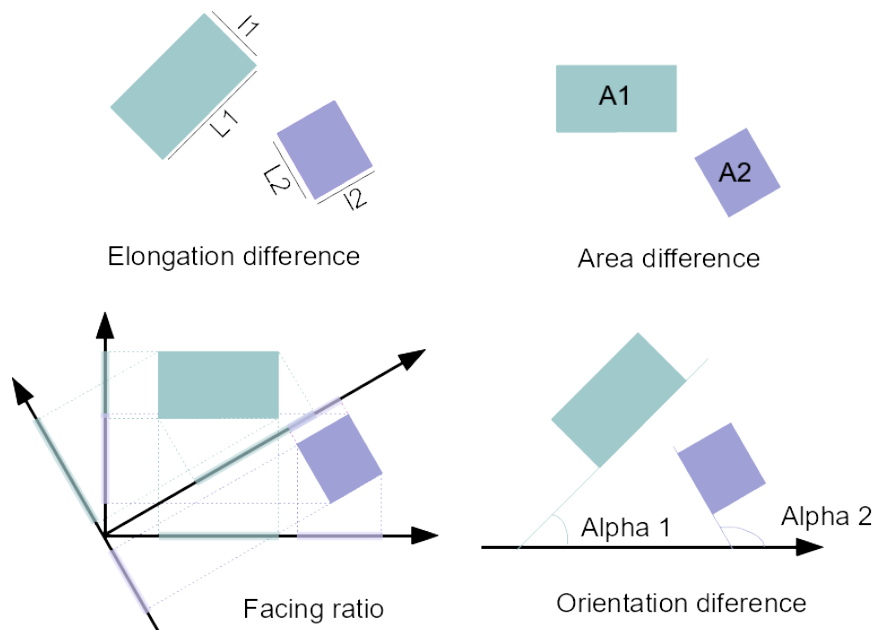


Figure C.29: Attributes that describe the difference or ratio of related buildings.

Buildings perceived as aligned often look similar, and thus have a similar shape descriptor; we choose to include a measure of three common shape descriptors: **area, elongation, and orientation**. The elongation of one building is the ratio between its longest direction and its shortest direction; thus, for each building, we computed a rectangular bounding box, and the elongation is the elongation of this bounding box. We use a similar structure for orientation, the orientation is the orientation of the building bounding box (Duchêne 2003). Thus, our measures of orientation and elongation are only relevant in the case of a quite regular building; however, in general, and in particular, in our dataset, the alignment mostly concerns the buildings with quite a simple shape. Once the indicators are calculated for

each building, it is the relation between these values for two adjacent buildings that interests us; we can characterize a link by a measure of the distance of these attributes; the two obvious distances are the absolute difference (eq. C.12) and the ratio (eq. C.13). For X an attribute of building A and B .

$$D(X) = |X_A - X_B| \quad (\text{C.12})$$

$$R(X) = \frac{|X_A - X_B|}{X_A + X_B} \quad (\text{C.13})$$

The difference is not bounded; it simply measures the gap between two values. The ratio allows us to compare the gap between values relative to their magnitude (i.e., a difference of area of two square meters is less important when the buildings are large). We choose to use the ratio for the area. However, in the case of orientation and elongation that already are bounded and where the sum is meaningless (e.g., the sum of building elongation is not the elongation of the set of both buildings), the absolute difference is preferred.

Finally, we calculate the **facing ratio** (from Wang and Burghardt 2019). This measure takes into account both the relative position and orientation of the buildings. It measures how buildings face each other: it is the maximum intersection of the projection of the rectangular bounding box of the building on the orthogonal marker parallel to each bounding box, divided by the union of the projection. The aligned buildings usually have a large facing ratio, while the non-aligned buildings usually have a small facing ratio.

Road network selection

Our dataset for road selection consists of roads from a map at the 1:25,000 scale and their generalization at the 1:50,000 scale in urban, suburban, and rural areas. In this experiment, geographic information is organized in a network; its representation as a graph is natural. The detailed road network contains 12,837 edges, 55% that should be kept, and 45% that should be deleted in its generalized version.

Content enrichment The graph itself is not sufficient to learn the proposed task. In fact, the selection of the road section is a mixture of its role in the road network and its attributes (e.g. nature, width, etc.). Moreover, the graph also misses some information on the structural role of the road in the network (e.g. stroke, i.e., set of continuous road section that may form a unique road for an human user). Thus, we chose to characterize the edges of our graph with the following attributes.

- **Segment or line length**
- **Importance:** an ordered category from this initial database for the symbolization on the map, derived from width and nature.
- **Stroke length:** measure the stroke to which each link belongs; the stroke is often a more relevant level of analysis for road selection (Thomson and Richardson 1999).

- **Betweenness centrality**(Jiang and Claramunt 2004): measure the number of shortest paths between two pairs of nodes that use each section. This centrality measure is really important to determine the usefulness of a section in a transportation network, and contrary to other centrality measures (e.g. degree centrality and closeness centrality), cannot be handled by convolution on the graph (as defined in GCN).
- **Face size**: measure for each face of the network its size and report for each edge the largest and smallest adjacent face. It gives information on the density of the network around the section; it also allows one to detect the city limits (e.g. a large face and a small one mean the network is dense on one side of the section and dense on the other side).
- **Sinuosity** road length divided by section base length. Very sinuous roads are often less useful for path construction; however, they can be important on a map, as they reflect a certain type of landscape (mountain road).

We tested the interest and consistency of all of these attributes for the desired task by training a random forest classification on the nodes of the graph using these attributes. This classification gives results worse than the graph classification because it does not use the structure and relation between elements. However, it allows one to see the contribution of each attribute in the classification decision for each attribute. We conclude that each the proposed attribute are relevant and usefull for the desired task.

Chapter's conclusions

Applying a deep learning model for map generalization requires setting up an adapted architecture and a training dataset. The choice of each of these elements can affect the quality of the prediction and must be adapted for the desired task.

First, the characteristics of a deep neural network to keep in mind when defining or choosing a deep neural network are its ability to abstract, preserve, transfer, deal with spatial information, its variability, and its gluttony.

Then the objective function guides the learning, this function may require an adaptation to be more relevant for the desired task. In particular we think the objectiv function must encourage the preservation of information and must not be calculated at element level (compare prediction and target element by element), but rather globally verify the target and prediciton are similare.

Finally, it is critical to create a dataset adapted for training. An example has to include all the necessary information for generalization (at least in an implicit way) and to be consistent. Representation must include as little bias as possible. Especially, the MAUP (Modifiable Areal Unit Problem) and the symbolization effect have to be avoided.

Part II

Exploration of deep learning for map generalization

Chapter D

Can graph neural networks model spatial relations?

The goal of this chapter is to illustrate the potential of graph-convolutional networks (GCNs) for tasks that ask a deep understanding of spatial relations. The ability to understand spatial relation is essential in map generalization as spatial relations structure the map. We explore two use cases: building pattern detection and road selection. Both tasks require some degree of understanding of the spatial relations between map objects.

D.1 Motivations

Spatial relations and map generalization automation Maps are made up of geographically related spatial objects. Spatial relations are key information on the map, as they support the understanding of geographic space. These relations are usually not explicitly encoded in the database. Consequently, conceiving a model or algorithm capable of automatically identifying and preserving these relations is important for the automation of map generalization (Mackaness and Edwards 2002). For example, building typification is a generalization operation that seeks to reduce the number of buildings while preserving the relation between and within homogeneous building groups (Regnauld 2001). Furthermore, Mustière and Moulin (2002) explain that spatial relations are parts of the spatial context necessary for the correct generalization of cartographic objects. For instance, Touya, Bucher, Falquet, Jaara, and Steiniger (2014) raise the following issue : "how do we handle the relation of a building alignment along a dead end street when moving the street is required by the city generalisation process?". Then, Duchêne, Ruas, and Cambier (2012) explain how relational constraints can be used for map generalization in an agent process. These constraints are designed to a) ensure the legibility of the map, b) ensure the preservation of the relations, and c) ensure the geographic coherence of map content.

Which spatial relation to learn ? Spatial relations are diverse; they can be one-to-one relations or link several objects together, like groups and patterns. All relations are not equally difficult to make explicit. For example, the recognition of topological relations is quite well processed since the 9-intersections model extended to the dimension (Clementini, Di Felice, and Oosterom 1993), while the detection

of groups is currently still experimented (Yan, Ai, Yang, and Tong 2020; Wang and Burghardt 2020). In fact, the current challenge is to identify non-topologic relations as they are perceived by human.

D.2 The use case of pattern recognition in a map

D.2.1 Presentation

Mackness and Edwards (2002) describe a pattern as a complex of primitive objects linked by relations; thus, the patterns can be seen as N-to-one partonomic relations (e.g., buildings are part of the city). Figure D.1 illustrates the interest of pattern-based generalization. The generalization of an object in a pattern must be consistent with the generalized shape of the pattern.

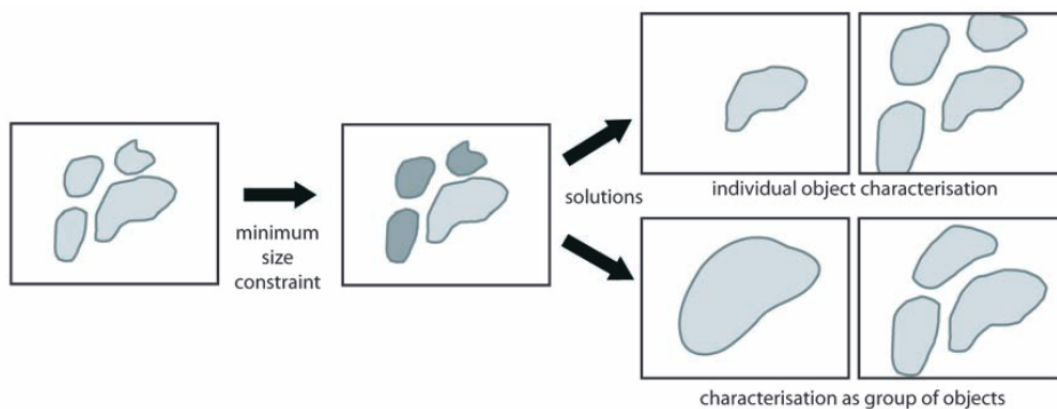


Figure D.1: Different generalization solutions when contextual relations are ignored (top-right) and observed (lower-right) (from Steiniger and Weibel 2007).

The identification of a pattern by a human is related to a remarkable or meaningful feature group with particular common characteristics and relations. This identification often relies on a Gestalt property (e.g. objects are close, similar, and in a common region; see Figure B.1). Several papers are interested in automating the identification of such patterns; most of them rely on an auxiliary structure, this structure links potentially related features, and then this structure is characterized and refined to extract a relation (e.g., Wang and Burghardt 2020).

In this use case, we experiment with pattern detection in the classical use case of building alignments. This is a pattern composed of three or more buildings belonging to the same block, relatively close, similar (especially with a similar size and orientation), and with a linear arrangement. An alignment can be colinear if the orientation and distance are regular or curvilinear if the orientation and/or distance smoothly vary (the axis of the alignment is a curve). This pattern is quite common in urban areas on a detailed map; its generalization must reduce the density of buildings while still conveying this idea of alignment. Figure D.2 illustrates the input, expected prediction, and building patterns in our experiments. Our goal is to demonstrate that deep learning can recognize spatial patterns, so we will train a graph neural network to classify the link in a proximity structure with labels: "curvilinear alignment" or "colinear alignment". This classification enables the alignment pattern to be located on a building block.

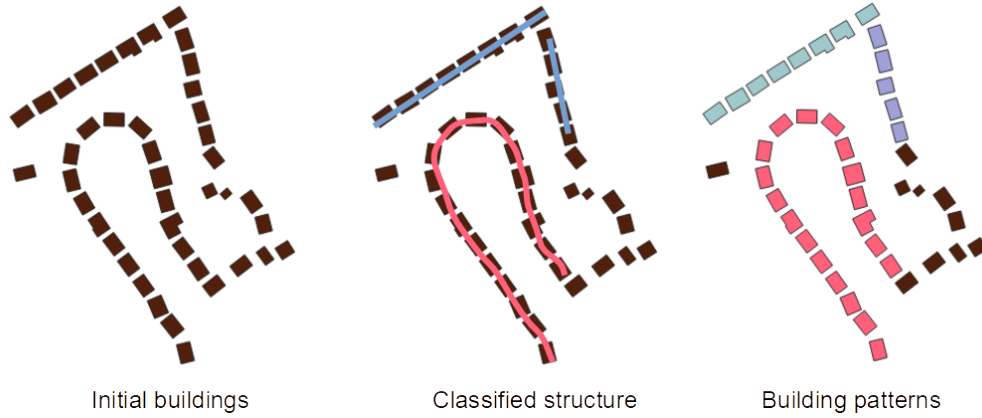


Figure D.2: The input, the prediction, and the goal of the alignments detection experiment. We predict the link between two building that are aligned, this link is characterised by the nature (colinear or curvilinear) of the alignment, from this prediction we can annotated with a same label building that belong to the same alignment.

D.2.2 Method

We use a graph structure to encode the input and target characteristics. We aim to predict the relation between buildings (colinear, curvilinear, or not an alignment). The general process is explained below (each point in the list on the right corresponds to the respective image in Figure D.3).

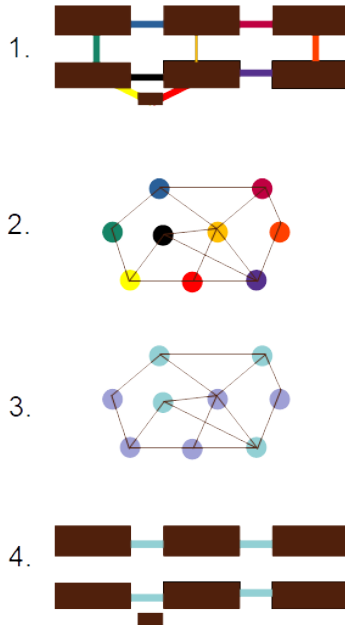


Figure D.3: Process for the prediction of alignments.

1. We construct a proximity structure that links the building of potential alignments.
2. We construct the associate graph where the nodes are the potential relation and the relation that concern the same building are linked; this graph is qualified with relation and building characteristics.
3. The model is trained to predict the classification of the nodes.
4. The node classification enables us to find the link between the aligned building and the proximity structure.

For this experiment, we tested a GCN (Kipf and Welling 2017), which is specialized in analyzing nodes and adjacencies in a graph to predict node or graph level information. As explained in Chapter C the GCN uses as input a feature description for every node, summarized in a feature matrix, with the adjacency matrix of the graph. We hypothesize that this characteristic would enable the model to make a

consistent decision across spatial relations. For example, an inconsistent decision would be saying that a link is an alignment if no neighbor is. We trained the PyTorch implementation of GCN¹ with our dataset for 200 epochs. Our training set is a graph with 17281 nodes. The construction of this dataset is detailed in Chapter C. We constructed the annotation of this dataset manually. The goal is to verify that the network can learn what a human perceives as an alignment.

D.2.3 Results

The model fails to recognize patterns from our example set: the result is really irregular across epochs, and no optimum is found. The confusion matrix (Table D.1) shows that the main problem lies in the distinction between non-alignment and alignment relations (e.g. a large portion of colinear and curvilinear edges are misclassified as non-alignment edges, and vice versa) and not the problem in the distinction between curvilinear and colinear alignments.

Prediction \ Annotation	Curvilinear	Colinear	No	Total
Curvilinear	1125	34	1412	2571
Colinear	66	1196	2662	3924
No	819	1746	9919	12484
Total	2010	2976	13993	18979

Table D.1: Confusion matrix for alignments detection using a GCN.

The visual interpretation of the results is consistent with this interpretation; the model is not able to detect which links belong to an alignment, but area with curvilinear alignment tend to contains only links curvilinear alignment and no alignment links and conversly.

D.2.4 Discussion

Can GNN learn spatial relations?

The results of this experiment are not clearly satisfying and are really different from the target. A quick analysis would lead to the conclusion that GNN cannot learn spatial relations. However, looking at these results in more detail may lead to some clues of the potential of GNN for learning spatial relations. Indeed, a building that has an alignment link is most likely to have its other link that is also in alignment. However, this element is not sufficient to say the model understand the spatial relation and our experiement has to be improved.

Probable causes of failure and possible improvements

The origin of this failure may be the dataset, model, or modeling. The following paragraphs investigate the limits of these three aspects in our experiment.

¹<https://github.com/tkipf/pygcn>

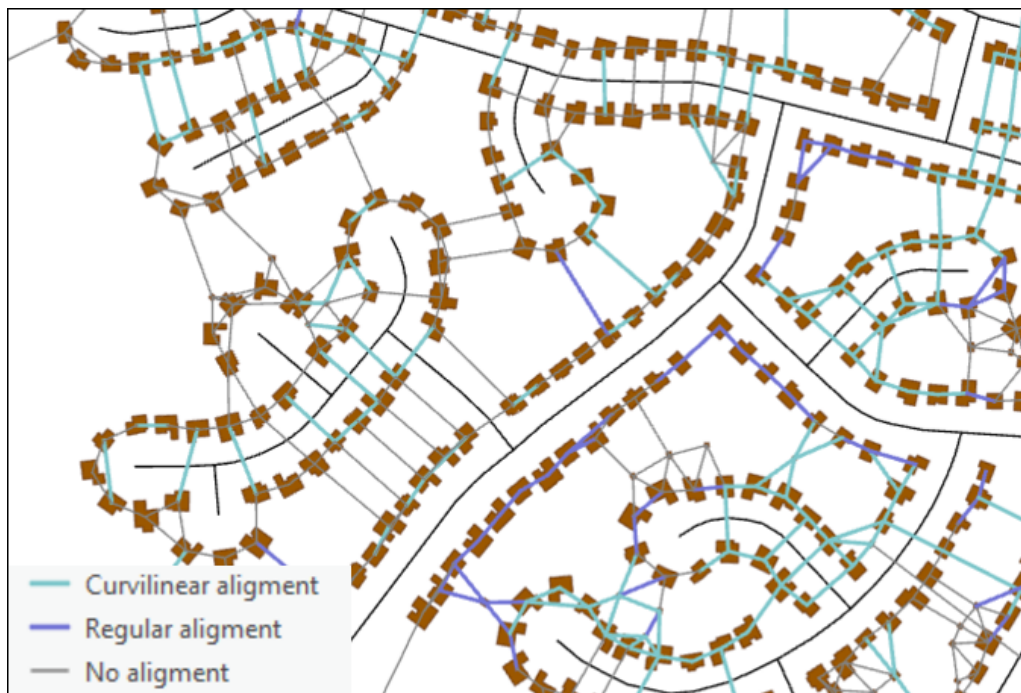


Figure D.4: Prediction of the model trained for alignment detection.

The training dataset may not be adapted for the expected task As explained in Chapter C, deep learning is based on examples and a bad example can compromise the learning. First of all, our number of examples seems reasonable compared to the other successful applications of the chosen model (Kipf and Welling 2017). Then, to determine the interest and sufficiency of the proposed features, we trained a random forest model for building classification. This quick experiment does not include the graph structure and considers each node individually. The results are better than with the GCN, which means that the information of graph structure are not correctly encoded or used. We conclude that our data modeling may not be adapted for this use case.

In this experiment, we proposed a custom proximity graph made from the refinement of a Delaunay triangulation between building center points; this structure aims to be the most complete as necessary to cover the aligned building while avoiding irrelevant links. We compare this proposal with other common proximity graphs constructed from the center of the building, this leads to similar or worst results. To go further, a structure that is constrained by a road network or based on other building points has to be investigated (Zhao, Ai, Yu, He, and Shen 2020).

Finally, we investigate the diversity of the examples: our examples are really imbalanced (most links are non-alignment links), this is a common limitation to deep learning. Thus, we tested to train the model with weighted examples: alignment link have a more important weight during the loss calculation. This experiment aims at testing the impact of the imbalance in our training set and bring no concluding results, the imbalance was probably not the main problem.

The model may be unadapted for the expected task The major advantage of GCN is its ability to learn from neighboring nodes. To ensure our GCN is adapted to our use case we explore different settings of the model. In particular, we make varying the depth, i.e. the number of convolutions, of the network. It reflects

the number of neighbors considered during information encoding for the use case of alignment. Nevertheless, this sensitivity study gives very few variations in the quality of predictions.

During our experiment, we only tested the GCN, mainly for time reasons. However, GCN is one particular model in the field of graph neural networks, and our experience does not allow us to conclude on the potential of all GNNs to learn spatial relations. Especially, some recent examples from the literature have shown the potential of other models for encoding relations in spatial data (Yan, Ai, Yang, Tong, and Liu 2020; Yan, Ai, Yang, and Tong 2020).

Our modeling may be unadapted In fact, alignment detection is more of a clustering problem than a classification one, but in our experiment we try to simplify it to classification with a neighborhood relation. The simplification may not be best way to model this pattern recognition problem. The alignment detection require many features and several kind of relations, while in our experiment, we only used a homogeneous graph. However, the ability to handle a heterogeneous graph is one of the main advantages of GNN and offers a wide opportunity to describe spatial relations (Iddianozie and McArdle 2021). For example, to detect alignments, a heterogeneous graph would be preferred that describes both the proximity between buildings and the road network, as alignments are often enclosed in urban blocks and along roads.

Then, the definition of neighboring in GCN allows the model to propagate information over neighboring nodes and make a consistent prediction. However, it relies on proximity in the graph, which is not exactly cartographic proximity. For example, in Figure D.5 the two highlighted roads are close to each other and have a similar role in the road network (join the city center to the highway), however, they are distant and different in the graph and thus not considered redundant during road selection learning. Similarly, the blue and purple buildings are not close enough to be in the same alignment, but the blue and pink buildings are even more distant neighbors in the same alignment.



Figure D.5: Illustration of the limitations of neighboring definition. The blue and purple buildings are not close enough to be in the same alignment, but the blue and pink buildings are even more distant neighbors in the same alignment. The two highlighted roads are close to each other and have a similar role in the road network, however, they are distant and different in the learning graph.

Moreover, the level of neighboring required for relation understanding in spatial data varies, while the GCN always uses a fixed number of neighbors (related to the model depth). For example, in rural areas where the road network is spare, looking at fourth-order neighboring can be sufficient for road selection, while in urban areas

the fourth-order neighbor can be a section of the same complex intersection! In our experiment, the building distance, and link length may represent the strength of the neighboring. However, this information is not enough to encode relative location of the object, this is an explanation for the unsatisfactory results. To overcome these limitations, it is necessary to encode the spatial characteristics of the node and edge; the training graph must be a spatial graph². Then we believe that a model that learn to look further when the edges are small, could be relevant. However to our knowledge a such model does not exist in the litterature. Another possible solution is to train separatly a model for dense area and one for rural areas.

D.3 The use case of the road network selection

D.3.1 Presentation

Network generalization is challenging because it is a multi-criteria decision that takes into account the semantic and geometric properties of the road section, the properties of spatially and topologically neighboring sections, and more globally, the role of the section in the road network. The generalization of the road network aims at removing the least significant sections while preserving the road networks global structure, preserving the main routes, and preserving coherence with other themes (e.g. if a point of interest is in the generalized map, the section that leads to this point has to be kept even if it is a dead end). Thus, to achieve network selection, it is necessary to encode and use the following relations: "section is connected with", and "belongs to a stroke", "belong to a pattern (e.g. round about)" but also "belongs to an important route", "serves", or is "redundant with". Finally, one of the challenges in road network generalization is that a road network includes complex patterns that require a specific generalization (e.g. roundabout, interchange, dual carriage ways, etc.). The reduction of road network density is a complex process that includes road section selection at some stage. For instance, Touya (2010) and Luan and Yang (2010) propose a road selection process based on the geographical context, where the first step is identifying such a pattern using network analysis; then strokes are used for the network selection. However, the generalized network is not a subset of the detailed network sections (see Figure D.6), and typification of particular structure have to be resolved otherwise.

We experiment with the prediction of a value for each road section; the objective is to learn from the example the probability to be selected for each edge of the graph (i.e. road section in the geographical database). This information depends on the characteristics of the section, but also on its relations with the rest of the network. The success of this task would contribute to the demonstration that the model has encoded the spatial relations. To simplify the problem, the model is trained to reproduce a classification of road sections according to whether they are "kept" or "erased" in the generalized network. This annotation is made on the detailed network and obtained by comparing the detailed and generalized networks (see Figure D.7). The generalized network is issued from a generalization process based on data enrichment and structure detection (Touya 2010) that takes into account the spatial relations. The generalized network is not an exact subset of the detailed one and includes schematization/typification of complex intersections; in

²Graph having nodes with spatial location (e.g. coordinates).

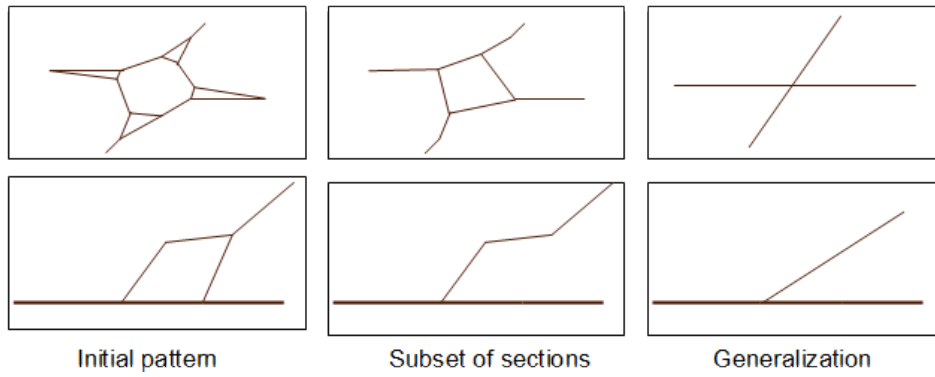


Figure D.6: The generalization of patterns in a road network is not a subset of initial sections.

our experiment, these cases are ignored, and sections of complex intersections are all annotated as kept to avoid the model learning to make disconnections.

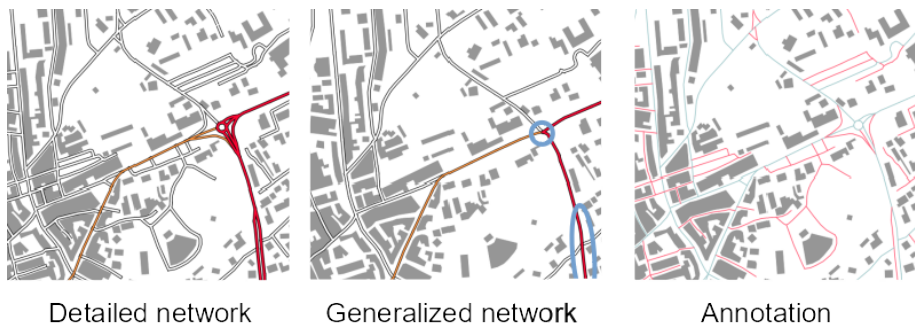


Figure D.7: The generalization of the road in the source databases (examples of pattern schematization in the generalization are circled in blue) and the annotations that the neural network aims to learn ("kept" section are symbolize in blue; "erased" one in pink).

D.3.2 Method

The road network is already a graph-like structure; however, the graph that supports the learning must represent the section as a node. Therefore, we use a graph complementary to the road network where the edges are the intersection and the nodes are the sections (see Figure D.8). Then, the model is trained for node binary classification; finally, when we apply the model instead of the prediction (label kept or erased), we extract the output of the last layer of the model value that corresponds to the probability of the positive value (kept label) for each a node (road section).

In this experiment, we used a GCN with the same experiment setting as previous use case. Our training set is a graph with 12,837 edges. The creation of this training set is detailed in Chapter C. We explore the interest of several features to characterize road sections. Then, the edges of the graph are describ using some attributes.

First, we include attributes that give direct information about the sections : **length**, **importance**, and **sinuosity**. The length is expressed in meters; the importance is a categorical attribute with a value between (1: highway and 6: local

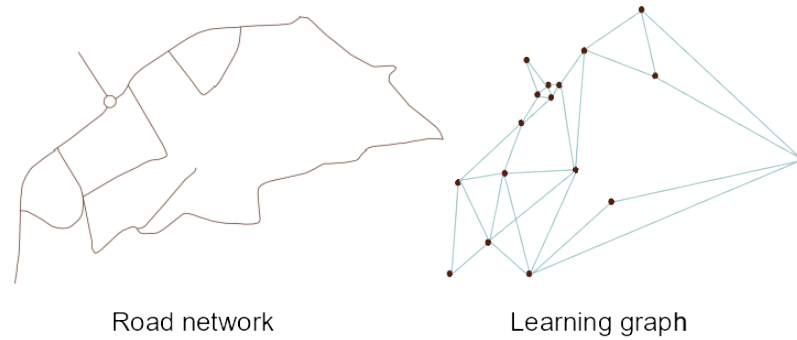


Figure D.8: A road network and the associated learning graph.

narrow path) issued from the initial database; the sinuosity is the ratio between the section length and base length (distance between start and endpoint). This information is interesting in road selection, as very sinuous roads are often less useful for path construction; however, they can be important on a map as they reflect a certain type of landscape (mountain road). This ratio is between 0 and 1.

Secondly, we include some attributes that describe the spatial relations internal to the road network and the context of the road sections:

- We measure the **length of the longest stroke** the section belongs to; the stroke is often a more relevant level of analysis for road selection (Thomson and Richardson 1999). Even a small road that belongs to a long stroke may be important.
- We measure the **area of adjacent faces of the network**. It gives information on the density of the network around the section; it also allows one to detect the city boundary (e.g. a large face and a small one mean that the network is sparse on one side and dense on the other side). We store two attributes, the size of the smallest and largest faces (the network may be able to discover the relevant combination of both these values).
- Finally, we added **betweenness centrality** for each section. The betweenness centrality measure is essential for road network selection, and we do not think it can be learned for the graph structure with the convolution definition in GCN. Indeed, it is the number of shortest paths in the network that uses this section. In our GCN, the convolution propagates node information to its neighbor, it does not consider shortest path and gives the same importance to each path in the network³.

D.3.3 Results

In this section, figures illustrate only extracts from our results (chosen as fairly as possible) the complete results are publicly available in Zenodo <https://doi.org/10.5281/zenodo.7244049>. The prediction of this experiment is the probability for each edge to be selected by the classification model. For evaluation we choose a threshold of 50% to perform a classification; this classification has a recall of around 83% and

³We think that the graph representation and definition of graph convolution in GCN nevertheless enable model to learn some other centrality measures (e.g., the degree centrality, or closeness centrality).

precision of around 72% (see confusion matrix, Table D.2). The model is under selecting the road network and kept too many undesired sections, however, it deletes very few undesired sections.

Prediction \ Annotation	Deleted	Kept	Total
Deleted	3098	1280	4378
Kept	2313	6246	8559
Total	411	7526	12937

Table D.2: Confusion matrix for the selection of road sections using a GCN.

Figures D.9 to D.11 allow us to visually compare this value (between 0% and 100%) with the selection annotation in the cartographic database in the city center, small town, and rural areas.

For the small town (the most common situation in our dataset), we observe that the model can clearly identify the main connections to the city (Figure D.9); these roads appear with a significant probability of being kept (in red). Some local roads also have a clear tendency to be deleted (in yellow). However, the dominant prediction is a medium probability to be selected and does not enable a clear separation between kept and deleted roads.

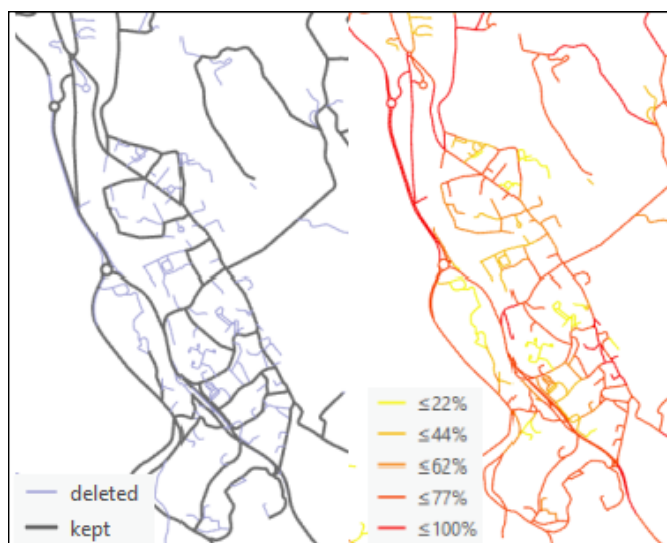


Figure D.9: Annotation and prediction of selection in a small town.

In the rural area (Figure D.10) the predictions are consistent with the annotation, and similarly, the prediction shows a mainly medium probability. The two hamlets are correctly classified. However, some probability discontinuities on an important road that can lead to disconnection are visible on the lower right of the area. Moreover, the classification of small dead ends includes errors.

Finally, in the urban area (Figure D.11) we observe more errors, the block structure of the road network is not learned at all, and the littoral path surrounding the city is not kept, while it structures the town. This defect comes from the small number of such a situation in our dataset and the more complex structure of the network in such an area. Moreover more features can be required for such complex situations.

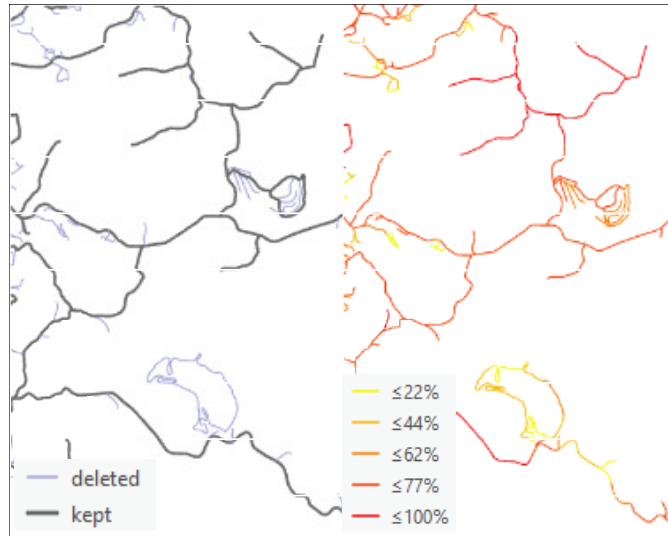


Figure D.10: Annotation and prediction of selection in a rural area.

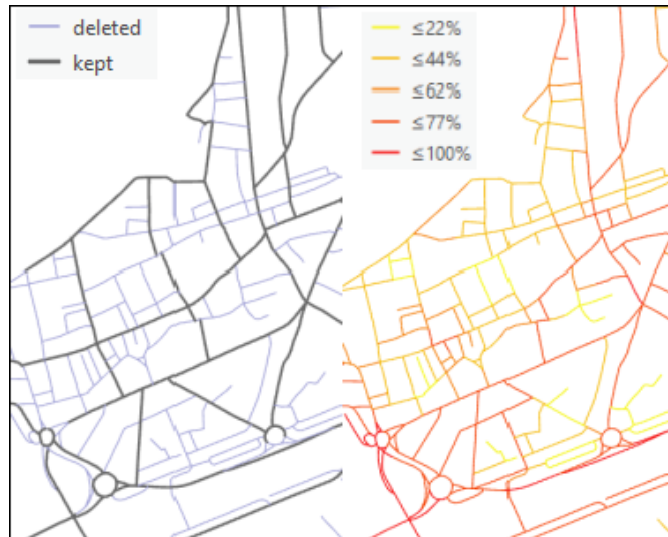


Figure D.11: Annotation and prediction of selection in the city center.

In fact, this classification does not tell us what to select but gives a hierarchy from which we can select the right number of routes, it indicates the negligible roads, the important ones, and the average ones. The integration of this information in a selection process is later discuss in next section.

D.3.4 Discussion

Can GNN learn spatial relations?

The results of this experiment are quite more satisfying. The predictions on the road section are credible and could mostly correspond to a correct map generalization. Some of the underlying relations are clearly understood, especially the consistency of a neighborhood. For example, the probability of selecting one section is similar to that of the other part of the road. Similarly to the previous use case, some improvement in the modeling of the use case may contribute to reinforce the demonstration

of GCN potential to use spatial relation in map generalisation task. In particular, the road selection process may require learning graph that explicitly describe the relations: "a road section belongs to an important route" or "a road section serves an important point of interest" with an heterogeneous graph.

The potential of GCN for map generalization

Our experiment has shown that GCN are able to make consistent predictions across a graph structure. Thus, such networks may be used to make explicit spatial relations for data enrichment prior to map generalization and for an integration in a global deep learning map generalization process. In this paragraph, we use predictions of alignments and road selection to explain how of such networks can be used in a generalization process. The road selection is the first step for topographic map generalization. The predicted importance of road sections from the GCN model could be used in several ways for road selection.

- The result of our model can be directly employed for selection, using a threshold (e.g., all sections with a probability up to 50% percent are kept). This simple solution is not the most efficient regarding our prediction distribution that gives a medium value to most of the section.
- The predicted value can be included as an attribute in a geometry-based or graph based process for road selection (e.g., Touya 2010).
- A mixed solution that combine network refinement by selection and stroke based simplification. For example, first we keep (and respectively erase) section with an extremely high (and extremely low) probability predicted. Then, we process selection based on stroke characterized by the maximum, or mean prediction of its sections (this extreme value would allow keeping even small paths that directly deserve an important section).

Chapter's conclusions

In this chapter, we presented two experiments that aimed to identify the potential of deep learning to learn spatial relations with a graph modeling. These experiments have shown potential to make predictions consistent with spatial relations. The road selection experiment is satisfying for most situations, but the results suffer from a lack of city center examples. The alignment detection is less satisfying and suffered from both dataset and model limitations. Both experiments have revealed challenges in the modeling as a graph and in the development of adapted GNN, in particular they have shown the necessity of encoding and using the location of the node and cartographic neighboring. Our experiment did not fully succeed in demonstrating the potential of deep learning to understand spatial relations. In next chapters we rather experiment with an image based approach and the understanding of spatial relation is no more the central question, but an ability required to perform the desired task. For example, to predict the shape of a generalized road in Chapter E, the continuity and noncoalescence of the road section has to be encoded; and for map generation explored in Chapter F the relative position of different map objects and patterns have to be preserved.

Chapter E

Convolutional Neural Networks for the generalization of roads

This chapter aims at the exploration of the potential of the image-based approach for roads generalization. As stated in Chapter B, learning graphic generalization operations can be made from various approaches; in this chapter we focus on the image-based approaches, which seem more natural and direct for this problem.

E.1 Materials and methods

E.1.1 Use case

To explore the potential of an image-based deep neural networks for graphic generalization, we choose the use case of the mountain road generalization. This use case is classical and challenging in map generalization because the sinuosity of roads causes symbol coalescence (Lecordix, Plazanet, and Lagrange 1997; Mustiere 1998; Duchêne 2014a). Moreover, roads play an important role in maps, they partition the map in blocks and their generalization is always the first step in the generalization of maps (Ruas and Morisset 1997). Thus, it may be impossible to learn a more complex generalization without first learning how to generalize roads. The generalization of mountain road shape involves simplification, caricature, typification, displacement, and smoothing operations (see Figure E.1), and frequently requires the combination of several algorithms (Roth, Brewer, and Stryker 2011). Mountain road generalization is often composed of two steps: first, road selection allows the elimination of insignificant roads, then graphical generalization adapts road line shape to be legible at a target scale, removes symbol coalescence, and smooths the geometries. Traditional methods for mountain road generalization are already efficient; however, they are difficult to apply, as they require important knowledge about road sections and important calculation time. We are not aiming at results with a better cartographic quality, but results of equivalent quality that could be obtained more easily. In these experiments, we focus on the second step, the graphical generalization that produces the shape of generalized roads.

In the literature review (Chapter A), we present two usual road simplification processes, (1) the global approaches that try to find the optimal shape at once (Bader and Barrault 2001; Sester 2005); (2) the iterative processes that separate roads in homogeneous sections, and apply several algorithms to each of them (Lecordix, Plazanet, and Lagrange 1997), and merge them together (Mustiere 1998; Duchêne

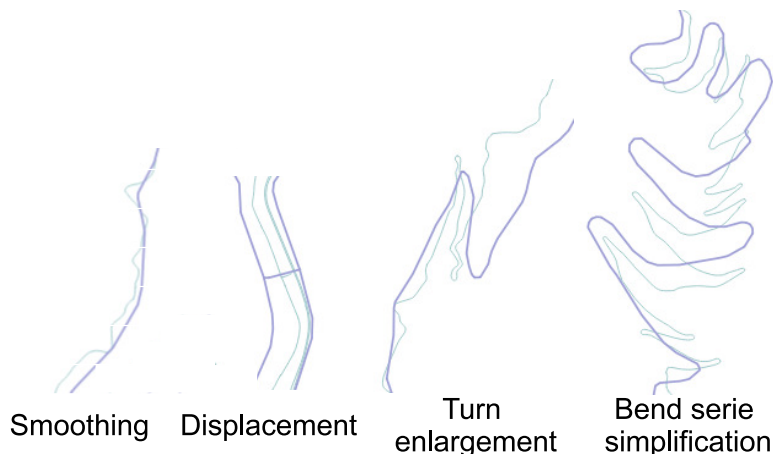


Figure E.1: Some examples of operations during mountain road generalization.

2014a). In the proposed deep learning approach the model is applied identically to each image of the test set, thus the approach is close to a global approach.

E.1.2 Approach

Our approach is framed as an image segmentation or generation problem using map images (raster data) as input. That is, using an image of detailed map tile as input to predict an image of the same size where each pixel value represents the probability of belonging to the generalized shape (Figure E.2).



Figure E.2: Simple illustration of our approach.

This approach is similar to those of (Feng, Thiemann, and Sester 2019) and (Du, Wu, Xing, Gong, and Yu 2021)¹ respectively for building and coastline generalization. These studies have demonstrated the potential of the segmentation approach for polygon simplification generalization². Our experiments are complementary to these two studies that learn the shape of a generalized polygon while we try to learn on polylines. The generation of polylines is challenging, as their narrow width is more sensitive to disconnections and dependencies between pixels are larger (very far pixels can belong to the same entity). Moreover, mountain road generalization is a complex task (see Figure E.1) that not only includes smoothing and symbol enlargement, but also bend enlargement, bend series typification/schematization, etc.

¹Realized after our experiment.

²In the coastline experiment the input is not only the line but segmentation of inland and sea parts, that are polygons.

Finally, in our experiment, and contrary to these first studies, we do not try to learn how to reproduce a generalization performed by an automatic generalization process, but we learn from a cartographic database that includes manual correction. This use case is thus a good complement to demonstrate the ability of an image-based approach to generate the shape of a generalized object using deep learning.

E.1.3 Experiment setting

In this section, we present our experiments, especially the dataset construction methods, our models, and the tested loss function for connectivity preservation improvement. We tested and compare the ability of several models to learn graphic generalization: one segmentation model and two image generation models.

Segmentation model We first tested a convolutional neural network designed for image segmentation called U-Net (Ronneberger, Fischer, and Brox 2015). This network has already shown potential to predict generalized shapes (Feng, Thiemann, and Sester 2019) and relies on a u-shape structure that encodes and decodes an image with convolutions. It includes skip connections to deal with multiscale relations in the image (a more detailed explanation of the U-net architecture is given in Figure C.5). We implemented a U-Net using Python 3 with the Keras library (Chollet 2018) in the Google Colaboratory platform with the available graphic processing unit (GPU) for standard licenses. Our U-net is designed to deal with images of size 256×256 ; it includes six up and down sampling layers. We then provide and compare a similar architecture made for images of size 512×512 .

Generation models We trained two generic GANs to generate images of generalized roads: CycleGAN and pix2pix. We used the PyTorch implementation of these models³. Each model has been trained over 200 epochs with default parameters. The choice of the number of epochs for a GAN is a difficult question as the model convergence is sometimes never reached (Arjovsky, Chintala, and Bottou 2017). This number of epochs gives a good level of satisfaction for both the loss value and the visual observation of the validation data. The visual analysis of the results shows that below this number of epochs, the results are still unsatisfactory, while they show no more progress after this value for both paired and unpaired models.

Datasets For these experiments we use the datasets presented in C.2.2. This dataset includes 1,223 pairs of images from the Alps and represents the road shape at 1:25,000 and 1:250,000 scales.

Connectivity loss Finally, one of our challenges is the preservation of the topology and connectivity of the road network, despite the image approach. We tested the connectivity loss proposed in C.1.4 as an additional part of the GAN objective function. It measures the preservation of connectivity information in images. The objective of our cycleGAN is thus eq. E.1, with L_{cycle} the cycle consistency loss, and L_{GAN} the adversarial losses for each side of the learning.

³Available here <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix.git>.

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, Y, X) + \lambda_1 L_{cycle}(G, F) + \lambda_2 L_{connectivity}(X, G) \quad (\text{E.1})$$

E.2 Results

In this section, the results are presented and evaluated according to visual evaluation focusing on the list of questions below. The conclusions were confirmed with some quantitative measures and a user test (see Chapter H). Reader can refer to Appendix A for more results images of main experiments.

Except for the contrary mention, in this section, all results are obtained using the dataset with fixed-size tiles of 2.5km where each pixel represents 10 meters on the ground, and the overlapping rate between tiles is 60 percent.

What is a good image prediction?

- **Q1:** Does the image look like a map and especially a road map? This question refers to the realism of the prediction. We trained the network with map images that follow cartographic rules, and the network should learn to generate images that look like a map.
- **Q2:** Are the geographic objects credible? Goodchild and Li (2012) propose to control spatial relations between objects. They develop this idea for maps produced using volunteer geographic information (VGI) where the expertise of contributors is unknown. This point is also important in the case of deep learning where the prediction process is a "black box" (Touya, Zhang, and Lokhat 2019). The key criteria for a road to be credible are its connectivity in the network and its linear and smooth shape.
- **Q3:** Is the image generalized enough to be legible at the target scale? In the use case of mountain roads, we have to take care of the smoothing level, and the reduction of symbol coalescence.
- **Q4:** Is the information sufficiently preserved and are errors created during the generalization process?

E.2.1 Segmentation approach

In this section, we present some results obtained with the U-Net segmentation of mountain road images. This section first includes an analysis of the overall quality obtained with this approach, and then we compare and discuss the creation of the training dataset.

Overall quality Our model reaches a segmentation accuracy on the validation (and training) dataset of around 60% (80%) according to the Dice measure and 50% (70%) according to the IOU measure (see Figure E.3). This Figure shows the progress of the model during the training phase but values calculation is based on the pixel accuracy that is not suitable for line generalization and not consistent with the

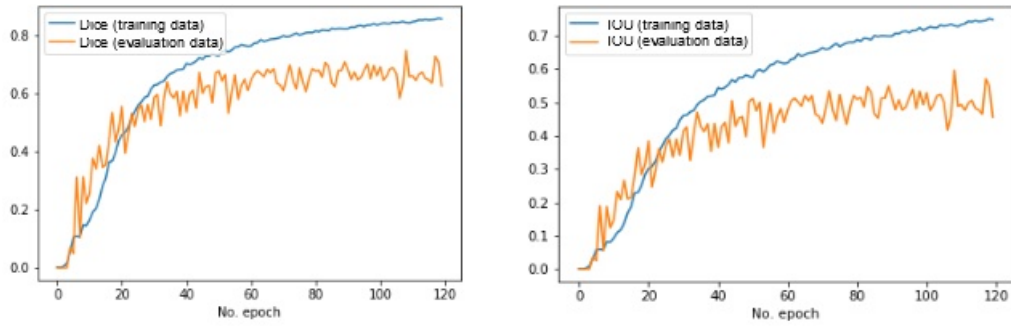


Figure E.3: Evolution of the evaluation value across epochs. (a) With Dice measure. (b) With IOU measure.

visual quality of the prediction. Nevertheless, the validation score shows that the model is not overfitting (validation and training score are relatively close), and finds consistency between input and target (the scores are increasing). This progression gets slower from the 50th epoch, which means the results may be correct from this point.

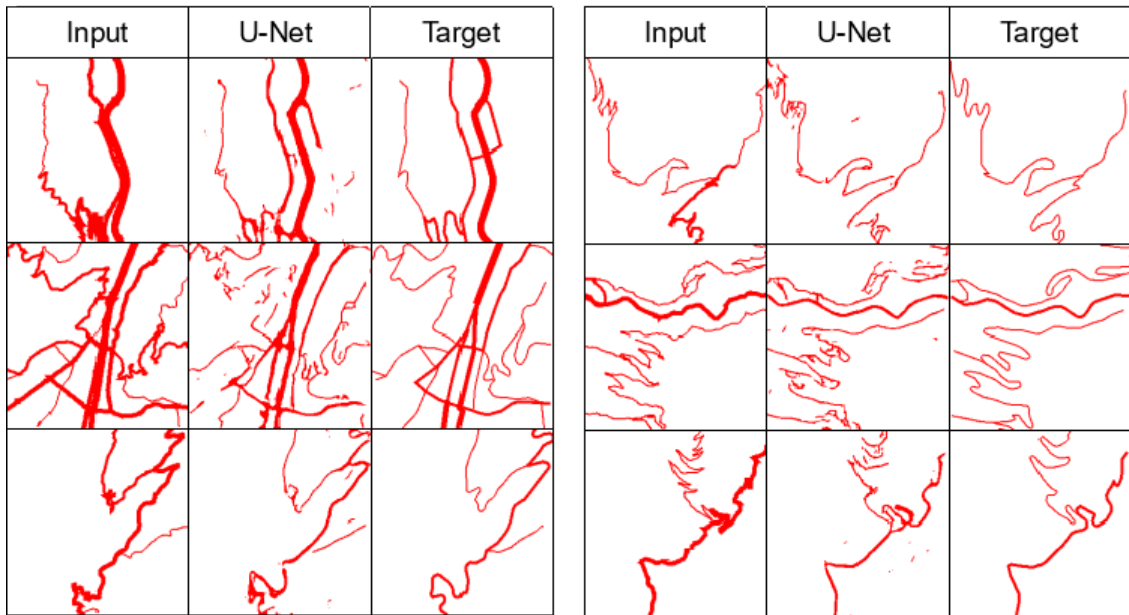


Figure E.4: Some results of the segmentation approach (input, prediction and target from the testing set).

Figure E.4 shows some of the results obtained on the images of the test set. In most of the images, the road shape is recognizable (Q1) and is very similar to the generalized image. However roads are often discontinuous and not credible (Q2) and some undesired part of road can be seen in some images. It is also often smoother than the initial image, even if it is sometimes less smooth than the generalized image (Q3). Bends are enlarged in most sinuous bend series, but the enlargement is sometimes not correct, which can lead to loops or U-shaped bends (Q3). The displacement performed to avoid parallel road to overlap is learned; however, when roads are displaced to avoid overlaps with other objects (river or rail line), as such data are not visible in our images, the displacement fails. Then, the main limitation

of the prediction is the numerous disconnections, that decrease the information preservation (Q4).

Results on training set constitution In this section, we compare the different proposed methods for creating the training dataset. First, Figure E.5 compares the two proposed approaches for tile generation (fixed scale window-based and object-based).

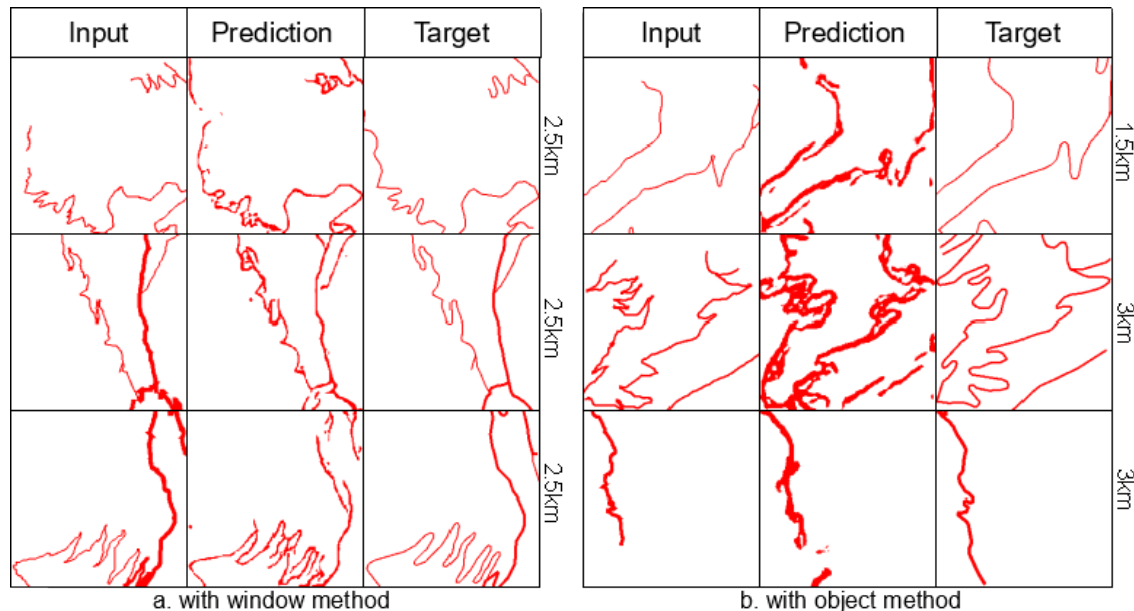


Figure E.5: Comparison of U-Net predictions for different tiling methods. **a.** Fixed-size tiles with an image size of 2.5km and an overlapping rate of 40%. **b.** Object-based tiles.

We observe that the fixed-scale tiling approach gives better results than the object-based approach. The object-based approach seems to systematically create loops in bends and enlarge the road symbols too much. It is probably due to its varying scale. Indeed, the coalescence in the initial image that prevents "seeing" the initial shape of the line, is more important in images of the objects approach.

Then, we tested to improve the object method with a mixed representation of roads in input images. The shape of the line is displayed in green with a width of 1 pixel, this shape includes no coalescence at all and thus allows the model to access the exact shape of the road. It is overlaid on top of the larger red symbol that coalesces and represents the width of the road on the map. Thus in this image the model may have at once a legible representation of the exact shape and the road symbol. Figure E.6 presents a comparison of the results given by this mixed representation. The mixed method is compared to an image where the symbol width is left unchanged and to an image where symbol width is exaggerated to convey the scale reduction due to the size of the road. It shows that the shape readability improves thanks to the green thin lines, and it does reduce the "loop" effect, but the results remain unsatisfactory.

Finally, the method comparison leads to the conclusion that despite randomly cutting roads and degrading context (i.e. the global shape of the road), the window approach is currently the most adapted for map generalization. Figure E.7

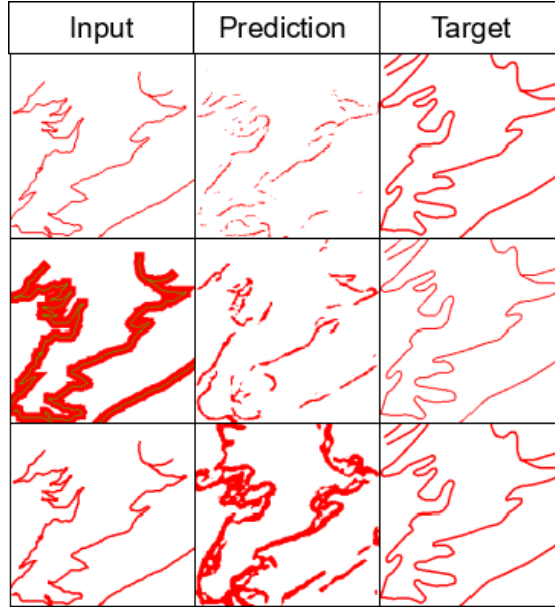


Figure E.6: Comparison of methods to avoid coalescence in the input image. a) without correction, the width of the road is fixed, b) moderate enlargement relative to the deformation rate, c) enlargement relative to the deformation rate (red) and the shape of the road (green)

investigates the optimal size for each example. It is a parameter that affects two characteristics: the context and the number of images. We measure this parameter in terms of the ground size represented by a 256×256 pixel image. It is not relevant to choose a size under 1 km (not enough context). As a result, all pixels are classified as "not road". In contrast, up to 5 km, the covered study zone is too small to have enough images (less than 150). We observe that sizes between 2.5 and 3 km are more relevant for our tasks. Moreover, we can see that smoothing is not well learned when the model is trained with large areas (5×5 km). The coalescence of small bends increases when the size of window increase, and this leads to bad results.

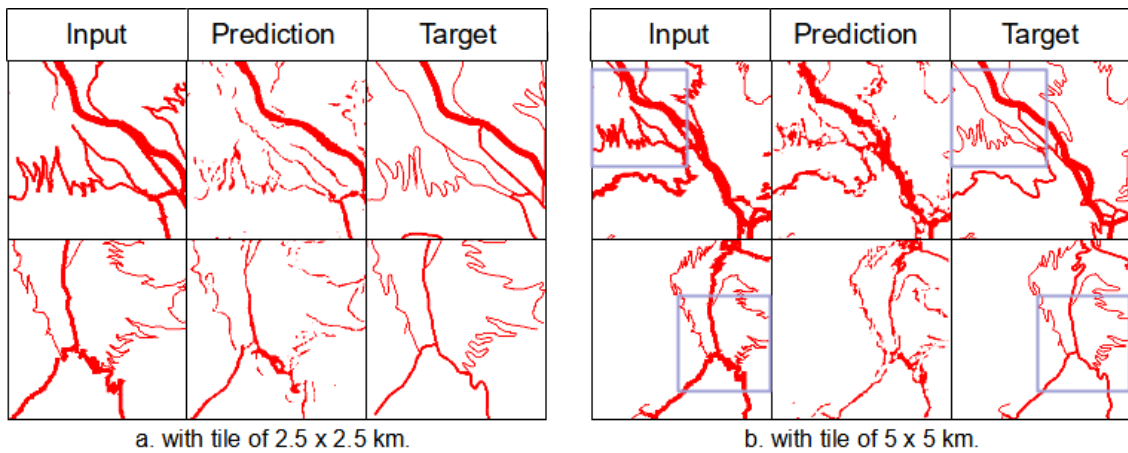


Figure E.7: Effect of window size on generalization. (a) 2.5×2.5 km tile (b) 5×5 km tile, the limits of the smaller size are represented in blue over the input and target image to improve readability.

The last parameters of the fixed-size tiling approach are the overlapping ratio between tiles, and the image resolution. The overlapping increases the number of tiles. The experiment with 40%, 50%, and 60% values shows good progress in both qualitative and quantitative evaluation when overlap increases. Finally, to reduce the amount of non-readable tiles, we increase the resolution of tiles to 512. This change increases the computation time required to train the model, but it greatly decreases the connectivity errors in the predictions. All the next predictions and those of Appendix A are made with this resolution.

E.2.2 Generation approach

In this section, we present some results obtained with pix2pix and cycleGAN to generate generalized mountain road images. We first analyse the overall quality obtained and compare the two models, then we discuss some model improvements, in particular the model setting, the addition of a connectivity loss, and the addition of examples.

Model comparison

We present and compare the results obtained by training cycleGAN and pix2pix. Figures E.9 to E.11 present some predictions from both models in different situations (for images that includes an important series of bends, for images that includes a main road and for images that includes an intersection). Figure E.8 illustrates and gives an example of the main defaults identified. We observe that both models produce images of roads (Q1) and lines are smoother than the input (Q3). However, the paths produced by pix2pix are neither legible nor credible because of loops and disconnections, that alterate the cartographic information (Q2 and Q4). This problem is less important in CycleGAN predictions (unexpected disconnections are rare, and the loop problem occurs only when the input is not legible enough). Then, cycleGAN does not enhance correctly the road shape in prediction; some bends are not enlarged enough and some edges not enough smoothed. In fact, it seems that when the model cannot conclude how to generalize a situation, it leaves the input shape. Finally, cycleGAN produces some small blue noise that does not visually affect the legibility of the road. This noise may be a consequence of the steganography effect of cycleGAN (Chu, Zhmoginov, and Sandler 2017); the reconstruction tends to teach the model to hide information in the prediction for a later iteration of the cycle. It can easily be post-processed.

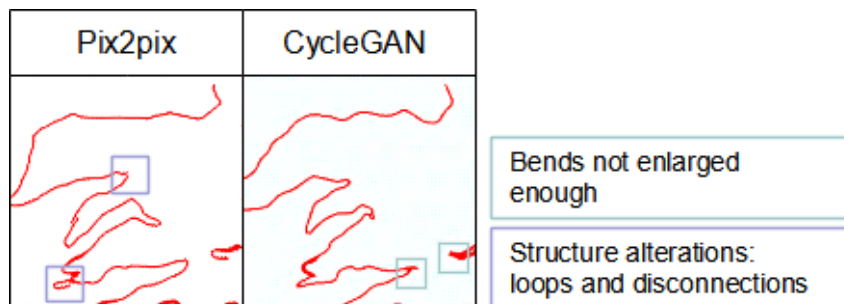


Figure E.8: Illustration of the main defaults of generation models with two predictions example.

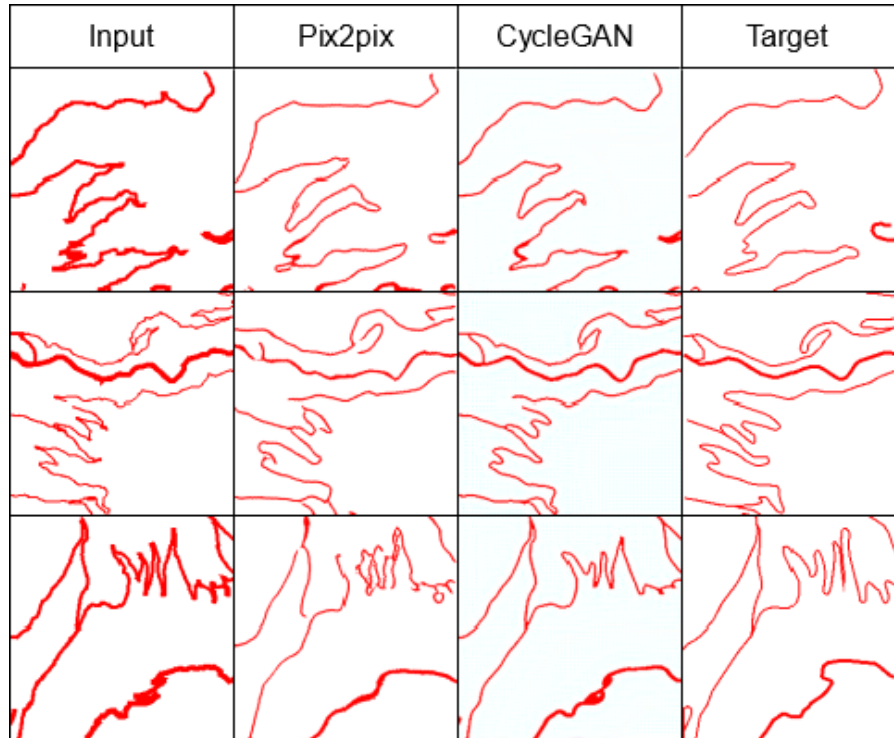


Figure E.9: Comparison of CycleGAN and pix2pix prediction in situations including an important bend series.

Figure E.9 focuses on bend series, which is the most challenging situation in mountain areas. They require bend enlargements, displacements, schematizations and simplifications. The first image shows that simple situations are as well generalized by both models. However, the next images show that despite the fact that it produces incredible images and provokes an important loss of information pix2pix proposes more enlargement and some simplifications, that cycleGAN do not generalize enough.

Figure E.10 focuses on intersections. The preservation of the main structure and shape of such an image seems learned in most cases; however, in the case of the second situation, the models fails because the input is not very legible at this resolution, and because of the presence of two bends close to the intersection. Pix2pix learns the intersection global shape, but, it creates a loop, while cycleGAN learns a solution that does not preserve the connection but has more road pixel correspondence with the input.

Figure E.11 represents situations near a major road; such situations are often those that require more displacement because the information on the map is more dense. This figure confirms that the displacement cannot be learned from tile datasets like ours because more context is needed.

Model improvement

The interest of a connectivity loss. Figure E.12 compares the prediction made with cycleGAN for the common objective function and for the objective function that includes our connectivity loss. The results of both models look similar. The global prediction quality appears to be not affected. Visual comparison of connectivity is not sufficient to say if the loss improves prediction. Most predictions are identical,

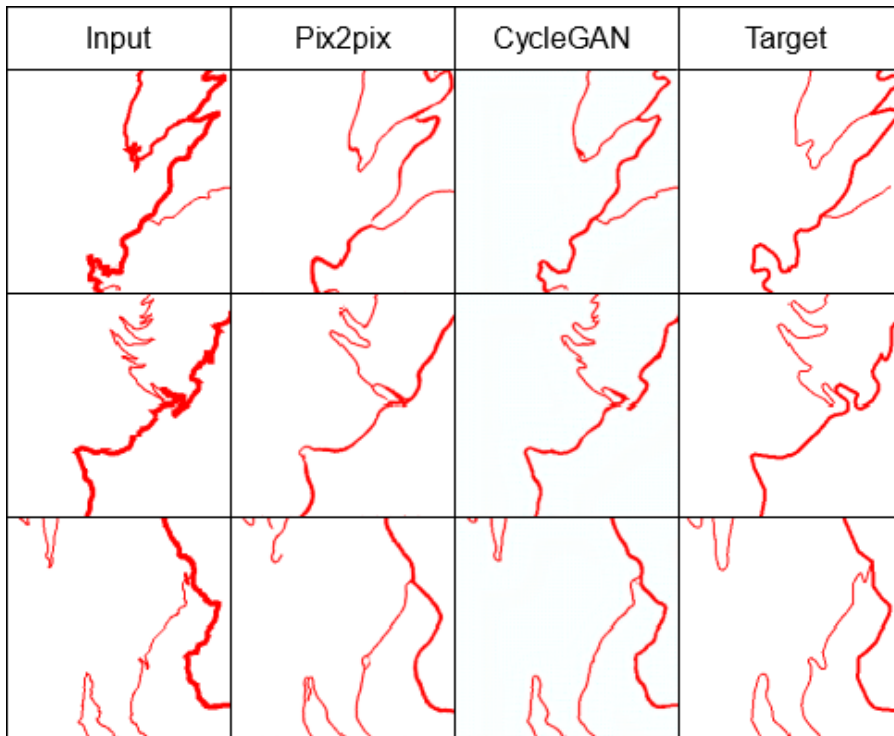


Figure E.10: Comparison of CycleGAN and pix2pix prediction in situations including an intersection.

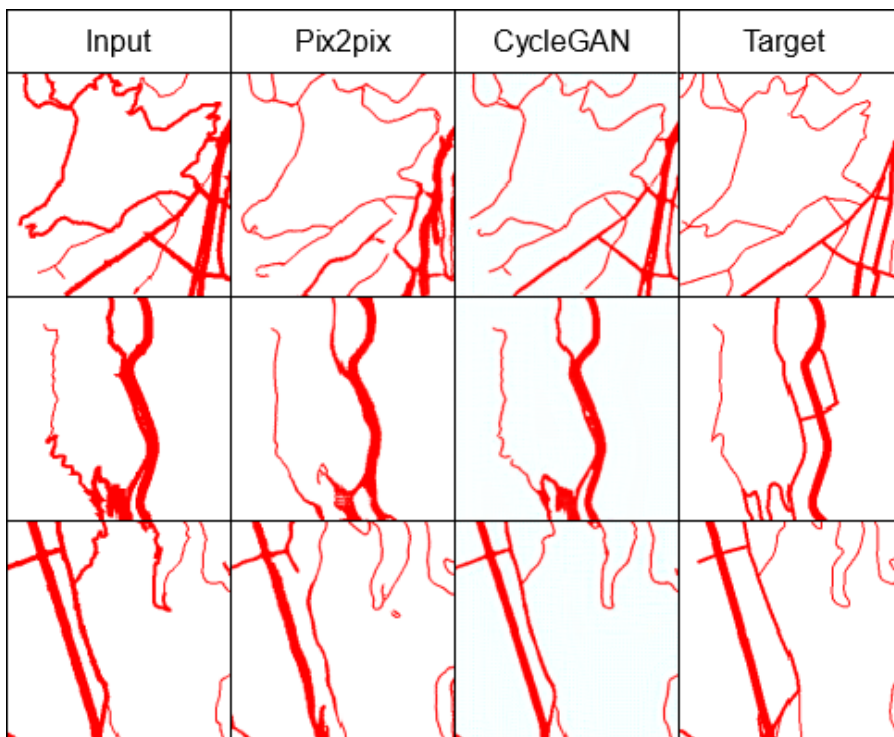


Figure E.11: Comparison of CycleGAN and pix2pix prediction in situations including a main road.

but the selected example in Figure E.12 shows some differences. For example, the connection of the main road with the image border of Situation 1 is better with the

connectivity loss, in Situation 2 the loop is avoided and one pixel is left between both roads.

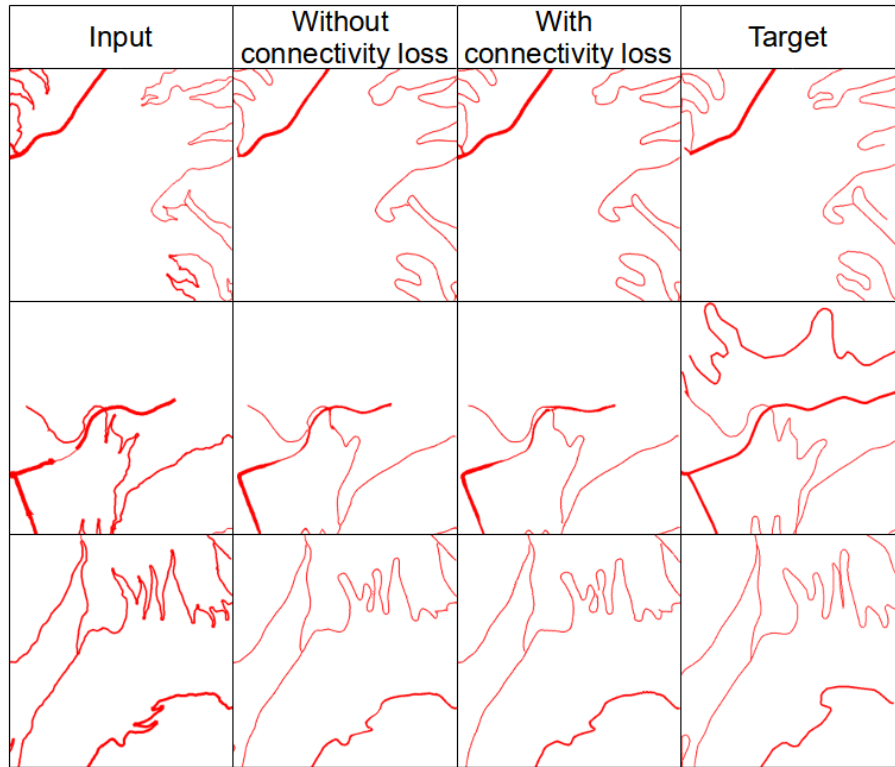


Figure E.12: Comparison of the model trained without and with the connectivity loss.

The measure of disconnections and loop creations decreases, but few changes are visible. This means that the loss instead of discouraging visible connectivity alteration discourages very small alterations that are not visible to humans (isolated road pixel on background or background on road). We also hypothesize that in some cases the number of background and roads pixel set does not reflect the connectivity alteration, as the change in background and in road sets can be balanced. For example, in situation 3 a loop and a disconnection are created, the loop increases the number of the background sets by one and the disconnection decreases this same number by one, so the measure detects no alteration while there are two. We think we could improve the loss by separate the two-aspect disconnection and loop creation and providing a more precise measure for each cases.

The interest of adding examples. Figure E.13 compares the models trained with 1223 images and 2271 images. We created the second dataset by manually matching and adding road images from several mountain areas in France. In this experiment, the benefit of adding training examples is not clear; models predictions are similar; we suppose that the differences in prediction are probably due to random effects during the learning process. In fact, we think that adding new images to the training set only improves the results if the images added are of good quality and add diversity to the training set that may not be the case in our experiment.

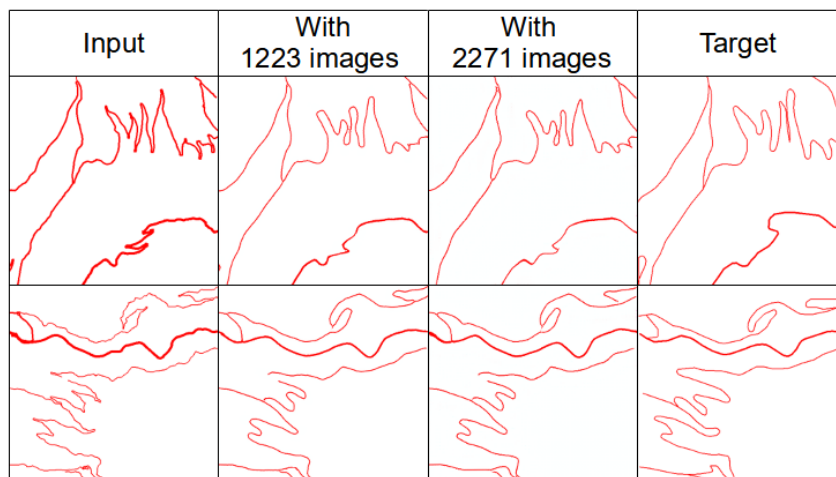


Figure E.13: Comparison of CycleGAN trained with different sizes of the training set.

E.3 Discussion

We discuss the learning strategy, the quality of the results, and the ability of deep learning to learn graphical generalization.

E.3.1 Learning strategy

This chapter contributes to highlighting the comparative interest of several learning strategies for object generalized shape prediction using an image-based approach.

Segmentation vs generation approach For the use case of mountain road, and unlike for building simplification (Feng, Thiemann, and Sester 2019), the GAN architecture gives much better results than a U-Net architecture. The discriminator forces the generated image to look like a road. The main difference between the two use cases is that buildings are mostly composed of angular shapes that GANs were unable to precisely reproduce, whereas in the road use case, a smooth shape is desired. Nevertheless, segmentation has some advantages, in particular, models train quicker and require fewer examples, they give a less fuzzy outline for objects and are more stable/easier to make converge.

Paired vs unpaired approach Another surprising result of our experiment is that the paired model (pix2pix) was less effective than the unpaired one (CycleGAN). In other words, it is easier to learn what a generalized map looks like than to learn how to generalize a map with mountain roads. The main flaw of images generated by the paired model is that the network structure is hardly preserved: the multiple loop creations and disconnections make the generated images highly unrealistic. In our training set, several parts of roads can appear in one image, several close roads and two disconnected parts of a unique road are really similar, this can cause the model to not learn to avoid disconnection. With the unpaired approach, the cycle consistency prevents generators from modifying the information conveyed in the image. Thus, this approach is less sensitive to disconnections and loop creation. Nevertheless, paired learning has two advantages: 1) it is able to learn

displacement, even if it cannot be well applied due to the lack of context of with the small tiles; and 2) it does not simplify the results too much. Finally, we hypothesize that a semisupervised approach, such as the one proposed in SMAPGAN (Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020), could help us get the advantages of both approaches, and maybe the most adapted for mountain road shape generalization.

E.3.2 What can we learn?

In this section, we go further to explore what can be learned by networks such as those employed in our experiment.

Can we learn the selection of roads? Figure E.14 compares the prediction of the segmentation model trained with initial images that contain all the roads and the prediction made with the same model trained with road selection already processes.

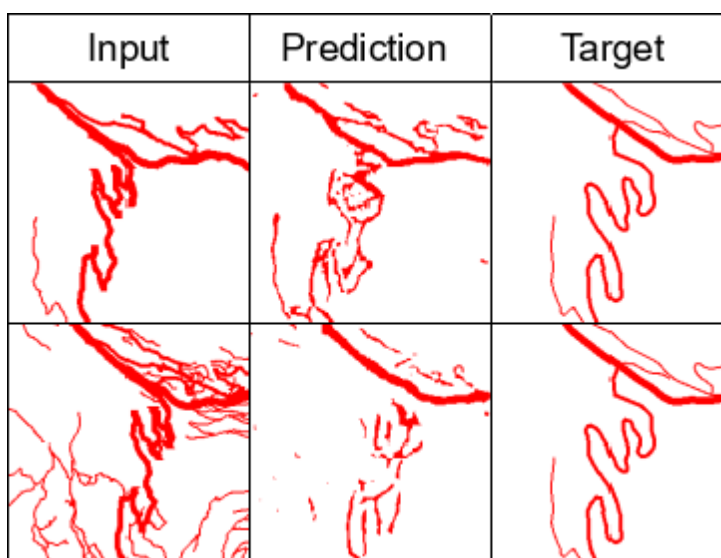


Figure E.14: Comparison of prediction when the model is trained (a) without the selection preprocess. (b) with the selection preprocess.

We observe that removing the unselected roads from the initial image seems to be a better solution, but the results in both cases are quite unsatisfactory for this road. We can note that the model learns to remove the unimportant roads, but does not do it properly (some road parts remain, and one important road on top is partially erased). Road selection is based on shape characteristics (Thomson and Richardson 1999), semantics features (importance, status, width, etc.), and the role played in the road network (Garcia-Balboa and Ariza-López 2008). Most of these elements are not accessible in a small image of the map, image representation is not really adapted for this step (Touya, Zhang, and Lokhat 2019). Moreover, many other methods succeed in combining these factors and properly select roads (e.g., Touya 2010; Benz and Weibel 2014, etc.).

Generalization of roads and rivers A comparison with previous experiments on buildings has shown that the appropriate deep learning architecture is closely

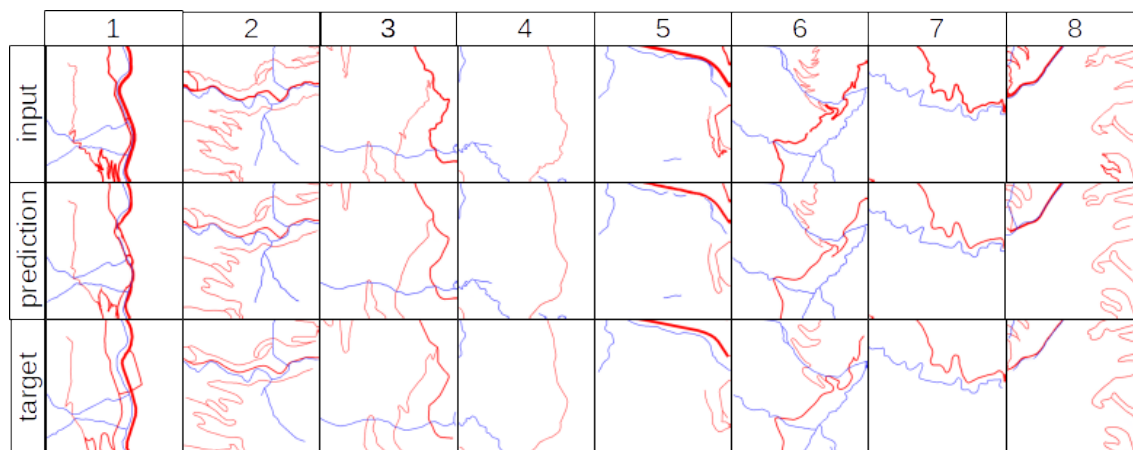


Figure E.15: Example of CycleGAN predictions for images with roads and rivers.

related to the use case, and our method cannot be directly applied to other cartographic objects. To assess the ability of the proposed method to generalize other objects, we experimented with the joint generalization of roads and rivers. The geometric generalization of both map themes has similarities, but rivers naturally have smoother shapes and require less generalization. Adding rivers should provide more context in tiles, which should explain the displacement frequently applied to roads to avoid overlaps between roads and rivers. However, more context will produce more complex situations and make the task more complex.

For this experiment, CycleGAN was trained with the same datasets, but the image tiles contained blue lines representing the rivers, under the road lines (red pixels are visible when roads and rivers intersect). Some results are presented in Figure E.15. We observe that roads are globally as well generalized as when they are alone. Then, rivers are really slightly generalized as expected, and the model learns that roads and rivers have to be generalized differently. We can see very few local errors due to a very coalescent input image, such as the road bend that is not enlarged enough in Situation 6, or the island creation in Situation 4. Then, the relations between roads and rivers seem to be a problem for the model. In Situation 1, the enlargement to avoid symbol overlap is not applied, and some parts of the rivers are erased. It seems that our model has learned that roads and rivers do not have to overlap, but did not learn how to avoid it. We can find an explanation for this problem in our input data: only half of our target training data present a significant displacement to avoid symbol overlap between roads and rivers (1,5,7), and the other half only propose a very small displacement (2), or no displacement at all (8). It reflects some generalization problems in our target dataset. The GAN logically learns to generate badly generalized maps from badly generalized training examples. A correction of the 1:250k dataset would be necessary to check if CycleGAN is able to learn how to generate images where roads and rivers have been displaced from each other. However, this use case has shown the ability of deep neural networks to learn the generalization of several different objects at the same time.

Finally, we believe that a graph approach is most adapted to learning the relation between objects, and especially this displacement while the image approach is more relevant to learning the shape of the generalized object. Thus, we think for a complete generalization, a mixed deep learning approach would benefit from image and graph information and generalize together several objects.

E.3.3 Remaining limitations

In this section, we discuss the remaining errors and try to explain their causes. The user test (see Chapter H) has shown that the images generated using cycleGAN are considered as legitimate as the reference map. Nevertheless, these results are not really adapted for cartographic production for the following reasons:

- Some important loops and disconnections remain in some tiles, while the preservation of road structure is a hard constraint compared to the other constraints well managed by the model.
- The output tiles are not a usable map, and post-processes are necessary to integrate these predictions in a map or in a more classical generalization process (see Chapter F).
- Some noise can occur; in particular, we observe small blue dots that are barely visible in the output images. This color noise has to be post-processed.
- Only roads are generalized in the experiment, and a complete topographic map is made of other topographic features, and these features (e.g., nearby buildings, relief, or rivers) may even influence how roads are generalized.

One of the main reasons for loops and disconnection is that training images contain too much detail for the image resolution. This complexity in our input images has at least three consequences: (1) we need more examples to train the model; (2) we need larger images or images with a better resolution; and (3) the network needs to be much deeper than our current setting to understand these images. Our proposed connectivity measure penalizes these errors without completely avoiding them.

The second problem is the displacement failure. First, to avoid overlapping of two roads (e.g. parralele roads), they have to be displaced from each other, but this case is too rare in our example set to be learnt. Then, to learn to move the roads from other obstacle, we need to know that there are other close objects on the map. Therefore, this information must be represented in tiles, and we need a representative example for this displacement.

The third problem is the alteration of visible connectivity in some images. It consists of the appearance of a few background pixels in the middle of some roads, which does not much alter the quantitative evaluation but clearly degrades the legibility of the road. This problem is less present if we augment the number of epochs but still occurs. Our proposed connectivity loss also penalizes such disconnections without completely avoiding them. Then, the only post-process that could restore connectivity would be to convert the roads back to vectors and to use a connectivity check (Touya 2010).

Our work also experienced limitations that are not errors.

- The evaluation measure and associated loss function should be improved; this may allow one to make the connectivity loss more efficient and avoid loop and disconnection in the output, or other measures used as the loss may improve the prediction in another manner.
- Our dataset has a limited size, and it does not contain enough examples of very narrow and sinuous bend series.

- We also faced computation time and memory limitations. Indeed, this thesis was done during the covid19 pandemic and the different lockdowns forced me to work several months from my personal computer and thus to give up some experiments.

Chapter's conclusions

In conclusion, the models we proposed correctly achieves smoothing, enlargement, and caricature operations in most of the cases. It shows that most of the knowledge behind road generalization can be captured by a convolution neural network. Thus, this chapter has shown the ability of deep learning methods to achieve generalization tasks from images derived from vector data. The modelisation as a segmentation tasks or as a generation tasks and between paired and unpaired image has shown different capabilities for generalization summarized in Table E.1.

Approach	Realism	Credibility	Legibility	Preservation
Segmentation	✓	✗	✗	✗
Paired generation	✓	✗	✓	✓
Unpaired generation	✓	✓	✓	✗

Table E.1: Comparison of approach for image-based graphic generalization.

The unpaired generative approach is preferred by the evaluators and give results considered as good as the reference map. However, some important local issues remain, and the generated images cannot be used as end-product topographic maps yet. The post-process and integration of the predicted road into a final map have to be investigated.

This chapter also contributes to illustrate the difficulty in creating an adapted training set and an adapted deep learning model, even with an image representation. In particular, the large gap of scale make tiles suffering from a legibility problem that blurs the input shape of the road. We proposed several tile creation and improvement methods. The training set improvement affects the global quality of prediction, while improving the loss function rather focus on correcting some specific errors.

In addition, this experiment can be extended by adding multiple geographic themes to the training datasets to predict a complete map. This perspective would have to deal with spatial relations in a dense context, such as urban areas and multiple representations of objects, and is explored in Chapter F.

Chapter F

The generation of topographic map with several cartographic themes

The challenges of deep learning-based generalization are numerous, but the first step is to address a more global approach with multiple cartographic themes generalized at the same time on the map. This chapter presents some first experiments to test the suitability of generative adversarial networks (GAN) for the generation of generalized maps. It is a step towards the generation of complete generalized maps.

F.1 Materials and methods

F.1.1 Approach

In our experiment, we trained GAN models to produce a generalized map from an image of detailed geographic information. This approach is illustrated in Figure F.1.

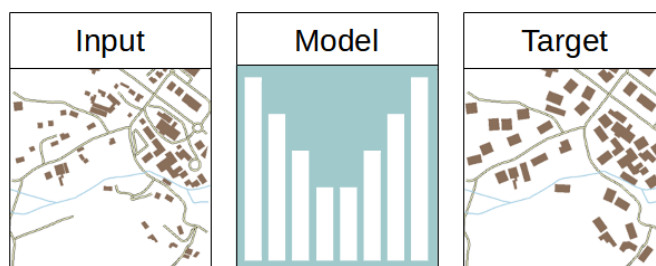


Figure F.1: Simple illustration of our approach for generalized map generation.

Our goal is to generate maps of urban areas at a medium-large scale (e.g. 1:50,000). This is a complex scale, with a difficult balance to find between the buildings enlargements and the constraints of minimum separation between buildings and with road symbols. At smaller scales, the buildings are no longer shown individually, and at the larger scales, they do not need to be generalized to much. The input data are extracted from detailed maps at 1:25,000, whereas the target scale is 1:50,000. These maps mainly contain buildings, roads, and rivers, and the representation of these elements should maintain an adequate level of detail to be legible and prevent symbol overlap. We chose a study area that includes a dense context of urban and rural areas, located near Saint-Jean-de-Luz, which covers 30*15 kilometers in the southwest of France. We use detailed buildings, roads, and

rivers from topographic datasets produced by the French National Mapping Agency (IGN) as input, and we expect as an output of the GAN a legible map of these elements at a target scale of 1:50,000.

This level of detail requires reducing the amount of information while preserving the patterns and main relations. Usually, roads and rivers are generalized first, because they partition and structure the space (Ruas and Plazanet 1996). They may be selected, and some patterns may be schematized. However, the shape and position vary very little at this scale. Then the buildings have to be enlarged and simplified to make them legible. As the space in each block is limited, the buildings also have to be displaced, typified (density is reduced while preserving patterns), amalgamated, etc. Finally, it may be necessary to cover dense inner cities with polygons of built-up areas when these areas cannot be represented by enlarged individual buildings (Touya and Dumont 2017).

The generalization can be performed by several algorithms, each focusing on some aspect and giving less or more importance to each constraint. Thus, we compare how deep learning is able to learn to reproduce the result of two alternative generalization processes. Both have been proposed to be progressive intermediate representations between 1:25,000 and 1:100,000 scales by (Touya and Dumont 2017). We compared the AGENT process (Barrault et al. 2001), that focuses on building end position, and the typification-based generalization (Burghardt and Cecconi 2007), that rather focus on building patterns and relation preservation.

In our use case, generalized road and river are really similar with the input. Then, network is selected and complexe interesections and situations are simplified (e.g. round-about), but the global shape does not change. Generalized buildings must satisfy the following classical constraints:

- (C_1) Buildings should be larger than a minimum size ($25mm^2$);
- (C_2) the smallest edge of the buildings should be greater than a minimum value ($2mm$);
- (C_3) The buildings should not be too close to the roads symbols ($2mm$);
- (C_4) The buildings should not be too close to each other ($2mm$);
- (C_5) The density of buildings in a block should remain stable;
- (C_6) Building patterns, such as alignments, should be preserved;
- (C_7) Topological relations should be preserved; for example, the buildings should remain in the same block.

Built-up areas must be areas in the city center, where the density of construction is important compared to the rest of the city, it might be one or more blocks and constitute a coherent area, which is interpreted by map users as a whole with a meaning of "city center".

F.1.2 Implementation

Dataset

We created square tiles of map image each of size 512x512 pixels, and representing 500x500m² on the ground (this means that each pixel represent around 0.98m). Table

F.1 represent the constraint threshold in pixel.

Table F.1: Table of constraint threshold in pixel for a display at 1:50 000 scale.

Name	Constraint	Threshold in mm on map	Threshold in pixel
Size threshold	C_1	25mm ²	256
Granularity threshold	C_2	2mm	102
Separation threshold	C_3 and C_4	2mm	102

These dimensions guarantee a legible situation for both input and target data, and it is small enough to build a training set with around 2,700 images. We randomly extracted 100 tiles to evaluate the model (test set), we make sure afterwards that the different building densities are represented in these 100 images.

Input representation The input tiles are created using the detailed dataset. For the sake of simplicity, we only extracted the main themes of the map (i.e., buildings, roads, and rivers). We proposed and tested two representations of input data as image tiles (see C.22):

- **symbolized representation:** the geographic information is represented using a cartographic style (the same as the target style),
- **layered representation:** the shape and position of each theme are encoded as a binary value for each pixel in one layer of the image.

Target The output data are manually corrected to remove errors due to the automatic processes used. In particular, some overlaps between buildings and rivers and buildings and roads were removed.

The output information is always symbolized. In a first quick test we tried to mimics the style of *planIGN* (Figure F.2), and use the same colors in our target scale images. However, we observe that the two browns used for buildings and built-up areas are very close, and model confused both elements. Therefore, we decided to use our own style with more contrast.

In our target style of map, we symbolized the roads with a borderline with sizes and colors that vary according to their importance, and the buildings are represented in brown or gray according to their nature. We did not represent the building outlines to reduce overlaps between symbols. For simplicity, we decided to preserve a unique symbol for each road (the yellow one) and building (the brown one). The covering of dense blocks (or graying) is symbolized in lighth brown.

Model

We explore and compare the potential of paired and unpaired GANs, using pix2pix and cycleGAN PyTorch implementation¹. We believe that the preservation of geographic information should be better learned by a paired model, while the increase in legibility can be learned from both methods, as they only need to learn how each domain looks like.

¹Available here <https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix.git>.



Figure F.2: Extract of Plan IGN, inspiration for our target representation.

Objective improvement Then these networks are not specifically designed for geographic data, so adjustments to default architectures may be required. We tested two loss functions that aim to improve the shape preservation in map (see C.1.4 Fu, Gong, Wang, Batmanghelich, Zhang, and Tao 2019; Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020).

Fusion GAN As mentioned in the dataset section we experimented with the addition of a context tensor. Hazırbaş, Ma, Domokos, and Cremers (2016) propose an architecture to fuse a colored image and a depth map and generate a segmentation of indoor scene. Our idea is to use this architecture for map generation: the architecture fuse the input location and shape of map objects with the additional information and then generate the generalized map. The architecture is composed of two encoding paths (one for each input tensor, linked by "fuse connections", which sum the encoded additional information to the input, at different stages. We used this architecture as the generator of our GAN to replace the U-Net or ResNet (which are traditional GAN generator). Thus, our FusionGAN uses the main and additional information to predict images that are evaluated by a discriminator trained to learn if the image of the generalized map looks realistic (see Figure F.3).

F.2 Results

In this section, we present the results of the experiment and visually evaluate (using the criteria C_1 to C_7 from F.1.1). The prediction presented are issues from our test set, the complete results are available here <https://doi.org/10.5281/zenodo.7244047>.

F.2.1 General results

Overall quality In the results presentation, the not dense and dense areas, where block covering is necessary, are separated in two figure (Figure F.4 for not dense areas and Figure F.6 for dense areas) because expectations and generalization mechanisms in both situations are really different.

In Figure F.4 we observe that the predicted images look like the target map, the style is consistent with our expectations, and the predictions are credible map

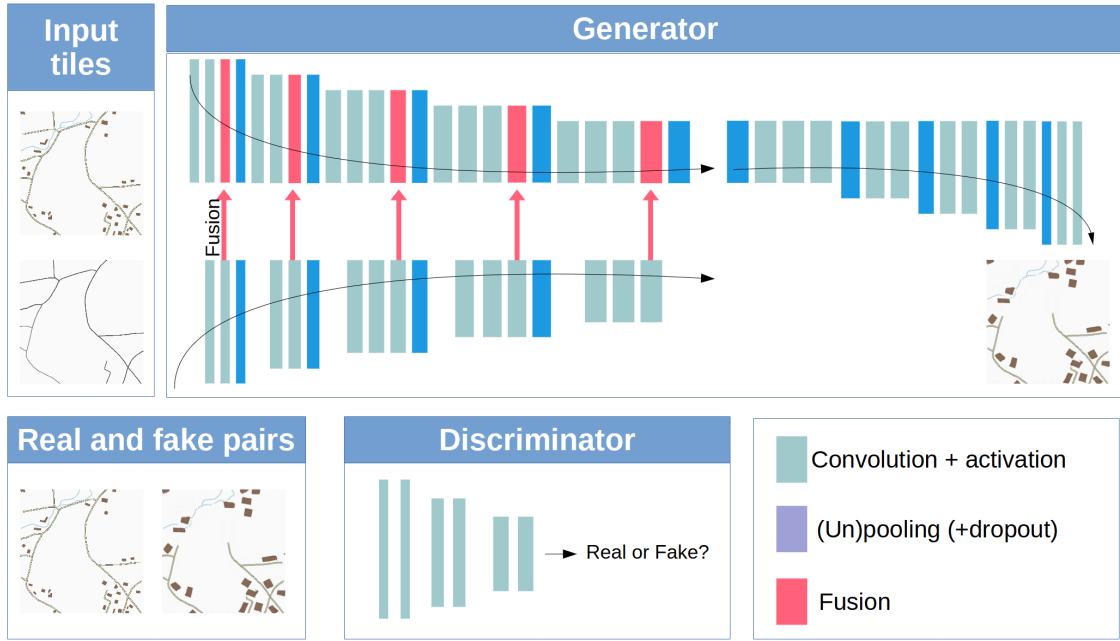


Figure F.3: The fusion GAN architecture.

images. However, some complex rivers suffer from fading (We observe in another test that this problem is simply solved by enlarging the water symbol on examples tiles). Roads are also globally preserved.

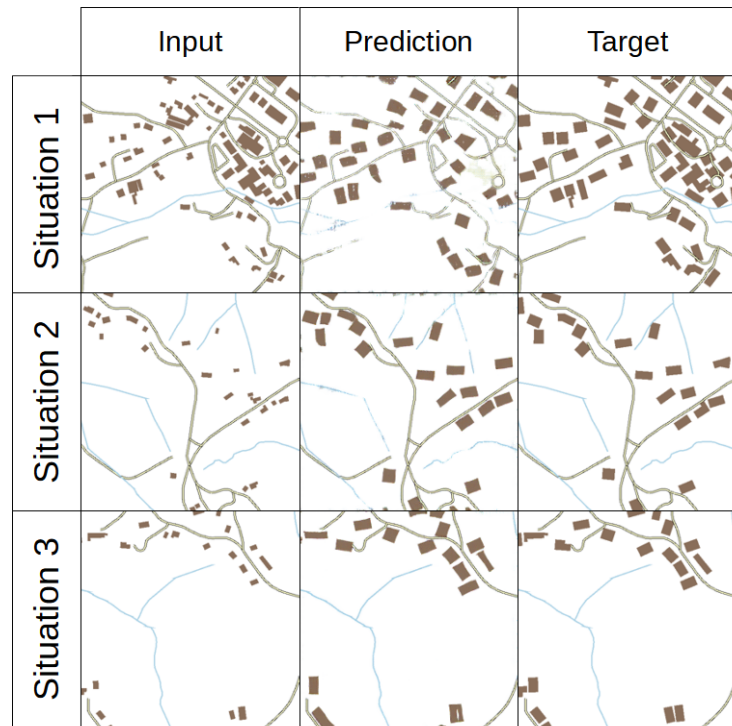


Figure F.4: Pix2pix prediction for generalized map generation.

First, at the individual building level, or micro-level according to (Barrault et al. 2001), we observe that the building size constraint ($C1$) is always satisfied, buildings are large enough, even a bit larger than in the reference. But the granularity

constraint ($C2$) is somewhat less satisfied: the shape is simplified, but sometimes blurred, which is consistent with previous attempts to generalize buildings with GANs and U-Nets (Feng, Thiemann, and Sester 2019; Kang, Rao, Wang, Peng, Gao, and Zhang 2020). We can also observe that some rectangles can be distorted and have an unrealistic shape. Moreover, some buildings are over-simplified and lose their distinctive shape: for example, some L-shaped buildings are transformed into a rectangle (see Figure F.5).

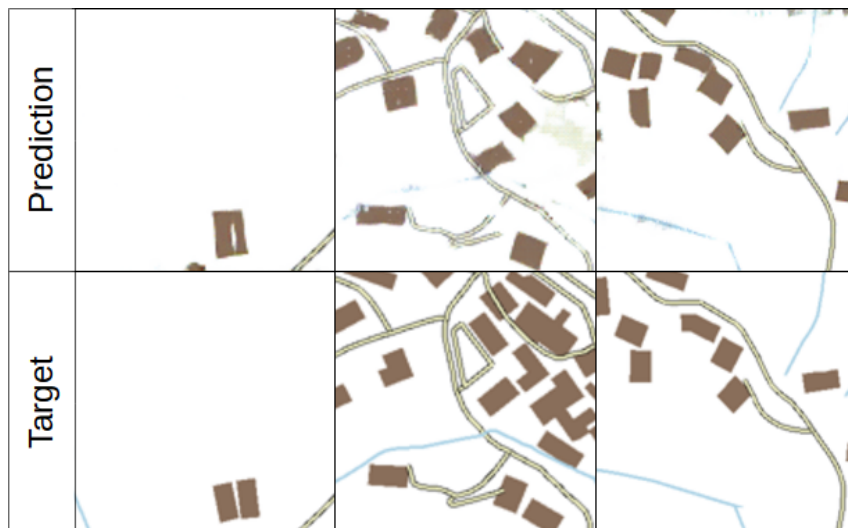


Figure F.5: Some zoomed image on prediction and target, the predicted building shape may be blurred, irregular or not credible.

The global legibility of the images is satisfactory: most of the buildings are sufficiently separated from the roads ($C3$) and do not overlap each other ($C4$). Building density is well preserved globally ($C5$). However, some parts of the map (e.g. in Situation 1) suffer from over-reduction of density. The remaining overlap problems are mainly induced by similar errors in the target images. Then, the most important challenge is to preserve the relations between different geographic objects ($C6$ and $C7$). For the road-building and river-building relations, the relative orientation, topology most of the time, and proximity relations are preserved. However, we do observe some building-road overlaps, and some inclusions in a small block disappear (i.e. some buildings change block). Grids and alignment are really rare in our dataset and thus it is difficult to evaluate the quality of their preservation, this point is later discussed in next section.

In city centers, the most important challenge is to transform dense urban blocks into built-up areas. Figure F.6 illustrates some prediction in a dense area where graying may be required. We observe that the model tends to erase most of the buildings but does not apply a gray area in the whole area. For some images, the predicted covering corresponds to the target (e.g., Situation 4), while for some other examples, unexpected parts of built-up areas are covered, and some covering are missing (Situations 5 and 6). We believe that there are three possible causes for this problem:

1. The image is not sufficient to learn the block covering. The density and shape of the buildings in a block are important criteria for a block-covering decision (Touya and Dumont 2017), and both are visible in the images, but they are not

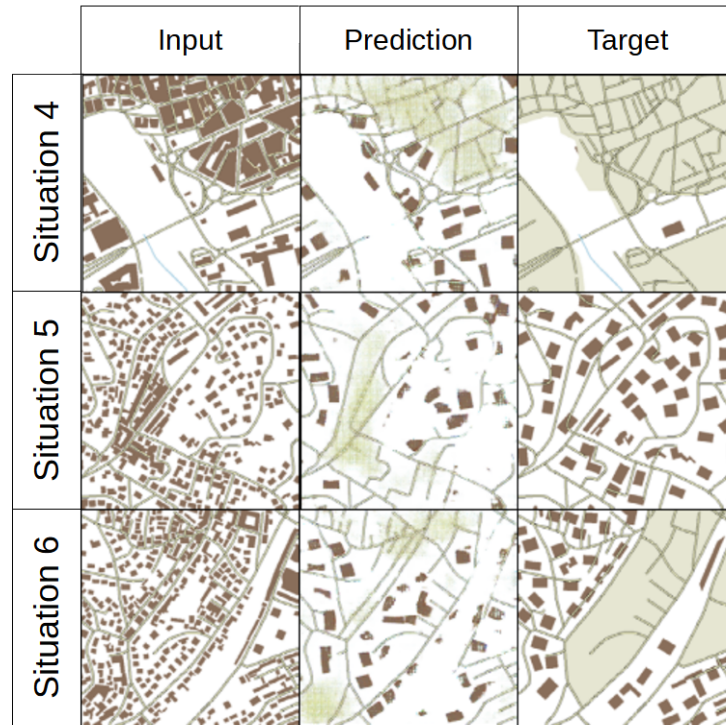


Figure F.6: Pix2pix prediction for generalized map generation in city center areas.

sufficient. For example, on Situation 6 the central block needs to be covered, while the block up to it has a similar aspect and does not. The covering decision might be due to touching blocks outside the image, which are also covered.

2. The image tiles might not show the complete block. Coverage is applied to a complete block, but nothing guarantees that the image shows the complete block. For example, in Situation 6 the graying is not predicted because only a short part of the expected covered area is visible, and this short part is not really dense to justify a covering decision.
3. There are not enough examples of block covering in the training set. Our training set is composed of 2,623 images with very few (57) containing block covering.

Paired vs unpaired approach Figure F.7 compares the prediction of paired (pix2pix) and unpaired (cycleGAN) models. The first column presents the input images, the second and third columns present the predicted image for each approach, and the last column presents the target generalization. We observe that the paired (pix2pix) and unpaired (CycleGAN) results are really similar. Contrary to the use case of mountain roads (see Chapter E), pix2pix seems more promising for this experiment, the shape of buildings is quite more regular with pix2pix and this model took less time to train. Thus, we use this model for the next experiments.

Typification vs Agent target In map generalization, there is often no unique optimal solution, as most of the time it is not possible to satisfy all constraints at the same time, and compromises are necessary. Different algorithms will focus

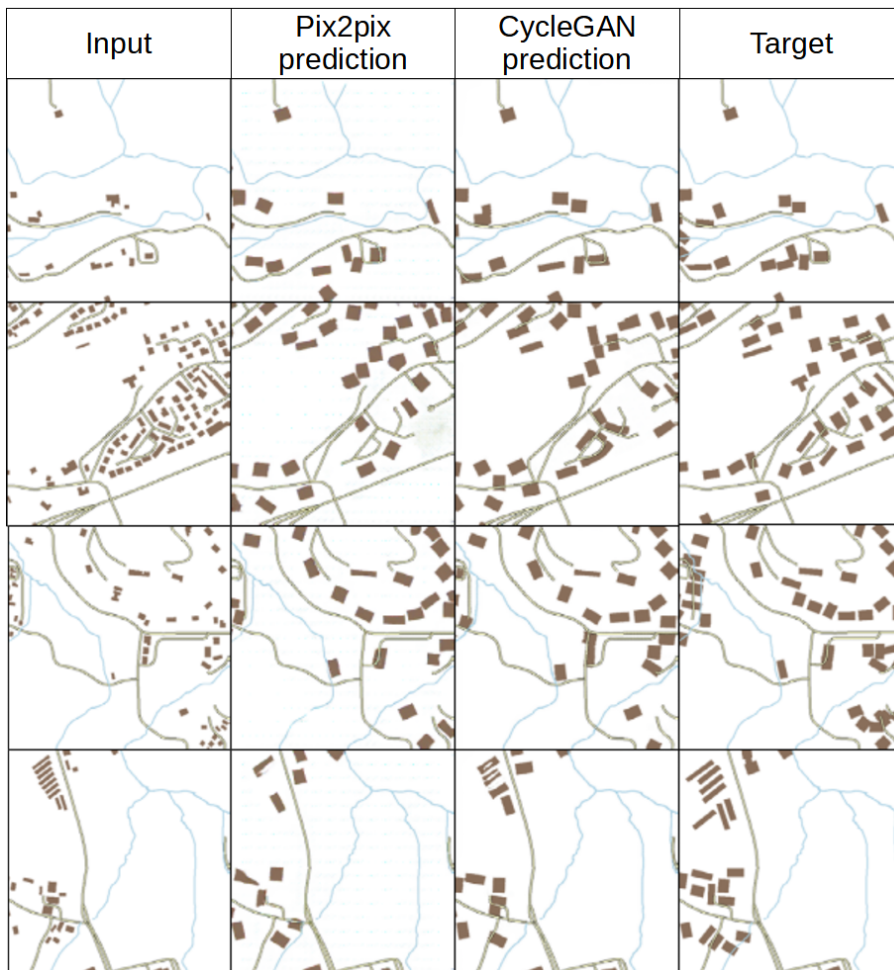


Figure F.7: Examples of results from pix2pix (paired) and CycleGAN (unpaired) to generate generalized maps.

on solving different constraints and will give different results. We want to verify whether the model can learn to focus on different generalization constraints when we train it with data produced with different generalization algorithms. Thus, we compare the prediction targets of the trained model with the agent-based (Barrault et al. 2001) and the typification-based (Burghardt and Cecconi 2007) generalization (Figure F.8).

We observe that predicted alignments and other building structures are better preserved by model trained with example from typification-based method; however, the shapes are less regular and the minimum separation between buildings is violated more frequently. The preservation of structures and alignments is one of the big strength of typification process. The fact that these elements are better preserved with the model trained from typification examples is a first hint to conclude that the network learns to reproduce generalization strategies from the examples and prioritizes the same constraints as the generalization used for the target data.

F.2.2 Improving the default models

In this section, we present the experimented models and datasets to improve the generation of generalized maps.

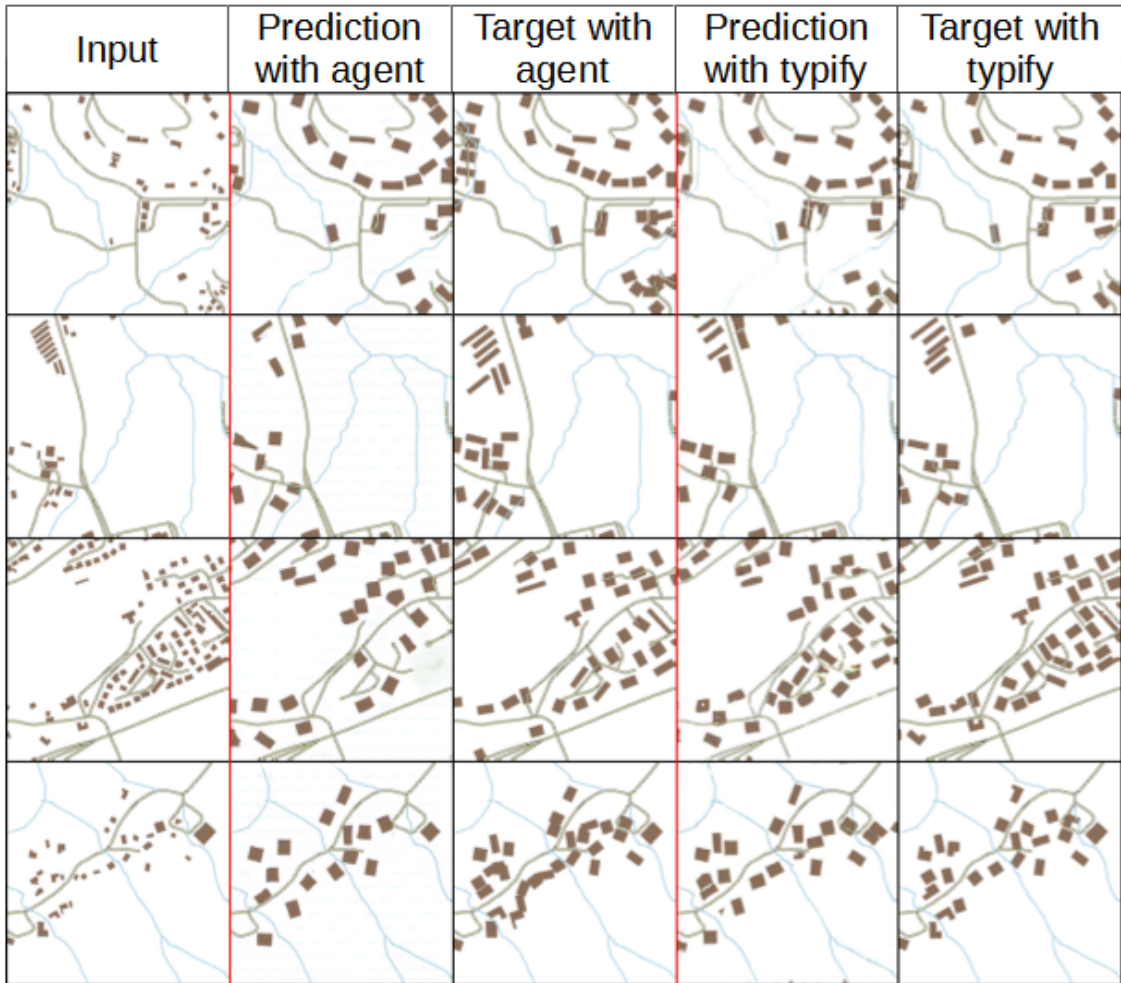


Figure F.8: Comparison of images generated by pix2pix trained with agent-based generalization, and typification-based generalization.

Input representation

First, we tested using a layered representation of the input representation that limits the image bias and simplifies the input preparation for a computer. Figure F.9 compares the symbolized and the layered representations to learn the generation of generalized maps. We only represented the symbolized input in the figure because it is more convenient for human readers.

We observe that the results are quite similar in most cases (e.g. Situation 1). However, in dense parts of the map (Situation 2), the prediction of layered method contains less undesired fading and overlapping, and the information is less blurred and more preserved. Finally, in Situation 3, the layered method reduces overlapping between buildings and roads, and the prediction is a generalization as relevant as the target. So, the layered representation seems to be useful in the ambiguous case, where the input information is dense or overlapping, and it does neither decrease the quality of prediction nor increase the time required for training (for a three-layer image, with roads, rivers and buildings).

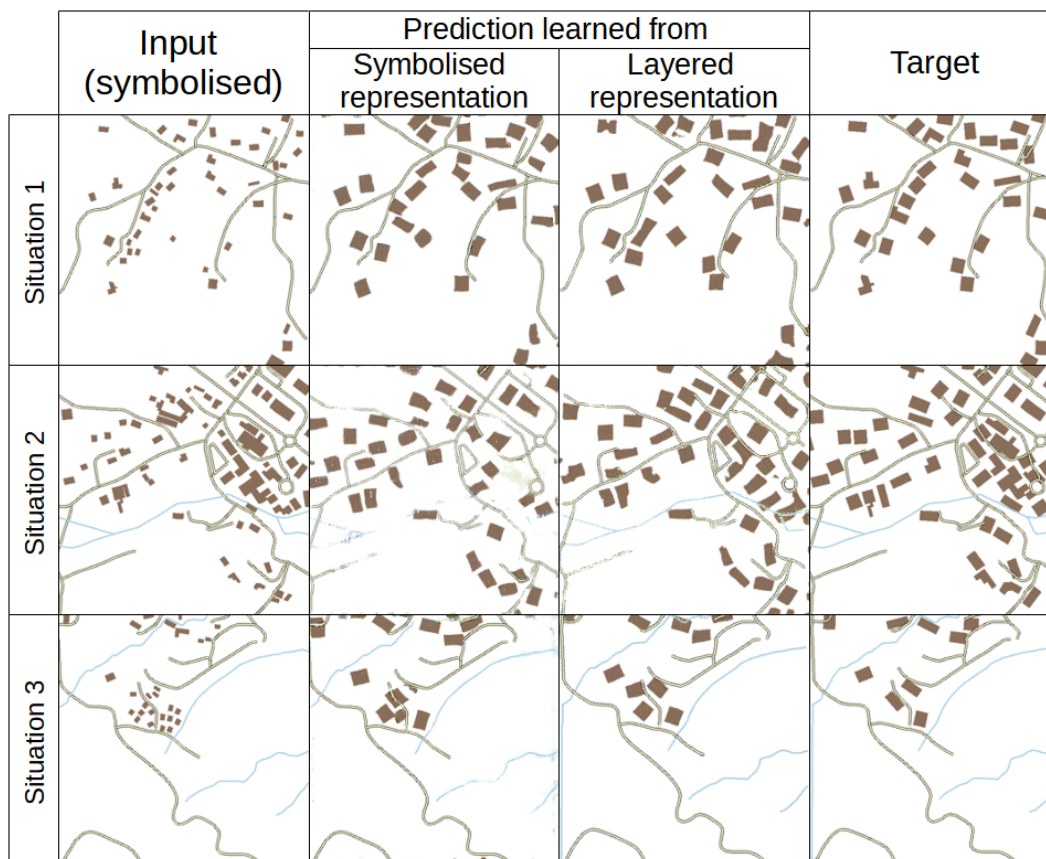


Figure F.9: Comparison of images generated by pix2pix trained with layered and symbolised representation as input.

Road selection

Our previous experiment, despite good results, suffers from a major limitation: the road network is not selected during the process. We did not experiment to learn this selection because we think that the image alone is not sufficient for this generalization, indeed it is a multicriteria decision based on road attributes, and section role in the global road network. Figure F.10 illustrates this problem, showing the prediction of a model trained to produce maps that include road selection. We observe that the model understands that it has to erase some roads, but not which one or what a complete road is (which pixel belongs to the same object). The other themes of the map are also badly generalized. Therefore, road selection widely increases the complexity of the target task.

To allow the model to also learn the selection of the roads, we integrated additional information that should give hints on which road to eliminate, and we use this additional information with a Fusion GAN. In our experiment, this additional information is the prediction of the probability of road selection from Chapter D, computed with a GCN. In Figure F.11 we present the prediction made by the Fusion GAN model. We observe that the results are more blurred and less legible than the initial model (when selection is not required), but more credible than the predictions without the additional information (when selection is required). The roads are not correctly/completely erased, but the model tends to erase more often the less important roads, leaving the important roads. It seems that the model learns which



Figure F.10: Prediction of a model trained to generate a generalized map that includes road selection.

pixels are more important to keep and erase, but it still does not learn to find the roads limits and to applies the same decision to a whole road.

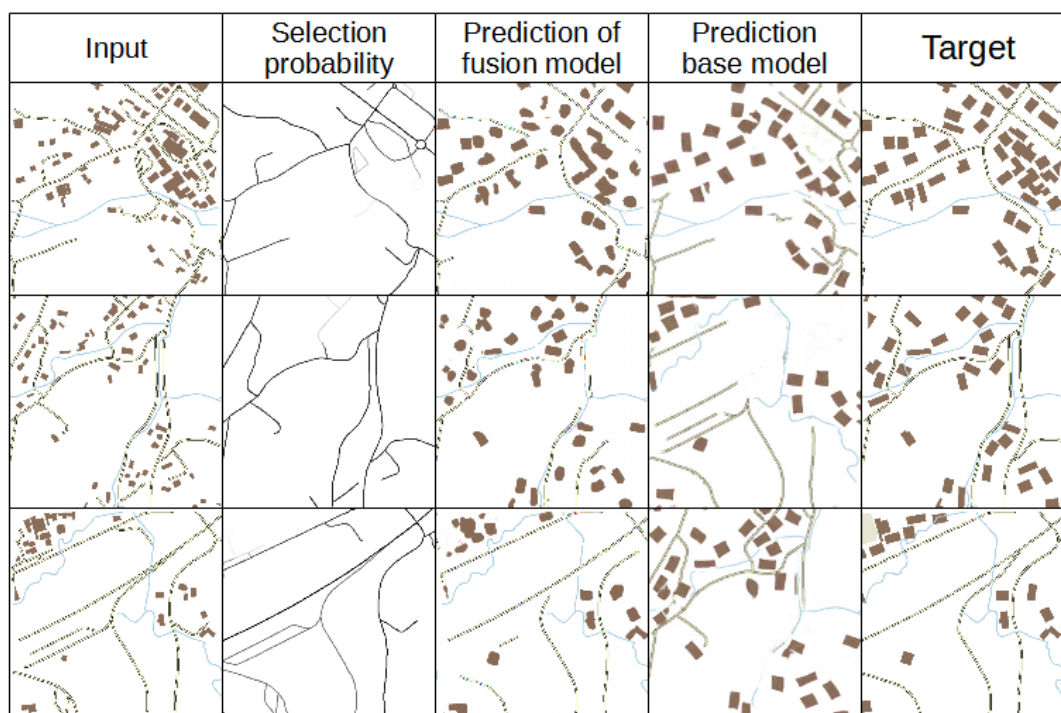


Figure F.11: Comparison of prediction of base model and fusion model with road selection probability information to generate a generalized map that includes road selection.

Alignments preservation

The preservation of alignments is a classical challenge in map generalization at this scale. As there are very few alignments in our dataset, we cannot strictly verify how the models learn alignment generalization. Thus, we try to apply the model trained on our dataset to a test set in another area with more alignment (see Figure F.12). We observe that the prediction does not preserve all alignments, the building orientations, shape, and spacing does not vary very much but they vary irregularly across the buildings making the alignment not salient anymore for humans. We present for comparison a manual generalization of both area that illustrate an acceptable generalization of the situation. We expect the model to preserve alignment and the classical struce of building in two lines in each blocks.

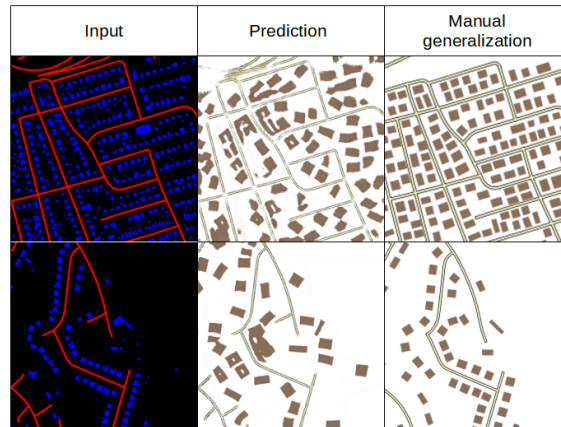


Figure F.12: Prediction on test area with many strong alignments.

To improve alignment preservation during the generation of a generalized topographic map, we first tested the fusion of additional information that would help locate the alignment (e.g., alignment axes, see Figure F.13). The alignments have been manually annotated to obtain this additional information. However, none of these representations improve the prediction, moreover, some decrease the quality of the generated maps. Unlike the selection probability that improve prediction, the different variants we proposed do not help the model to know how to generalize the alignments. Indeed, the alignment location contrary to our other proposal of the additional information, does not aim to add information not visible in the image such as context or attribute value. It is more an indication of areas that require more care during the generation process.

Thus, we tried several strategies to teach the model to give more importance to such areas.

First, we tested a weighted loss that gives more importance to errors in aligned areas. We define for each input image X_i the map W_i of the pixel that belongs to an alignment with eq. F.1. Then we train the model with $WL1$ (see eq. F.2) loss instead of $L1$ (see eq. C.2)

$$W_i = \begin{cases} 1.25, & \text{if } i \text{ belong to an alignment} \\ 1, & \text{otherwise.} \end{cases} \quad (\text{F.1})$$

$$WL1 = \sum_{i=1}^n W_i \times |Y_i - \hat{Y}_i| \quad (\text{F.2})$$

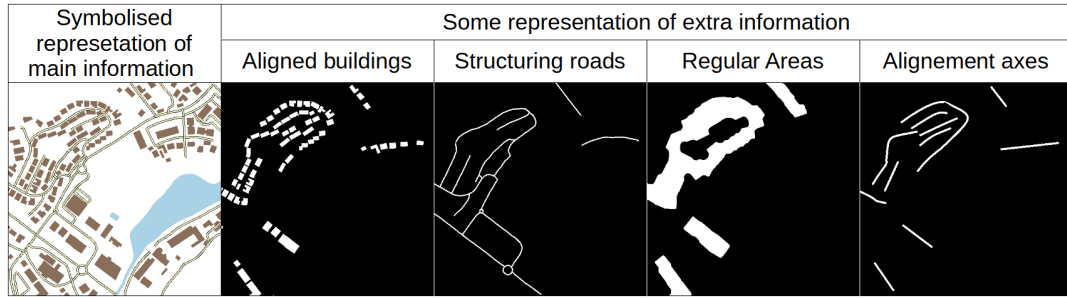


Figure F.13: Some proposals to represent the alignment on an additional image.

Secondly, we try to fine-tune the model with only tiles that contains alignments (i.e., continuing to train the model for a few only epochs with the examples of the particular situation we want to fit to).

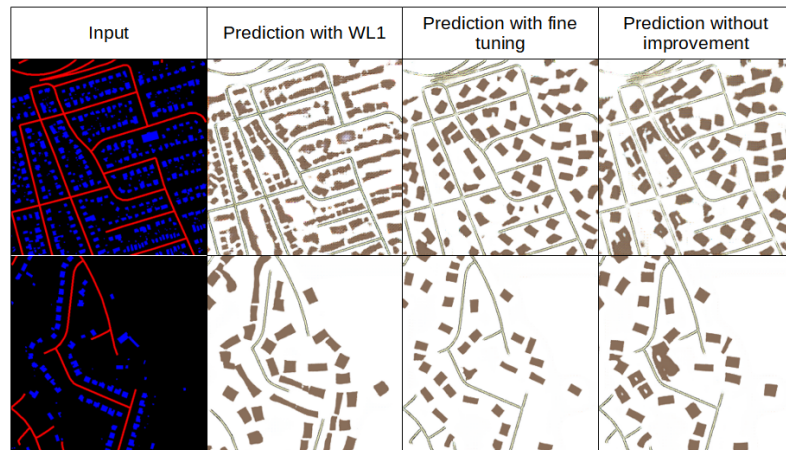


Figure F.14: Prediction of the two proposed alignment improvement methods.

Figure F.14 presents the results of both of these improvements. We observe that both solutions work well to improve alignment prediction quality in the not-dense case. For the very dense case (Situation 1) the weighted model is less efficient because such images are very rare in our input data, the errors are more linked to the extreme density than to the presence of alignments. Both solutions make the model able to learn the alignment generalization; however, both have limitations. First, the weight objective requires annotated training data; second, the fine-tuning method requires to train the model a few times instead of once and to have some reference generalized data with the desired characteristics (here alignments).

An adapted objective function for map generation?

Finally, we tested the interest of two loss functions that aim to improve the prediction of maps from the literature review.

Geometry consistency First, we tested the geometry consistency loss (from Fu, Gong, Wang, Batmanghelich, Zhang, and Tao 2019) that encourages simple geometric transformations to not change the semantic structure of the image. This loss is presented with more detail in Chapter A. It is coupled with an unpaired GAN has

shown the potential to make a more regular predictions of map (in style of google map without generalization) and simplify learning by reducing the space of possible solutions. This loss is based on the comparison of several generalizations for one building (with different orientations); thus, we hope to force this irregular prediction to converge to a regular shape. Therefore, the fuzzy outline of the buildings may be reduced. Figure F.15 presents the results of our model trained with this loss.

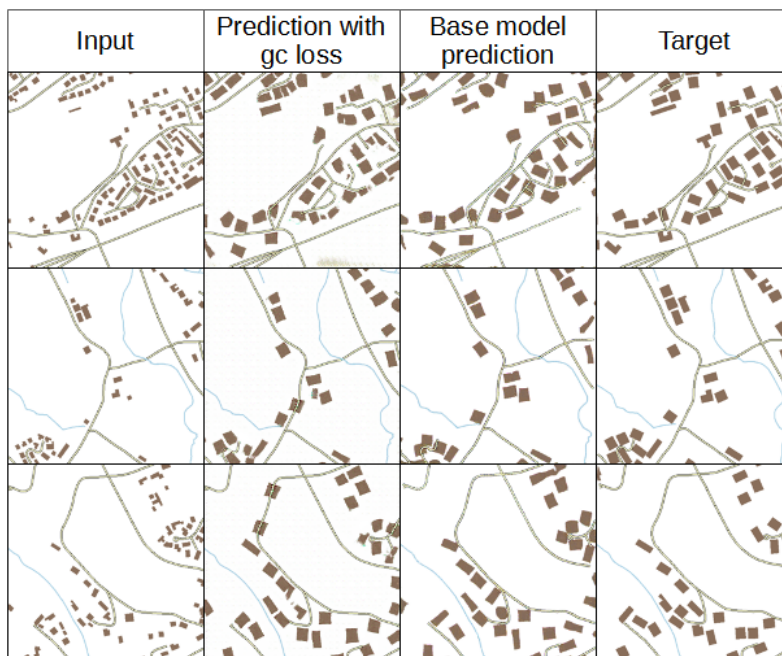


Figure F.15: Prediction made with the geometry consistency loss.

We do not observe that the building outlines are less blurred than without the loss. Building shapes are neither regular nor credible. More buildings are kept in the prediction; however, the overlaps between the roads and the buildings are also more frequent. Then the model tends to predict some nonexistent buildings on isolated roads as if the model is learning that the buildings must be along the roads (even if there is no building in the input). Finally, the generated road appears to be more blurred and poorly generalized compared to the prediction without this loss. The predicted image seem to suffer from a too important complexity that degraded the results, especially the road generalization does not benefit from the geometry consistency loss. Thus, why not only applying this loss only to themes where it is beneficial? We think that such a loss can be interesting in a simpler case (i.e. for the generalization of buildings only).

Structure loss We tested the loss of image gradient structure (from Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020) designed for map style transfer with preservation of global topological relations and detailed edge curves of objects, which are important in cartography. The interest of this loss for map generalization is explained in Chapter C. We tested this loss for data without road selection because road selection involves a large change of image structure that the loss penalizes. Figure F.16 presents the results of our model trained with this loss.

We observe that for simple situations the results are similar to the prediction without this loss, the edges of objects are neither more regular nor more credible. In

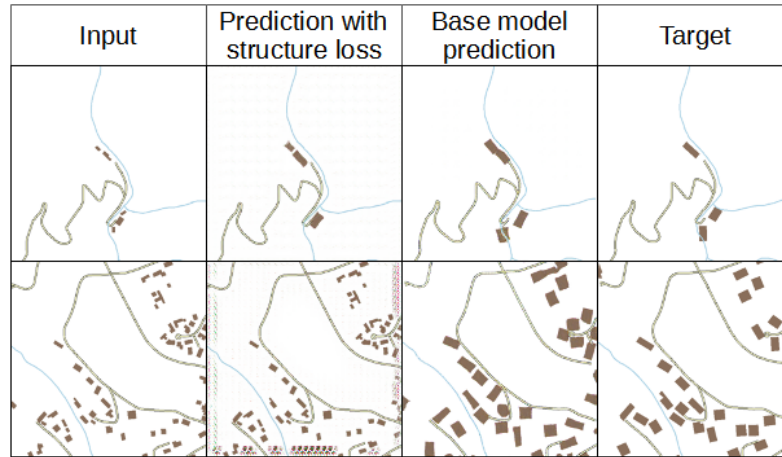


Figure F.16: Results of the model trains with the structure loss.

the most complex situations (e.g. line 2), the result is not satisfying, the model learns to predict an image with very few changes from the input (except a small undesirable noise), and no generalization is performed. Indeed, in these cases, we expect the loss to encourage the preservation of topological relations between different elements (i.e., maintaining disjoint relation between a road and a building, inclusion between a block and a building, or disjoint relation between two different buildings, etc.). However, the loss rather measures the structural variation between one element and its generalized version, and this comparison, contrary to the symbolization task, is not relevant because the generalization of the dense areas involves an important change in the shape and location of objects. Such loss seems to be not adapted to a map generation task that includes a change in the level of abstraction between the target and the input map.

F.3 Discussion

This section presents some discussions about our experiments on generalized map generation. In particular, we discuss the interest of the fusion GAN architecture; the additional information required to generalize a complete map at the 1:50,000 scale, and the possibility of a model for complete map generation.

F.3.1 Are fusion models useful?

First, we discuss why a fusion architecture is necessary for the models to benefit from additional information. During traditional generalization, cartographers do not use all the information on geographic objects at the same time; they first look at the position and shape of the object, find conflicts and then use other information (the importance, meaning of objects, etc.). The fusion architecture is based on the same principle; the model first encodes the shape and position of the cartographic object and then needs additional information to resolve the conflict, but this information does not require the same encoding as the main information. To demonstrate the interest of such architecture, we compared its predictions with an architecture that sees all information in the same tensor, and the main and additive tensors are stacked together. Figure F.17 presents the results of this experiment. We observe that the

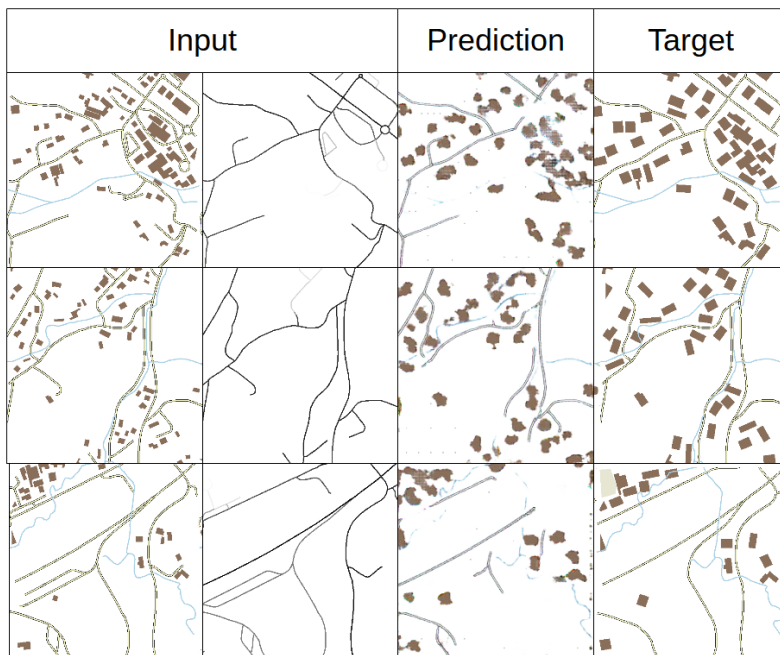


Figure F.17: Prediction of the model trained with the stack of the main and additional information for map generation that includes road selection.

prediction of stacked information is similar to the prediction without additional information, random road parts are selected, many disconnections appear and the result is not legible. The model does not understand the information hierarchy and does not benefit from the additional information.

Then, the results predicted with the fusion GAN have a degraded quality: including more disconnection, more blurred boundaries, etc. than the model trained with the base architecture of pix2pix. This degradation of quality is partly due to the increase of the task complexity, but mainly due to the model that has to be improved. Indeed, the Fusion GAN we propose is just a Pix2pix where we replaced the U-Net with a FuseNet. Indeed, this replacement is bound to include additional information, but fusetnet structure does not include the ability to deal with multi-scale phenomena unlike the u-net that include long-range skip connection (see Figure C.5). Moreover, the depth, initialization method, and optimization method of both these models are different, and thus have to be investigated.

Finally, in our architecture the discriminator is unchanged and trained with input, output and input predictions paires. However, the GAN discriminator may also requires to "see" the additional information to judge if the predictions are realistic. We can assume an architecture that uses the triplet: main input, additional input, output or prediction, might be more suitable.

Which information to fuse? If the optimal method to fuse additional information is not yet defined, the interest of this information is not a question and the more complex the target map, the more additional information may be required. We think that the attribute information of map objects and the context around the tiles are the more relevant additional information, but we did not have time to study the benefit of many additional information sources. Thus, in this section, we give some examples of potential improvement with additional information.

- A meaning-aware generalization of building: in some cases, it can be interesting to apply different generalization rules to some buildings due to their nature, e.g. a church often has a cross shape, and preserving this shape even if it breaks the minimum side constraint may be interesting as this shape is meaningful and may serve as a landmark in the city. However, such a generalization can be learned only with information on the nature of the building.
- The artificial and natural waterways are often really different to generalize, and this information may be provided as additional information.
- The Horton order of a river or river stroke can be useful for its generalization (Touya 2007), as it describes the context in the watershed that is not visible in an image.
- The density of a block can be a contextual information as the block can be heterogeneous in its spatial distribution and not completely visible. This information may influence the block covering decision and building placement in the block.

F.3.2 How to learn the preservation of spatial relations?

The main limitation of our results is the preservation of spatial relations. Our networks seem to succeed in preserving the relative orientation and relative position of element pairs in most situations. However, they fail to preserve the patterns of particular group of map objects. For example, it fails to select roads while preserving the connectivity of the road network.

The main difference between road density reduction and building density reduction seems to be the scale of the change: erasing a road would impact several tiles, while erasing a building may only impact a part of a tile. The tile scale may not be adapted to represent the context necessary for road selection. Furthermore, the absence of attribute information on the road is more important for roads than for buildings. For such relations we think the image approach is not the most adapted, or have to be combined with other approaches (e.g., graph based learning).

F.3.3 Is it possible to generate a complete generalized map?

The last and most important question brought up by this experiment is the possibility of generalizing a map that combines several cartographic themes using these networks.

First, we observe the following limitations:

- The predicted buildings have a blurred outline and not regular shapes. Some buildings are not credible. This problem is common with all GAN architectures tested in the literature and some model improvements for geographical data or regular data have to be investigated.
- Some overlaps between roads and buildings remain. This problem has two main causes: (1) the under-selection of the road network which makes it difficult to find a generalization solution that avoids overlapping and (2) the presence of such overlaps in the example from the reference. The correction of the training set is time-consuming but required to improve the results.

-
- Disconnections appear in the road network, or the road network is poorly generalized. In the first step, this problem was due to a lack of context. Then we provided context with the Fusion GAN, but the model has to be adapted to use it.
 - the block graying fails. We think that our input data do not include enough graying to allow us to learn this process, and our tiles are not adapted, a tile that includes the whole block may be required.

Secondly, we learnt from this experiment the following guidelines for the use of GANs for map generalization:

- The input and target domains should not be too different with a supervised architecture. Consequently, we believe that only a small-scale gap is possible (i.e. initial and target scales should be close enough).
- Input and target tiles must be legible; thus, the information presented in a symbolized representation must be limited, or a layered representation has to be used.
- A sufficient context has to be visible in the tile, so the scale of the images has to be adapted, or an additional information tensor has to be provided; this information has to be relevant and correctly encoded to benefit the learning.

These constraints on task definition and tile creation reduce the possibility offered by GAN. Moreover, some other techniques like attention-based architectures (Vaswani, Shazeer, Parmar, Uszkoreit, Jones, Gmoez, and Kaiser 2017) may resolve some context-relative issues. Finally, model improvement can benefit some cartographic themes and degrade others. Therefore, we believe that it is currently more reasonable to design a process that learns to generalize different independent elements of the map separately. Such an architecture is explored in Chapter G, our proposal separates generalization processes by theme, similarly to traditional generalization methods that treats roads first and then buildings in the fixed generalized road network; etc.

Chapter's conclusions

This experiment shows that GANs can generate map-looking images that include generalization. Our results succeed in reproducing the simplification, enlargement, and typification of buildings. However, more context is needed to achieve road selection.

This experience also highlights some challenges in the data organization. We proposed an organization of geographic information that seems promising for map generation. This information is organized as two tensors (see Figure F.18), one representing the shape and position of the cartographic object by theme, and one representing additional information required for map generalization. The layered representation limits bias in the creation of the dataset makes the input information more accessible for a computer, and the output is more complete in dense areas. Then, the additional information characterizes the map objects (attributes or context) and helps their generalization. This additional information must be encoded by the generator network to benefit from the prediction. For this task, we propose using a FuseNet as a generator, but our proposal is not yet satisfactory.

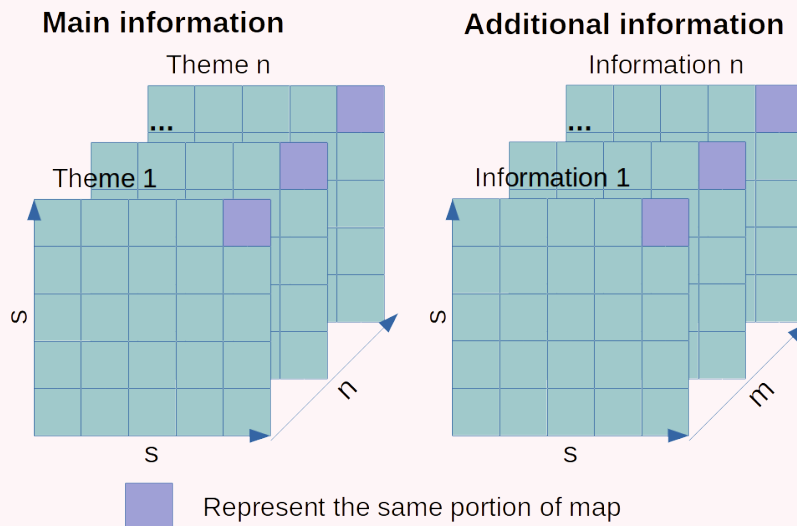


Figure F.18: The recommended organization of input data to learn generalized map generation.

Finally, we think that GAN are more adapted for map style transfer tasks than map generalization. The generation of a complete map may require much context information and hard work on tile preparation. We believe that a more promising approach is the successive learning of theme generalization combined with a GAN that makes a map from raw generalized tiles. This process is explored in Chapter G.

Part III

The future of map generalization with deep learning

Chapter G

Usages of deep learning models for map generalization

This chapter deals with the main possible perspectives for using the deep learning models in map generalization. In particular, we discuss the ability of deep learning models to be applied in new unseen situations. Then we describe how both the integration of deep learning in pre-existing map generalization process and its usage as an end-to-end process are challenging.

G.1 Transfer learning

Once a deep learning model has been trained, it aims to be applied and fitted to unseen situations. This ability of a model to be relevant for various unseen situations is called transferability.

In this section, we illustrate this ability in case of map generalization models for some models presented earlier in this thesis.

G.1.1 Location transfer

First, we tested the transfer capacity for other geographic areas. Indeed, location can impact both the map content (e.g. some area are denser or more regular than others) and cartography (the requirement, scale, level of detail or feature of the map varies in space and time); and as our model is trained with examples from the same geographic area as the test set, we would like to verify that it can be in other areas. We tested the transferability of the graphic generalization learned for the mountain road use case described in Chapter E. Models were trained with images from the Alps, we applied them to roads in other mountain areas of France (see Figure G.1).

Each new testing area was separated in testing images created in the same manner as training images, and contains at least 200 new images with various road configurations. Figure G.2 shows a very small extract of these test sets for each of these areas. We observe that the prediction is as good as for the initial test set (in the north Alps). Quantitative evaluation (presented in the next chapter) also confirms this result. Thus, this model is correctly transferred to mountain roads from the same country.

This use case is quite simple as the input images in the test set are quite similar to those of the training set but some other cases are more complex. For example, in next section we present experiment for transferring model to another country. A

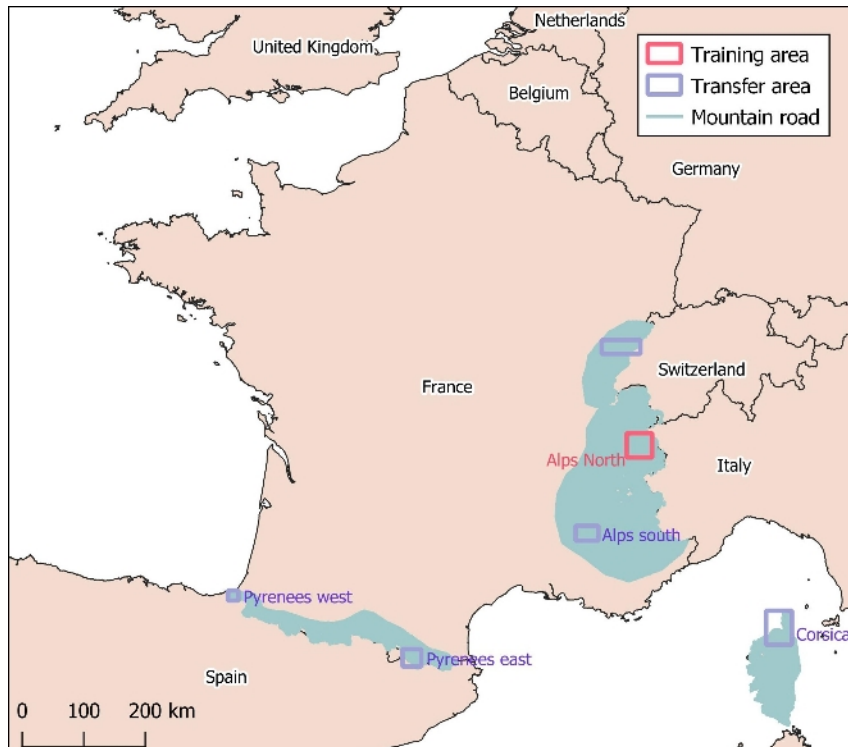


Figure G.1: Situation map of the training and tested transfer areas.

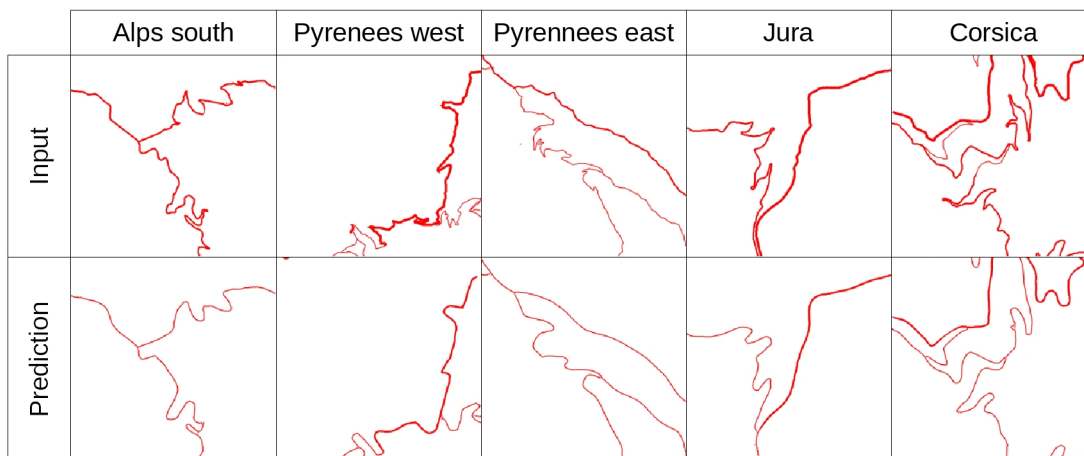


Figure G.2: Prediction of the model trained in the north Alps for tiles from other regions in the same country.

such transfer require to fine tune the model as the input image may be created from other data sources or in a different manner, and the characteristics of represented situation can be very different.

G.1.2 Transfer to geographic information from a new data source

Secondly, we experiment with how the model trained for topographic map generalization (from Chapter F) can be transferred to some maps with other data sources. The input tiles are created using OSM data in the USA. The level of detail of the

input is not homogeneous and more detailed than the input tiles of the training set (we estimate that the level of detail for selected areas would be adapted for cartography between 1:10,000 and 1:25,000; see Figure G.3). The second challenge of the chosen dataset is that it is very dense compared to the French one and contains many small and aligned buildings. We have shown that the model transfer to USA data is not efficient unless retraining a few epochs of the model with a few examples from the target areas. We consider that the objective of the model is not exactly the same as when it is applied to french data: the input data are constructed differently and the scale gap is greater. Moreover, the very dense situation are common in the new test area and rare in the initial trainging set, they may be hardly generalized without fine-tuning.

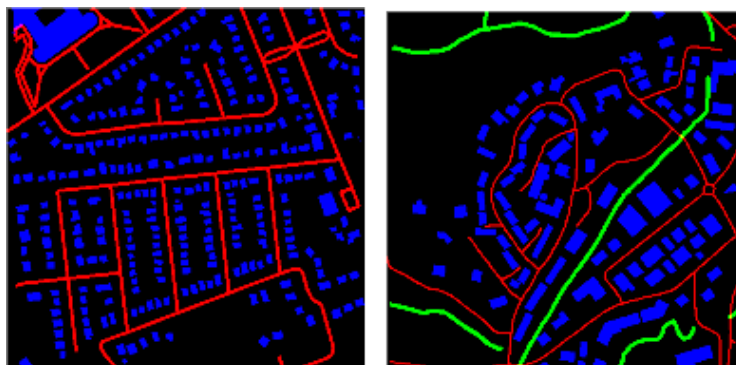


Figure G.3: Input tile a) in the USA; b) in France.

G.2 Usage as a tool in a mixed process

Map generalization models are traditionally composed of operators (that construct new objects or geometries from existing entities) and indicators or data enrichment (that describe existing objects, make explicit the implicit knowledge necessary to map generalization and help to choose and parameterize operators). According to its formulation (see Chapter B) a deep learning model can play both roles; for example, the detection of alignments (proposed in Chapter D) is an indicator of building group, which can be used to improve the generalization of these groups, while the segmentation of road shape (proposed in Chapter E) is an operator as it produces a new road shape. Moreover, some models, such as road section classification, can be employed both as an indicator of the importance of road sections in a road network selection process or directly applied as an operator to select edges in the road network. In the next paragraphs, we detail the issues of using deep learning as an indicator and as an operator for in a map generalization process.

G.2.1 Deep enrichment

First, many deep learning tasks can be seen as data enrichment for map generalization. Concretely; it can be employed similarly to traditional data enrichment prior to map generalization, the information is made explicit, registered as a characteristic for the database object, and then used during the generalization process. For example, the following enrichment can be brought about by a deep learning model and is a benefit to map generalization.

- Urban block classification could be performed using a deep neural network and can be beneficial for progressive block graying (Touya and Dumont 2017).
- The segmentation of highway interchange (Touya and Lokhat 2020; Yang, Jiang, Yan, Ai, Cao, and Chen 2022) can be beneficial for the generalization of such patterns;
- Building shape classification (Yan, Ai, Yang, and Tong 2020) may allow models to choose the relevant operator for building simplification.
- Alignment detection (Chapter D) can be used to assess the generalization of the building group or to determine the groups for the generalization of the buildings.

The data enrichment prior to map generalization is classical, and the use of deep learning for such a task does not imply new/specific integration issues. Knowledge is made explicit and can be used for map generalization processes (eventually with a form adaptation¹). In this case, the contribution of deep learning is to make accessible knowledge that was previously not (except by a manual annotation).

G.2.2 Deep generalization

The usage of deep learning as an operator is promising, as it allows one to generate generalized entities and to use them in a generalized map. For now, the creation of such a model is more challenging, and in most identified approaches (see Chapter B) the prediction cannot be in vector format and require post-process to be integrating into a map generalization process. Indeed, in the construction of a map generalization process that includes a deep learning operator, the output of the deep learning process must match with the expected input in the next step of the generalization process (which are commonly entirely vector).

Graph to vector The conversion from graph to vector is often not a problem, as graph construction methods are a bijection (one element in the graph corresponds to one element in the vector), and the graph structure is not altered during learning (even removal of an object is not properly a removal, but the prediction of a removal label that will lead to suppression of the associated object in the generalized output). For example, if a shape graph (i.e., a graph that encode the shape of an entity, see Chapter B) is used to predict a generalized shape, one corner corresponds to one node, and the existence and displacement of each node is predicted, then this information applied to the corner point allows the generation of the vector generalized shape. Similarly, for the selection of the road network in Chapter D the graph and the road network are linked by a bijection and the selection of the road section (according to a chosen threshold) can be applied to the road network to obtain the vector selection of the road network.

¹The image approach may require more form adaptation to find the object concerned by each images.

Image to vector The integration of generalized objects in image format is more challenging, very few raster data are used in the map generalization process, and the conversion has important limits (Shunbao, Zhongqiang, and Yan 2012). In fact, the vectorization process creates an envelope of consecutive pixels that represent the same object. This process depends on a pixel clustering method (to determine which pixel describes the same vector object) and the pixel resolution. For example, for the generalized mountain road shape in Chapter E, we proposed the following steps²:

1. Projecting images. During the image creation, we stored the coordinates of the input image and used this information to project the corresponding output image of the map.
2. Constructing a road mask. We identify the pixels representing a road in the image (i.e. whose color is closer to the color of the roads in the target style than any other color of the target style.)
3. Polygonize. We use the `gdal_polygonize` function that creates vector polygons for all connected regions of pixels in the raster that share a common pixel value.
4. Skeleton. The polygon skeleton is used to find the road axes and linearize the shape.
5. Smoothing. The artifacts of linearization are eliminated.

The second challenge is the assembly of tiles. Indeed, tiles are generalized independently and roads are arbitrarily separated by the image border, but at the end, we expect a coherent road vector network without disconnection at the tile limits. Moreover, with the tile overlapping and the independent generalization of each tile, several generalizations are provided for one part of an object, and the assembly process has to match different generalizations in a unique shape. Several approaches to merge shapes are possible from the choice of one (the best) of the predicted shapes; to the creation of a new 'average' shape of all predictions. Finally, the assembly may be performed before or after the vector conversion according to the use case.

For mountain roads we choose to make an average summary of all shapes, as the most important criterion was continuity preservation between lines in different tiles and the mean line may be closer than two independent best generalizations from different tiles (we think the best generalization assemblies are better adapted for polygon objects like buildings); then we tested both approaches of assembly in raster and vector mode, and finally proposed a combined approach that may improve the quality of assembly:

1. The roads are arbitrary separated in sections.
2. Section are matched using a distance criteria.
3. The shape of matched section is optimized using a least square optimization, to induce a "mean" shape for each road portion.

The optimized constraints are the following:

²The work about mountain road tiles vectorization and assembly is the results of a student project by Ziad Boukbir, Quentin Courtiade, and Juliette Rabbe and that we supervised.

- Minimization of the movement of each point.
- Preservation of angles.
- Minimization of the Hausdorff distance between two analogous segments.

This process results in a vector road network that suffers from some disconnections and important distortions at some parts where two matched routes are very far apart, either because the matching is wrong or because the two generalizations are very different. Moreover, the displacement is homogeneous along the section, while we may want some points to be fixed (e.g., the intersection or dead end). Thus, we propose a second approach for tile assembly but before vectorization: we try to detect points of interest in the image and use them for distorting the image so that analogous roads cover each other. Such a process is promising for maintaining structure and connectivity of the network while assembling the tiles, but, we were not able to find both enough relevant points of interest and an adapted deformation function for tile assembly. This work was done in a limited time and that it deserves to be explored further. Our last suggestion is to combine both approaches and apply a smooth distortion to the image to assemble the points of interest, and then vectorize and match the other segment point.

G.3 Usage as an end-to-end process

The second approach is to use deep learning models in a holistic deep learning approach for generalized map generation. Chapter F has shown the complexity of learning to generate a generalized topographic map image from a detailed geographic database, using an image approach. Indeed, it seems complex to encode all the necessary information in a single tensor that is the input of a neural network, and the complexity of the task makes the model design complex. Moreover, we have shown that some architectures are suitable for some sub-tasks of the generalization, but not for others. No architecture proved effective for all required tasks.

In this section, we propose to combine several simpler generalization models in an end-to-end process. Indeed, the separation into simpler tasks allows us to simplify the formulation, the research of relevant deep neural networks, and its adaptation and evaluation. For example, the additional information that can improve the generalization of the road network in Chapter F also degrade the prediction in the case of a complete map generalization. Bias can be avoided in the case of a separate process. Thus, we propose, experiment, and discuss a deep neural network workflow for generalized map generation, called DeepMapScaler.

G.3.1 DeepMapScaler: a workflow of deep neural networks for the generation of generalized maps

We propose to separate the problem into several simpler tasks in order to apply the relevant deep learning architecture to the small problem it is able to solve (U-Net, FuseNet, GAN, Graph Convolution Network, etc.). Our contribution is organized as a workflow that describes an end-to-end process for generalized map generation, where deep learning is used in most steps.

The first part of the workflow is called data adaptation for a reduced scale (or in a shortened way scale adaptation); it encodes the detailed shape and location of

each desired geographic theme and predicts a possible generalization. This stage is responsible for adapting the level of detail of objects and preserving the main structures in each theme, it include both the enrichment prior to generalization and the individual generalization of themes. Then, the second stage assembles generalized objects in a symbolized map; it is responsible for the preservation of inter-theme spatial relations, and for the global legibility of the map. The general organization is illustrated in Figure G.4, then G.5 illustrate an application for this workflow on a map containing roads, rivers and buildings. In the following paragraphs, we describe the process with more detail.

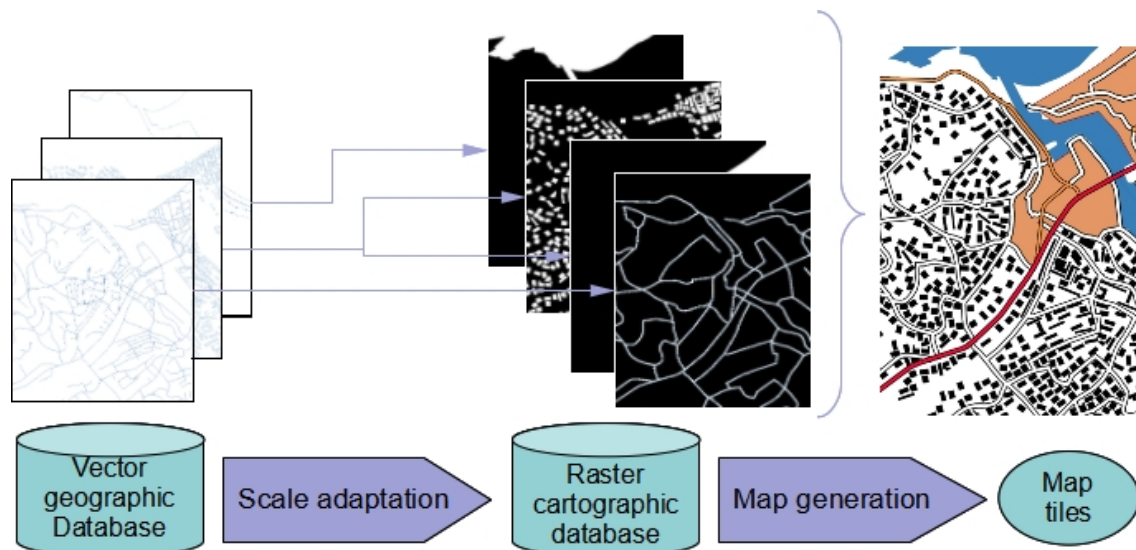


Figure G.4: Illustration of the main steps of DeepMapScaler workflow.

Data adaptation to the scale reduction The scale adaptation manages the level of detail of the geographical data, but does not exactly perform its complete generalization, as the inter-theme relation and resulting displacements are not considered. It may be assimilated to the creation of a first draft of cartographic database for a specific scale. The goal of scale adaptation is to produce the masks used as input for map generation. This shape satisfies legibility constraints at the target scale; it may respect constraint that concern individual, pair or group of object from the same class (e.g. the granularity, pattern preservation constraint, etc.).

The scale adaptation of each cartographic theme is processed independently in one or several steps (including both some deep enrichment and operator steps). It can have diverse forms according to the themes depicted in the map and the scale; the more complex a theme generalization at a scale, the more steps may be needed. We think it is possible and beneficial to mix diverse representations of geographic information during the scale adaptation of one theme (e.g., graph and raster). The input of the first step must be derived from the detailed database, then the information of the next step can be derived from both the detailed database and/or the predictions of previous steps.

Map generation This step leads to the prediction of a map image at the target scale, it is composed of only one model that is trained to generate a symbolized map from cartographic data, it mainly affect the style of images. Cartographic data

already have a level of detail adapted to the scale of the target map, but it also includes a generalization step to displace objects that overlap once symbolized and gather in map. We propose to repeat our proposal of layered representation and use as input a stack of masks that represent the position and shape of the generalized map item organized with one cartographic theme on each layer.

G.3.2 Experiment

To illustrate the interest of such a workflow, we propose an instantiation for the use case of topographic map generation at 1:50,000 scale, already presented in Part II.

Implementation

The use case and base vector database are the same as for the experiment presented in Chapter E. Thus, we can compare the results of both approaches. The target maps include roads, buildings, and hydrography, at 1:50000 scale. The global implementation of the workflow for this use case is illustrated in Figure G.5 and then detailed in next paragraphs.

We propose to experiment with the scale adaptation of roads and buildings and then with the generation of maps. We did not experiment with the water scale adaptation, as it is really simple at the chosen scale (no shape simplification is performed, the water containers are deleted and a slight selection of water lines based on attribute is performed). We implemented the deep learning models of the workflow using PyTorch, and the code for raster tile creation is published as a plugin of CartAGen open-source map generalization software. Each model is trained after the validation of the previous one in the workflow.

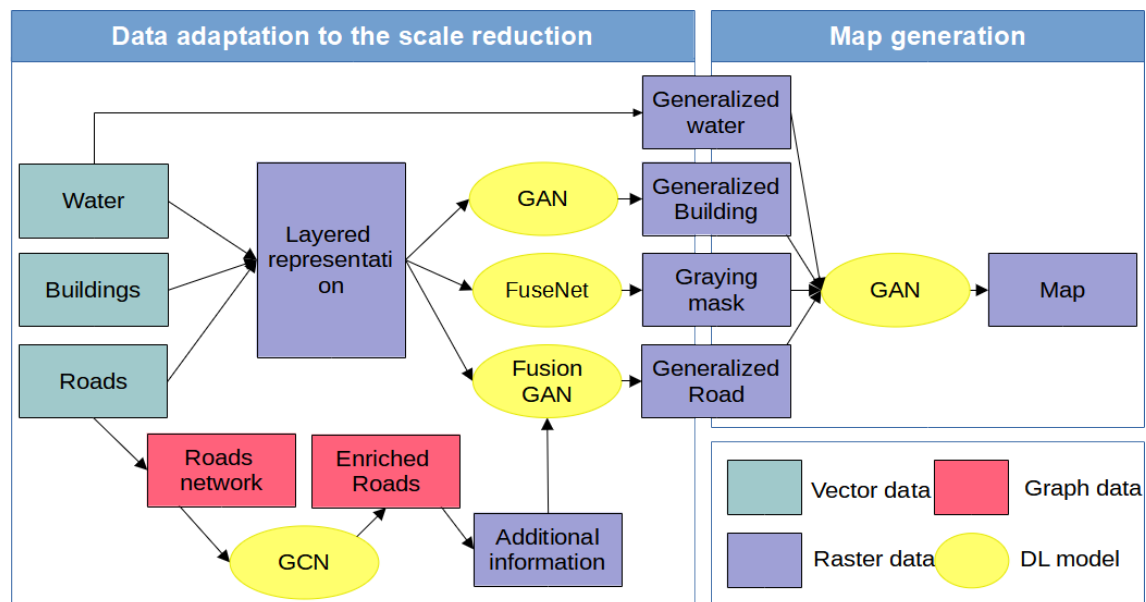


Figure G.5: Illustration of an application for our workflow on map containing roads, rivers and buildings.

Road generalization

For the road generalization, we tested a proposal in two steps: first, a GCN predicts the probability for each road to be selected at target scale from the attribute and context of each road sections (from Chapter D). Then, a GAN is responsible for interpreting this information and generating images of the generalized road. For this step, we use the proposed fusionGAN from Chapter F, with a binary mask of the presence or absence of roads as output. This part of the workflow is illustrated in Figure G.6. We also tested the addition of a connectivity loss (adapted from the connectivity loss proposed in Chapter F) and the topology loss from (Chen, Chen, Xu, Yin, Peng, Mei, and Li 2020). The connectivity loss measures the preservation of the number of road sets after generalization to avoid disconnections. Contrary to the experiment at the 1:250 000 scale, the loop creation and undesired coalescence of two road parts are very rare, so we did not measure the preservation of the number of background sets. This measure can be computed even in the more complex images; however, the overlapping of the road and other themes create disconnections that does not affect human perception of continuity and makes it less relevant in the case of map generation. The topology loss computes the structure of the predicted image and compares it with the input image. This comparison seems relevant to use for the case of roads image because the global structure of the road network does not change much. However, we do not think it is adapted for most other themes (e.g. building) that also contribute to the structure of the image and may vary in position in an important range.

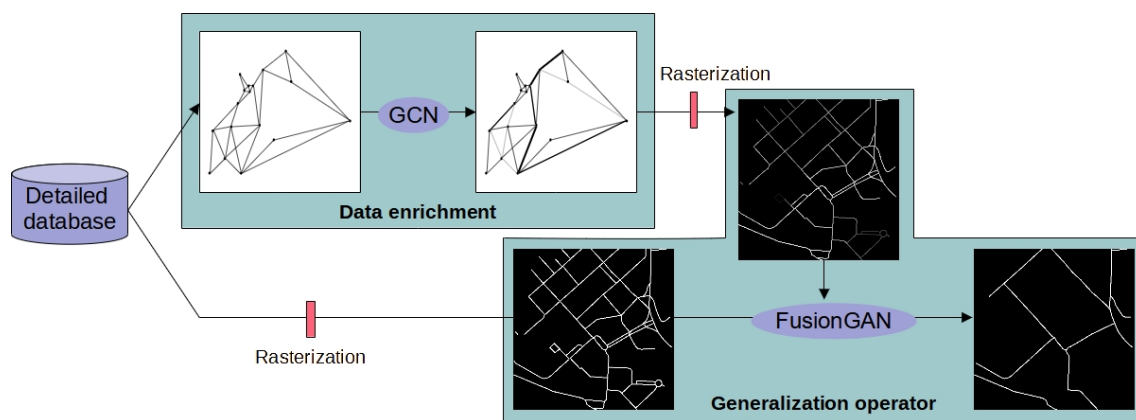


Figure G.6: Proposed steps for scale adaptation of road network at 1: 50 000 scale.

The results are presented in Figure G.7. We observe that the continuity of the prediction is still the major issue, the model is able to understand the selection and select appropriate edges but it does not preserve them entirely, and intersections are disconnected. The proposed improvements are not conclusive; the topology loss gives a similar quality of results for most situations and creates new disconnections for complex intersections where simplification (and thus change in the image structure) is required. The connectivity loss slightly reduces the number of disconnections, but the numerous artifacts make the image hardly legible.

We believe that despite the obvious failure of our attempts to maintain the continuity of the road network in the predictions. The image approach may be adapted for the road generalization, however, further research on the fusion model and the loss function are necessary.

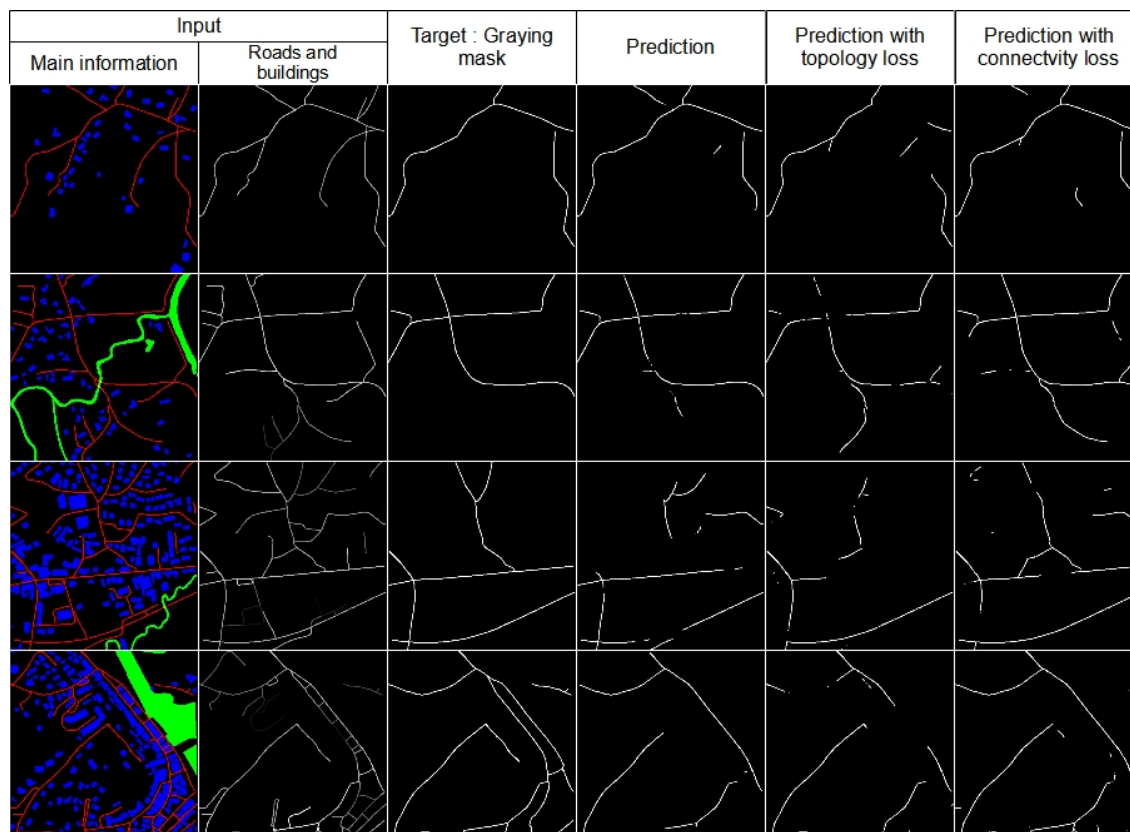


Figure G.7: Prediction of the scale adaptation for the road at 1:50 000.

Building generalization

The experiment on map generation from Chapter F lets us consider building generalization as simple enough to be learned in one step with satisfying results. Thus, we train a supervised GAN to predict the position and shape of the generalized building in a mask. The input of the model is the layered representation of the detailed buildings, roads, and hydrography because we think that the other theme of the map helps to identify crowded and free areas and gives an indication for the generalization of buildings. We tested the classical implementation of pix2pix and one that included the geometry consistency loss from (Fu, Gong, Wang, Batmanghelich, Zhang, and Tao 2019). This loss may allow us to generate more regular shapes and reduce the number of epochs required for generating acceptable results (see Figure F.15).

Figure G.8 presents the results of these steps. The evaluation of such buildings is visually difficult, the both results seems satisfying, with realistic and generalized buildings. However a quantitative evaluation of building shape regularity is necessary to study the interest of the geometry consistency loss. This evaluation is based on conversion to vector and vector constraint evaluation and detailed in Section H.4. It leads to the conclusion that the geometry consistency loss does not bring a significant improvement of regularity in predicted building shape.

Block graying

Finally, we try to learn the graying of the city center. This task is essential for legibility when the situation is too dense at this scale. This process requires the road

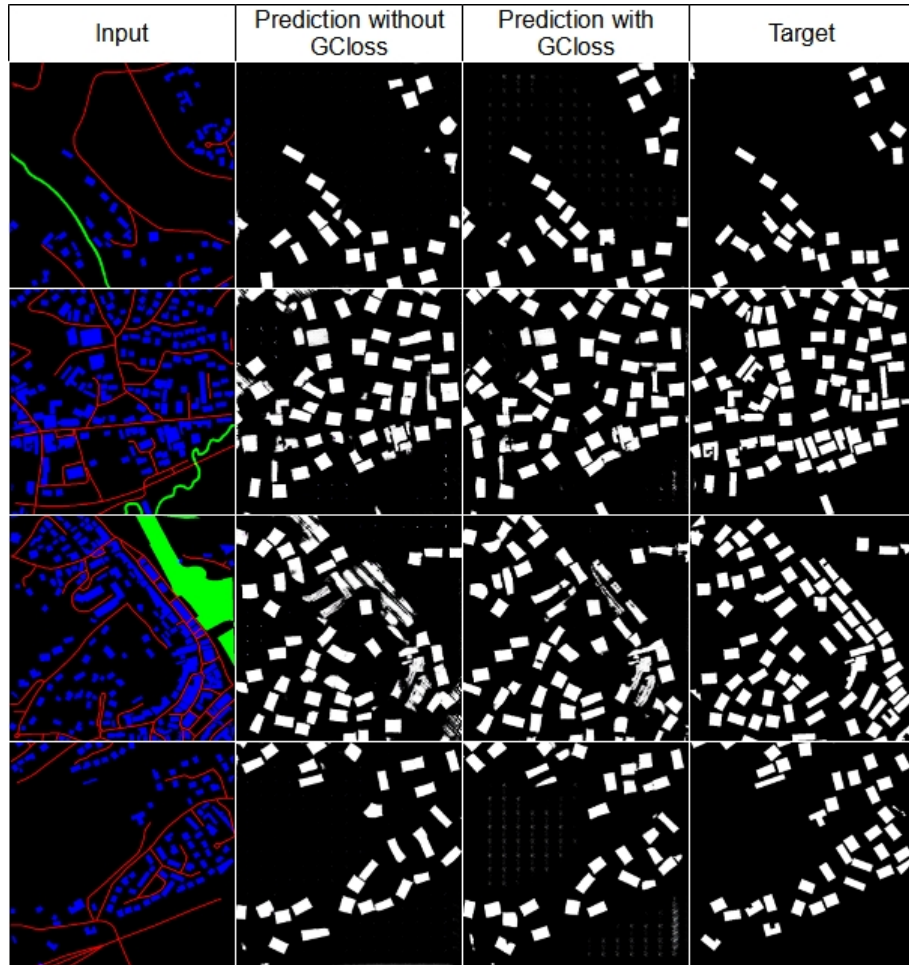


Figure G.8: Prediction for building shape adaptation.

network (to define the urban block), the detailed shape and position of the buildings (to determine the cluttered areas), and some semantic information that summarize the properties and the spatial neighbourhood of each block. For this experiment we reuse the organization of geographic information proposed in Figure C.21: roads and buildings shape and locations are the main information, a grayscale category map that describes with values from 0 to 255 for categories from rural to inner-city is the additional information. This additional information is obtained using the multi-criteria classification described in (Touya 2021). The resulting dataset was not really relevant because the graying case is very rare (representing less than 20% of our tiles) in such mostly rural areas and at this scale; thus, we only kept tiles that include urban pixels and added examples from other cities in France until we obtained a reasonable dataset of 1200 images. The graying was manually annotated. We trained a FuseNet to segment pixels that may require to be grayed. Figure G.9 presents one tile of input, target, and prediction. We observe that the model learned to cover the entire urban envelope rather than just the dense blocks of the center. We did not have time to further improve the prediction of graying, but we think that the tile scale and the provided information are not adapted. The classification results are not correlated with target graying (e.g. some block are classified as sub-urban and grayed and conversly some urban block are not grayed). Moreover we think the size of tile is not adapted for block analysis, and images that allow to see

several blocks or even the whole city would be more relevant for this task.

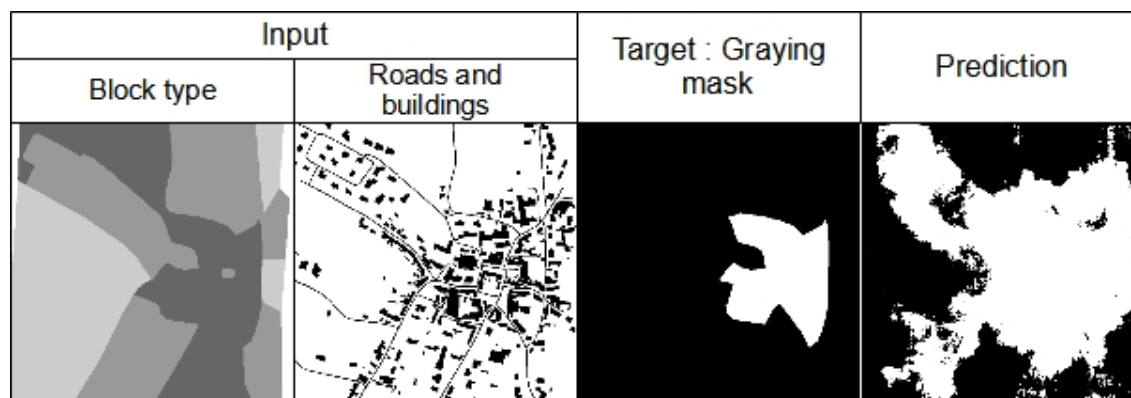


Figure G.9: Prediction made by the model trained for graying segmentation.

Map generation

Finally, we experiment with the generation of legible maps without overlapping data at the target scale from the scale adapted cartographic data (output of the previous steps). The goal of our experiment is to test the interest of DeepMapScaler principles. As the quality of road network and graying prediction are not good enough, we only use the buildings generalized by a GAN and faked the scale adaptation of roads and hydrography (at this step the roads and hydrography input tile are not the output of the previous step but a generalization obtained using a traditional map generalization process and manual corrections). The input and target of this step are illustrate in Figure G.10.

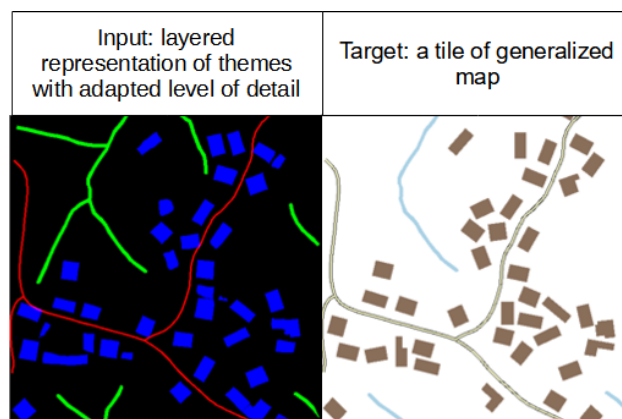


Figure G.10: Input and target of map generation step.

We choose an unpaired model (cycleGAN) because we want to learn what a generalized map looks like and not how to convert from one generalization (the one performed by previous steps) to another (the target one). This approach had worked quite well for mountain roads in Chapter E. Figure G.11 presents the results of this generation experiment. The output is in a satisfying style and quite credible. The images presented on the left are particularly satisfying because the prediction outperforms the generation made in Chapter E and demonstrates the ability of CycleGAN

to reproduce new generalization operators: in purple, we observe some amalgamations when the initial buildings are too close and have a complex shape; in green, we observe some cases where the buildings are displaced or distorted to avoid overlaps with the roads; in pink we observe some displacements to avoid the overlap of two or more buildings. However, the prediction still has some limitations (illustrated on the right part of the figure), in particular in the prediction, the road-river relation and road-building relations are sometimes not correctly processed and undesired overlaps occur (circled in red); and some building-to-building proximity relations are not correctly handled as the model does not choose between amalgamation or displacement, and tries to do both, which results in a shape that is not credible. We explain the buildings-roads overlap by the rarity and complexity of the situation, buildings are in a very thin block and where it is not possible to maintain buildings in the blocks while increasing their size. The roads-river overlaps is due to a bad generalization of such situation in our example set. For the second problem, the amalgamation and displacement of buildings, we think it is mostly an under-training problem; we confirmed this hypothesis by training the model for a few more epochs and observing a reduction of such situations.

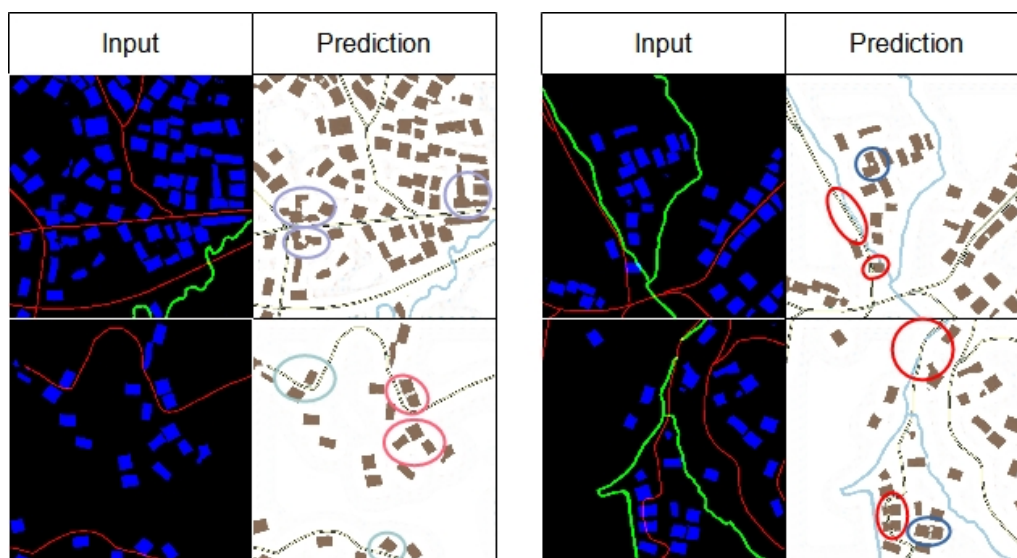


Figure G.11: Predictions of the map generation steps; good predictions are presented on the left and more controversial ones on the right.

G.3.3 Discussion

Advantages and limitations of the workflow Our experiment aims to show that a workflow that combines several deep neural networks is more promising than the approach of Chapter E (i.e., training a unique model to achieve map generalization at once). The proposed approach has the following advantages:

- It allows for a simpler evaluation and comparison of prediction quality for each task.
- It allows the independent training of different tasks: some tasks are more complex and require a different number of steps (e.g., map generation takes

300 epochs, while building generalization only needs between 150 and 200 according to the model).

- As the objective is more clearly defined, the model and loss adaptations are also simpler.
- Some tasks (e.g. road network selection) require different representations, and such a model allows the combination of representations.
- We can choose and use different input data for each task.
- Post-processes and corrections between steps are more easy than with one model for holtics generalization (e.g. the blue spot noise in cycleGAN prediction can be simply corrected in buildings or road images as no blue pixels are expected, while in a complete map they can be noise or rivers pixels).

However, our experiment is limited by the chosen model for each scale adaptation; in fact, if we are convinced that the separation by theme will allow the design of a specific architecture, and thus the prediction of better generalization, we were unable to propose a model that visually improvement of the prediction is significant and not just link to the randomness of learning processes. In particular, road connectivity is still a problem and may require an improved fusion architecture. Furthermore, the map generation step provides an improvement. In particular, it has shown that the unpaired GAN can learn displacement and amalgamation and seems more efficient in producing a legible map than the unique generalization model. Finally, the proposed workflow has limitations in its definition.

- Data acquisition is more complex than for the unique model;
- It requires more time and storage capacity than the unique model;
- The propagation of errors is more probable.

Transition between steps Then we discuss the transition between steps. In our first implementation, we decided to train each step independently and to include a data representation between each step to make the transition between one model output and the next input. This choice is beneficial for research purposes where we are still searching for an adapted model for each step, and we can and want to evaluate each of the steps visually. This choice is the simpler one for us, but other solutions are possible and may be of benefit. In particular, for a step that has a data enrichment function (e.g. prediction of elimination probability), an approach with an encoding-only model that provides an embedding representation of attributes or context information to the next model directly may be adapted. The idea is that the decoding step is to render the information in a form that is understandable to a human. But as in the case of our workflow it is directly another learning model that will use it and not a human, it could be possible to transmit directly the encoded information (embedding) to the next model (see Figure G.12). We imagine that this should reduce the eventual degradation of the information related to its successive encoding and decoding.

More than reducing the cost and bias of decoding/representation such transition would also allow to chain processes based on any kind of object representation, while

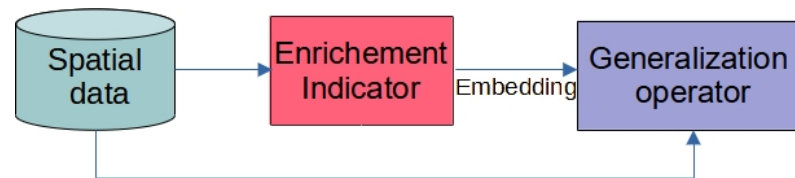


Figure G.12: Illustration of the alternative transition between steps based on embedding.

for now the decoding transition only allows some transition (e.g. the graph-to-raster transition is possible but not raster-to-graph), and would give more freedom on the steps of the workflow.

Chapter's conclusions

In this chapter, we show that the use of deep learning may allow us to apply deep learning-based generalization model to a large dataset. However, this application may require fine-tuning in the case where it must be applied to very different situations or to use case where the objectives varies (e.g. the input data is from another source, the target style is quite different, the scale gap has changed, etc.).

We also explore the possible usage of deep learning based model in a generalization process. We assume that the usage of deep learning alone is currently not suitable and models have to be integrated in larger workflows. We show that models can serve as an enrichment indicator or a generalization operator in a traditional generalization process or in a fully learning-based map generation workflow. In both cases, integration is challenging, mainly due to the data format (that is not vector) and the deep learning evaluation (see Chapter H).

Chapter H

Evaluation of deep learning predictions

As shown in Chapter F, the evaluation of generalization performed using deep learning is a challenge. This chapter introduces the main notions and issues about evaluation and then proposes three evaluation methods adapted for deep-learning based map generalization.

H.1 The specificities of evaluation of deep learning-based generalization

H.1.1 Why evaluating deep learning predictions?

The evaluation of generalization is traditionally an important research problem; evaluation is used to tune the parameters of generalization processes (e.g. setting constraint threshold and importance), to control its orchestration (e.g. choose the relevant operator or level of analysis), and to assess its results (e.g. verifying the legibility and informatio presevation) (Mackaness and Ruas 2007). In addition, deep learning processes also require evaluations at several steps (Figure H.1) and the evaluation often differs according to its objective.

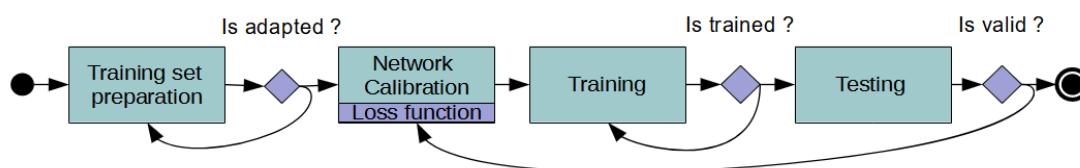


Figure H.1: The main evaluation steps when designing a deep learning model (the learning process is colored in green, and the evaluation steps are colored in purple).

First, the evaluation of the training examples is used to ensure that the examples are good enough for learning or to filter examples of lower quality. Then, evaluation can be used to validate a model, choose the most adapted configuration of a network, or guide the post-processing (e.g. identifying a noise to eliminate). Finally, the main interest of an evaluation measure calculated using images is to guide the learning process, i.e. including a particular evaluation measure in the objective function of the neural network. This evaluation requires a very small computation time and

cost, as it is repeated for each image in the dataset at each epoch of the training. On the contrary, the evaluation at the end of the process can be more time consuming, as it would be calculated only once; however, this evaluation must be as quick as possible.

In next paragraph we summarize the evaluation objectives for deep learning based map generalization.

Tuning In deep learning the main parametre we can tune are not the model paramter that are learnt but the training data set from which they are learnt. Thus, the evaluation processed on the training set at the beginning of the experiments, can be seen as a kind of tuning evaluation. This step aims at identifying un-desired examples and estimating the quality and diversity of examples. Identifying undesired examples and estimating the quality of the training set require an evaluation methods that identify and quantify errors in each image. Then, describing the diversity of a training set requires a more global indicator, such as the measure of image complexity and similarity to the detailed map¹.

Controlling An evaluation measure can be used to control the learning process. At the end of each epoch the measure is used to verify the progress of the model and to choose when to stop learning. For now the loss value and a pixel wise comparison with the reference are used in most training process. The evaluation for controlling the learning requires a measure that globally describes the quality of the prediction; moreover, the calculation of such an evaluation has to be done automatically in a reasonable time, because it is repeated many times during the training process.

Setting An evaluation measure can also be used as a part of the objective function during training. The learning objective (loss function) is the function that measures the error of a model and is optimized during learning. The choice of an adapted learning objective for map generalization is detailed in Section C.1.4.

Validating The most common reason for evaluation is the validation of a model. An evaluation of this type has to make comparisons with other generalization methods or with a reference generalized map. The aim is to verify that the produced map is a functioning map. We think such measures are essential for the development of deep learning-based map generalization, as this allows us to formally demonstrate the potential of such an approach and promote the comparison, and thus the progression of models.

Post processing Finally, evaluation can be used to identify errors after generalization and choose a post-processing method that corrects them. For example, we can apply a filter to correct a noise using the relevant evaluation. Consequently, the goal of this evaluation is to fully describe errors in a map image.

¹The evaluation measure can be used for tuning the training set but the training set tuning is not necessarily based on an evaluation measure (e.g., the sinuosity based filtering, and the diversity measure for mountain road in Chapter E).

H.1.2 Evaluation in map generalization

The objective of the evaluation is to determine whether the map has the characteristics necessary to be used for a given task. This evaluation can be performed visually by an expert or automatically by identifying and measuring the respect (or violation) of a set of constraints. Constraints are essential for many tasks in cartography, including map generalization evaluation (Werder 2009; Touya, Duchêne, and Ruas 2010). For example, Figure H.2 is an illustration of the main constraints for buildings on a topographic map.

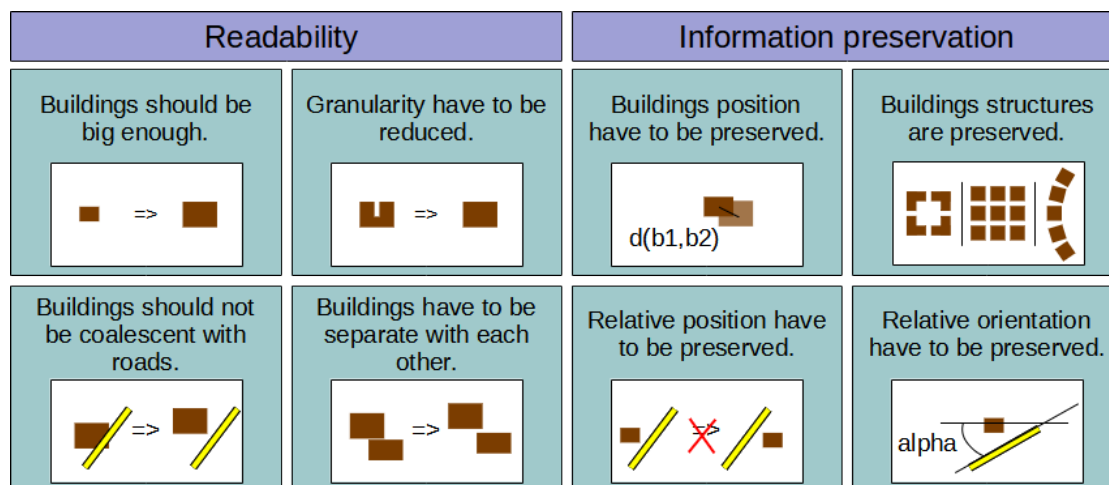


Figure H.2: Illustration of the main specifications of the buildings on a topographic map.

Evaluating a map or a generalization process from a set of constraints is challenging and raises the following issues:

1. Identification of relevant constraints/desired characteristics of the map. Both legibility and preservation should have similar weights in the evaluation to avoid caricatured results (Zhang 2012).
2. Identification of the relevant level of analysis for each constraint. This is a key issue in constraint definition: a constraint on each section of the road is different from those on a complete road or on the entire road network. Ruas (1999) identifies three levels of constraints: *micro* which relates to an individual object of the database used to make the map; *meso* which relates to a group of objects, and *macro* which relates to a population (e.g. all roads of the map).
3. The definition of a measurement method that fits the visual ranking (Tailandier and Gaffuri 2009; Harrie, Stigmar, and Djordjevic 2015). The complexity of the measure depends on the constraint. For instance, constraints on object size can be easily measured, while others require advanced spatial analysis, e.g. alignment preservation (Zhang, Stoter, Ai, Kraak, and Molenaar 2013), line position accuracy (Skopeliti and Tsoulos 2001), etc.
4. The calculation of a level of satisfaction. The goal of this step is to determine whether the result of the measure is acceptable. To do this, we can define an

evaluation function (Bard 2004). For example, for the constraint "the road position must be preserved", the measure could be the distance between two points from the previous and the new positions, and the evaluation function would evaluate how the constraint is respected or violated given the measured distance.

5. The fusion of all constraints into a synthetic value would allow the comparison of map generalization methods. However, generalization constraints often express conflicting ideas and cannot be totally satisfied at the same time: in a dense area, the legibility constraint demands information selection and/or object displacement, which violates the preservation constraints. Several strategies can be used for the fusion of constraints (Touya 2012). For instance the general objective can be to minimize the value of the worst satisfied constraint, or to reduce the number of critical measures (measure with a very bad value (under a certain threshold), to reduce the number of conflict areas (area where one or several constraint cannot be satisfied), or to increase the average satisfying of all constraints using a mean value, etc.

Many constraints and measures have been proposed for map generalization. However, most of them cannot be directly employed to evaluate the prediction from deep learning for three reasons:

- First, the output format is not adapted: such constraints and measurements are defined for vector objects, while our prediction may be a raster. Although conversion to vector format is possible, it can complicate the whole end-to-end process. Sometimes it is even not suitable. For example, when evaluation is used to guide the learning process during training.
- Second, the evaluation of a map can require important computation time, while deep learning evaluation may be computed many times for large example sets (in particular the evaluation for controlling and setting the learning).
- Finally, deep learning processes tend to hide causal mechanisms and do not guarantee a result that follows cartographic principles. Traditional map generalization evaluation may hypothesize that the evaluated map follow cartographic principles. Thus, the evaluation of deep learning prediction may requires to examine whether and how the map objectives are satisfied with an independent evaluation process, before the evaluation of generalization.

H.1.3 Evaluation in deep learning research

The evaluation of deep learning models differs according to the problem formulation; in this subsection, we describe these different methods and the challenge for the different formulations described in Chapter B.

Classification The evaluation of any classification (even without deep learning) is commonly performed using a confusion matrix and some scores calculated on its values. For example, Table H.1 presents the evaluation of road selection using a GCN (from Chapter D).

This evaluation is simple and relevant as long as classes are not fuzzy and a validation set can be provided. It is informative on the quality of prediction and

Prediction \ Annotation	Deleted	Kept	Total
Deleted	3098	1280	4378
Kept	2313	6246	8559
Total	5411	7526	12937

Table H.1: Confusion matrix for the selection of road section at 1:50 000 using a GCN.

most of the time adapted for the evaluation of the desired task: it answers questions like "How well the selection corresponds to the target?". However, it may not be sufficient to determine if the generalization is good; for example, in the use case of road selection, it does not say if the selection preserves the connectivity of the road network or if the model is more efficient in some situation or very bad in another (like city center). Finally, our validation set does not represent the only good solution for generalization. Thus, we can have solutions for which the confusion matrix is bad, but which are cartographically correct because the differences in selection are for small routes, and the two selections are cartographically equivalent.

Segmentation Segmentation is a kind of classification at the pixel level; consequently, classification measures can be applied to each image; however, the number of correctly classified pixels is not really significant. Most papers on deep learning segmentation evaluate the quality of the segmentation using the intersection over union (IOU) measure, which quantifies the number of common pixels for each class in the ground truth and the prediction divided by the among of pixels in the union of these two images.

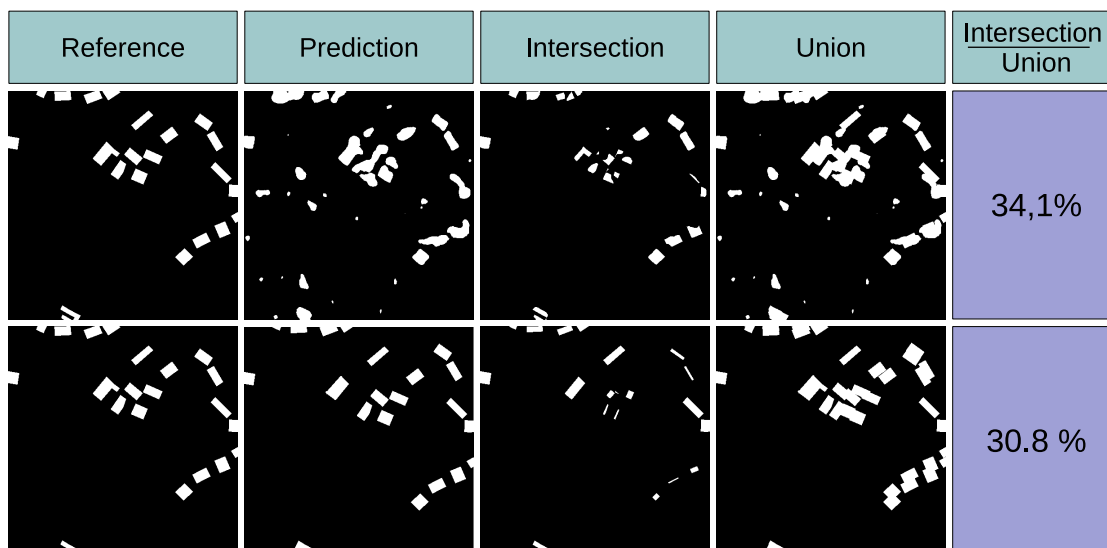


Figure H.3: Illustration of the limits of intersection over union. The value is almost the same for two different generalizations: one very bad at the top and one correct at the bottom.

This measure is quite efficient in estimating the accuracy of the predicted shape occupied by a class. However, it does not really match for evaluating map generalization results because map generalization involves displacement that reduces the

intersection and increases the union. Therefore, in Figure H.3 the second generalization is better: the shape, position and relation constraints on building are satisfied, but this generalization has a worst IOU value because the position preservation constraint is not respected. Moreover, this measure is only adapted for evaluating the segmentation of surface objects, the intersection of a line is almost always null if the prediction and reference are not exactly the same. If the IOU is not the only measure used for the evaluation of segmentation, deep learning measure are often base on pixel-wise comparison and suffer from the same bias (e.g. pixel accuracy and Dice coefficient).

Image generation and style transfer The evaluation of images generated using a GAN is an important issue in the literature. In such generations, the problem is similar to the evaluation of map generalization, the pixel accuracy is less important than the global appearance of the images. Moreover, the evaluation of the generated images often deals with problems where a good solution does not exist or where the good solution is not unique. This is also true for map generalizations: several different generalizations can be acceptable. There is no unique global measure for generalization quality (Touya 2012), and there is no easy way to measure how much an image looks like a real map. The questions that evaluation of GANs seeks to answer are the following: "Is the prediction realistic?", "Does it look like to the target domain?", "Does it have the expected style?", etc.

For the evaluation of such generated images, computer scientists propose evaluating the similarity with a reference or conducting a user preference test. On the one hand, the similarity measures are often not correlated with a visual evaluation of the prediction; this bias may also be important in map generalization, where results that are really similar to the target can be very bad. On the other hand, user tests focus on the realism of the output and try to demonstrate that people would not be able to determine which image is generated, and which one is real. This show the requirement of specific evaluation for map generated using GANs that question both the image realism and the map generalization quality.

H.1.4 Proposed evaluation methods

Neither the map generation evaluation method nor those of deep learning seem to be adapted to evaluate the prediction of deep learning-based map generalization. In most cases, the prediction assessment requires a visual examination of results; however, we want a more formal and automatic measure of quality to complement this visual examination. Table H.2 summarizes the evaluation methods from the literature on deep learning and map generalization. In the next sections, we propose to detail the two approaches highlighted in Table H.2 we selected to assess the prediction of map generalization based on deep learning.

	Automatic	Manual
Map generalization	Constraint violation measure	Expert evaluation
Deep learning	Similarity with target measure	User preference test

Table H.2: Possible and tested evaluation methods.

H.2 User test

We proposed a user test to evaluate and compare the generalization of the mountain roads. The design and results of this test are described in this section.

H.2.1 Design of the user test

We designed a test that aims to attest and quantify the observation made in Chapter E, i.e. that the GAN model produce image of map that cannot be distinguish from semiautomatics generalization and the interest proposed improvement. In particular, we compare the prediction of five models (see Table H.3). We hypothesize that:

- H1: The predictions should not be distinguished from the reference.
- H2: Model B should perform better than Model D.
- H3: The predictions for Model C should be more appreciated.

Name	Strategy	Model	Sample image
Model A	Unsupervised	CycleGAN	1223 from Alps only
Model B	Unsupervised	CycleGAN	2271 from several mountain regions
Model C	Unsupervised	CycleGAN	1223 with connectivity loss
Model D	Supervised	Pix2pix	1223 from Alps only

Table H.3: List of the models compared during the user test.

Task The test is composed of two different tasks: an absolute evaluation (marking task) and a relative evaluation (ranking task). For both tasks, users do not know how each image has been produced. In the absolute evaluation, participants have to observe an input-generalization pair and answer to three questions (listed below). For each question, a Likert scale with four values from "very good" to "very bad" is proposed.

1. How well do you think roads are generalized? The goal of this question is to evaluate whether the generalization operations have been applied properly.
2. How much do you think the generalized image seems realistic? This question evaluates whether the images look like a map that has been traditionally produced.
3. How well do you think the generalized image is readable? This question assesses the legibility of images.

Then, the ranking task presents an input image and five generalization images for the corresponding area. Participants have to drag and drop the generalization images by preference (from the most appreciated to the least appreciated). This task is common in GAN prediction evaluations (Wang, Girshick, Gupta, and He 2018) because it verifies that predictions cannot be distinguished from real images. Moreover, the ranking task seems to be an easier task for people who are not familiar with the expected output. Finally, the combination of these two tasks should avoid fatigue for participants.

Participants We received 113 responses during the two weeks of the test from nonpaid volunteers. The participant population is made up of men at 60 %; only 10 % of the participants are aged under 25 and more than 40 % are aged over 50. Most of the respondents use maps weekly or more often, but have never heard of generalization. The "expert population" is composed of twelve generalization experts, twelve geo-visualization or cartography experts, and 25 other informed people (who declared having already heard about map generalization).

Procedure The user test was designed as a web questionnaire; the link has been sent to two different e-mail lists: one composed of generalization experts and one of the researchers from various domains and the employees of our university. The test is designed to take around 20 minutes. The first page of this survey gives information on the objective of the test, the images presented, and the generalization of the mountain road. This introduction page can be reached again at any time in the test if the user has doubts. The last page asks some general information questions, such as their level of expertise in maps and generalization. For the main part of the test, we selected 30 situations (24 for the marking task and 6 for the ranking task), and we distributed the images (for each situation and method) in 8 tests, which were randomly assigned to the respondents. Each test is organized as follows: 5 marking tasks; 1 ranking task; 5 marking tasks; 1 ranking task.

H.2.2 Results

We obtained between 2 and 20 responses for each image. This difference is due to the random distribution of the tests. All images with more than five responses are included in the evaluation. Marking ranges from 0: "very bad" to 3: "very good".

Overall results Figure H.4 represents the average mark given to all images for each method. We observe that the CycleGAN model and its improvements (Models A, B, and C) are rated better than the reference image for the generalization question (1), and at least as good as the reference for the realism (2) and legibility (3) questions. We tested the impact of population expertise; indeed, in map general-

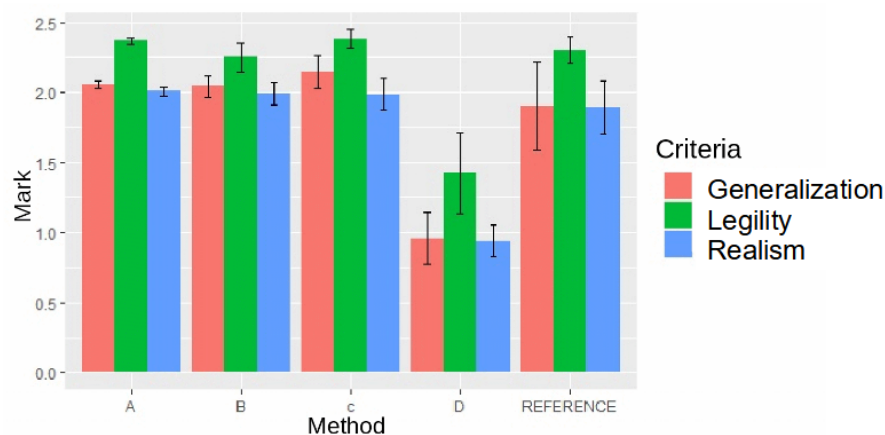


Figure H.4: Evaluation of the different methods according to user marks.

ization, the visual evaluation is often performed by a cartographer who knows the cartographic rules, while in deep learning the user evaluation rather focuses on the

user. In our case, we ask both cartography experts and map users (almost everyone is at one moment a map user) to answer our test. Then, we compare the responses given by the two populations, to verify if a non-expert population that does not know map generalization and cartography rules is able to assess the efficiency of a map. We observe that the results are similar for both populations.

Statistic verification of hypothesis In the following statistical test M_r represents the average realism mark (between 0 and 3) and M represents the average sum of each mark (realism, generalization and legibility, between 0 and 9). The average realism marks for the images from reference ($M_r=1.90$) and from Model A ($M_r=2.05$) are similar. A paired t-test on the realism average marks for each image shows no significant difference ($t(8) = 0.26$, $p = 0.79$) and validates (H1). Model A gives better quality results ($M=6.43$) than Model D ($M=3.32$), a unilateral t-test allows us to confirm H2 (the null hypothesis: learning approach does not affect generalization quality is rejected $t(7) = 7.70$, $p < 0.01$). The ranking evaluation also confirms this idea; Model D is classified as the last for 72.5% of the trials. Finally, the average image marks for Models A ($M=6.43$) and C ($M=6.51$) are similar. A paired unilateral t-test on the marks shows that Model C is not significantly better than Model A ($t(6) = 0.80$, $p = 0.77$) and to refute (H3). The ranking evaluation gave similar results: the location of the reference in the ranking is the one varying the most, and most of the time Models A, B, and C are well placed.

Variation in evaluation In addition to the statistical validation of the hypothesis, we study the variation in evaluation. In fact, each model may be more or less adapted to certain particular situations. For example, we observe that the standard deviation is lower for CycleGAN images than for the reference, i.e., there are more differences in quality perception between different situations when they are traditionally generalized than when they are generalized using deep learning. Deep learning seems to have an averaging effect that smooths generalization. The higher deviation for the improved model (Models B and C) can be explained by the error correction on some images and some other images where errors are introduced compared to the standard CycleGAN.

Extreme values The study of some controversial images (with the larger difference in marks) helps to more precisely explain this deviation. These images are grouped in Figure H.5. Image 25 is the image where people are less agreeing: 25% of the participants rank the reference as the best and 45% as the worst generalization result. The left of this image is a very good generalization, whereas the right contains inconsistent information compared to the input image. This inconsistency in our data, due to different update processes in input and output data, is avoided in the deep learning process. The same kind of problem is visible in Image 166. Then, with regards to the realism question, there are some surprising answers. Image number 294 has a very low realism for Model C only, we cannot explain this difference. Finally, Image 214 have a low realism for each method it reveals that the real shape of the road in this image may not seem to user. In general the evaluation of legibility has a less extreme low value, user consider each image equally legible.

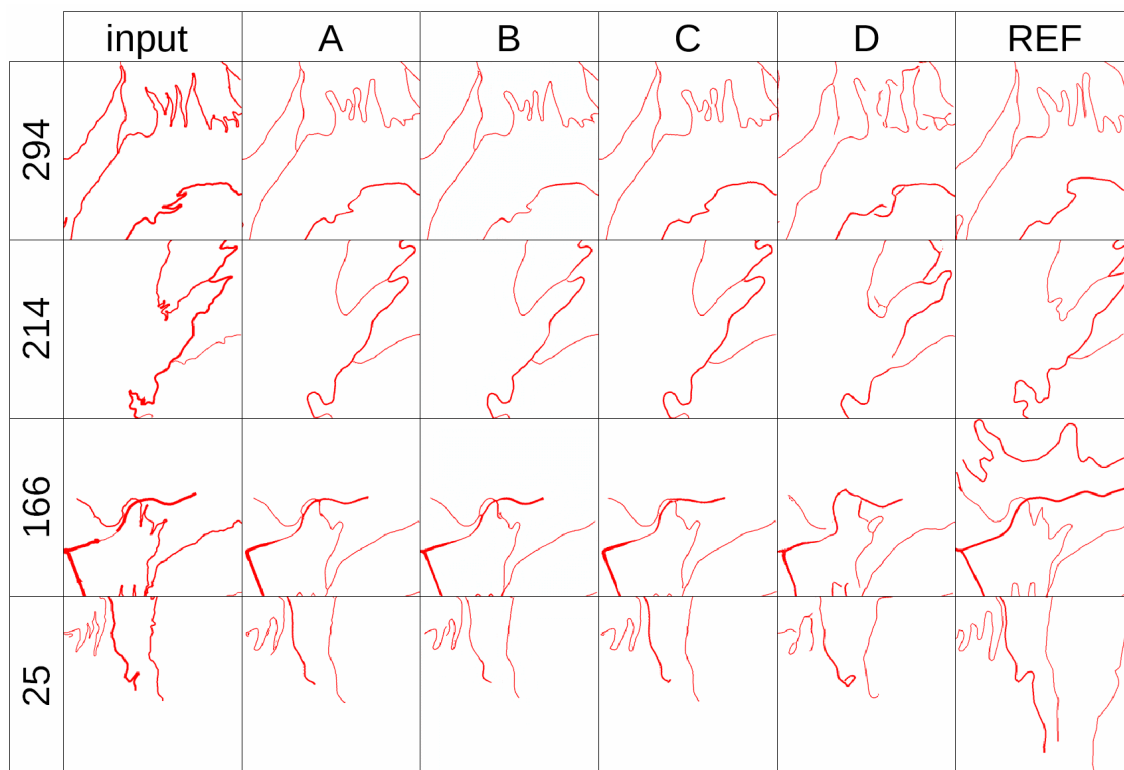


Figure H.5: Examples of images badly evaluated from the validation set. All tiles represent a part of the map of 2.5 km^2 . Each line represents a situation and in each column the employed learning method.

H.2.3 Limitations

Our user test presents the following limitations:

- The images presented are not maps, but maps extracts of $2.5 \times 2.5 \text{ km}$. The evaluation of the usability of such an image in a concrete task (e.g., following a routes) may be more relevant ranking and marking. We think that the evaluation of tiles is necessary to attest their quality and usability in but not sufficient. Moreover, a more complete evaluation of map functioning is not possible without the map reconstruction.
- The number of images for each method is limited and can constitute a bias in model evaluation. We choose the images of the user test manually to be representative of most situations from the test set. However, this image set may contain less simple situation than the global study area. Increasing the number of images presented may require increasing the number of participants to maintain a sufficient number of responses for each image.
- The number of evaluations for each image is limited and not fixed; the random attribution of the test makes the number of responses by image varying, which is not practical for evaluation and comparison between images (if we had to do it again, we would correct this point).
- The questions are imprecise. In particular, some problems in the predictions can be interpreted by some users as information alteration, while considered as a legibility problem by others.

- The topological correctness of the prediction is not evaluated, while it is a major characteristic of road generalization.

H.3 Constraints for raster based evaluation

The definition of constraints is common in map generalization (Beard 1991), but most of the measures of constraint violation are adapted to vector data (Mackaness and Ruas 2007), and not to raster data. Traditionally, the constraint-based evaluation of map generalization is based on a constraint violation measure calculated on vector data. In this section, we describe how some constraints can be adapted to be applied to generalized mountain road images produced using deep learning.

H.3.1 Method

In this section, we present how raster-based constraints can be defined, configured, and applied to the specific case of mountain roads. We focus on evaluating mountain-road graphic generalization for a display at the 1:250,000 scale. The goal is to be able to evaluate the images generated in Chapter E.

The main challenge of this use case is the raster format of maps. In vector map generalization, road segments, faces, and networks are used to set up constraints. In contrast, a raster is a grid of pixels valued, where the limit of each object is not explicitly defined. Consequently, constraints should not hold to the notion of object shape or limits, and other levels of analysis must be explored, such as the pixel, a set of pixels, or the image. These levels are presented in Table H.4. We can construct a pixel set using the following rule: two pixels that are 4-connected and have the same color should belong to the same pixel set. Consequently, a road map is made up of sets of road and background pixels.

Table H.4: Table of possible levels of analysis for vector and raster data.

Level of analysis	Vector data	Raster data
Micro	Roads	Pixel
Meso	Road network faces	Set of background pixels
	Stroke	
	Other sets of road	Set of road pixels
Macro	Road network	Image

First, we directly derived some constraints from the usual requirements of mountain road map generalization and road map generalization (Mustiere 1998; Duchêne 2014a). We expect the mountain road generalization to reduce coalescence in a series of bends and smooth the shape of the road while preserving the position, connectivity, and structure of the network. All these constraints are classical, and we only need to adapt the way they are measured. In addition, the observation of the example images of our use case validates our hypothesis that specific measures are required for deep learning results. We observe that some predictions are unrealistic, blurred, or noisy. Consequently, we decided to add some realism constraints that aim to check that the model produces images that look like real maps.

To validate and configure the proposed measures, we use a *validation set* with both *real data* (images extracted from our deep learning results presented above)

and *fake data* (hand-drawn images of interesting road configurations). For example, Table H.5 lists interesting fake situations constructed by modifying the shape of a random road.

Table H.5: Alterations are applied to create fake test images.

Legibility alteration	Information alteration
Enlarged bends	Appearance of a road
Enlarged road width	Noise in the image
Small displacement	Big structure alteration
Loop creation	Continuity alteration
Blurred road	Continuity alteration at border
Only background in the image	Only roads

H.3.2 Proposed constraints

In this section, we define each chosen constraint and associated measures, validation, and limitations. We implemented all our measures in Python using the open-source libraries PIL and Image².

Clutter Reduction

Clutter can be defined as too much and disorganized information in an image and can be seen as a proxy for the complexity of the image. The clutter has been shown to make image use more difficult (Rosenholtz, Li, and Nakano 2007). This is why we introduce a constraint to avoid the clutter that could arise during generalization. This constraint may be adapted for images of generalized maps, not only for images of generalized mountain roads.

Clutter measures are image-level measures that aim to evaluate the complexity of an image. Their adaptation to assess map legibility remains a question (Dumont, Touya, and Duchêne 2016). We have tested four different measures for estimating the legibility of our road images: edge density, the subband entropy (Rosenholtz, Li, and Nakano 2007), which is similar to JPEG compression, the quad-tree method, which measures the number of homogeneous square cells in the image (Touya, Decherf, Lalanne, and Dumont 2015), and the object segmentation-based measure (Bravo and Farid 2008). Most of these measures represent a quantity; for more clarity, we propose to convert these quantities into a ratio using the formula (eq. H.1).

$$clutter_generalisation = \frac{|clutter_{initial} - clutter_{prediction}|}{clutter_{initial}} \quad (\text{H.1})$$

Figure H.6 presents the value of the edge density measure for the validation images. This measure mainly distinguishes the noisy image from other images. This observation is similar to all the clutter calculation methods tested. The measure succeeds in detecting an increase in complexity in the images. However, some images that a human could judge as quite more complex than the others only show a small difference in clutter value.

²The code is available in the DeepMapGen repository <https://github.com/umrlastig/DeepMapGen>, and should run on any square image, but most measures are only relevant for images made up of roads (pixels close to red) and backgrounds (pixels close to white).

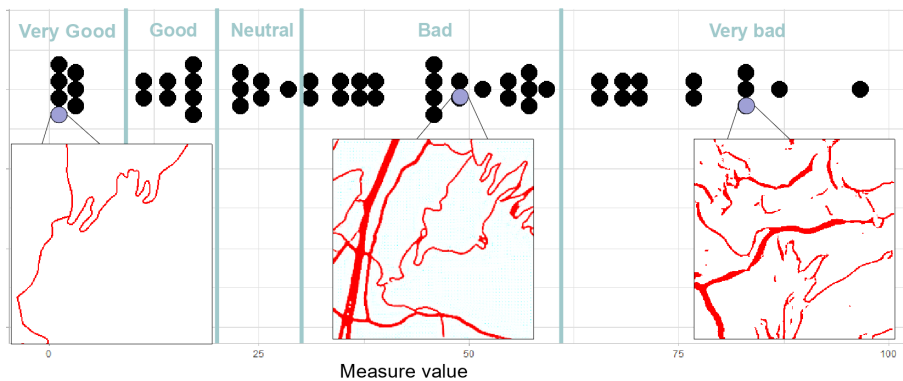


Figure H.6: Distribution and examples of the value for edge-density clutter measure. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

Smoothness

The smoothness constraint is a legibility measure that should make the generalized road smoother than the input road. The fewer irregularities are visible, the more legible the shape is. This constraint is designed for images of generalized roads in general.

To measure smoothness, we have to detect and measure the number of irregularities around the roads. We detect irregularities at the pixel level using mathematical morphology. We simplify the problem by transforming the image from colored (RGB) to a black-and-white map that represents the presence or absence of a road. A closing (dilation followed by erosion) of road pixel with an adapted size should allow us to fill in the concave irregularities of the road shape. Then, the removal of the road shape ($irregular_{pixels} = closing_{pixels} - road_{pixels}$) isolates the irregular pixels near the road, as illustrated in Figure H.7. Then we propose a measure at the image level: the ratio of the initial roughness removed by generalization (H.2).

$$smoothness = \frac{irregular_{initial} - irregular_{prediction}}{irregular_{initial}} \quad (H.2)$$

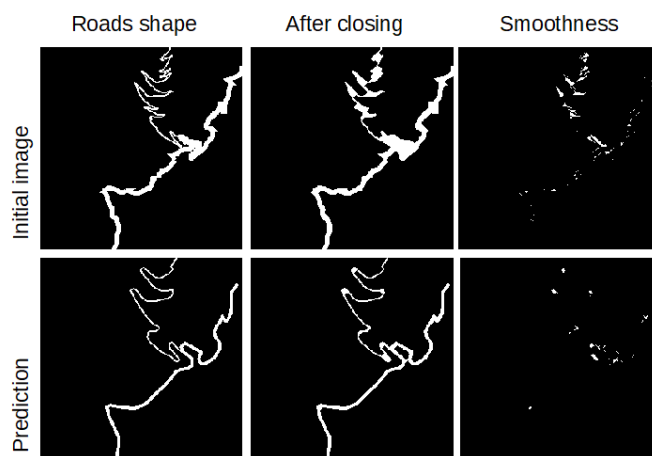


Figure H.7: Effect of the smoothing operation on a road and its generalization.

The size of the closing determines the size of the biggest irregularity that should be smoothed by generalization; if we choose a threshold too large, concave portions that are not irregularities might be wrongly filled. This measure seems to work the best at the resolution and scale of our use case with a closing value of three pixels (30 m). Our experiment has shown that closing with less than three pixels detects very few irregularities, while closing with more pixels fills the bends and tiny regular background areas.

Validation and Limits Figure H.8 represents the distribution and some example values of the smoothness measure. This measure is globally well correlated with our visual perception of smoothness. However, we observe the following limitations:

- The accuracy of irregularity detection has to be questioned; depending on the threshold, irregularities can be missed, and/or some sharp bends can be wrongly identified as smoothness problems.
- The measure depends on the resolution and scale of the road; other use cases might require different parameters.
- If the initial image is not irregular, the ratio measure cannot be computed (this case occurs in 10 percent of images in our test set).

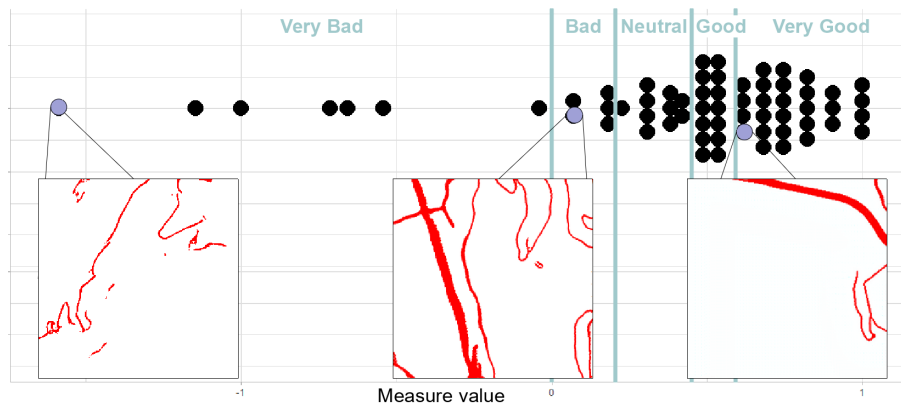


Figure H.8: Distribution and examples of values for the smoothness measure. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

Coalescence Reduction

The coalescence constraint is a legibility constraint. Coalescence is the fact that the symbol of a line overlaps itself in the interior of a sharp bend (Mustiere 1998). The generalized image should not contain symbol coalescence. This constraint is designed for generalized mountain roads, where coalescence might occur.

To measure coalescence, we also use mathematical morphology operations to detect pixels that belong to a coalescent part of the road. This measure is also based on black-and-white images. The measure is also at the pixel level if we consider the map of the coalescent pixels and aggregated at the image level, as we consider the ratio of pixels in an image. First, a dilation of roads (of n pixels) simulates

the coalescence that occurs when the symbol of the road is enlarged due to scale change; then an erosion of the size $n + w$ (the maximum width of the roads) erases the parts that are not coalescent; and finally, a dilation of w pixels evaluates the number of pixels in the coalescent areas. This process is illustrated in Figure H.9. Finally, similarly to the smoothness constraint, we propose to measure the evolution of coalescence instead of the number of coalescent pixels using (Eq. H.3).

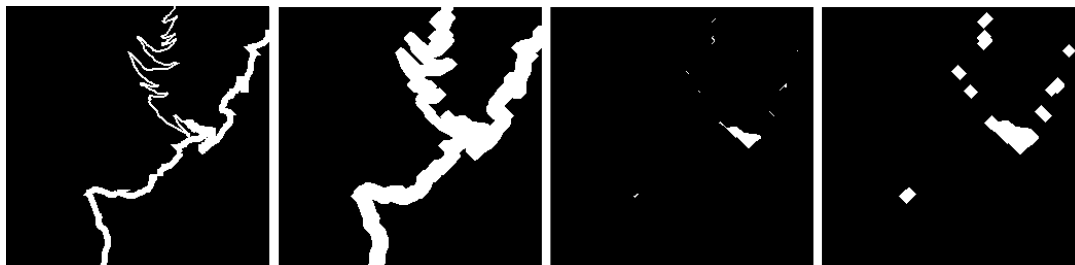


Figure H.9: Steps for detecting coalescent pixels, from left to right: initial image, dilation, erosion, and dilation.

$$Coalescence = \frac{coalescent_pixel(initial) - coalescent_pixel(prediction)}{coalescent_pixel(initial)} \quad (H.3)$$

In this process, the size n represents the maximal distance between two coalescent parts of the road from the map specifications. For our use case (at the 1:250k scale), a threshold of five pixels is the most adapted, and it corresponds to a distance of 50 m. This value allows the detection of most of the coalescent areas. However, it also detects some non-coalescent areas, as presented in Figure H.10, while a thinner threshold would miss many coalescent areas.

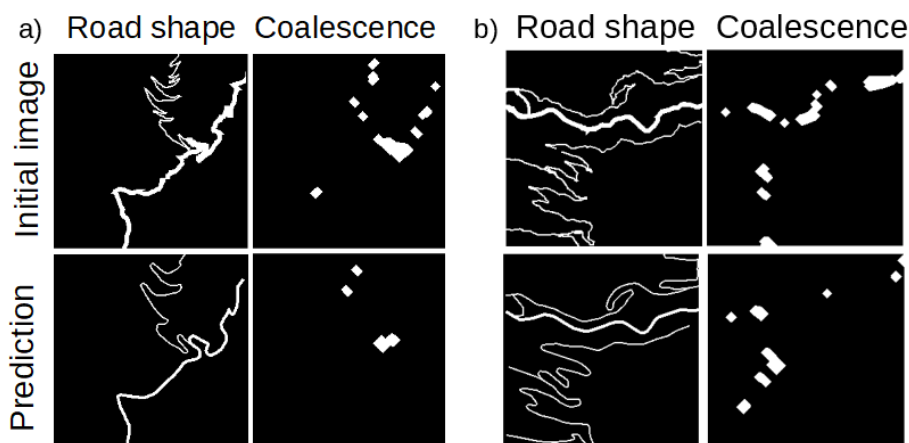


Figure H.10: Effect of coalescence detection on a road and its generalization. a) Example with a few over-detections. b) Example with an over-detection of coalescence in the generalized image.

Figure H.11 presents the distribution and some examples of values for the coalescence reduction measure. This measure is more questionable than the others. First of all, it does not exactly measure the coalescent parts but the pixels that could be coalescent with a larger symbol. This definition implies that the overlapping

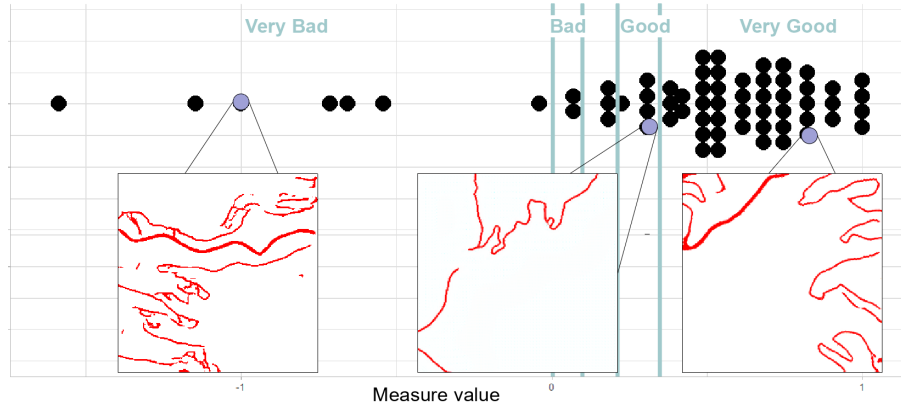


Figure H.11: Distribution and examples of values for the coalescence reduction measure. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

parts of the coalescent road are identified as a unique road and are not detected as coalescent. The observed ratio value is most of the time consistent with our visual evaluation of coalescence, but some high coalescence values are explained by two parallel roads that are separated by fewer pixels than the threshold (detected as coalescent by the measure), and this is not coalescence, as these are different roads, but just symbol overlap. These situations are limited to 5% of the cases in our real test set.

Position Preservation

The generalized road must be as close as possible to the initial one. This constraint is designed for an image of generalized roads, in general.

We measure this constraint using the intersection over union (IoU) of the road pixels of the initial and generalized images. Here, we also based our computation on transformed black-and-white images. The intersection between the road pixels in the initial image and the road pixels in the generalized image is computed, and the intersection area is divided by the union area, that is, the pixels that belong to the generalized or the initial roads (Eq. H.4). This is a classic measure for assessing image segmentation and determining which parts of the road pixels are common in the initial image. Figure H.12 represents the intersection and union of a road and its generalization.

$$IoU = \frac{road_pixel(initial) \cap road_pixel(prediction)}{road_pixel(initial) \cup road_pixel(prediction)} \quad (H.4)$$

We propose to use a buffer of size n around roads instead of just the road pixels. This threshold should model the area where the roads are allowed to move or be distorted without considering the position as too far. This threshold can be very large because the displacement offset can be large for mountain roads at this scale (1:250k). In our image set, the position is always visually preserved. However, the measured position accuracy is usually less than 50%, so introducing this buffer with a size of 20 pixels (200 m) better reflects this preservation and the variation between the two generalization methods: some cause more displacement than others.

Figure H.13 represents the distribution and some examples of values for position

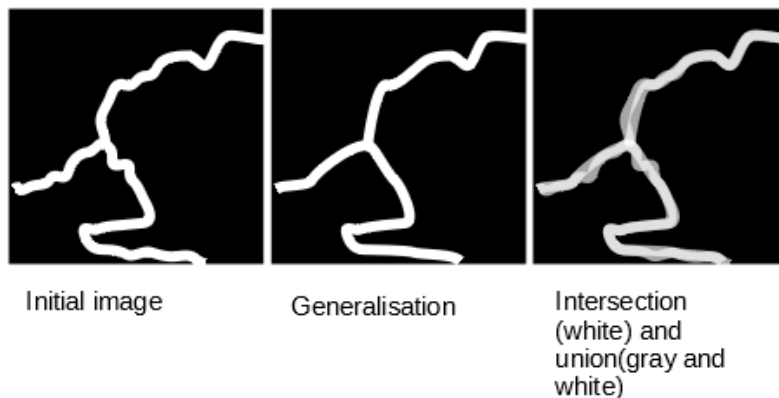


Figure H.12: Intersection and union of roads and their generalization.

measurement. Despite the great dissimilarity of the IoU values in our image set, we observe that all real images visually respect position accuracy, even those with a very low IoU threshold. Moreover, the measure is not regular: with a similar displacement, the IoU of two large objects will be more important than the IoU of two small objects. Then this measure quantify the displacement but does not identify whether the displacement is relevant. Road displacements and distortions are allowed when they tend to increase legibility. Finally, the accuracy of the road position is a minor constraint: it can be seen as a soft constraint whose satisfaction does not need to be maximized.

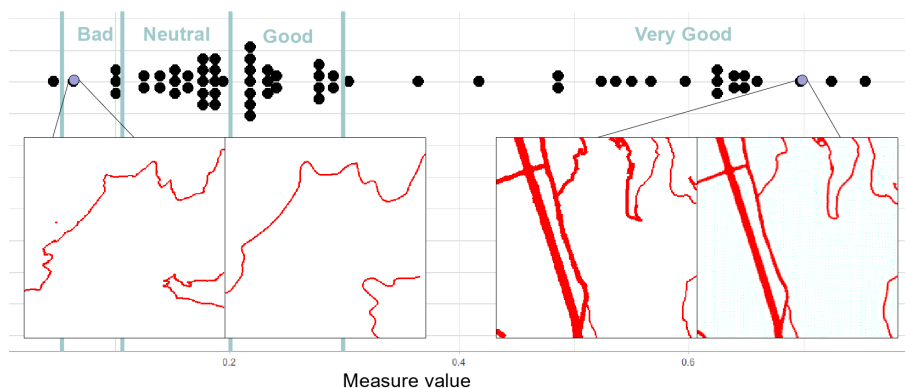


Figure H.13: Distribution and examples of values for the position measurement. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

Road Connectivity Preservation

Then, we think that more than the position, the preservation of the structure of the road network is a key point for road maps. With road connections, map readers can follow routes, which is the most important function of such maps. This constraint is adapted to images of generalized roads in general.

For the connectivity preservation constraint, we measure connections and disconnections using the number of sets of contiguous road pixels and contiguous background pixels. This measure is based on the same principle as the connectivity loss proposed in C.1.4: a disconnection will reduce the number of background pixel sets

and increase the number of road pixel sets. This measure also detects problems such as the creation of a loop, which deteriorates both the correctness and legibility of the information. Equation H.5 shows how we combine these numbers of contiguous parts to measure continuity preservation for (x, y) a pair of detailed and generalized images.

$$structure_alteration = \frac{|nback(x) - nback(y)| + |nroad(x) - nroad(y)|}{nback(x) + nroad(x)} \quad (\text{H.5})$$

We choose to use the 4-connectivity to construct the sets of pixels; pixels of the same value connected by diagonal only do not belong to the same group of pixels. We also include a threshold for set size: the groups composed of one or two pixels are not considered, consequently the appearance or disappearance of such group is not measured.

Figure H.14 gives the distribution and some examples of values for the connectivity measure, each pair represents an initial situation and the corresponding generalization. This measure is quite consistent with our observation, but it seems to overvalue the structure changes. This problem is explained by the fact that all changes have equal weight in the calculation, even the addition of very small sets or the disconnection of one pixel, which is not visible at the target scale. Then, this measure is sensitive to border effects (illustrated in the last column of Figure ??). When a road has bends at the border of an image, these bends can be connected or disconnected by a small displacement, without a real change in the structure.

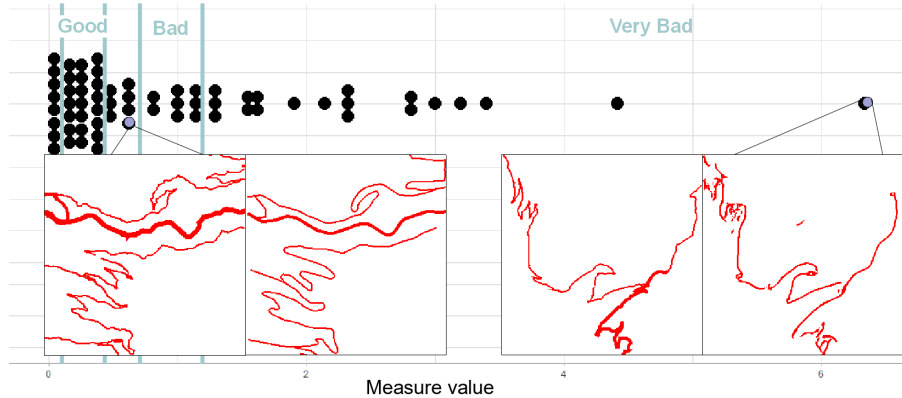


Figure H.14: Distribution and examples of values for the connectivity alteration measure. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

Color Realism

The next two constraints are not adapted from conventional vector-based constraints but are specific to images generated by deep learning models. They aim to quantify errors that do not exist in traditional map generalization. The goal of the color constraint is to ensure that the generated images do not convey unexpected information. On a map, a different color represents a different object. This constraint is adapted for all images of maps generated using deep learning.

We propose to measure if there is color noise: in our use case, it means pixels that are not roads (red) or backgrounds (white). We have decided to use an implementation of the CIEDE2000 distance between colors (Sharma, Wu, and Dalal 2005) because of its consistency with human color perception. Then, we counted the number of pixels that are visibly too distant from red or white. We choose to consider a pixel white or red when the distance with pure red (255,0,0) or pure white (255,255,255) is greater than 9.

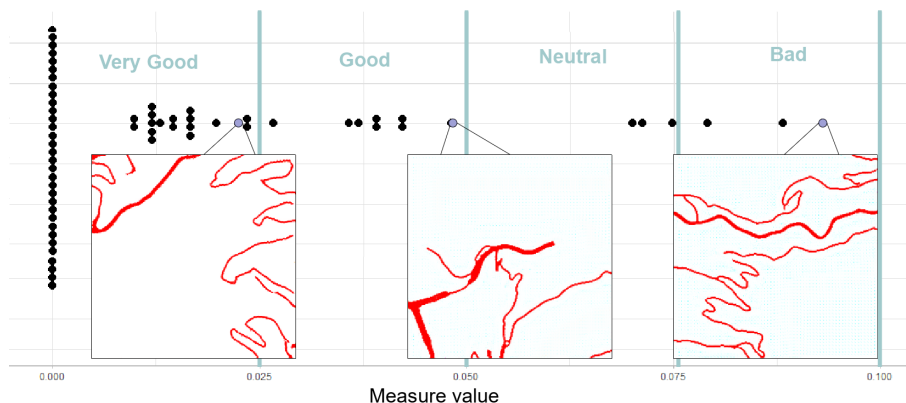


Figure H.15: Distribution and examples of values for the color measure. Each point represents an image from our testing set, the corresponding level of satisfaction of the constraint is indicated in green.

Figure H.15 represents the distribution and examples of values for the color measure. This measure has allowed quantifying the observed noise of CycleGAN images and verifying that images from other methods do not suffer from this problem.

Noise Absence

We also found that some deep learning models can generate noisy pixels: for example, isolated road pixels that are not roads. This is because the loss function optimized in the training process is global and might not be fully optimized after all iterations of the training. This constraint is designed for all images of maps generated using deep learning, not just generalized mountain roads.

We just count the number of sets of contiguous road pixels that are too small to be real roads. Similarly to the road connectivity preservation constraint, we use 4-connectivity to define sets of contiguous pixels. The noise detection threshold depends on the resolution of the image. For our use case, we have fixed a size below 6 pixels. Figure H.16 presents the distribution and some examples of values for this measure. This measure is quite satisfying, as it permits us to distinguish noise in both the fake images that we created with noise in our validation set and in the images generated by a U-Net model, which often contain noise pixels.

H.3.3 Interpretation and usage of constraint values

In this section, we present how the proposed constraint can be interpreted and used for validating a model, identifying and solving its quality issues, and comparing deep learning models.

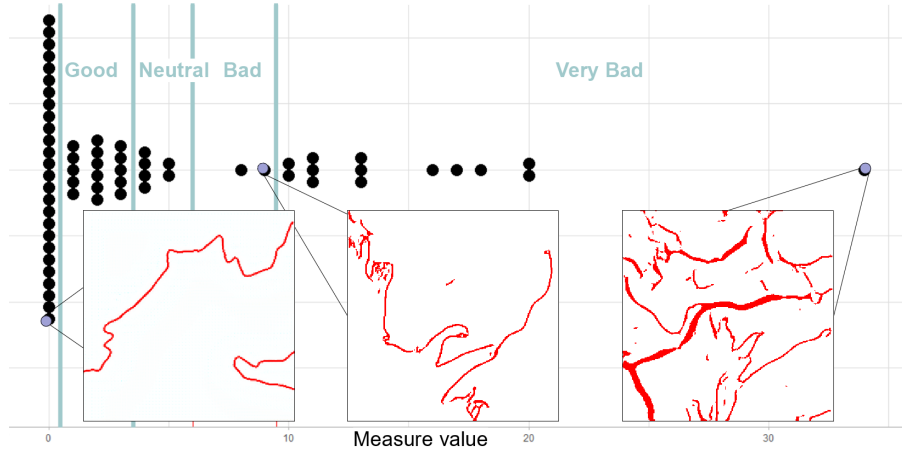


Figure H.16: Distribution and examples of values for the noise measure. Each point represents an image from our testing set; the corresponding level of satisfaction of the constraint is indicated in green.

Constraint values interpretation and model validation

We choose to define a qualitative level of constraint satisfaction: "very good", "good", "neutral", "bad", "very bad". We propose the definition of a threshold for each satisfaction level based on the examination of the validation set of our experiment. The satisfaction values for the constraints are presented in Table H.6.

Table H.6: Interpretation of constraints in a level of satisfaction (at the image level all values are in percent except noise that is in pixel).

Constraint	Very good	Good	Neutral	Bad	Very bad
Color	$x < 2.5$	$2.5 \leq x < 5$	$5 \leq x < 7.5$	$7.5 \leq x < 10$	$x \geq 10$
Connectivity	$x < 10$	$10 \leq x < 40$	$40 \leq x < 70$	$70 \leq x < 100$	$x \geq 100$
Clutter	$x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 60$	$x \geq 60$
Smoothness	$x > 60$	$60 \geq x > 40$	$40 \geq x > 20$	$20 \geq x > 0$	$x \leq 0$
Coalescence	$x > 30$	$30 \geq x > 20$	$20 \geq x > 10$	$10 \geq x > 0$	$x \leq 0$
Position	$x > 30$	$20 < x < 30$	$10 < x < 20$	$5 < x < 10$	$x < 5$
Noise	$x = 0$	$x < 3$	$3 \leq x < 6$	$6 \leq x < 9$	$x \geq 9$

To validate a model, we propose to verify that all its predictions are evaluated above a defined threshold of satisfaction for each constraint (e.g., "good" or "very good"). Then a model with a majority of valid images is valid (Touya 2012).

Identifying and solving the quality issues of a model

In this section, we explain how constraint evaluation can be used to detect quality issues of a given deep learning model and then eventually improve the model. In this experiment, we provide two levels of constraints: at the pixel level (e.g., smoothness and coalescence identify the irregular or coalescence pixels) and at the image level (e.g., clutter measure or position accuracy).

At the image level, the constraints provide a comprehensive indicator of quality for the constrained property.

- If a characteristic is often badly generalized, it means that the model is not adapted to learn to reproduce this characteristic (e.g., disconnections in our experiment). Since this default has been identified, it can be solved by exploring the adapted loss functions.
- If a situation or a kind of situation is often poorly generalized, it might reflect an imbalance in the training dataset. For example, in our experiments, the cases with two close parallel roads are badly generalized because they are under-represented. Then, adaptations of the training set can be performed.

On the other hand, at the pixel level, constraints identify conflicts and situations that are badly generalized within each tile, which enables post processes with other generalization techniques.

Comparing Models

In this section, we propose to explain how evaluation can be used to compare different deep learning models or similar models used with different parameters. A good generalization solution is a balance between several criteria, so it is not easy to determine if a solution is better than another. Constraints could contribute to the evaluation of which network configuration gives better results in an ambiguous case where it is visually difficult to determine a preference. For comparison purposes, a good combination of different measures is needed to give a synthetic measure of the global quality of a tile image. Finally, the average values are often not sufficient for this task, and the worst or last deciles are interesting to find the weakness of a model. We propose in Figure H.17 a visual comparison of the average and worst constraint values for several models for mountain road segmentation. It is a convenient representation for model strength and weakness comparison and confirms that cycleGAN gives as good results as the reference (see Chapter E).

H.3.4 Discussion

Limitations

Image representation allows for a visual evaluation of generalization. Automating this evaluation is a complex combination of multiple constraints. We described how some constraints can be adapted for a calculation based on raster data. However, we do not believe that all existing constraints can be converted using this method. Especially, the relation preservation constraints seem to be challenging as the raster format does not encode object boundaries and spatial relations.

Moreover, the evaluation of a map must be related to the map function. The automatic evaluation of map efficiency is still challenging for both raster and vector data, as it relates to human perception. Finally, the proposed measures still suffers from multiple limitations.

- The measure is sometimes not representative of a real difference in quality or of the desired characteristic (clutter and position accuracy measures).
- The measure is sometimes not accurate and can miss or overestimate some errors. This is the case for the smoothness, connectivity, and coalescence measures.

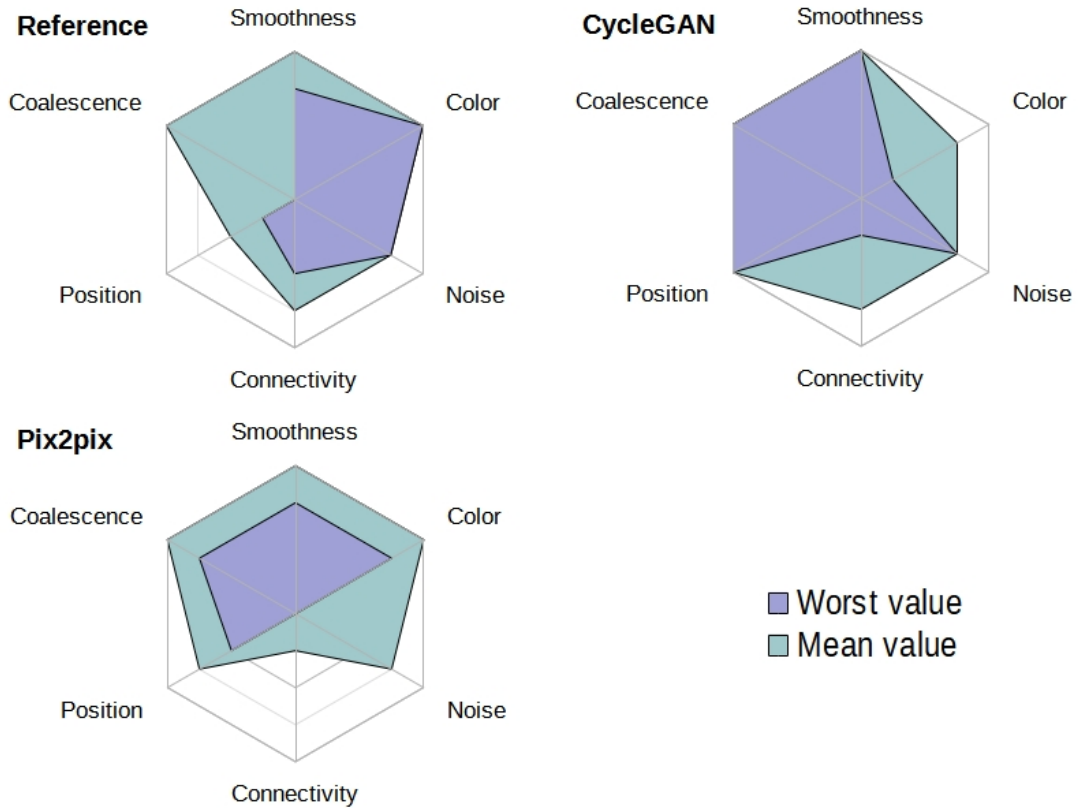


Figure H.17: Comparison of models strengths and weakness, with a visual representation of constraint satisfaction.

- Some measures are sensitive to resolution and scale (smoothness, coalescence, and position measures).
- Some measures vary more according to the initial situation than according to the quality of the prediction (smoothness and coalescence values).
- Some measures are not calculable (division by zero) for some images (smoothness and coalescence measures).
- The connectivity measure may suffer from some border effects.

All of these limitations mean that the proposed constraint-based evaluation is not sufficient for analysis without visual exploration of predictions. Therefore, the evaluation of the prediction provided by this evaluation needs to be nuanced. An improvement and completion of our constraint may be needed for evaluation tasks that require fully automatic evaluation, such as controlling and tuning deep learning processes.

Do we need a meso level of analysis?

Some of the major limitations identified above may be due to an unadapted level of analysis. In this paragraph, we discuss the interest in using a meso-level of analysis.

First, the border effect can be reduced by the use of a sliding window over the image to create a new sub-image or a patch that can be better adapted for the calculation of some constraints. This process can permit locating problems in the

image for constraints that are only measurable at the image level. For example, the connectivity constraint of an image that suffers from border effects would have a high value for a patch at the border but not for a patch in the middle of the tile (see Figure H.18).

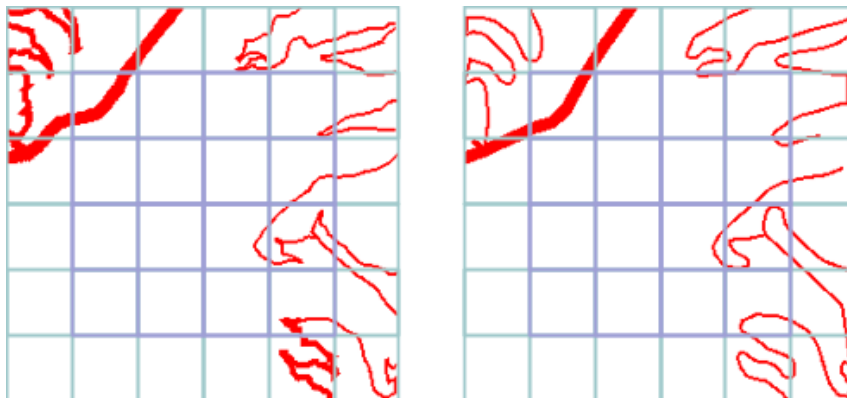


Figure H.18: Illustration of the patches to detect border effects: green patches would have a greater value than purple ones.

Then, we believe that it could be useful to define an intermediate level of constraint. It would synthesize the satisfaction of several pixel-level constraints for each pixel contained in a meso object (Ruas 1999). For example, instead of measuring the ratio of the pixels that are not smooth in the whole image, we could synthesize the smoothness at the level of a pixel set (e.g., set of contiguous pixels with the same value, patches, or image segment from image processing). Then, an image is smooth if all or most of its pixel sets are not irregular. This definition first designed for agent models is also adapted to guide deep learning processes, where the loss is a statistical summary of the values of pixels or patches.

Can these methods be adapted to a multi-theme map?

Then, a challenge is to adapt this constraint-based evaluation to maps with several themes, not just roads. In these maps, the spatial relations are even more important and concern several different objects. For example, buildings are aligned along a road or a road is crossing a forest. The preservation of these relations is an important point in the preservation of information, as they contribute to the global structure and understanding of the map (Mackness and Edwards 2002) and help determine the location of the map.

However, the raster representation makes conflict detection harder on target and prediction map, as each pixel can have only one value even if there is a coalescence of several elements at this place. Finally, the identification of groups of objects can be done using a distance criterion using buffers. We propose to estimate the difficulty of the raster adaptation of the constraints (Stoter et al. 2009) in the Appendix B. We identify four levels of adaptability of a constraint: some constraints can be measured directly in the image (e.g. noise quantity, size of objects); some may require image processing (e.g. granularity, squariness), some may require a new manner of calculating (e.g. orientation measure, topology preservation measure), and some seem impossible to adapt without vectorization (e.g. spatial distribution preservation measure).

Finally, we think that raster evaluation can be useful at a stage where vector conversion is not ideal. However, the vectors remain more accurate and practical and permit comparison with classical generalization approaches or validation.

H.4 Vector based evaluation

In this last section, we explore an evaluation that includes vector conversion. Such a process can be used to compare our results with a classical map generalization approach or to evaluate some aspects of predicted maps, where the raster adaptation is not possible. For illustrating this evaluation method we use building generalization as a case study. Indeed, we observe in Chapter G that building shape is sometimes not credible. However, the image format does not allow the quantification of this problem.

H.4.1 Vector conversion for tiles of generalized buildings

Building tiles to vector

To vectorize the buildings on tiles, we propose the following process:

1. Projecting images: during image creation, we stored the coordinates of the input image, and we use this information to project the corresponding output image of the map.
2. Constructing a building mask. We identify the pixels representing a building in prediction (e.i. whose color is closer to the color of the buildings in the target style than any other color of the target style.)
3. Polygonize. We use the `gdal_polygonize` function that creates vector polygons for all connected regions of pixels in the raster that share a common pixel value.
4. Pixel smoothing. Vectorization cannot be evaluated directly due to the notched shape due to the pixel limits. Thus, we apply a Douglas-Peucker filter (threshold = 2m) to simplify the shape (see Figure H.19). This simplification does not generalize the building and globally preserves the generalization problems of the building, but it smooths the pixel border.
5. Elimination of the truncated building at the border of the image.

Evaluation

Then we propose to evaluate the generalization of the shape of the building. In particular, we want to quantify the not-credible shape irregularities. We propose several measures to compare the quality of the prediction without and without the geometry consistency loss.

- **Granularity** We measure the length of the smallest building edge for each building, this measure indicates the granularity of a building.

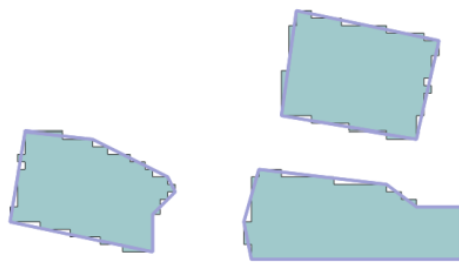


Figure H.19: Vectorization of the building (green), and a Douglas Peucker simplification (purple).

- **Number of significant orientations** We clustered edges according to their orientation, the contribution of an edge to orientation is its length. Then we count the number of significant orientations i.e., above a certain threshold (Duchêne, Bard, Barillot, Ruas, Trévisan, and Holzapfel 2003). The more orientation a building has, the less credible its shape is, in fact, even a very complex real building often has very few different wall orientations.
- **Rectangularity coefficient** We measure the difference in area between the building shape and its minimal bounding box; this measures how far a shape is from a rectangle (in a similar way we could compare it with the maximal rectangle included in the shape). However, it is only relevant in quasi-rectangular-shaped buildings, and in the case of more complex buildings, this measure is not relevant.
- **Deviation to the most credible angle** Finally we sum up how far the angle is from a flat or right angle, for each pair of the consecutive edge of the building (Touya and Lokhat 2016).

We compare the distribution of these measures for both buildings generalized with and without the geometry consistency (see Table H.7). This comparison shows really similar results that could not prove the interest of geometry consistency loss for building generalization. If we could not show a significant difference between models using this evaluation, it has been shown a significant difference between entities and in particular between several generalizations of one building. In fact, the overlapping between tiles conduces to several generalizations for each building and the proposed evaluation could help to choose which is the best assembly tiles.

Table H.7: Table of statistiques on regularity measures for prediction of model with and without geometry consistency (GC).

Constraint	With GC				Without GC			
	min	max	mean	median	min	max	mean	median
Granularity (metre)	2	32	7.7	4.5	2	38	8	4.5
Number of edges	4	70	7.6	6	4	70	7.5	6
Rectangular coefficient (percent)	0	2.3	0.3	0.2	0	1.7	0.2	0
Deviation coefficient (degree)	0	45	15.6	14.0	0	45	15.7	14.0

Chapter's conclusions

The development and integration of deep learning into the map generalization process is conditioned to the existence of common and adapted evaluation methods. The evaluation of deep learning-based map generalization is a challenge. Table H.8 summarizes for each of the main evaluation objectives identified, if the proposed evaluation is adapted.

Objective	Constraint adaptation	User test	Vector conversion
Tuning	✓	✗	✗
Controlling	✓	✗	✗
Setting	✓	✗	✓
Descriptive evaluation	✓	✓	✓
Evaluation for grading	✗	✓	✓
Evaluation for post-processing	✓	✗	✗

Table H.8: Applicability of proposed evaluation methods to main evaluation objectives.

The human perception of a prediction can be assessed with a user test. Such evaluation is closer to the real expectation of generalization (we do a generalized map for the user). However, it is not practical and the presentation of small tiles is biased. Measuring constraints is a more formal evaluation, but, it is sometimes not correlated with human perception. Moreover, the calculation of traditional constraints cannot be performed directly on the raster. We experiment with both strategies: adapting constraint measures for vector calculation and converting to vector prior evaluation. The first option is more practical when the evaluation has to be integrated in a fully raster pipeline but can require image processing (segmentation, mathematical morphology statistical analysis, etc.) or a new level of analysis (e.g., considering a group of pixels instead of an object) and even sometimes it seems not possible to evaluate. The second option is useful when compared with the non-vector approach required or when adaptation is not possible.

Conclusion

Reminder of the objective and approach

The objective of this thesis was to explore the potential of deep learning to contribute to map generalization research.

We conducted this exploration using four use cases. For each use case we highlighted potentials and issues of deep learning and current use of deep learning models are still far from practical map generalization and cartography.

- **Road network selection and alignment detection**, both these experiments assess the potential of deep learning to encode spatial relations with graph modeling. They have shown potential to make predictions that are consistent with spatial relations. The road selection experiment is satisfactory for most situations, but the results suffer from the lack of city center examples. The alignment detection is less satisfying and suffers from both dataset and model limitations.
- **Road shape generalization**, this experiment shows that most of the knowledge for shape generalization can be captured from an image. The unpaired generative model gives results considered as good as the reference map. It correctly achieves smoothing, enlargement, and caricature operations.
- **Holistic map generation**. This experiment shows that GANs can generate map-like images that include generalization. Our results succeed in reproducing the simplification, enlargement, and typification of buildings. However, more context is needed for road selection and graying in the city center. We proposed an organization of geographic information that seems promising for map generation.

Can deep learning contribute to map generalization?

All these experiments reveal the ability of deep learning models to learn, to interpret, to abstract, and to represent geographic information from images or graphs. It leads to the conclusion that deep learning is promising for the automation of map generalization.

First, the use of deep learning as an indicator for data enrichment prior to map generalization is already possible and beneficial. Indeed, the quality of prediction in our experiments and in the literature review is satisfactory, and the integration and

evaluation of such models is simpler than when they are used as a generalization operator.

Then, the use of deep learning models as generalization operators seem possible. This usage would require further development of specific models for map generalization, in particular for the encoding of spatial geometry and relations. Moreover, for integrating deep learning-based operators in a traditional generalization process, the explanation and evaluation of such models have to be investigated.

Finally, the holistic approach is more a long-term goal than a short-term objective. Both explored approaches (with a unique end-to-end model and with a workflow in several steps) are promising but require still lacks the mechanism to better integrate the spatial context required in most map generalization tasks. The workflow approach seems to offer more possibilities, as each step may have specific needs in terms of the number of examples; input information representation; architecture; and loss function.

Contributions

Our exploration leads to the proposal of several solutions for each use case. Thus our contribution include datasets and models for each proposed use cases. In addition, we propose:

- The **layer representation**, a method for representing geographic information in an efficient training dataset for learning map generalization from images. It allows to encode two different levels of information (main and additional information) in a machine-readable way. The code to generate tiles with this representation from vector geodatabases is available as a plugin of the open source generalization platform CartAgen <https://github.com/IGNF/CartAgen>.
- **DeepMapScaler**, a deep neural network workflow for generalized map generation. In this workflow, we organize and divide the generation of generalized maps into simpler tasks that deep learning now seems capable of solving. Unlike approaches in the literature that attempt to train a single model for map generation, this approach precisely tailors the dataset and model component for each step of the generalization and makes the generation of generalized maps possible.
- The **connectivity loss**, a measure of road network connectivity alteration that can be used as a part of deep neural network objective function <https://github.com/ACourtial/DeepMapGen.git>.
- **FusePix**, an architecture able to use the additionnal and main information in a GAN <https://github.com/ACourtial/FusePix.git>.
- **AlpineBends**, a benchmark for the graphical generalization of mountain roads that includes a training set and several test sets to study the transferability of the model (available on Zenodo: [10.5281/zenodo.5257686](https://zenodo.org/record/5257686)).
- An evaluation method based on adapting the traditional map constraint for an image-based measure. This includes a proposal for six constraints for evaluation of the AlpineBends benchmark (code available in DeepMapGen's Github

repository: <https://github.com/umrlastig/DeepMapGen>); and the study of the adaptability of twelve other classical map generalization constants from the EuroSDR framework (see Appendix B).

- An evaluation method to assess and compare the regularity of building shape in deep learning predictions.
- A user test to evaluate GAN prediction for mountain road generalization.
- A method for reconstructing and assembling independently generalized tiles into a vector database.

Lessons learned

Due to its exploratory approach, this thesis deals with various research questions, and the lessons learned from the exploration are varied. Our exploration provides first lessons on the potential of deep learning for map generalization:

- **Can graph convolution networks encode the spatial relations that are important for cartography?** In our experiment on road selection, GCN reveal an ability to understand/encode some relations in a graph. However, the definition of neighbors in graph convolution and cartography is different and the construction of a graph adapted for learning has to be investigated.
- **Can image-based approaches learn graphic generalization?** Our experiments on mountain roads show that image based model can learn to reproduce most generalization operator that affect the shape of a road (simplification, schematization, bends enlargement etc.), the predicted image are close to our expectations and lead us to think that image approach is relevant for graphic generalization.
- **Can GAN learn the generation of generalized maps?** Our experiments show that it is possible to generate maps that look like the target map and respect the target level of detail.

The exploration then provides responses to the research questions identified in Chapter A.

How to automatically construct an adapted training set from IGN data?

The construction of an adapted training set is mainly based on the quantity, diversity, and quality of the examples. The examples must not include errors and have to be legible and adapted to the target task. They must contain the necessary information to allow the model to learn the desired task (in particular, map generalization requires several levels of contextual information, and the representation of this context is an issue). The main conclusions of our experiments on dataset creation are the following: 1) the more complex the task is, the more examples would be needed; 2) the choice of the study area for the example extraction affects both the scope of the possible situations resolved by the model and the transferability of the model; 3) in general, the more complete and simple the representation of information is, the easier the learning would be. Moreover we observed that the IGN data was not perfect, and that it must be manually corrected to avoid inducing the model to reproduce the residual errors.

How to conceive an adapted deep neural network? In Chapter C, we identify seven criteria for the choice of an adapted deep neural network: abstraction ability, interpolation ability, preservation ability, transferability, spatial ability, variability, and gluttony. Furthermore, our experiments have shown that the adapted model may require to encode multiple level of spatial information and context. Subsequently, we believe that if the generic models tested in our exploration are able to demonstrate the potential of deep learning for map generalization, it is now necessary to develop specific and adapted models. This would allow to reach a level of prediction quality, that would further unlock the integration of the deep learning model in the map production process. In conclusion, we suggest that research on the conception of such a model has to be directed towards an architecture with several encoding path (for the different levels of information and context) that would focus on improving one or more of the seven identified abilities.

How to deal with the specific aspects linked to geographic information and map generalization? This research question is about the spatial ability of deep neural networks. Some elements of common neural network tend to fit spatial characteristics by definition. For example, convolution series first encode small-range relations and progressively search for long-range relations, this encoding is consistent with the principle of geography (close things are more related than far ones). On the contrary, we identify some aspects that are not adapted to spatial data. For example, most of the loss function calculations are based on a pixel level comparison. However, in cartography and map generalization, a more global comparison of the image is needed.

How to evaluate the results of deep learning-based map generalization? The evaluation of the deep learning-based map generalization must be more complete than the classical map generalization because we have to ensure that the prediction follows cartographic rules and is a valid map. In addition to the visual assessment of our results, we propose three different approaches for evaluation (adapted constraint, user test, and constraint evaluation on converted data). We believe that the most promising method is the completion and adaptation of the generalization constraints.

How could the deep neural network collaborate with previous techniques for automatic map generalization? Each problem formulation (Chapter B) requires a different integration in automatic map generalization process. Globally we identify two kind of usages of the deep learning models: as a data enrichment step and as an operator. We also found that the difficulty of the integration lies mainly in the format conversion because, for now, most automatic map generalization processes are vector-based and this representation cannot be used in deep learning-based map generalization.

To what extent is it possible to transfer or apply the models learned with deep learning to large areas? The transferability seems sufficient for the transfer to a similar situation in other areas; however, fine-tuning may be required in cases where the input or target are too different from those presented in the example set (different data sources, different geographic configuration, different scales, etc.).

The use of image representation is a major obstacle to the encoding of spatial relations. In fact, the use of image representation is not mandatory for deep learning-based map generalization and other approaches are also promising (see Chapter B). The image representation, even if it does not model cartographic objects explicitly, seem able to encode and preserve relations between objects and seem to include most of the necessary knowledge for map generalization.

Future research

During our exploration, we also identified some new research questions that must be addressed in a future research.

Experiments on spatial relations

The experiments on the spatial relations (in Chapter D) have shown the complexity of the creation of the best graph for the task. In this paragraph, we present some perspectives to improve alignment and selection predictions.

Improving the graph for road selection The graph used for road selection is the road section graph. However, the selection of the road network is often made on strokes, i.e. groups of connected road sections that follow the Gestalt principle of good continuity. Because a stroke should be entirely selected or eliminated, we could improve this use case by constructing and learning on the stroke relation graph (graph where node are strokes and link represent stroke connections (Touya 2010)) instead of the section graph.

Improving the graph for the detection of alignments The alignment detection is performed on a proximity graph, and the method chosen for the construction of this graph does not seem satisfying to us. In particular, we would like to compare more construction methods, including a triangulation constrained by roads and a triangulation not based on the center point but on several building points. We also think that the number of examples and the consistency of the examples for this use case could be improved, but it requires time to manually annotate alignments.

Simplifying the tasks Urban and rural areas respond to very different rules for both the tested tasks (alignment detection and road selection). In particular, urban data are denser and then require an implicit relation between the farthest objects, therefore we think that it can be beneficial to train two separate models with different depths for urban and rural data.

Experiments on mountain roads

Despite the good quality of our predictions, they face a certain number of limitations that make them undesirable for production. The main problem is the alteration of the structure of the road network (loops and disconnections). Among other more minor problems, we can cite the absence of context that causes errors in displacement; the noise in the image and the difficulties in tile assembly.

Training set improvement One of the reasons for loops and disconnection is that the input of training images contain too much detail for the resolution, that makes them few legible. Thus, it is necessary to improve the training set. On one hand, we believe that increasing the number of examples (and especially example of very narrow and sinuous bend series that are underrepresented) would benefit to the model. On the other hand, increasing the image quality seem difficult: larger tiles or tiles with a better resolution would require to widely increase the computation ability and seem not suitable; and constructing tiles with a more relevant splitting method may be beneficial, but impossible. Then, the lack of displacement causes road overlaps, but, to learn that the roads should be moved, we first need to know where are other close objects on the map. Therefore, we want to experiment with the layered representation of main map theme for learning road generalization that includes displacement.

Model improvement We suggest testing the use of a more adapted model for a geographic image (e.g., Deng, Tian, and Newsam 2021) for making a stronger prediction. In a similar way, we think the use of a specific term in the objective function can be a good solution to encourage (or discourage) a specific characteristic in prediction. However, our proposed connectivity measure penalizes these errors without completely avoiding them. We think this measure can be improved to emphasize visible connectivity alteration and the border effect. Moreover, other kinds of complementary loss terms can be experimented to penalize other problems in prediction (e.g. fuzzy border, isolated road section, etc.). Then, we would like to experiment with the depth of the model and the size of the convolution kernel to encourage the model to look at long-range relations between pixels in the images. Indeed, in road images, two far pixels may be related (belong to the same road, while in a building, the pixels belonging to the same entity are closer).

Work on post processes Finally, we proposes a method for post-processand assembly of tiles generalized independently. It imply to match roads that can have very different shapes, and finding a solution that creates a correct connection between edges both this aspect require further research. Then, a line connectivity check (Touya 2010) must be used at image border.

Experiments on map generation

During our exploration, we experimented with the generation of generalized maps in two manners: first, a GAN is trained for holistic generation of generalized map from detailed data with an image representation (Chapter F), then a workflow combines several models to perform a part of the generalization (Chapter H). Both approaches give map-looking results and succeeded in generalizing and combining several cartographic themes into a credible map. The second approach allows us to make the training set architecture and/or objective function adapted for the specific task of the model, and, it resolves the overlap between roads and buildings. However, models still have important limitations and require further experiments before considering using them for map production: they fail to generate road selection without disconnection, and block are not properly grayed. We also observe that the predicted buildings have a blurred outline and not regular/credible shapes.

Understanding spatial context The main limitation of our proposal is due to the tiles that lack the context for map generalization. We proposed a fusionGAN model that can provide additional information on attributes of the tile features and on the context of the tile and inform the encoding path of the spatial context. The results of this model are not yet satisfying: they allow some generalization operations that were not possible without context (e.g. road selection); however, the proposed model degrades the global quality of the predicted map tiles (predictions include more disconnections and more blurred boundaries). We believe that this model could be improved. First, in our implementation, we did not experiment with the depth, initialization method, and optimization method of the model before using it as a generator, and these parameters have to be explored. Then, in its current form, the model does not include skip connections (while it replaces a U-Net in the GAN), since (Feng, Thiemann, and Sester 2019) showed that they are essential to encode multi-scale relations in map images. We think a fusion model with such a connection must be tested. A second hypothesis is that the connection between main and additional encoding path are not adapted (e.g. too frequent, thus fusions have to be done only at a certain level of encoding or require a dropout to make the prediction less dependent on the additional information). Another perspective is to use another model. For example, recent approach to integrate a vision transformer mechanism to image approach, it promise to focus on some parts of a large image and resolve some context-related issues. In principle, they have better capability than convolution network in capturing long-range dependence between pixels/objects on an input image, but requires far more data to train. Indeed, such methods separate the image into a patch and examine the sequence of the patches (rather than the image). This allows the model to find the parts of the image that are most important in learning and encoding a distant or heterogeneous relation in the image (between patches).

Training set improvement Our training set also includes some limitations: the number of examples and the diversity of situations could be improved, in particular, city center situations are too rare (we think it is the main explanation for block graying problems with the lack of context). Despite the manual correction of the target generalization, some badly generalized features remain in our target images, which can encourage the generation of unclear shapes in the prediction.

Workflow improvement DeepMapScaler is a sequence of models trained separately, and we think that a deeper collaboration between the models needs to be studied. In fact, in the current version of DeepMapScaler, each model encodes and decodes the information in a result that can be interpreted by a human, but we could use as input of a step the embedding of the previous step instead of the decoded output that is only suitable for human interpretation (see Figure G.12). Then, each step of the workflow can be independently improved to better fit to the specific constraints it aims to resolve.

Diversification of applications

The diversity of map objects, generalization operators, input scales, target scales, styles, formulations (see Chapter B) induces a great diversity of possible use cases. In this thesis, we explore four use cases among the many possible. We chose our

use cases with three criteria: 1) the scientific interest of the use case (e.g. the mountain roads generalization and the alignment detection are challenging tasks; the generation of a complete generalized map at once is a moonshot objective of map generalization, etc.); 2) the existence of data at our disposal; and 3) the existence of one or several state-of-the-art solution to compare with; but we could have chosen many others.

First, many other experiment can be used to demonstrate for the potential of deep learning for the choosen tasks of our exploration (relation encoding, shape generalization, and map generation). For instance, we only experiment the graph approach for two kinds of spatial relations, while they are numerous in maps, especially this thesis does not address the study of the relations between different types of map object with heterogeneous graph representation (Iddianozie and McArdle 2021). We further encourage the experiment with a graph that encodes several kinds of spatial relations useful form map generalization (Touya, Balley, Duchêne, Jaara, Regnaud, and Gould 2012) may be more efficient. Moreover, in Chapter F we restrict our work to simple maps where only roads and rivers are represented, while the generalization of various other themes (e.g. vegetation, elevation, etc.) may be interesting. Moreover, we only experimented use cases with a small scale gap between input and target. We think that the more the input and target are different the more complex would be the learning. To verify this statement we would like to replicate our experiment with other target scales. Nevertheless, we considere the generalization of widely generalized map more promizing using several successive models, each perfoming a small generalization. Finally, we claim that the the proposed layered representation is essential for map generation, but we only tested it for a few examples of possible information.

Secondly, Chapter B includes many ideas of deep learning usage for map generalization that could resolve several other research questions. In particular, after our exploration, we estimate that the following applications are the most important to explore:

- Graph-based shape generalization, consists in encoding the shape of an object in a graph and predicting position in the generalized object of each edges. Such a method, allows us to better control the shape of objects and avoid blurred shapes. Moreover it is more convenient for the integration in a generalization process and for the generalization evaluation.
- Prediction of label placement in map and schematics map is a common and important problem in cartography and map generalization for years. Both the graph (label are node of a graph that represent map relations, the model perform a regression of a position for each label) and image approach (segmentation of place where it is possible to set a label) may be investigate for this task.
- Making spatial relations explicit from a map using link prediction or scene graph generation. Indeed, recent models from computer vision has shown potential to generate a graph that describs the relation between objects in an image (e.g. the image depict a man on a bike, the bike in behind the car). Our goal would be to apply such interpretation to a map image and produce a description of spatial relations useful for map generalization (e.g. the road is connected to the highway; it leads to a forest, etc.).

Promoting the usage of deep learning for cartography

To encourage the use deep learning for map generalization, we believe that the promizing results of this thesis must be completed with the exploration of the following research question.

Explainability The default of explainability of deep learning is the first fear for considering the use of deep learning. This default is a common limitation to many applications of deep learning where we want to be sure and understand predictions to trust them (e.g. medicine). Thus, it is a complete field of research (Barredo Arrieta et al. 2019), and we believe that improvement in the explainability of deep neural networks is needed. For example, for a convolution neural network, it is already possible to visualize some explanations, such as the overlaid heat map or gradient-based localization (Selvaraju, Cogswell, Das, Vedantam, Parikh, and Batra 2019) that indicates which image regions are most important for a prediction. This explanation can be beneficial to some use cases of cartography, in particular, the detection of map objects and classification of map types. However, for the use case where we want to predict a map object (e.g. a generalized map object in the Chapter E) we do not want to see which input pixels impact each class but which input pixels impact each output pixel.

Computation cost Then the important computational requirement of deep learning may be a limitation to the use of deep learning for cartography. In particular, we think that it is essential to ensure that the environmental and economical cost of such a model is suitable for map production. This cost is mainly impacted by how much the models have to be trained (epochs and number of examples) before providing satisfying results for map production. Then model can be applied at a small cost to a large area. In comparison, the agent-based processes require few resources at when they are configured, but are expensive to apply to large areas.

Evaluation Finally, the use of deep learning for cartography and map generalization cannot spread without a strong evaluation and comparison with traditional processes. In this thesis, we proposed evaluation methods and shared our datasets to encourage the comparison. We also co-organized a workshop on benchmarks in map generalization. We believe that research must continue in this direction and propose new benchmarks and evaluation methods. In particular, our proposed raster constraints must be extended to other cartographic themes, corrected to limit bias, and improved to better fit the human perception of generalization quality.

The future of cartography with map generalization based on deep learning

This thesis is part of the global objective of the artificial geo-intelligence: "Design[ing] a software agent that takes a user's GIS-related domain question, understands how to gather the required data, how to analyze them, and how to present the results in a suitable form" (Janowicz, Gao, McKenzie, Hu, and Bhaduri 2020, p. 631). Indeed, cartography is still the most suitable solution for presenting to the human user geographic information or phenomena, and we think that the use of deep learning could unlock the search for a model capable of generating such

maps. We already observe that deep learning offers an answer to some of the current limitations in map generalization automation listed in Chapter A. For example, the map generation approach deals with map stylization and map generalization all together, and it can deal with heterogeneous data. We demonstrate in Chapter G that a model trained with a data source can be transferred and applied to an image constructed with other data sources (with a fairly similar level of detail) with very few changes. An important perspective of this thesis is to test how other current challenges can be dealt with using a deep learning approach.

Parametrization of map generalization is challenging. Pre-trained deep learning models can be used without any parametrization very quickly. The difficulty is not in parametrization, but more on identifying the case where the model can apply. Moreover, in traditional map generalization, the parametrization can allow the production of different variants of generalization. Thus, the research question can be reformulated as: how to make possible this variability of solutions in deep learning-based map generalization. The simplest solution is in the training set characteristic; indeed it is not the model choice or hyper-parameter that will determine how the output will be generalized but rather the training examples. However, this solution is not efficient: we do not want to re-create an example set and re-train (or fine-tune) the model to obtain a small variation that can be obtained by just changing one parameter in traditional map generalization (e.g; changing the minimum display size of an element, or encouraging typification rather than aggregation). We think research must focus on a model that can produce several different solutions for one input like multi-modal GANs (Huang, Liu, Belongie, and Kautz 2018).

User requirements for map generalization are difficult to handle. Our opinion is that user requirements are simpler to formulate using examples (the user wants a generalization similar to this or that map)(Hubert 2002), a perspective is to explore how deep learning could use such input to predict a generalization that fits the user requirement expressed in the example. Our first idea is to explore the usage of reference-guided image synthesis (Choi, Uh, Yoo, and Ha 2019).

On-the-fly map generalization. We think that it would be interesting to study the feasibility of an application that interactively produces a generalized map, on the same idea as (Guérin, Digne, Galin, Peytavie, Wolf, Benes, and Martinez 2017) that proposes an interactive platform for quickly, easily and intuitively authoring realistic terrain models by using sketches of crest lines, rivers, or iso-contours (i.e. map element). The task for this application is an interpolation task, while on the contrary the task for map generalization is an abstraction task; thus we have to explore how the input data must be prompt. Then, the input data would be represented in the appropriate form and serve as input of a model or workflow that generates map tiles, finally the tiles are assembled and displayed together.

Multi-scale map generation. This generation is more than the generation of one representation for each desired representation level, it requires multiple consistent intermediate representations and smooth transitions. We already demonstrate that deep learning is able to generate one intermediate generalization (see Chapter F),



Figure 2: Multi-scale map generation as a smooth interpolation problem: comparison between example of smooth interpolations generated by (Liu, Sangineto, Chen, Bao, Zhang, Sebe, Lepri, Wang, and Nadai 2021) methods, and an extract of the geoportail.

and we would like to explore how the model can be trained to generate a consistent series of generalizations that together can form a multi-scale map. In the Chapter B we already mention that a formulation as a video problem can be used for this problem; in addition, we propose the perspective of the model that predicts smooth interpolations between domains (see Figure 2).

Bibliography

- Ai T., Shuai Y., and Li J.** (2008). “The Shape Cognition and Query Supported by Fourier Transform”. *Headway in Spatial Data Handling*. Ed. by A. Ruas and C. Gold. Lecture Notes in Geoinformation and Cartography. Berlin, Heidelberg: Springer, pp. 39–54. ISBN: 978-3-540-68566-1. DOI: [10.1007/978-3-540-68566-1_3](https://doi.org/10.1007/978-3-540-68566-1_3).
- Ai T. and Zhang X.** (2007). “The Aggregation of Urban Building Clusters Based on the Skeleton Partitioning of Gap Space”. In: *The European Information Society*. Ed. by W. Cartwright, G. Gartner, L. Meng, M. P. Peterson, S. I. Fabrikant, and M. Wachowicz. Lecture Notes in Geoinformation and Cartography. Berlin, Heidelberg: Springer, pp. 153–170. ISBN: 978-3-540-72384-4 978-3-540-72385-1. DOI: [10.1007/978-3-540-72385-1_9](https://doi.org/10.1007/978-3-540-72385-1_9).
- Ai T., Zhang X., Zhou Q., and Yang M.** (2015). “A vector field model to handle the displacement of multiple conflicts in building generalization”. *International Journal of Geographical Information Science* 29.8, pp. 1310–1331. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/13658816.2015.1019886](https://doi.org/10.1080/13658816.2015.1019886).
- Ai T., Zhou Q., Zhang X., Huang Y., and Zhou M.** (2014). “A Simplification of Ria Coastline with Geomorphologic Characteristics Preserved”. *Marine Geodesy* 37.2, pp. 167–186. ISSN: 0149-0419, 1521-060X. DOI: [10.1080/01490419.2014.903215](https://doi.org/10.1080/01490419.2014.903215).
- Allouche M. K. and Moulin B.** (2005). “Amalgamation in cartographic generalization using Kohonen’s feature nets”. *International Journal of Geographical Information Science* 19.8-9, pp. 899–914. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/13658810500161211](https://doi.org/10.1080/13658810500161211).
- Arjovsky M., Chintala S., and Bottou L.** (2017). “Wasserstein Generative Adversarial Networks”. *Proceedings of the 34th International Conference on Machine Learning*. Ed. by D. Precup and Y. W. Teh. Vol. 70. Proceedings of Machine Learning Research. PMLR, pp. 214–223.
- Aslan S., Bildirici I., Simav O., and Cetinkaya B.** (2012). “An incremental displacement approach applied to building objects in topographic mapping”. *13th ICA workshop on generalisation and multiple representation*. Istanbul.
- Bader M. and Barrault M.** (2001). “Cartographic Displacement in Generalization: Introducing Elastic Beams”. *4th workshop on progress in automated map generalisation*. Beijing, China: ICA.
- Bader M., Barrault M., and Weibel R.** (2005). “Building displacement over a ductile truss”. *International Journal of Geographical Information Science* 19.8-9. Publisher: Taylor & Francis, pp. 915–936. ISSN: 1365-8816. DOI: [10.1080/13658810500161237](https://doi.org/10.1080/13658810500161237).
- Baeijs C., Demazeau Y., and Alvares L.** (1996). “SIGMA: Application of Multi-Agent Systems to Cartographic Generalization”. *Agents Breaking Away*.

- Ed. by W. Van de Velde and J. W. Perram. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 163–176. ISBN: 978-3-540-49621-2. DOI: [10.1007/BFb0031854](https://doi.org/10.1007/BFb0031854).
- Balley S. and Regnaud N.** (2011). “Models and Standards for On-Demand Mapping”. *25th International Cartographic Conference*. Paris, France.
- Bard S.** (2004). “Quality Assessment of Cartographic Generalisation”. *Transactions in GIS* 8.1, pp. 63–81. ISSN: 1467-9671. DOI: [10.1111/j.1467-9671.2004.00168.x](https://doi.org/10.1111/j.1467-9671.2004.00168.x).
- Barrault M. et al.** (2001). “Integrating multi-agent, object-oriented, and algorithmic techniques for improved automated map generalisation”. *20th International Cartographic Conference*. Vol. 3. Beijing, China: ICA, pp. 2110–2116.
- Barredo Arrieta A. et al.** (Dec. 1, 2019). “Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI”. *Information Fusion* 58. DOI: [10.1016/j.inffus.2019.12.012](https://doi.org/10.1016/j.inffus.2019.12.012).
- Beard M. K.** (1991). “Constraints on rule formation.” In: *Map Generalization: Making Rules for Knowledge Representation*. Ed. by B. P. Buttenfield and R. B. McMaster. Longman, London, pp. 121–135.
- Benaim S. and Wolf L.** (2017). “One-Sided Unsupervised Domain Mapping”. *Advances in Neural Information Processing Systems* 30.
- Bengio Y., Courville A., and Vincent P.** (2013). “Representation learning: a review and new perspectives”. *IEEE transactions on pattern analysis and machine intelligence* 35.8, pp. 1798–1828. ISSN: 1939-3539. DOI: [10.1109/tpami.2013.50](https://doi.org/10.1109/tpami.2013.50).
- Benz S. A. and Weibel R.** (2014). “Road network selection for medium scales using an extended stroke-mesh combination algorithm”. *Cartography and Geographic Information Science* 41.4. Publisher: Taylor & Francis, pp. 323–339. ISSN: 1523-0406. DOI: [10.1080/15230406.2014.928482](https://doi.org/10.1080/15230406.2014.928482).
- Brassel K. E. and Weibel R.** (1988). “A review and conceptual framework of automated map generalization”. *International Journal of Geographical Information Systems*. Publisher: Taylor & Francis. DOI: [10.1080/02693798808927898](https://doi.org/10.1080/02693798808927898).
- Bravo M. J. and Farid H.** (2008). “A scale invariant measure of clutter”. *Journal of Vision* 8.1, pp. 23.1–9. ISSN: 1534-7362. DOI: [10.1167/8.1.23](https://doi.org/10.1167/8.1.23).
- Burghardt D. and Cecconi A.** (2007). “Mesh simplification for building typification”. *International Journal of Geographical Information Science* 21.3, pp. 283–298. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/13658810600912323](https://doi.org/10.1080/13658810600912323).
- Burghardt D., Duchêne C., and Mackaness W.** (2014). *Abstracting Geographic information in a Data Rich World: methodologies and applications of map generalisation*. Publications of the International Cartographic Association (ICA). Springer International Publishing. ISBN: 978-3-319-00202-6.
- Burghardt D. and Neun M.** (2006). “Automated Sequencing of Generalisation Services Based on Collaborative Filtering”. In: *Geographic Information Science - 4th International Conference GIScience*. Ed. by M. Raubal, H. J. Miller, A. U. Frank, and M. F. Goodchild. IFGI prints. Münster, Germany: IFGI prints, pp. 41–46. ISBN: 3-936616-25-6.
- Burghardt D., Schmid S., and Stoter J.** (2007). “Investigation on cartographic constraint formalisation”. *Workshop of the ICA Commission on Generalisation and Multiple Representation*, p. 16.
- Cecconi A.** (2003). “Integration of Cartographic Generalization and Multi-Scale Databases for Enhanced Web Mapping”. Doctoral Thesis. ETH Zurich, p. 155. DOI: [10.3929/ethz-a-004553772](https://doi.org/10.3929/ethz-a-004553772).

- Chen W., Wu A., and Biljecki F.** (2021). “Classification of urban morphology with deep learning: Application on urban vitality”. *Computers, Environment and Urban Systems* 90, p. 13. DOI: [10.1016/j.compenvurbsys.2021.101706](https://doi.org/10.1016/j.compenvurbsys.2021.101706).
- Chen X., Chen S., Xu T., Yin B., Peng J., Mei X., and Li H.** (2020). “SMAP-GAN: Generative Adversarial Network-Based Semisupervised Styled Map Tile Generation Method”. *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–19. ISSN: 1558-0644. DOI: [10.1109/TGRS.2020.3021819](https://doi.org/10.1109/TGRS.2020.3021819).
- Chen X., Yin B., Chen S., Li H., and Xu T.** (2022). “Generating Multiscale Maps From Satellite Images via Series Generative Adversarial Networks”. *IEEE Geoscience and Remote Sensing Letters* 19, pp. 1–5. ISSN: 1558-0571. DOI: [10.1109/LGRS.2021.3129285](https://doi.org/10.1109/LGRS.2021.3129285).
- Chen Y., Carlinet E., Chazalon J., Mallet C., Dumenieu B., and Perret J.** (2021). “Vectorization of historical maps using deep edge filtering and closed shape extraction”. *16th International Conference on Document Analysis and Recognition (ICDAR'21)*. Ed. by Springer. Vol. 12824. Lecture Notes in Computer Science. Lausanne, Switzerland, pp. 510–525. DOI: [10.1007/978-3-030-86337-1_34](https://doi.org/10.1007/978-3-030-86337-1_34).
- Cheng B., Liu Q., Li X., and Wang Y.** (2013). “Building simplification using backpropagation neural networks: a combination of cartographers’ expertise and raster-based local perception”. *GIScience & Remote Sensing* 50.5, pp. 527–542. ISSN: 1548-1603. DOI: [10.1080/15481603.2013.823748](https://doi.org/10.1080/15481603.2013.823748).
- Cho K., Merrienboer van B., Gulcehre C., Bahdanau D., Bougares F., Schwenk H., and Bengio Y.** (2014). “Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation”. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, pp. 1724–1734. DOI: [10.3115/v1/D14-1179](https://doi.org/10.3115/v1/D14-1179).
- Choi Y., Uh Y., Yoo J., and Ha J.-W.** (Dec. 4, 2019). “StarGAN v2: Diverse Image Synthesis for Multiple Domains”. *arXiv:1912.01865 [cs]*. arXiv: [1912.01865](https://arxiv.org/abs/1912.01865).
- Chollet F.** (2018). *Deep learning with Python*. ISBN: 978-1-61729-443-3 1-61729-443-8.
- Christophe S. and Ruas A.** (2002). “Detecting Building Alignments for Generalisation Purposes”. *Advances in Spatial Data Handling*. Ed. by D. E. Richardson and P. van Oosterom. Berlin, Heidelberg: Springer, pp. 419–432. ISBN: 978-3-642-56094-1. DOI: [10.1007/978-3-642-56094-1_31](https://doi.org/10.1007/978-3-642-56094-1_31).
- Chu C., Zhmoginov A., and Sandler M.** (2017). “CycleGAN, a Master of Steganography”. *arXiv:1712.02950 [cs, stat]*.
- Clementini E., Di Felice P., and Oosterom van P.** (1993). “A small set of formal topological relationships suitable for end-user interaction”. *Advances in Spatial Databases*. Ed. by D. Abel and B. Chin Ooi. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 277–295. ISBN: 978-3-540-47765-5. DOI: [10.1007/3-540-56869-7_16](https://doi.org/10.1007/3-540-56869-7_16).
- Courtial A., Touya G., and Zhang X.** (2021a). “Generative Adversarial Networks to Generalise Urban Areas in Topographic Maps”. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLIII-B4-2021. ISSN: 1682-1750. Copernicus GmbH, pp. 15–22. DOI: [10.5194/isprs-archives-XLIII-B4-2021-15-2021](https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-15-2021).

- Courtial A., El Ayedi A., Touya G., and Zhang X.** (2020). “Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation”. *ISPRS International Journal of Geo-Information* 9.5. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute, p. 338. DOI: [10.3390/ijgi9050338](https://doi.org/10.3390/ijgi9050338).
- Courtial A., Touya G., and Zhang X.** (2020). “Constraint Based Evaluation of Generalized Images Generated by Deep Learning”. *23rd ICA Workshop on Map Generalisation and Multiple Representation*. Delft, Netherlands: ICA.
- Courtial A., Touya G., and Zhang X.** (2021b). “AlpineBends - A Benchmark for Deep Learning-Based Generalisation”. Vol. 4. Copernicus Publications. DOI: [10.5194/ica-abs-4-1-2022](https://doi.org/10.5194/ica-abs-4-1-2022).
- Courtial A., Touya G., and Zhang X.** (2021c). “Can Graph Convolution Networks Learn Spatial Relations?” *Abstracts of the ICA* 3, pp. 1–2. DOI: [10.5194/ica-abs-3-60-2021](https://doi.org/10.5194/ica-abs-3-60-2021).
- Courtial A., Touya G., and Zhang X.** (2022a). “Constraint-Based Evaluation of Map Images Generalized by Deep Learning”. *Journal of Geovisualization and Spatial Analysis* 6. DOI: [10.1007/s41651-022-00104-2](https://doi.org/10.1007/s41651-022-00104-2).
- Courtial A., Touya G., and Zhang X.** (2022b). “Deriving map images of generalised mountain roads with generative adversarial networks”. *International Journal of Geographical Information Science* 0.0. Publisher: Taylor & Francis, pp. 1–30. ISSN: 1365-8816. DOI: [10.1080/13658816.2022.2123488](https://doi.org/10.1080/13658816.2022.2123488).
- Courtial A., Touya G., and Zhang X.** (2022c). “Representing Vector Geographic Information As a Tensor for Deep Learning Based Map Generalisation”. *AGILE 2022*. Ed. by E. Parseliunas, A. Mansourian, P. Partsinevelos, and J. Suziedelyte-Visockiene. Vol. 3. Vilnius, Lithuania: Copernicus Publications, p. 32. DOI: [10.5194/agile-giss-3-32-2022](https://doi.org/10.5194/agile-giss-3-32-2022).
- Daley N., Goodenough D., Bhogal A. S., Bradley Q., and Yin Z.** (1997). “Comparing raster and object generalization”. *IGARSS'97. IEEE International Geoscience and Remote Sensing Symposium Proceedings. Remote Sensing - A Scientific Vision for Sustainable Development 2*, 677–679 vol.2. DOI: [10.1109/IGARSS.1997.615221](https://doi.org/10.1109/IGARSS.1997.615221).
- Deng X., Tian Y., and Newsam S.** (2021). “Generalizing Deep Models for Overhead Image Segmentation Through Getis-Ord G_i^* Pooling”, p. 14. DOI: [10.48550/arXiv.1912.10667](https://doi.org/10.48550/arXiv.1912.10667).
- Douglas D. H. and Peucker T. K.** (1973). “Algorithms for the reduction of the number of points required to represent a digitized line or its caricature”. *Cartographica: The International Journal for Geographic Information and Geovisualization* 10.2, pp. 112–122. ISSN: 0317-7173, 1911-9925. DOI: [10.3138/FM57-6770-U75U-7727](https://doi.org/10.3138/FM57-6770-U75U-7727).
- Du J., Wu F., Xing R., Gong X., and Yu L.** (2021). “Segmentation and sampling method for complex polyline generalization based on a generative adversarial network”. *Geocarto International*. Publisher: Taylor & Francis, pp. 1–23. ISSN: 1010-6049. DOI: [10.1080/10106049.2021.1878288](https://doi.org/10.1080/10106049.2021.1878288).
- Duchêne C.** (2003). “Automated map generalisation using communicating agents.” *Proceedings of the 21st International Cartographic Conference*.
- Duchêne C.** (2014a). *Individual Road Generalisation in the 1997-2000 AGENT European project*. Tech. rep. Saint-Mandé, France: IGN, COGIT lab.
- Duchêne C.** (2014b). “Making a map from “thematically multi-sourced data”: the potential of making inter-layers spatial relations explicit”. Vienna, p. 8.

- Duchêne C., Baella B., et al.** (2014). “Generalisation in Practice Within National Mapping Agencies”. In: *Abstracting Geographic Information in a Data Rich World*. Ed. by D. Burghardt, C. Duchêne, and W. Mackaness. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 329–391. DOI: [10.1007/978-3-319-00203-3_11](https://doi.org/10.1007/978-3-319-00203-3_11).
- Duchêne C., Bard S., Barillot X., Ruas A., Trévisan J., and Holzapfel F.** (Jan. 2003). “Quantitative and qualitative description of building orientation”.
- Duchêne C., Ruas A., and Cambier C.** (2012). “The CartACom model: transforming cartographic features into communicating agents for cartographic generalisation”. *International Journal of Geographical Information Science* 26.9. Publisher: Taylor & Francis, pp. 1533–1562. ISSN: 1365-8816. DOI: [10.1080/13658816.2011.639302](https://doi.org/10.1080/13658816.2011.639302).
- Duchêne C., Touya G., Taillandier P., Gaffuri J., Ruas A., and Renard J.** (2018). *Multi-Agents Systems for Cartographic Generalization: Feedback from Past and On-going Research*. Tech. rep. DOI: [10.13140/RG.2.2.35489.92006](https://doi.org/10.13140/RG.2.2.35489.92006).
- Dumont M., Touya G., and Duchêne C.** (2016). “Assessing the Variation of Visual Complexity in Multi-Scale Maps with Clutter Measures”. *19th ICA Workshop on Generalisation and Multiple Representation*. Helsinki, Finland. DOI: [10.13140/RG.2.1.4775.3849](https://doi.org/10.13140/RG.2.1.4775.3849).
- Edwardes A. and Mackaness W. A.** (2000). “Intelligent Generalization of Urban Road Networks”. *GIS Research UK Conference (GISRUK'2000)*, pp. 81–85.
- Egenhofer M. and Herring J.** (1990). “A mathematical framework for the definition of topological relationships”. *Proceedings of the 4th International Colloque on Spatial Data Handling*, pp. 803–813.
- Fairbairn D., Gartner G., and Peterson M. P.** (2021). “Epistemological thoughts on the success of maps and the role of cartography”. *International Journal of Cartography* 7.3. Publisher: Taylor & Francis, pp. 317–331. ISSN: 2372-9333. DOI: [10.1080/23729333.2021.1972909](https://doi.org/10.1080/23729333.2021.1972909).
- Feng Y., Thiemann F., and Sester M.** (2019). “Learning cartographic building generalization with deep convolutional neural networks.” *ISPRS International Journal of Geo-Information* 8.6. ISSN: 2220-9964. DOI: [10.3390/ijgi8060258](https://doi.org/10.3390/ijgi8060258).
- Fu H., Gong M., Wang C., Batmanghelich K., Zhang K., and Tao D.** (2019). “Geometry-Consistent Generative Adversarial Networks for One-Sided Unsupervised Domain Mapping”. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, CA, USA: IEEE, pp. 2422–2431. ISBN: 978-1-72813-293-8. DOI: [10.1109/CVPR.2019.00253](https://doi.org/10.1109/CVPR.2019.00253).
- Fukushima K.** (1980). “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”. *Biological Cybernetics* 36.4, pp. 193–202. ISSN: 1432-0770. DOI: [10.1007/BF00344251](https://doi.org/10.1007/BF00344251).
- Gaffuri J.** (2007). “Field deformation in an agent-based generalisation model: the GAEL model.” Vol. 30.
- Gaffuri J.** (2008). “Généralisation automatique pour la prise en compte de thèmes champ : le modèle GAEL”. These de doctorat. Paris Est.
- Galanda M. and Weibel R.** (2002). “An agent-based framework for polygonal subdivision generalization.” *Advances in spatial data handling, 10th international symposium on spatial data handling*. DOI: [10.1007/978-3-642-56094-1_10](https://doi.org/10.1007/978-3-642-56094-1_10).
- Gan Z., Chen L., Wang W., Pu Y., Zhang Y., Liu H., Li C., and Carin L.** (2017). “Triangle Generative Adversarial Networks”. In: *Advances in Neural Information Processing Systems 30*. Ed. by I. Guyon, U. V. Luxburg, S. Bengio,

- H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett. Curran Associates, Inc., pp. 5247–5256.
- Ganguli S., Garzon P., and Glaser N.** (2019). “GeoGAN: A Conditional GAN with Reconstruction and Style Loss to Generate Standard Layer of Maps from Satellite Images”. *arXiv:1902.05611 [cs]*.
- Garcia-Balboa J. L. and Ariza-López F. J.** (2008). “Generalization-oriented road line classification by means of an artificial neural network”. *Geoinformatica*. DOI: [10.1007/s10707-007-0026-z](https://doi.org/10.1007/s10707-007-0026-z).
- Gatys L. A., Ecker A. S., and Bethge M.** (2016). “Image Style Transfer Using Convolutional Neural Networks”. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, pp. 2414–2423. ISBN: 978-1-4673-8851-1. DOI: [10.1109/CVPR.2016.265](https://doi.org/10.1109/CVPR.2016.265).
- Girres J.-F. and Touya G.** (2014). “Cartographic Generalisation Aware of Multiple Representations”. *GIScience 2014*. Ed. by K. Stewart, E. J. Pebesma, G. Navratil, P. Fogliaroni, and M. Duckham. Vol. 40. Extended abstract proceedings of the GIScience 2014. Vienna, Austria: TU Vienna, pp. 289–291.
- Goodchild M. F. and Li L.** (2012). “Assuring the quality of volunteered geographic information”. *Spatial Statistics* 1, pp. 110–120. ISSN: 2211-6753. DOI: [10.1016/j.spasta.2012.03.002](https://doi.org/10.1016/j.spasta.2012.03.002).
- Goodfellow I., Bengio Y., and Courville A.** (2016). *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press.
- Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., and Bengio Y.** (2014). “Generative Adversarial Networks”. *arXiv:1406.2661 [cs, stat]*.
- Guérin É., Digne J., Galin É., Peytavie A., Wolf C., Benes B., and Martinez B.** (2017). “Interactive example-based terrain authoring with conditional generative adversarial networks”. *ACM Transactions on Graphics* 36.6, pp. 1–13. ISSN: 07300301. DOI: [10.1145/3130800.3130804](https://doi.org/10.1145/3130800.3130804).
- Guilbert E.** (2013). “Multi-level representation of terrain features on a contour map”. *GeoInformatica* 17.2, pp. 301–324. ISSN: 1384-6175, 1573-7624. DOI: [10.1007/s10707-012-0153-z](https://doi.org/10.1007/s10707-012-0153-z).
- Gulgen F. and Gokgoz T.** (2008). “Selection of roads for cartographic generalization”, p. 7.
- Hangouet J.-F.** (1998). “Approche et méthodes pour l’automatisation de la généralisation cartographique ; application en bord de ville”. PhD thesis. Université de Marne la Vallée.
- Harrie L.** (2000). “Generalisation Of Vector Data Sets By Simultaneous Least Squares Adjustment”. *International Archives of Photogrammetry and Remote Sensing*. Vol. Vol. XXXIII. Part B4. Amsterdam, p. 8.
- Harrie L.** (2003). “Weight-Setting and Quality Assessment in Simultaneous Graphic Generalization”. *The Cartographic Journal* 40.3, pp. 221–233. ISSN: 0008-7041, 1743-2774. DOI: [10.1179/000870403225012925](https://doi.org/10.1179/000870403225012925).
- Harrie L., Stigmar H., and Djordjevic M.** (2015). “Analytical Estimation of Map Readability”. *ISPRS International Journal of Geo-Information* 4.2, pp. 418–446. DOI: [10.3390/ijgi4020418](https://doi.org/10.3390/ijgi4020418).
- Hauert J.-H. and Sester M.** (2008). “Area Collapse and Road Centerlines based on Straight Skeletons”. *GeoInformatica* 12.2, pp. 169–191. ISSN: 1384-6175, 1573-7624. DOI: [10.1007/s10707-007-0028-x](https://doi.org/10.1007/s10707-007-0028-x).

- Hauert J.-H. and Wolff A.** (2010). “Optimal and topologically safe simplification of building footprints”. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '10*. San Jose, California: ACM Press, p. 192. ISBN: 978-1-4503-0428-3. DOI: [10.1145/1869790.1869819](https://doi.org/10.1145/1869790.1869819).
- Hazırbağ C., Ma L., Domokos C., and Cremers D.** (2016). “FuseNet: Incorporating Depth into Semantic Segmentation via Fusion-Based CNN Architecture”. ISBN: 978-3-319-54180-8. DOI: [10.1007/978-3-319-54181-5_14](https://doi.org/10.1007/978-3-319-54181-5_14).
- He K., Zhang X., Ren S., and Sun J.** (2016). “Identity Mappings in Deep Residual Networks”. *Computer Vision – ECCV 2016*. Ed. by B. Leibe, J. Matas, N. Sebe, and M. Welling. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 630–645. ISBN: 978-3-319-46493-0. DOI: [10.1007/978-3-319-46493-0_38](https://doi.org/10.1007/978-3-319-46493-0_38).
- He X., Hua Y., Song T., Zhang Z., Xue Z., Ma R., Robertson N., and Guan H.** (2019). “Unsupervised Video Summarization with Attentive Conditional Generative Adversarial Networks”. *Proceedings of the 27th ACM International Conference on Multimedia*. MM '19. New York, NY, USA: Association for Computing Machinery, pp. 2296–2304. ISBN: 978-1-4503-6889-6. DOI: [10.1145/3343031.3351056](https://doi.org/10.1145/3343031.3351056).
- Henry C. J., Storie C., Alagappan M. P., Alhassan V., Swamy M., Aleshinloye D., Curtis A., and Kim D.** (2019). “Automated LULC map production using deep neural networks”. *International Journal of Remote Sensing* 40, pp. 1–25. DOI: [10.1080/01431161.2018.1563840](https://doi.org/10.1080/01431161.2018.1563840).
- Hu Y., Gui Z., Wang J., and Li M.** (2021). “Enriching the metadata of map images: a deep learning approach with GIS-based data augmentation”. *International Journal of Geographical Information Science*, pp. 1–23. DOI: [10.1080/13658816.2021.1968407](https://doi.org/10.1080/13658816.2021.1968407).
- Huang B., Zhao B., and Song Y.** (2018). “Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery”. *Remote Sensing of Environment* 214, pp. 73–86. ISSN: 0034-4257. DOI: [10.1016/j.rse.2018.04.050](https://doi.org/10.1016/j.rse.2018.04.050).
- Huang X., Liu M.-Y., Belongie S., and Kautz J.** (2018). “Multimodal Unsupervised Image-to-Image Translation”. In: *Computer Vision – ECCV 2018*. Ed. by V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss. Vol. 11207. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 179–196. ISBN: 978-3-030-01218-2 978-3-030-01219-9. DOI: [10.1007/978-3-030-01219-9_11](https://doi.org/10.1007/978-3-030-01219-9_11).
- Hubert F.** (2002). “Map Samples to Help GI Users Specify their Needs”. In: *Advances in Spatial Data Handling*. Ed. by D. E. Richardson and P. van Oosterom. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 533–545. ISBN: 978-3-642-62859-7 978-3-642-56094-1. DOI: [10.1007/978-3-642-56094-1_39](https://doi.org/10.1007/978-3-642-56094-1_39).
- Iddianozie C. and McArdle G.** (2021). “Towards Robust Representations of Spatial Networks Using Graph Neural Networks”. *Applied Sciences* 11.15. Publisher: Multidisciplinary Digital Publishing Institute, p. 6918. DOI: [10.3390/app11156918](https://doi.org/10.3390/app11156918).
- Iddianozie C. and McArdle G.** (2021). “Transferable Graph Neural Networks for Inferring Road Type Attributes in Street Networks”. *IEEE Access* 9. Conference Name: IEEE Access, pp. 158331–158339. ISSN: 2169-3536. DOI: [10.1109/ACCESS.2021.3128839](https://doi.org/10.1109/ACCESS.2021.3128839).

- Isola P., Zhu J.-Y., Zhou T., and Efros A. A.** (2017). “Image-to-Image Translation with Conditional Adversarial Networks”. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 1063-6919, pp. 5967–5976. DOI: [10.1109/CVPR.2017.632](https://doi.org/10.1109/CVPR.2017.632).
- Janowicz K., Gao S., McKenzie G., Hu Y., and Bhaduri B.** (2020). “GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond”. *International Journal of Geographical Information Science* 34.4, pp. 625–636. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/13658816.2019.1684500](https://doi.org/10.1080/13658816.2019.1684500).
- Jenny B., Heitzler M., Singh D., Farmakis-Serebryakova M., Liu J. C., and Hurni L.** (2021). “Cartographic relief shading with neural networks”. *IEEE Transactions on Visualization and Computer Graphics* 27.2. Publisher: IEEE, Institute of Electrical and Electronics Engineers, pp. 1225–1235. ISSN: 1077-2626. DOI: [10.1109/TVCG.2020.3030456](https://doi.org/10.1109/TVCG.2020.3030456).
- Jiang B. and Claramunt C.** (2004). “A Structural Approach to the Model Generalization of an Urban Street Network”. *GeoInformatica* 8.2, pp. 157–171. ISSN: 1384-6175. DOI: [10.1023/B:GEIN.0000017746.44824.70](https://doi.org/10.1023/B:GEIN.0000017746.44824.70).
- Kang Y., Gao S., and Roth R. E.** (2019). “Transferring Multiscale Map Styles Using Generative Adversarial Networks”. *International Journal of Cartography* 5.2-3, pp. 115–141. ISSN: 2372-9333, 2372-9341. DOI: [10.1080/23729333.2019.1615729](https://doi.org/10.1080/23729333.2019.1615729).
- Kang Y., Rao J., Wang W., Peng B., Gao S., and Zhang F.** (2020). “Towards Cartographic Knowledge Encoding with Deep Learning”. *Autocarto*. Redlands, California, p. 6.
- Karsznia I. and Weibel R.** (2018). “Improving settlement selection for small-scale maps using data enrichment and machine learning”. *Cartography and Geographic Information Science* 45.2, pp. 111–127. ISSN: 1523-0406. DOI: [10.1080/15230406.2016.1274237](https://doi.org/10.1080/15230406.2016.1274237).
- Kazimi B., Thiemann F., and Sester M.** (2019). “Semantic segmentation of manmade landscape structures in digital terrain models”. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* IV-2/W7, pp. 87–94. ISSN: 2194-9050. DOI: [10.5194/isprs-annals-IV-2-W7-87-2019](https://doi.org/10.5194/isprs-annals-IV-2-W7-87-2019).
- Kilpelainen T.** (2000). “Knowledge Acquisition for Generalization Rules”. *Cartography and Geographic Information Science* 27.1, pp. 41–50. ISSN: 1523-0406. DOI: [10.1559/152304000783547993](https://doi.org/10.1559/152304000783547993).
- Kinghorn P., Zhang L., and Shao L.** (2017). “Deep learning based image description generation”. *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 919–926. DOI: [10.1109/IJCNN.2017.7965950](https://doi.org/10.1109/IJCNN.2017.7965950).
- Kipf T. N. and Welling M.** (2017). “Semi-Supervised Classification with Graph Convolutional Networks”. *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net.
- Knura M.** (2021). “Deep Learning for Map Generalization: Towards a new Approach using Vector Data”. *Abstracts of the ICA*. Vol. 3. Florence, Italy, pp. 1–2. DOI: [10.5194/ica-abs-3-152-2021](https://doi.org/10.5194/ica-abs-3-152-2021).
- Krizhevsky A., Sutskever I., and Hinton G. E.** (2012). “ImageNet Classification with Deep Convolutional Neural Networks”. *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc.

- Lamy S., Ruas A., Demazeau Y., Jackson M., and Mackaness W. A.** (1999). “The application of agents in automated map generalisation.” *19th International Cartographic Conference*.
- Lan T., Li Z., Wang J., Gong C., and Ti P.** (2022). “An ANNs-Based Method for Automated Labelling of Schematic Metro Maps”. *ISPRS International Journal of Geo-Information* 11.1. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 36. DOI: [10.3390/ijgi11010036](https://doi.org/10.3390/ijgi11010036).
- Lapaine M., Midtbø T., Gartner G., Bandrova T., Wang T., and Shen J.** (2021). “Definition of the Map”. *Advances in Cartography and GIScience of the ICA* 3. Publisher: Copernicus GmbH, pp. 1–6. DOI: [10.5194/ica-adv-3-9-2021](https://doi.org/10.5194/ica-adv-3-9-2021).
- Lecordix F., Plazanet C., and Lagrange J.-P.** (1997). “A Platform for Research in Generalization: Application to Caricature”. *GeoInformatica* 1.2, pp. 161–182. DOI: [10.1023/A:1009736628698](https://doi.org/10.1023/A:1009736628698).
- Lecun Y., Bottou L., Bengio Y., and Haffner P.** (1998). “Gradient-based learning applied to document recognition”. *Proceedings of the IEEE* 86.11. Conference Name: Proceedings of the IEEE, pp. 2278–2324. ISSN: 1558-2256. DOI: [10.1109/5.726791](https://doi.org/10.1109/5.726791).
- LeCun Y., Bengio Y., and Hinton G.** (2015). “Deep learning”. *Nature* 521.7553. Number: 7553 Publisher: Nature Publishing Group, pp. 436–444. ISSN: 1476-4687. DOI: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Li B., Wilkinson G. G., and Khaddaj S.** (2001). “Cell-based Model For GIS Generalization”. *Geocomputation*. Vol. 1, p. 6.
- Li J., Chen Z., Zhao X., and Shao L.** (2020). “MapGAN: An Intelligent Generation Model for Network Tile Maps”. *Sensors* 20.11. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute, p. 3119. DOI: [10.3390/s20113119](https://doi.org/10.3390/s20113119).
- Li W., Ai T., Shen Y., Yang W., and Wang W.** (2020). “A Novel Method for Building Displacement Based on Multipopulation Genetic Algorithm”. *Applied Sciences* 10.23. Number: 23 Publisher: Multidisciplinary Digital Publishing Institute, p. 8441. DOI: [10.3390/app10238441](https://doi.org/10.3390/app10238441).
- Li Z.** (1994). “Mathematical Morphology in Digital Generalization of Raster Map Data”. *Cartography* 23.1, pp. 1–10. ISSN: 0069-0805. DOI: [10.1080/00690805.1994.9713969](https://doi.org/10.1080/00690805.1994.9713969).
- Li Z. and Su B.** (1995). “Algebraic Models for Feature Displacement in the Generalization of Digital Map Data Using Morphological Techniques”. *Cartographica: The International Journal for Geographic Information and Geovisualization* 32.3. Publisher: University of Toronto Press, pp. 39–56. ISSN: 0317-7173. DOI: [10.3138/H7K9-2160-2765-8230](https://doi.org/10.3138/H7K9-2160-2765-8230).
- Liqiang Z., Hao D., Dong C., and Zhen W.** (2013). “A spatial cognition-based urban building clustering approach and its applications”. *International Journal of Geographical Information Science* 27.4. Publisher: Taylor & Francis, pp. 721–740. ISSN: 1365-8816. DOI: [10.1080/13658816.2012.700518](https://doi.org/10.1080/13658816.2012.700518).
- Liu Y., Sangineto E., Chen Y., Bao L., Zhang H., Sebe N., Lepri B., Wang W., and Nadai M. D.** (2021). “Smoothing the Disentangled Latent Style Space for Unsupervised Image-to-Image Translation”. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 2575-7075, pp. 10780–10789. DOI: [10.1109/CVPR46437.2021.01064](https://doi.org/10.1109/CVPR46437.2021.01064).
- Liu Y., Guo Q., Sun Y., and Ma X.** (2014). “A Combined Approach to Cartographic Displacement for Buildings Based on Skeleton and Improved Elastic

- Beam Algorithm”. *PLoS ONE* 9.12. Ed. by R. Ji, e113953. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0113953](https://doi.org/10.1371/journal.pone.0113953).
- Long J., Shelhamer E., and Darrell T.** (2015). “Fully Convolutional Networks for Semantic Segmentation”. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Los Alamitos, CA, USA: IEEE Computer Society, pp. 3431–3440. DOI: [10.1109/CVPR.2015.7298965](https://doi.org/10.1109/CVPR.2015.7298965).
- Luan X. and Yang B.** (2010). “Generating strokes of road networks based on pattern recognition”. *Proceedings of the 13th workshop of the ICA commission on generalisation and multiple representation*. Zurich, Switzerland.
- Mackaness W. A. and Purves R. S.** (2001). “Automated Displacement for Large Numbers of Discrete Map Objects”. *Algorithmica* 30.2, pp. 302–311. ISSN: 0178-4617, 1432-0541. DOI: [10.1007/s00453-001-0007-9](https://doi.org/10.1007/s00453-001-0007-9).
- Mackaness W. and Beard K.** (1993). “Use of Graph Theory to Support Map Generalization”. *Cartography and Geographic Information Science - CARTOGR GEOGR INF SCI* 20, pp. 210–221. DOI: [10.1559/152304093782637479](https://doi.org/10.1559/152304093782637479).
- Mackaness W. and Edwards G.** (2002). “The Importance of Modelling Pattern and Structure in Automated Map Generalisation. Join ISPRS/ICA workshop on multi-scale representations of spatial data”. *Proceedings of the Joint ISPRS/ICA Workshop on Multi-Scale Representations of Spatial Data*, pp. 7–8.
- Mackaness W. A.** (2007). “Chapter 1 - Understanding Geographic Space”. In: *Generalisation of Geographic Information*. Ed. by W. A. Mackaness, A. Ruas, and L. T. Sarjakoski. International Cartographic Association. Amsterdam: Elsevier Science B.V., pp. 1–10. ISBN: 978-0-08-045374-3. DOI: [10.1016/B978-008045374-3/50007-7](https://doi.org/10.1016/B978-008045374-3/50007-7).
- Mackaness W. A. and Ruas A.** (2007). “Chapter 5 - Evaluation in the Map Generalisation Process”. In: *Generalisation of Geographic Information*. Ed. by W. A. Mackaness, A. Ruas, and L. T. Sarjakoski. International Cartographic Association. Amsterdam: Elsevier Science B.V., pp. 89–111. ISBN: 978-0-08-045374-3. DOI: [10.1016/B978-008045374-3/50007-7](https://doi.org/10.1016/B978-008045374-3/50007-7).
- Mahasseni B., Lam M., and Todorovic S.** (2017). “Unsupervised Video Summarization with Adversarial LSTM Networks”. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 1063-6919, pp. 2982–2991. DOI: [10.1109/CVPR.2017.318](https://doi.org/10.1109/CVPR.2017.318).
- Mai G., Janowicz K., Hu Y., Gao S., Yan B., Zhu R., Cai L., and Lao N.** (2022). “A Review of Location Encoding for GeoAI: Methods and Applications”. *International Journal of Geographical Information Science*, pp. 1–35. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/13658816.2021.2004602](https://doi.org/10.1080/13658816.2021.2004602).
- Maudet A., Touya G., Duchêne C., and Picault S.** (2014). “Representation of interactions in a Multi-Level Multi-Agent model for cartography constraint solving.” *Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection*. Ed. by Y. Demazeau, F. Zambonelli, J. M. Corchado, and J. Bajo. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 183–194. ISBN: 978-3-319-07551-8. DOI: [10.1007/978-3-319-07551-8_16](https://doi.org/10.1007/978-3-319-07551-8_16).
- McCulloch W. S. and Pitts W.** (1943). “A logical calculus of the ideas immanent in nervous activity”. *The bulletin of mathematical biophysics* 5.4, pp. 115–133. ISSN: 1522-9602. DOI: [10.1007/BF02478259](https://doi.org/10.1007/BF02478259).
- McMaster R. B. and Monmonier M.** (1989). “A conceptual framework for quantitative and qualitative raster-mode generalization”. *Proceedings of GIS/LIS'89*.

- Orlando, Florida: American Society for Photogrammetry and Remote Sensing, pp. 390–403.
- Méneroux Y., Dizier V., Margollé M., Van Damme M. D., Kanasugi H., Le Guilcher A., Saint Pierre G., and Kato Y.** (2018). “Convolutional Neural Network for Traffic Signal Inference based on GPS Traces”. *Spatial Big Data and Machine Learning in GIScience*, p. 9.
- Miller H. J.** (2004). “Tobler’s First Law and Spatial Analysis”. *Annals of the Association of American Geographers* 94.2. Publisher: Routledge, pp. 284–289. ISSN: 0004-5608. DOI: [10.1111/j.1467-8306.2004.09402005.x](https://doi.org/10.1111/j.1467-8306.2004.09402005.x).
- Monmonier M. S.** (1983). “Raster-mode area generalization for land use and land cover maps”. *Cartographica: The International Journal for Geographic Information and Geovisualization* 20.4. Publisher: University of Toronto Press, pp. 65–91. ISSN: 0317-7173. DOI: [10.3138/X572-0327-4670-1573](https://doi.org/10.3138/X572-0327-4670-1573).
- Mustiere S.** (1998). “GALBE : Adaptative generalization. the need for an adaptative process for automated generalisation an exemple on roads.” *Proceding of 1st GIS PlaNet conference*.
- Mustière S. and Moulin B.** (2002). “What is Spatial Context in Cartographic Generalization”. *Symposium on Geospatial Theory, Proceeding and Applications*.
- Mustière S., Zucker J.-D., and Saitta L.** (2000). “Abstraction-Based Machine Learning Approach To Cartographic Generalisation”. *Proceedings of 9th International Symposium on Spatial Data Handling*. Vol. 1a. Beijing, China: IGU, pp. 50–63.
- Neun M., Burghardt D., and Weibel R.** (2009). “Automated processing for map generalization using web services”. *GeoInformatica* 13.4, pp. 425–452. ISSN: 1384-6175, 1573-7624. DOI: [10.1007/s10707-008-0054-3](https://doi.org/10.1007/s10707-008-0054-3).
- Olteanu-Raimond A.-M., Mustière S., and Ruas A.** (2015). “Knowledge Formalization for Vector Data Matching Using Belief Theory”. *Journal of Spatial Information Science* 10, pp. 21–46. DOI: [10.5311/JOSIS.2015.10.194](https://doi.org/10.5311/JOSIS.2015.10.194).
- Oosterom van P., Meijers M., Stoter J., and Šuba R.** (2014). “Data Structures for Continuous Generalisation: tGAP and SSC”. In: *Abstracting Geographic Information in a Data Rich World: Methodologies and Applications of Map Generalisation*. Ed. by D. Burghardt, C. Duchêne, and W. Mackaness. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 83–117. ISBN: 978-3-319-00203-3. DOI: [10.1007/978-3-319-00203-3_4](https://doi.org/10.1007/978-3-319-00203-3_4).
- Pantazis D., Karathanasis B., Kassoli M., Koukofikis A., and Stratakis P.** (2009). “Morphing techniques: Towards new methods for raster based cartographic generalization”. *Proceedings of the 24th International Cartographic Conference*. Santiago, Chile: ICA.
- Peter B. and Weibel R.** (1999). “Using vector and raster-based techniques in categorical map generalization”. *Third ICA Workshop on Progress in Automated Map Generalization*.
- Peuquet D. J.** (1979). “Raster Processing: An Alternative Approach to Automated Cartographic Data Handling”. *The American Cartographer* 6.2. Publisher: Taylor & Francis, pp. 129–139. ISSN: 0094-1689. DOI: [10.1559/152304079784023159](https://doi.org/10.1559/152304079784023159).
- Plazanet C., Bigolin N. M., and Ruas A.** (1998). “Experiments with Learning Techniques for Spatial Model Enrichment and Line Generalization”. *GeoInformatica* 2.4, pp. 315–333. ISSN: 1573-7624. DOI: [10.1023/A:1009753320636](https://doi.org/10.1023/A:1009753320636).

- Porta S., Crucitti P., and Latora V.** (2006a). “The network analysis of urban streets: A dual approach”. *Physica A: Statistical Mechanics and its Applications* 369.2, pp. 853–866. ISSN: 0378-4371. DOI: [10.1016/j.physa.2005.12.063](https://doi.org/10.1016/j.physa.2005.12.063).
- Porta S., Crucitti P., and Latora V.** (Oct. 2006b). “The Network Analysis of Urban Streets: A Primal Approach”. *Environment and Planning B: Planning and Design* 33.5. Publisher: SAGE Publications Ltd STM, pp. 705–725. ISSN: 0265-8135. DOI: [10.1068/b32045](https://doi.org/10.1068/b32045).
- Potié Q., Mackaness W. A., and Touya G.** (2022). “When is a Ring Road a ‘Ring Road’? A Brief Perceptual Study”. *AGILE: GIScience Series* 3. Publisher: Copernicus GmbH, pp. 1–7. DOI: [10.5194/agile-giss-3-54-2022](https://doi.org/10.5194/agile-giss-3-54-2022).
- Raposo P. and Samsonov T.** (2014). “Towards general theory of raster data generalization”. *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation, Vienna, Austria*, p. 10.
- Raposo P., Touya G., and Bereter P.** (2020). “A Change of Theme: The Role of Generalization in Thematic Mapping”. *ISPRS International Journal of Geo-Information* 9.6. Number: 6 Publisher: Multidisciplinary Digital Publishing Institute, p. 371. DOI: [10.3390/ijgi9060371](https://doi.org/10.3390/ijgi9060371).
- Rathore A., Nagar P., Arora C., and Jawahar C.** (2019). “Generating 1 Minute Summaries of Day Long Egocentric Videos”. *Proceedings of the 27th ACM International Conference on Multimedia*. MM ’19. Nice, France: Association for Computing Machinery, pp. 2305–2313. ISBN: 9781450368896. DOI: [10.1145/3343031.3350880](https://doi.org/10.1145/3343031.3350880).
- Regnauld N.** (2001). “Contextual Building Typification in Automated Map Generalization”. *Algorithmica* 30.2, pp. 312–333. ISSN: 0178-4617, 1432-0541. DOI: [10.1007/s00453-001-0008-8](https://doi.org/10.1007/s00453-001-0008-8).
- Regnauld N.** (2005). “Spatial structures to support automatic generalisation”. *Proceedings of 22 Nd International Cartographic Conference, La Coruña*.
- Regnauld N. and Revell P.** (2007). “Automatic Amalgamation of Buildings for Producing Ordnance Survey[®] 1:50 000 Scale Maps”. *The Cartographic Journal* 44.3, pp. 239–250. ISSN: 0008-7041, 1743-2774. DOI: [10.1179/000870407X241782](https://doi.org/10.1179/000870407X241782).
- Ren and Malik** (2003). “Learning a classification model for segmentation”. *Proceedings Ninth IEEE International Conference on Computer Vision*, 10–17 vol.1. DOI: [10.1109/ICCV.2003.1238308](https://doi.org/10.1109/ICCV.2003.1238308).
- Ricker B. and Roth R.** (2018). “Mobile Maps and Responsive Design”. *Geographic Information Science & Technology Body of Knowledge* 2018. DOI: [10.22224/gistbok/2018.2.5](https://doi.org/10.22224/gistbok/2018.2.5).
- Ronneberger O., Fischer P., and Brox T.** (2015). “U-Net: Convolutional Networks for Biomedical Image Segmentation”. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi. Vol. abs/1505.04597. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 234–241. ISBN: 978-3-319-24574-4. DOI: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Rosenholtz R., Li Y., and Nakano L.** (2007). “Measuring visual clutter”. *Journal of vision* 7, pp. 17.1–22. ISSN: 1534-7362. DOI: [10.1167/7.2.17](https://doi.org/10.1167/7.2.17).
- Roth R. E., Brewer C. A., and Stryker M. S.** (2011). “A typology of operators for maintaining legible map designs at multiple scales”. *Cartographic Perspectives* 68, pp. 29–64. ISSN: 1048-9053. DOI: [10.14714/CP68.7](https://doi.org/10.14714/CP68.7).

- Ruas A.** (1998). “A method for building displacement in automated map generalisation”. *International Journal of Geographical Information Science* 12.8, pp. 789–803. ISSN: 1365-8816, 1362-3087. DOI: [10.1080/136588198241509](https://doi.org/10.1080/136588198241509).
- Ruas A.** (1999). “Modèle de généralisation de données géographiques à base de contraintes et d'autonomie”. PhD thesis. Université de Marne la Vallée.
- Ruas A.** (2002). *Les problématiques de l'automatisation de la généralisation cartographique*. Ed. by A. Ruas. *Traité IGAT information géographique et aménagement du territoire Géomatique*. Paris: Hermès science, Chap. 4, pp. 75–90. ISBN: 2-7462-0494-0.
- Ruas A. and Morisset B.** (1997). “Simulation and Agent Modelling for road selection in generalisation”. *18th International Cartographic Conference*. Vol. 3. ICA, pp. 1376–1380.
- Ruas A. and Plazanet C.** (1996). “Strategies for automated generalization”. *Proceedings of 7th International Symposium on Spatial Data Handling*. Vol. 1. Issue: 6, pp. 319–336.
- Satari R., Kazimi B., and Sester M.** (2021). “Extraction of linear structures from digital terrain models using deep learning”. *AGILE: GIScience Series 2*, pp. 1–14. DOI: [10.5194/agile-giss-2-11-2021](https://doi.org/10.5194/agile-giss-2-11-2021).
- Schnürer R., Sieber R., Schmid-Lanter J., Öztireli C., and Hurni L.** (2020). “Detection of Pictorial Map Objects with Convolutional Neural Networks”. *Cartographic Journal The*. DOI: [10.1080/00087041.2020.1738112](https://doi.org/10.1080/00087041.2020.1738112).
- Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., and Batra D.** (2019). *Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization*. arXiv:1610.02391 [cs]. DOI: [10.1007/s11263-019-01228-7](https://doi.org/10.1007/s11263-019-01228-7).
- Sester M.** (2000a). “Generalization Based On Least Squares Adjustment”. *International Archives of Photogrammetry and Remote Sensing*.
- Sester M.** (2000b). “Knowledge acquisition for the automatic interpretation of spatial data”. *International Journal of Geographical Information Science* 14.1, pp. 1–24. DOI: [10.1080/136588100240930](https://doi.org/10.1080/136588100240930).
- Sester M.** (2005). “Optimization approaches for generalization and data abstraction”. *International Journal of Geographical Information Science* 19.8. Ed. by P. Fisher, M. Gahegan, and B. Lees, pp. 871–897. DOI: [10.1080/13658810500161179](https://doi.org/10.1080/13658810500161179).
- Sharma G., Wu W., and Dalal E. N.** (2005). “The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations”. *Color Research & Application* 30.1, pp. 21–30. ISSN: 0361-2317, 1520-6378. DOI: [10.1002/col.20070](https://doi.org/10.1002/col.20070).
- Shen Y., Ai T., and He Y.** (2018). “A New Approach to Line Simplification Based on Image Processing: A Case Study of Water Area Boundaries”. *ISPRS International Journal of Geo-Information* 7.2. Publisher: Multidisciplinary Digital Publishing Institute, p. 41. DOI: [10.3390/ijgi7020041](https://doi.org/10.3390/ijgi7020041).
- Shen Y., Ai T., and Li C.** (2019). “A simplification of urban buildings to preserve geometric properties using superpixel segmentation”. *International Journal of Applied Earth Observation and Geoinformation* 79, pp. 162–174. ISSN: 1569-8432. DOI: [10.1016/j.jag.2019.02.008](https://doi.org/10.1016/j.jag.2019.02.008).
- Shen Y., Ai T., Li W., Yang M., and Feng Y.** (2019). “A polygon aggregation method with global feature preservation using superpixel segmentation”. *Computers, Environment and Urban Systems* 75, pp. 117–131. ISSN: 0198-9715. DOI: [10.1016/j.compenvurbsys.2019.01.009](https://doi.org/10.1016/j.compenvurbsys.2019.01.009).

- Shen Y., Ai T., and Zhao R.** (2021). “Raster-based method for building selection in the multi-scale representation of two-dimensional maps”. *Geocarto International* 0.ja. Publisher: Taylor & Francis, pp. 1–18. ISSN: 1010-6049. DOI: [10.1080/10106049.2021.1943007](https://doi.org/10.1080/10106049.2021.1943007).
- Shen Y., Li J., Wang Z., Zhao R., and Wang L.** (2022). “A raster-based typification method for multiscale visualization of building features considering distribution patterns”. *International Journal of Digital Earth* 15.1. Publisher: Taylor & Francis, pp. 249–275. ISSN: 1753-8947. DOI: [10.1080/17538947.2021.2023668](https://doi.org/10.1080/17538947.2021.2023668).
- Shunbao L., Zhongqiang B., and Yan B.** (2012). “Errors Prediction for Vector-to-Raster Conversion Based on Map Load and Cell Size”. *Chinese Geographical Science*. DOI: [10.1007/s11769-012-0544-y](https://doi.org/10.1007/s11769-012-0544-y).
- Simo-Serra E., Iizuka S., and Ishikawa H.** (2017). “Mastering Sketching: Adversarial Augmentation for Structured Prediction”. *ACM Trans. Graph.* 37.1. ISSN: 0730-0301. DOI: [10.1145/3132703](https://doi.org/10.1145/3132703).
- Simo-Serra E., Iizuka S., Sasaki K., and Ishikawa H.** (2016). “Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup”. *ACM Trans. Graph.* 35.4. DOI: [10.1145/2897824.2925972](https://doi.org/10.1145/2897824.2925972).
- Simonyan K. and Zisserman A.** (2015). “Very Deep Convolutional Networks for Large-Scale Image Recognition”. *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. arXiv. DOI: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556).
- Skopeliti A. and Tsoulos L.** (2001). “A methodology for the assessment of generalization quality”. *Fourth Workshop on Progress in Automated Map Generalization*, pp. 2–4.
- Stanislawski L. V., Briat M., Punt E., Howard M. A., Brewer C. A., and Buttenfield B. P.** (2012). “Density-stratified thinning of road networks to support automated generalisation for the national map.” *ICA Commission on Generalisation and Multiple Representation*. Istanbul.
- Stanislawski L. V., Buttenfield B. P., Bereuter P., Savino S., and Brewer C. A.** (2014). “Generalisation Operators”. In: *Abstracting Geographic Information in a Data Rich World*. Ed. by D. Burghardt, C. Duchêne, and W. Mackaness. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, pp. 157–195. DOI: [10.1007/978-3-319-00203-3_6](https://doi.org/10.1007/978-3-319-00203-3_6).
- Steiniger S., Burghardt D., and Weibel R.** (2006). “Recognition of island structures for map generalization”. *Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems*, pp. 67–74. DOI: [10.1145/1183471.1183484](https://doi.org/10.1145/1183471.1183484).
- Steiniger S., Lange T., Burghardt D., and Weibel R.** (2008). “An Approach for the Classification of Urban Building Structures Based on Discriminant Analysis Techniques”. *Transactions in GIS* 12.1, pp. 31–59. ISSN: 1361-1682. DOI: [10.1111/j.1467-9671.2008.01085.x](https://doi.org/10.1111/j.1467-9671.2008.01085.x).
- Steiniger S. and Weibel R.** (2007). “Relations among Map Objects in Cartographic Generalization”. *Cartography and Geographic Information Science* 34, pp. 175–197. DOI: [10.1559/152304007781697866](https://doi.org/10.1559/152304007781697866).
- Stoter J. et al.** (2009). “Methodology for evaluating automated map generalization in commercial software”. *Computers, Environment and Urban Systems. Geoinformation Generalisation and Multiple Representation* 33.5, pp. 311–324. ISSN: 0198-9715. DOI: [10.1016/j.compenvurbsys.2009.06.002](https://doi.org/10.1016/j.compenvurbsys.2009.06.002).

- Taillandier P. and Gaffuri J.** (2009). “Objective Function Designing Led by User Preferences Acquisition”. *International Conference on Information Technology and Applications*. Hanoi, Vietnam.
- Thomson R. C. and Richardson D.** (1995). “A Graph Theory Approach to Road Network Generalisation”. *17th International Cartographic Conference*. ICA, pp. 1871–1880.
- Thomson R. C. and Richardson D.** (1999). “The "Good Continuation" principle of Perceptual Organization applied to the Generalization of Road Networks”. *19th International Cartographic Conference*. ICA.
- Toussaint G.** (1980). “The relative neighbourhood graph of a finite planar set”. *Pattern Recognit.* DOI: [10.1016/0031-3203\(80\)90066-7](https://doi.org/10.1016/0031-3203(80)90066-7).
- Touya G.** (2007). “River network selection based on structure and pattern recognition.” *International Cartographic Conference*. Moscow.
- Touya G.** (2010). “A Road Network Selection Process Based on Data Enrichment and Structure Detection”. *Transactions in GIS* 14.5, pp. 595–614. DOI: [10.1111/j.1467-9671.2010.01215.x](https://doi.org/10.1111/j.1467-9671.2010.01215.x).
- Touya G.** (2011). “Le Modèle CollaGen: collaboration de processus automatiques pour la généralisation cartographique de paysages hétérogènes”. PhD thesis. Université Paris-Est.
- Touya G.** (2012). “Social Welfare to Assess the Global Legibility of a Generalized Map”. *Geographic Information Science*. Ed. by N. Xiao, M.-P. Kwan, M. F. Goodchild, and S. Shekhar. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 198–211. ISBN: 978-3-642-33024-7. DOI: [10.1007/978-3-642-33024-7_15](https://doi.org/10.1007/978-3-642-33024-7_15).
- Touya G.** (2015). “Lessons Learned From Research on Multimedia Summarization”. *Proceedings of 18th ICA Workshop on Generalisation and Multiple Representation*. Rio de Janeiro, Brazil.
- Touya G.** (2017). “Vers l’automatisation de la production de cartes”. Habilitation à diriger des recherches. Université Paris Est, p. 96.
- Touya G.** (2019). “Finding the Oasis in the Desert Fog? Understanding Multi-Scale Map Reading”. *Abstraction, Scale and Perception Workshop ICC 2019 Workshop*. Tokyo, Japan.
- Touya G.** (2021). “Multi-Criteria Geographic Analysis for Automated Cartographic Generalization”. *Cartographic Journal*. Publisher: Maney Publishing. DOI: [10.1080/00087041.2020.1858608](https://doi.org/10.1080/00087041.2020.1858608).
- Touya G., Balley S., Duchêne C., Jaara K., Regnauld N., and Gould N.** (2012). “Towards an Ontology of Spatial Relations and Relational Constraints”. *15th ICA workshop on generalisation and multiple representation*. Istanbul, Turkey.
- Touya G., Brisebard F., Quinton F., and Courtial A.** (2020). “Inferring The Scale And Content Of a Map Using Deep Learning”. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLIII-B4-2020. ISSN: 1682-1750. Copernicus GmbH, pp. 17–24. DOI: [10.5194/isprs-archives-XLIII-B4-2020-17-2020](https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-17-2020).
- Touya G., Bucher B., Falquet G., Jaara K., and Steiniger S.** (2014). “Modelling Geographic Relationships in Automated Environments”. In: *Abstracting Geographic Information in a Data Rich World: Methodologies and Applications of Map Generalisation*. Ed. by D. Burghardt, C. Duchêne, and W. Mackaness. Lecture Notes in Geoinformation and Cartography. Cham: Springer International

- Publishing, pp. 53–82. ISBN: 978-3-319-00203-3. DOI: [10.1007/978-3-319-00203-3_3](https://doi.org/10.1007/978-3-319-00203-3_3).
- Touya G. and Courtial A.** (2021). “BasqueRoads: A Benchmark for Road Network Selection”. Vol. 4, 2021. Copernicus Publications. DOI: [10.5194/ica-abs-4-5-2022](https://doi.org/10.5194/ica-abs-4-5-2022).
- Touya G., Decherf B., Lalanne M., and Dumont M.** (2015). “Comparing Image-Based Methods for Assessing Visual Clutter in Generalized Maps”. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*.
- Touya G., Duchêne C., and Ruas A.** (2010). “Collaborative Generalisation: Formalisation of Generalisation Knowledge to Orchestrate Different Cartographic Generalisation Processes”. *Geographic Information Science*. Ed. by S. I. Fabrikant, T. Reichenbacher, M. van Kreveld, and C. Schlieder. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 264–278. ISBN: 978-3-642-15300-6. DOI: [10.1007/978-3-642-15300-6_19](https://doi.org/10.1007/978-3-642-15300-6_19).
- Touya G. and Dumont M.** (2017). “Progressive Block Graying and Landmarks Enhancing as Intermediate Representations between Buildings and Urban Areas”. *Proceedings of 20th ICA Workshop on Generalisation and Multiple Representation*. Washington DC, USA.
- Touya G., Lobo M.-J., Mackaness W. A., and Muehlenhaus I.** (2021). “Please, Help Me! I Am Lost in Zoom”. *30th International Cartographic Conference*. Vol. 4. Proceedings of the International Cartographic Association (ICA) 2021. Backup Publisher: International Cartographic Association. Florence, Italy: Copernicus Publications, p. 107. DOI: [10.5194/ica-proc-4-107-2021](https://doi.org/10.5194/ica-proc-4-107-2021).
- Touya G. and Lokhat I.** (Dec. 2016). “Enhancing building footprints with squaring operations”. *Journal of Spatial Information Science* 13. DOI: [10.5311/JOSIS.2016.13.276](https://doi.org/10.5311/JOSIS.2016.13.276).
- Touya G. and Lokhat I.** (2020). “Deep learning for enrichment of vector spatial databases: application to highway interchange.” *ACM Trans. Spatial Algorithms Syst.* 6.3, p. 21. ISSN: 2374-0353. DOI: [10.1145/3382080](https://doi.org/10.1145/3382080).
- Touya G., Zhang X., and Lokhat I.** (2019). “Is deep learning the new agent for map generalization?” *International Journal of Cartography* 5.2-3. Publisher: Taylor & Francis, pp. 142–157. ISSN: 2372-9333. DOI: [10.1080/23729333.2019.1613071](https://doi.org/10.1080/23729333.2019.1613071).
- Usery E., Arundel S., Shavers E., Stanislowski L., Thiem P., and Varanka D.** (2021). “GeoAI in the US Geological Survey for topographic mapping”. *Transactions in GIS*. DOI: [10.1111/tgis.12830](https://doi.org/10.1111/tgis.12830).
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gmoez A., and Kaiser L.** (2017). “Attention Is All You Need”. *31st Conference on Neural Information Processing System*.
- Visvalingam M. and Williamon P. J.** (1995). “Simplification and Generalization of Large Scale Data for Roads: A Comparison of Two Filtering Algorithms”. *Cartography and Geographic Information Systems* 22:4, pp. 264–275.
- Wang X., Girshick R., Gupta A., and He K.** (2018). “Non-local Neural Networks”. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7794–7803. DOI: [10.1109/CVPR.2018.00813](https://doi.org/10.1109/CVPR.2018.00813).
- Wang X. and Burghardt D.** (2019). “A Mesh-Based Typification Method for Building Groups with Grid Patterns”. *ISPRS International Journal of Geo-Information* 8.4. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, p. 168. DOI: [10.3390/ijgi8040168](https://doi.org/10.3390/ijgi8040168).

- Wang X. and Burghardt D.** (2020). “Using stroke and mesh to recognize building group patterns”. *International Journal of Cartography* 6.1. Publisher: Taylor & Francis, pp. 71–98. ISSN: 2372-9333. DOI: [10.1080/23729333.2019.1574371](https://doi.org/10.1080/23729333.2019.1574371).
- Weibel R. and Dutton G.** (1999). “Generalising spatial data and dealing with multiple representations”, p. 32.
- Weibel R., Keller S., and Reichenbacher T.** (1995). “Overcoming the knowledge acquisition bottleneck in map generalization: The role of interactive systems and computational intelligence”. In: *Spatial Information Theory A Theoretical Basis for GIS*. Ed. by A. U. Frank and W. Kuhn. Vol. 988. Lecture Notes in Computer Science. Berlin Heidelberg: Springer, pp. 139–156. DOI: [10.1007/3-540-60392-1_10](https://doi.org/10.1007/3-540-60392-1_10).
- Weiss R. and Weibel R.** (2014). “Road network selection for small-scale maps using an improved centrality-based algorithm”. *Journal of Spatial Information Science* 9.9, pp. 71–99. ISSN: 1948-660X. DOI: [10.5311/JOSIS.2014.9.166](https://doi.org/10.5311/JOSIS.2014.9.166).
- Werder S.** (2009). “Formalization of Spatial Constraints”. *12th AGILE International Conference on Geographic Information Science, Leibniz Universität Hannover, Germany*, p. 13.
- Wu A. and Biljecki F.** (2021). “GANmapper: geographical content filling”. *arXiv:2108.04232 [cs.CV]*.
- Xu Y., Chen Z., Xie Z., and Wu L.** (2017). “Quality assessment of building footprint data using a deep auto encoder network”. *International Journal of Geographical Information Science* 31.10. Publisher: Taylor & Francis. ISSN: 1365-8816. DOI: [10.1080/13658816.2017.1341632](https://doi.org/10.1080/13658816.2017.1341632).
- Yan H., Weibel R., and Yang B.** (2008). “A Multi-parameter Approach to Automated Building Grouping and Generalization”. *GeoInformatica* 12.1, pp. 73–89. ISSN: 1573-7624. DOI: [10.1007/s10707-007-0020-5](https://doi.org/10.1007/s10707-007-0020-5).
- Yan J., Guilbert E., and Saux E.** (2017). “An ontology-driven multi-agent system for nautical chart generalization”. *Cartography and Geographic Information Science* 44.3, pp. 201–215. ISSN: 1523-0406, 1545-0465. DOI: [10.1080/15230406.2015.1129648](https://doi.org/10.1080/15230406.2015.1129648).
- Yan X. and Ai T.** (2018). “Analysis of Irregular Spatial Data with Machine Learning: Classification of Building Patterns with a Graph Convolutional Neural Network”. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern Saarbruecken, Germany, p. 7. DOI: [10.4230/LIPICS.GISCIENCE.2018.69](https://doi.org/10.4230/LIPICS.GISCIENCE.2018.69).
- Yan X., Ai T., Yang M., and Tong X.** (2020). “Graph convolutional autoencoder model for the shape coding and cognition of buildings in maps”. *International Journal of Geographical Information Science* 0.0. Publisher: Taylor & Francis, pp. 1–23. ISSN: 1365-8816. DOI: [10.1080/13658816.2020.1768260](https://doi.org/10.1080/13658816.2020.1768260).
- Yan X., Ai T., Yang M., Tong X., and Liu Q.** (2020). “A graph deep learning approach for urban building grouping”. *Geocarto International* 0.0. Publisher: Taylor & Francis, pp. 1–24. ISSN: 1010-6049. DOI: [10.1080/10106049.2020.1856195](https://doi.org/10.1080/10106049.2020.1856195).
- Yang J., Lu J., Lee S., Batra D., and Parikh D.** (2018). “Graph R-CNN for Scene Graph Generation”. *arXiv:1808.00191 [cs]*.
- Yang M., Jiang C., Yan X., Ai T., Cao M., and Chen W.** (2022). “Detecting interchanges in road networks using a graph convolutional network approach”. *International Journal of Geographical Information Science* 0.0. Publisher: Taylor & Francis, pp. 1–21. ISSN: 1365-8816. DOI: [10.1080/13658816.2021.2024195](https://doi.org/10.1080/13658816.2021.2024195).

- Yasin H. M. and Abdulazeez A. M.** (2021). “Image Compression Based on Deep Learning: A Review”. *Asian Journal of Research in Computer Science*, pp. 62–76. ISSN: 2581-8260. DOI: [10.9734/ajrcos/2021/v8i130193](https://doi.org/10.9734/ajrcos/2021/v8i130193).
- Yi Z., Zhang H., Tan P., and Gong M.** (2017). “DualGAN: Unsupervised Dual Learning for Image-to-Image Translation”. *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, pp. 2868–2876. ISBN: 978-1-5386-1032-9. DOI: [10.1109/ICCV.2017.310](https://doi.org/10.1109/ICCV.2017.310).
- Yuan J., Chowdhury P. K. R., McKee J., Yang L. H., Weaver J., and Bhaduri B.** (2018). “Exploiting deep learning and volunteered geographic information for mapping building in Kano, Nigeria.” *Scientific Data* 5.1, p. 180217. ISSN: 2052-4463. DOI: [10.1038/sdata.2018.217](https://doi.org/10.1038/sdata.2018.217).
- Zhang X.** (2012). “Automated evaluation of generalized topographic maps”. PhD thesis. University of Twente.
- Zhang X., Ai T., Stoter J., Kraak M.-J., and Molenaar M.** (2013). “Building pattern recognition in topographic data: examples on collinear and curvilinear alignments”. *GeoInformatica* 17.1, pp. 1–33. ISSN: 1573-7624. DOI: [10.1007/s10707-011-0146-3](https://doi.org/10.1007/s10707-011-0146-3).
- Zhang X. and Guilbert E.** (2011). “A multi-agent system approach for feature-driven generalization of isobathymetric line.” In: *Advances in Cartography and GIScience, volume 1 of Lecture Notes in Geoinformation and Cartography*. Springer. Berlin Heidelberg, pp. 477–495.
- Zhang X., Stoter J., Ai T., Kraak M.-J., and Molenaar M.** (2013). “Automated evaluation of building alignments in generalized maps”. *International Journal of Geographical Information Science* 27.8. Publisher: Taylor & Francis, pp. 1550–1571. ISSN: 1365-8816. DOI: [10.1080/13658816.2012.758264](https://doi.org/10.1080/13658816.2012.758264).
- Zhang Y., Yin Y., Zimmermann R., Wang G., Varadarajan J., and Ng S.-K.** (2020). “An Enhanced GAN Model for Automatic Satellite-to-Map Image Conversion”. *IEEE Access* 8. Conference Name: IEEE Access, pp. 176704–176716. ISSN: 2169-3536. DOI: [10.1109/ACCESS.2020.3025008](https://doi.org/10.1109/ACCESS.2020.3025008).
- Zhang Z., Liu Q., and Wang Y.** (2018). “Road Extraction by Deep Residual U-Net”. *IEEE Geoscience and Remote Sensing Letters* 15.5. Conference Name: IEEE Geoscience and Remote Sensing Letters, pp. 749–753. ISSN: 1558-0571. DOI: [10.1109/LGRS.2018.2802944](https://doi.org/10.1109/LGRS.2018.2802944).
- Zhao B., Li X., and Lu X.** (2019). “Hierarchical Recurrent Neural Network for Video Summarization”. *arXiv:1904.12251 [cs.CV]*.
- Zhao R., Ai T., Yu W., He Y., and Shen Y.** (2020). “Recognition of building group patterns using graph convolutional network”. *Cartography and Geographic Information Science* 47.5, pp. 400–417. ISSN: 1523-0406, 1545-0465. DOI: [10.1080/15230406.2020.1757512](https://doi.org/10.1080/15230406.2020.1757512).
- Zheng C., Cham T.-J., and Cai J.** (2021). “The Spatially-Correlative Loss for Various Image Translation Tasks”. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA: IEEE, pp. 16402–16412. ISBN: 978-1-66544-509-2. DOI: [10.1109/CVPR46437.2021.01614](https://doi.org/10.1109/CVPR46437.2021.01614).
- Zhou Q. and Li Z.** (2017). “A Comparative Study of Various Supervised Learning Approaches to Selective Omission in a Road Network”. *The Cartographic Journal* 54.3, pp. 254–264. DOI: [10.1179/1743277414y.0000000083](https://doi.org/10.1179/1743277414y.0000000083).
- Zhou X., Li W., Arundel S. T., and Liu J.** (2018). “Deep Convolutional Neural Networks for Map-Type Classification”. *arXiv:1805.10402 [cs, stat]*.

- Zhu J.-Y., Park T., Isola P., and Efros A. A.** (2017). “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks”. *2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, pp. 2242–2251. ISBN: 978-1-5386-1032-9. DOI: [10.1109/ICCV.2017.244](https://doi.org/10.1109/ICCV.2017.244).
- Zhu X. X., Tuia D., Mou L., Xia G.-S., Zhang L., Xu F., and Fraundorfer F.** (2017). “Deep learning in remote sensing: a review”. *IEEE Geoscience and Remote Sensing Magazine* 5.4, pp. 8–36. ISSN: 2168-6831, 2473-2397. DOI: [10.1109/MGRS.2017.2762307](https://doi.org/10.1109/MGRS.2017.2762307).

Publications

Courtial A., El Ayedi A., Touya G., and Zhang X. (2020). “Exploring the Potential of Deep Learning Segmentation for Mountain Roads Generalisation”. en. *ISPRS International Journal of Geo-Information* 9.5. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute, p. 338. DOI: [10.3390/ijgi9050338](https://doi.org/10.3390/ijgi9050338).

Courtial A., Touya G., and Zhang X. (2020). “Constraint Based Evaluation of Generalized Images Generated by Deep Learning”. *23rd ICA Workshop on Map Generalisation and Multiple Representation*. Delft, Netherlands: ICA.

Touya G., Brisebard F., Quinton F., and Courtial A. (2020). “Inferring The Scale And Content Of a Map Using Deep Learning”. English. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLIII-B4-2020. ISSN: 1682-1750. Copernicus GmbH, pp. 17–24. DOI: [10.5194/isprs-archives-XLIII-B4-2020-17-2020](https://doi.org/10.5194/isprs-archives-XLIII-B4-2020-17-2020).

Courtial A., Touya G., and Zhang X. (2021a). “Generative Adversarial Networks to Generalise Urban Areas in Topographic Maps”. English. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XLIII-B4-2021. ISSN: 1682-1750. Copernicus GmbH, pp. 15–22. DOI: [10.5194/isprs-archives-XLIII-B4-2021-15-2021](https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-15-2021).

Touya G. and Courtial A. (2021). “BasqueRoads: A Benchmark for Road Network Selection”. en. Vol. 4, 2021. Copernicus Publications. DOI: [10.5194/ica-abs-4-5-2022](https://doi.org/10.5194/ica-abs-4-5-2022).

Courtial A., Touya G., and Zhang X. (2021b). “AlpineBends - A Benchmark for Deep Learning-Based Generalisation”. en. Vol. 4. Copernicus Publications. DOI: [10.5194/ica-abs-4-1-2022](https://doi.org/10.5194/ica-abs-4-1-2022).

Courtial A., Touya G., and Zhang X. (2021c). “Can Graph Convolution Networks Learn Spatial Relations?” *Abstracts of the ICA* 3, pp. 1–2. DOI: [10.5194/ica-abs-3-60-2021](https://doi.org/10.5194/ica-abs-3-60-2021).

Courtial A., Touya G., and Zhang X. (2022c). “Representing Vector Geographic Information As a Tensor for Deep Learning Based Map Generalisation”. *AGILE 2022*. Ed. by E. Parseliunas, A. Mansourian, P. Partsinevelos, and J. Suziedelyte-Visockiene. Vol. 3. Vilnius, Lithuania: Copernicus Publications, p. 32. DOI: [10.5194/agile-giss-3-32-2022](https://doi.org/10.5194/agile-giss-3-32-2022). **Best short paper award.**

Courtial A., Touya G., and Zhang X. (2022a). “Constraint-Based Evaluation of Map Images Generalized by Deep Learning”. *Journal of Geovisualization and Spatial Analysis* 6. DOI: [10.1007/s41651-022-00104-2](https://doi.org/10.1007/s41651-022-00104-2).

Courtial A., Touya G., and Zhang X. (2022b). “Deriving map images of generalised mountain roads with generative adversarial networks”. *International Journal of Geographical Information Science* 0.0. Publisher: Taylor & Francis, pp. 1–30. ISSN: 1365-8816. DOI: [10.1080/13658816.2022.2123488](https://doi.org/10.1080/13658816.2022.2123488).

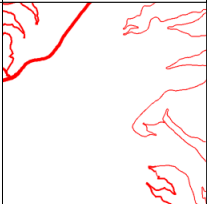

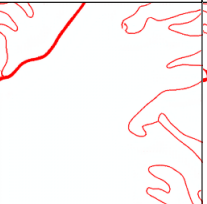
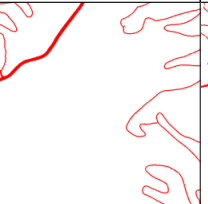


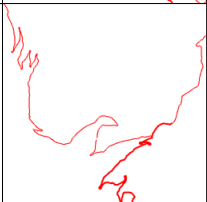



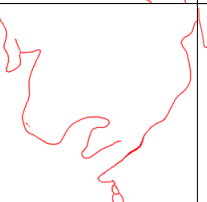


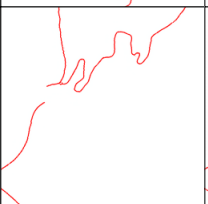
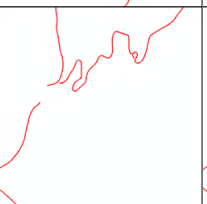
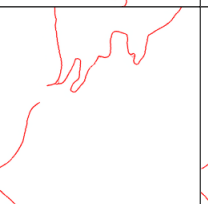
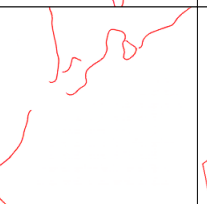

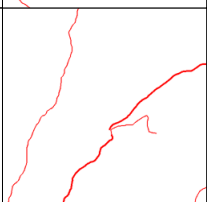


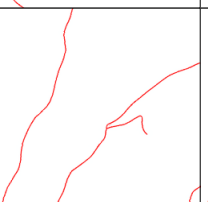
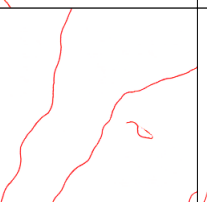


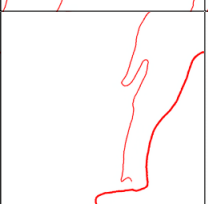



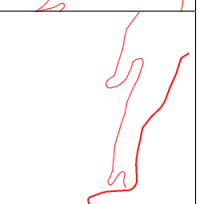
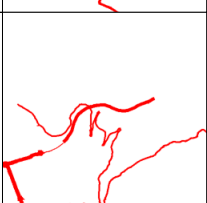



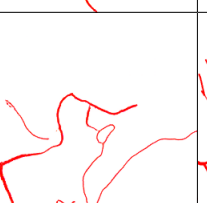
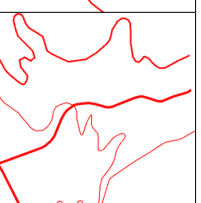
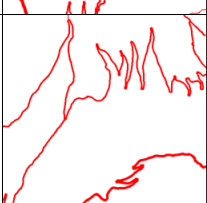

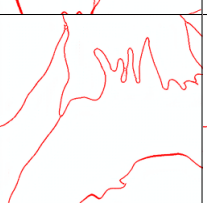

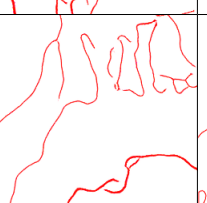
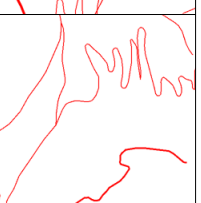
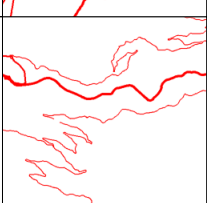
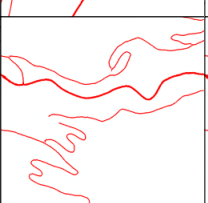
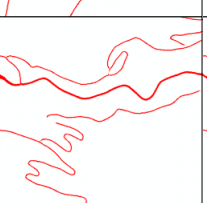
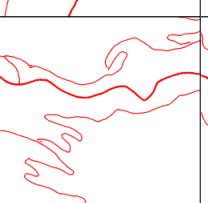

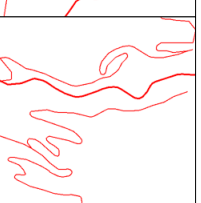
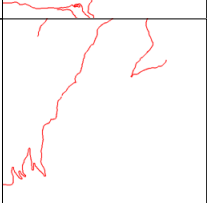


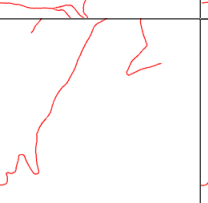
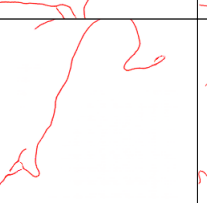

Appendices




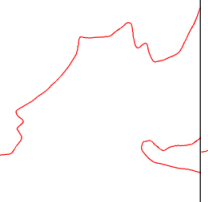


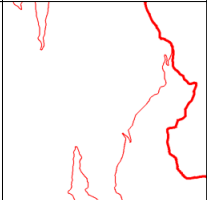





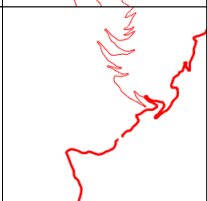
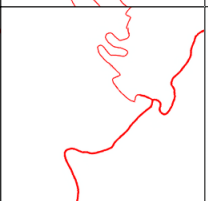

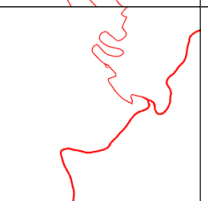
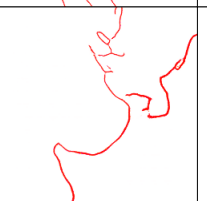

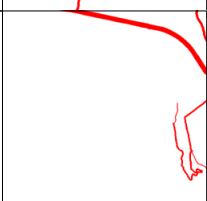


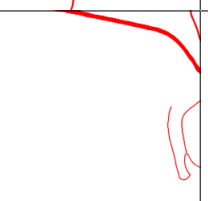
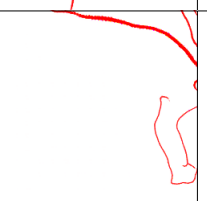
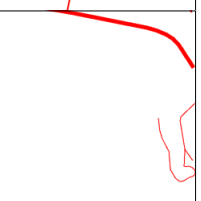
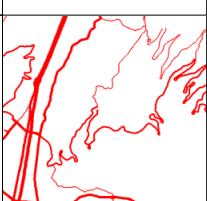
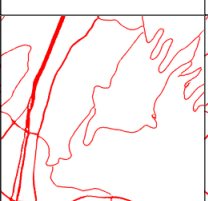
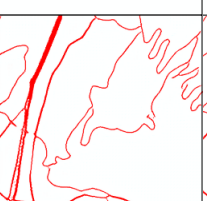
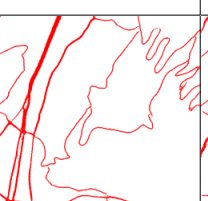
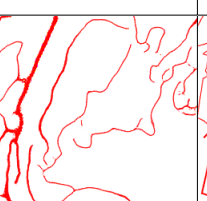
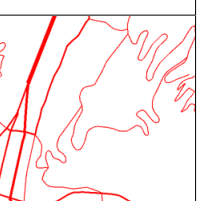


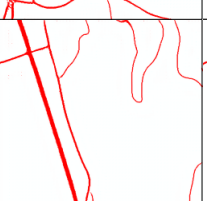
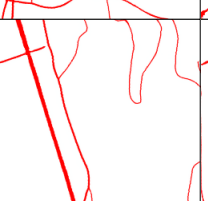
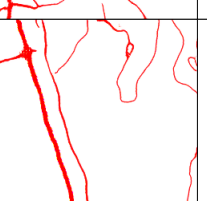

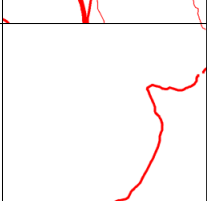
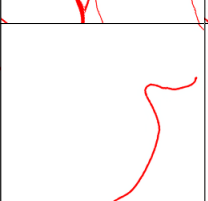
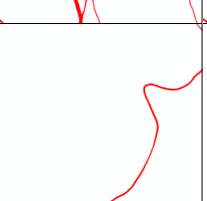
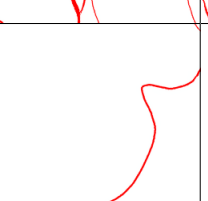
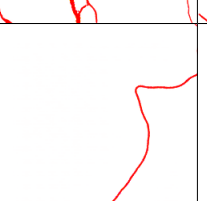





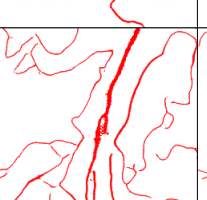





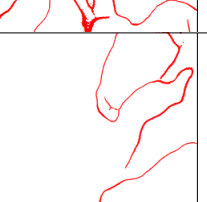

A Comparison of models for road shape generalization

We propose a comparison of the different models proposed in Chapter B. We compare the reference (i.e. state-of-the-art vector algorithms generalization and manual corrections) with the prediction of five variations of the proposed model for 18 situations from the test set. All tiles represent a portion of the 2.5 km² map.

Name	Strategy	Model	Sample image
Model A	Unsupervised	CycleGAN	1223 (from Alps only)
Model B	Unsupervised	CycleGAN	2271 (from several mountain regions)
Model C	Unsupervised	CycleGAN	1223 (+ connectivity loss)
Model D	Supervised	Pix2pix	1223 (from Alps only)

Table 1: List of the experiments compared in this appendix.

	input	A	B	C	D	reference
1						
2						
3						
4						
5						
6						
7						
8						
9						

	input	A	B	C	D	reference
10						
11						
12						
13						
14						
15						
16						
17						
18						

B Estimation of constraint adaptability for raster calculation

In Chapter H, we observe that it is challenging to adapt constraint-based evaluation to maps with several themes. In these maps, the spatial relations are even more important and concern several different objects, e.g. buildings are aligned along a road, or a road is crossing a forest. The preservation of relations contributes to the global structure and understanding of the map (Mackaness and Edwards 2002) and help to determine map location. However it seem difficult to evaluate.

Moreover, the raster representation makes conflict detection harder, as each pixel can have only one value even if there is a coalescence of several elements at this place. Finally, the identification of groups of objects can be done using a distance criterion using buffers, but the identification with similarity criteria could be more difficult. In this appendix, we propose to estimate the difficulty of the raster adaptation of constraints from (Stoter et al. 2009). We identify four levels of adaptability of measure:

- ++ Can be measured directly in the image (i.e., involves only a measure on the pixel values, the image statistics, the patch, or the pixel set).
- + May require image processing before the measurement calculation; the measure is calculated on a derived image.
- - May require a new paradigm. For example, for an overlapping measure, we propose not to measure existing overlapping (this seems impossible in raster), but rather measuring "almost overlapping" (i.e., pixel that would be overlap if the image is slightly zoomed out) that is simpler to measure in raster, and hypothesize that these two measures (overlapping and almost overlapping) are correlated. Similarly, for orientation rather than comparing all directions for each element, we identify some direction (horizontal and vertical) and compare the edge distribution for these directions, supposing that the edge distribution for all directions is correlated with the edge distribution for these directions.
- – We think there is no possible adaptation and that the measure may require vectorization. For these points we did not prove the in-adaptability and some un-evisaged solution may exist, and thus suggestions are welcome.

Id	Type	Geometry	Property	Vector measure	Adaptation	Feasability
1-1	Minimal dimension	Polygon	Area	Target area $> x$ map mm^2	Size of each polygon pixel set	++
1-4	Minimal dimension	Polygon	Area of any hole in a polygon	Target area of hole $> x$ mm^2	Size of each hole (set in a polygone pixel set) pixel	+
1-5	Minimal dimension	Line/polygon	Length of an edge/line	target length $> x$ map mm	?	--
1-13	Shape	Polygon	Squareness	Target angles = 90	Squareness measure from image processing	+
1-18	Orientation	Any	General orientation	target orientation = initial orientation $x\%$	Preservation of horizontal and vertical edges (detected using a Sobel filter)	-
2-1	Minimal dimensions	Any - any	Minimal distance	target distance $> x$ map mm	Buffer based measure	++
2-3	Orientation	Pairs of line/polygon	Orientation	Object must be parallel to object	Image processing for gap detection & image processing for measure of gap squareness	-
2-9	Topology	Any - any	Topological consistency	Target overlapping = initial overlapping	Mathematics morphology to identify almost overlapping pixels	-
3-4	Topology	Line and polygon	Intersection	No other-intersections must be created	Preservation of number of background set	++
3-5	Topology	Line and polygon	Connectivity	Connectivity must remain	Preservation of number of line or polygon set	++
3-7	Topology	Polygon	Spatial distribution	Target distribution should be similar to initial distribution in alignments	?	--
3-8	Topology	Polygon	Spatial distribution	Target distribution should be similar to initial distribution for buildings surrounded by minimal cycle of roads	?	--

Résumé substantiel : Exploration du potentiel de l'apprentissage profond pour la généralisation cartographique.

Contexte

Depuis 20 ans, l'usage des cartes a évolué de façon importante, les cartes sont désormais majoritairement en ligne, multi-représentations, et multi-échelles. Les utilisateurs attendent un contenu spécifique à leurs besoins, à leurs appareils et interactif, où les informations sont accessibles de façon directe et lisible. Néanmoins, actuellement les cartes en lignes ne sont ni spécifiques aux besoins de l'utilisateur ni simple à naviguer. Par exemple, le saut de représentation entre les échelles d'une carte produit de la confusion pendant le zoom (TOUYA 2019). D'autre part, la multiplication des sources de données a entraîné l'augmentation du besoin de synthèse, d'harmonisation et d'abstraction de l'information géographique sous forme d'une carte. Nous pensons que le développement de processus cartographiques qui limitent cette confusion, et répondent à ces nouveaux besoins est conditionné à de grandes avancées de la recherche en automatisation de la généralisation cartographique. Il s'agit du processus d'abstraction qui vise à adapter le niveau de détail de l'information géographique pour la rendre lisible sur une carte à une échelle plus petite. L'automatisation de ce processus est complexe mais essentielle à la production de cartes répondant à la fois aux usages traditionnels et à ces nouveaux usages. La recherche sur l'automatisation de la généralisation cartographique a ralenti ces dernières années tandis que le besoin lui n'a pas diminué. Nous pensons qu'elle pourrait bénéficier des avancées récentes en apprentissage profond qui rendent possible la résolution de tâches de plus en plus complexes, grâce à l'utilisation de nombreux exemples. L'usage de méthodes issues de l'apprentissage automatique pour la généralisation cartographique a déjà été expérimenté dans quelques cas d'étude isolés et limités. Le contexte actuel, et l'apprentissage profond sont prometteurs :

- les dynamiques open-data et l'ère du big data constituent de nombreuses sources pour construire d'importants jeux d'apprentissages ;
- la promesse de l'apprentissage profond d'apprendre de données brutes, nous permettrait de se passer de l'étape d'extraction de variables, qui était la principale limite à l'utilisation de l'apprentissage automatique en généralisation cartographique ;
- l'état de l'art en apprentissage profond montre une capacité accrue à résoudre des tâches similaires à la généralisation cartographique dans les domaines des sciences de l'information géographique et de la visualisation d'informations ;

Objectif et approche

L'approche par apprentissage profond de la généralisation est encore expérimentale, et par ailleurs elle est à l'opposé des approches traditionnelles de la généralisation

cartographique, où l'on formalisait d'abord des spécifications cartographiques puis on appliquait des opérations pour s'y conformer. En effet, un modèle d'apprentissage profond va plutôt apprendre à reproduire des cartes généralisées sans explicitement définir les spécifications de la carte ou les opérateurs employés, mais en déduisant ce à quoi devrait ressembler le résultat de la généralisation. Le premier objectif de cette thèse est donc de démontrer que l'approche est possible et prometteuse, puis nous proposons d'explorer via divers cas d'étude le potentiel et les questions de recherche soulevées par l'usage de l'apprentissage profond pour la généralisation cartographique. Notre exploration est construite à travers quatre cas d'études qui décrivent chacun un problème particulier de généralisation cartographique. Les trois premiers cas d'étude correspondent à des étapes classiques de processus de généralisation cartographique (généralisation graphique, détection de pattern, sélection d'entités) ; tandis que le dernier correspond à une approche plus globale pour la génération d'une carte généralisée qui combine plusieurs opérateurs de généralisation. Ce manuscrit collecte et organise les leçons apprises de ces expériences. Nous discutons donc les questions de recherches liées à la mise en œuvre concrète des modèles d'apprentissage profonds pour la généralisation cartographique (jeu de données et architecture), leurs usages et limites, l'évaluation de tels modèles, et leur intégration dans des processus cartographiques préexistants. En particulier, nous cherchons :

- à identifier le potentiel et les défis d'une telle approche (Partie I) ;
- à concrètement expérimenter l'apprentissage de quelques tâches de généralisation cartographique importantes (Partie II) ;
- à donner des indications pour les futures utilisations de l'apprentissage profond pour la généralisation cartographique (Partie III).

Partie I : Un nouveau paradigme pour la généralisation cartographique

Cette partie a pour objectif d'identifier le potentiel et les défis de l'approche par apprentissage profond pour la généralisation cartographique. Elle est constituée d'un chapitre d'état de l'art qui présente la recherche en généralisation cartographique, et en apprentissage profond, puis deux chapitres mobilisent cet état de l'art pour identifier, d'une part, les tâches de généralisation cartographique qui peuvent être résolues par une approche par apprentissage profond, et d'autre part, les méthodes d'apprentissage profond pouvant être mobilisées pour résoudre des problèmes de généralisation cartographique.

L'état de l'art (Chapitre A) nous permet d'identifier le potentiel et les défis d'une approche par apprentissage profond pour la généralisation cartographique :

- + l'apprentissage profond est capable de résoudre des tâches similaires à la généralisation cartographique (par exemple la génération de résumé à partir de vidéo, la simplification de ligne à main levée en dessin, ou la génération d'image de carte) ;
- + il est adapté pour résoudre des problèmes graphiques où l'information nécessaire est implicite dans les données (et c'est le cas de la généralisation cartographique) ;
- + de nombreux exemples de cartes détaillées et généralisées sont disponibles pour servir d'exemples d'apprentissage ;

- la représentation image est l'approche principale en apprentissage profond, mais l'utilisation d'images pour la généralisation cartographique est considérée moins efficace, car elle implique une perte d'information par rapport aux bases de données géographiques vectorielles ;
- les modèles d'apprentissage profonds sont complexes à paramétrer et nécessiteront probablement d'être adaptés pour traiter des informations cartographiques ;
- l'apprentissage profond donne des résultats sans explication qui peuvent être considérés comme non fiables ou difficiles à intégrer dans un processus semi automatique ;
- l'entraînement de modèles d'apprentissage profond demande des ressources importantes qui peuvent être difficilement accessibles aux chercheurs en cartographie, ou avoir un coût économique ou environnemental trop important.

Ensuite, dans le Chapitre B nous identifions comment des opérations de généralisation cartographique peuvent être formulées en une tâche qu'un modèle d'apprentissage profond peut résoudre. En particulier, nous relevons qu'une grande partie des tâches de généralisation cartographique peuvent être exprimées comme un problème de classification, segmentation ou régression sur des images ou des graphes (ou les éléments d'une image ou d'un graphe). Finalement, nous notons que la combinaison de plusieurs approches semble plus prometteuse encore. Par exemple pour la simplification d'une forme, une approche image est possible, ou bien une approche qui combine sélection et régression sur les nœuds d'un graphe pour prédire les nouvelles position des sommets.

Enfin, dans le Chapitre C, nous explorons comment mettre en œuvre des modèles d'apprentissage profond pour la généralisation cartographique. En particulier, nous proposons : des critères pour le choix d'un modèle de la littérature en vision par ordinateur ; des adaptations pour améliorer l'apprentissage de la généralisation cartographique ; des lignes directrices pour la création de jeux de donnée d'apprentissage adaptés et un modèle d'organisation des données géographiques en jeu d'exemples pour l'apprentissage de tâche liée à la cartographie et à la généralisation.

Partie II : L'apprentissage profond peut-il apprendre la généralisation cartographique ?

Cette partie a pour objectif d'expérimenter le potentiel de l'apprentissage de quelques tâches de généralisation cartographique importantes. Elle est organisée en trois chapitres décrivant chacun une série d'expériences visant à répondre à une problématique : l'apprentissage profond peut-il comprendre des relations spatiales / apprendre la généralisation cartographique / générer des cartes topographiques généralisées.

Le Chapitre D teste la capacité des réseaux d'apprentissage sur les graphes à comprendre des relations spatiales, pour cela nous explorons la capacité de ces modèles à réaliser des tâches de généralisation qui reposent sur des relations spatiales complexes et implicites dans les graphes. Les cas d'études sont la détection d'alignement de bâtiments et la sélection du réseau routier. Ces expériences montrent des résultats perfectibles mais révèlent une capacité à faire des prédictions cohérentes avec les relations spatiales que nous (humains) observons dans les données.

Le Chapitre E propose d'éprouver la capacité des modèles d'apprentissage basés sur des images à apprendre la généralisation graphique des entités cartographiques via le cas d'étude des routes de montagne pour un saut d'échelle important. Cette expérience compare plusieurs approches (par segmentation et avec un GAN) et révèle que les modèles sont capables de reproduire le lissage des lignes, l'élargissement de virages serrés, et la caricature, pour la plupart des situations testées. Dans cette expérience la préservation de la structure du réseau routier (et en particulier éviter la création de boucle et les déconnexions) est un enjeu majeur et les modèles génératifs sont plus prometteurs sur ce point. Nous introduisons également une mesure de ces altérations de structure, et proposons de l'utiliser comme terme de la fonction d'objectif pour les minimiser. Finalement nous obtenons des prédictions jugées aussi bonnes que la généralisation de l'état de l'art obtenue par des processus semi-automatiques.

Dans le Chapitre F, nous démontrons que les réseaux adverses génératifs (GAN) sont capables de générer des images ressemblant à des cartes à une échelle donnée. Nous générons des images de cartes généralisées à partir de cartes détaillées comportant : des bâtiments, les réseaux routier et hydrographique. Nos résultats reproduisent la simplification, l'agrandissement et la typification de bâtiments, mais la représentation en tuiles est limitante pour les tâches nécessitant plus de contexte comme la sélection du réseau routier et le grisage des centres ville.

Partie III : Le futur de la généralisation cartographique par apprentissage.

Les résultats de la seconde partie nous laissent penser que l'utilisation de l'apprentissage profond pour la généralisation cartographique pourrait être un tournant dans l'automatisation de la généralisation cartographique. Mais pour cela, les modèles doivent être insérés et comparés avec les processus traditionnels de généralisation cartographique. L'objectif de cette partie est de discuter les enjeux de l'utilisation d'approche par apprentissage profond pour la généralisation.

Dans le Chapitre G, nous discutons la capacité de transfert des modèles d'apprentissage profond et l'illustrons pour les modèles du Chapitre E et F. Puis, nous étudions l'intégration d'un modèle, d'une part dans un processus traditionnel (comme indicateur pour l'enrichissement des données géographique ou comme opérateur de généralisation) et d'autre part dans un processus constitué entièrement de modèles d'apprentissage. Enfin, nous proposons un exemple de processus de généralisation qui intégrerait les modèles proposés et testés dans la Partie II pour la génération de cartes topographique au 1 :50 000.

Dans le Chapitre H nous expliquons les enjeux d'évaluation de la généralisation cartographique réalisée par apprentissage profond. Nous illustrons trois méthodes d'évaluation en les appliquant à l'évaluation des modèles proposés en Partie II. En particulier, nous montrons que l'adaptation des contraintes de généralisation cartographique pour une mesure basée sur des tuiles permet une évaluation automatique et assez fine des différents aspects de la généralisation, mais toutes les contraintes proposées dans le passé pour des cartes topographiques ne semblent pas adaptables. La reconstruction d'une base de données vectorielle avant l'évaluation, quant à elle, permet d'appliquer les méthodes traditionnelles d'évaluation de la généralisation cartographique, mais demande un temps de calcul important et peut altérer la qualité des prédictions, ce qui n'est pas approprié dans de nombreux cas. Finalement,

le test utilisateur est le seul moyen totalement fiable de s'assurer que la prédiction correspond à ce qui est attendu par l'utilisateur, néanmoins, il ne permet pas l'évaluation automatique.

Contributions et conclusions

Pour conclure, les contributions de cette thèse sont d'une part des enseignements sur l'utilisation de l'apprentissage profond pour la généralisation cartographique et d'autre part des modèles, benchmarks et un processus de généralisation cartographique par apprentissage profond. En particulier, nous proposons :

- un jeu de données pour la généralisation graphique de routes de montagne incluant un jeu d'entraînement et plusieurs jeux tests pour étudier la transférabilité du modèle (disponible sur Zenodo) et un ensemble de mesures pour valider les tuiles prédites (disponibles dans le dépôt Github DeepMapGen) ;
- des modèles entraînés pour chaque cas d'études proposé ;
- le code pour les adaptations proposées des modèles utilisés notamment la fonction d'objectif de connectivité et le fusionGAN ;
- le processus DeepMapScaler qui intègre les modèles proposés dans cette thèse pour générer des cartes généralisées ;
- une méthode pour générer des jeux de donnée d'apprentissage adaptés à la généralisation cartographique et le code associé (disponible sous la forme d'un plugin de la plateforme open source de généralisation CartAgen) ;
- une méthode pour la reconstruction et l'assemblage de tuiles généralisées indépendamment en une base de données vecteur.

L'objectif de notre thèse était d'explorer le potentiel de l'apprentissage profond pour la généralisation cartographique. Les résultats de nos expériences sont prometteurs et révèlent la capacité des modèles d'apprentissage profond à apprendre à interpréter, abstraire et représenter l'information géographique depuis des images ou des graphes. Nos résultats sont suffisamment performants pour un usage comme indicateur (c'est-à-dire pour l'enrichissement de l'information géographique avant la généralisation), mais la qualité et les difficultés d'évaluation/intégration, des modèles qui visent à apprendre un ou plusieurs opérateurs ne sont pour l'instant pas suffisant pour remplacer ou compléter les opérateurs traditionnels dans un processus de généralisation. Nous pensons que pour rendre possibles de tels usages, l'amélioration des architectures et jeux d'apprentissage est nécessaire et en particulier pour une meilleure utilisation du contexte spatial.

Aussi, notre thèse s'inscrit dans l'objectif global de la geo-intelligence artificielle, qui est de concevoir un programme capable de comprendre une question d'un utilisateur relative à l'information géographique, et de regrouper, analyser et présenter les données requises pour répondre à la question de l'utilisateur dans une forme adaptée (JANOWICZ, GAO, MCKENZIE, HU et BHADURI 2020, p. 631). En effet, la carte est bien souvent le mode de communication de la solution la plus souhaitable pour communiquer une information géographique à un utilisateur humain, et requiert l'automatisation et l'adaptabilité des processus de généralisation. L'usage de l'apprentissage profond pourrait débloquer la recherche d'un modèle capable de générer de telles cartes.