



École Polytechnique Fédérale de Lausanne

A generic method for cartographic realignment using local feature matching: towards a computational urban history

by Maxime François Jan

Master Thesis

Prof. Dr. Frédéric Kaplan
Thesis Advisor

Dr. Julien Perret
External Expert

Rémi Petitpierre
Paul Guennec
Thesis Supervisors

Keywords : Urban history, historical maps, semantic segmentation, map realignment, georeferencing, map analysis

EPFL-CDH-DHLAB
INN 140 (Bâtiment INN)
Station 14
CH-1015 Lausanne

August 25, 2022

Abstract

Despite the massive number of historical maps that have been scanned and made public in recent years, it is impossible to manually analyse such large corpora. In this work, we propose a two-step pipeline to automatically georeference, i.e., to realign urban maps on contemporary geodata, and to enable computational analysis. First, we trained an OCRNet to segment historical maps and extract their road networks. Our model achieved state-of-the-art, thus improving the mean Intersection over Union (IoU) score of previous work by 12% on a reference benchmark. We examined the performance of SIFT and Superpoint+Superglue, two approaches that rely on local feature detection and matching to realign the segmented maps on today's road network. On a corpus of Parisian maps, 58% were correctly georeferenced with SIFT, with an average precision of 12m, while Superpoint+Superglue correctly realigned 78% of these maps, but with an average precision of 28m. By using SIFT as a model for fine-tuning, this latter inaccuracy could be reduced to 24m. Additionally, we demonstrated the genericity of our methods by applying them to 20 additional maps of five different cities. The results revealed a similar accuracy to that of the Parisian corpus. Finally, we used this pipeline to create a dataset of 346 realigned maps of Paris, covering 190 years, and proved the practicality of a computational urban analysis by answering three historical research questions. We began by computing a series of shortest pathways within Paris and demonstrated that travel distances decreased over time. The average number of connections at intersections was then examined, and it was shown that the Haussmannian transformations had no effect on this metric. Finally, by plotting the number of intersections and the length of the road networks as a function of the distance from the city centre, we identified the second half of the XIXth century and the beginning of the XXth century as the periods that showed the greatest expansion in our dataset.

Résumé

Malgré le grand nombre de cartes historiques qui ont été numérisées et rendues publiques ces dernières années, il est impossible d'analyser manuellement de tels corpus. Dans ce travail, nous proposons une pipeline en deux étapes pour géoréférencer automatiquement, c'est-à-dire pour réaligner des cartes urbaines sur leurs géodonnées contemporaines, et pour permettre une analyse computationnelle. Tout d'abord, nous avons entraîné un OCRNet à segmenter les cartes historiques et à extraire leurs réseaux routiers. Notre modèle a atteint l'état de l'art, améliorant ainsi le score moyen d'Intersection over Union (IoU) des travaux précédents de 12% sur un benchmark de référence. Nous avons examiné les performances de SIFT et de Superpoint+Superglue, deux approches qui reposent sur la détection et la mise en correspondance de caractéristiques locales, pour réaligner les cartes segmentées sur le réseau routier actuel. Sur un corpus de cartes parisiennes, 58% ont été correctement géoréférencées avec SIFT, avec une précision moyenne de 12m, tandis que Superpoint+Superglue a correctement réaligné 78% de ces cartes, mais avec une précision moyenne de 28m. En utilisant SIFT comme modèle de réglage fin, cette dernière imprécision a pu être réduite à 24m. En outre, nous avons démontré la généralité de nos méthodes en les appliquant à 20 cartes supplémentaires de cinq villes différentes. Les résultats ont révélé une précision similaire à celle du corpus parisien. Enfin, nous avons utilisé notre pipeline pour créer un jeu de données de 346 cartes réalignées de Paris, couvrant 190 ans, et avons prouvé la viabilité d'une analyse urbaine computationnelle en répondant à trois questions de recherche historique. Nous avons commencé par calculer une série de chemins les plus courts dans Paris et avons démontré que les distances de déplacement diminuaient avec le temps. Nous avons ensuite examiné le nombre moyen de connexions aux intersections, et montré que les transformations haussmanniennes n'avaient aucun effet sur cette métrique. Enfin, en traçant le nombre d'intersections et la longueur des réseaux routiers en fonction de la distance du centre-ville, nous avons identifié la seconde moitié du XIXe siècle et le début du XXe siècle comme les périodes qui ont montré la plus grande expansion dans notre ensemble de données.

Contents

Abstract (English/Français)	2
1 Introduction	6
2 Previous work	9
2.1 Semantic segmentation of historical maps	9
2.2 Map realignment	10
2.3 Computational map analysis	10
3 Methods	12
3.1 Semantic segmentation of historical maps	12
3.1.1 Dataset and data augmentation	12
3.1.2 Network architecture and configuration	13
3.2 Map realignment	13
3.2.1 Local feature detection, description, and matching	14
3.2.2 Homography fine-tuning	15
3.2.3 Quantitative and qualitative evaluation	15
3.3 Computational map analysis	16
3.3.1 Dataset	16
3.3.2 Travel distances	17
3.3.3 Number of roads at intersections	18
3.3.4 Rythm of expansion of the Parisian road network	19
4 Results and discussion	20
4.1 Semantic segmentation of historical maps	20
4.2 Map realignment	21
4.2.1 SIFT	21
4.2.2 Superpoint+Superglue	23
4.2.3 Homography fine-tuning	23
4.2.4 Qualitative assessment	24
4.3 Computational map analysis	25
4.3.1 Travel distances	25
4.3.2 Number of roads at intersections	25

4.3.3	Rythm of expansion of the Parisian road network	26
5	Conclusion	29
	Bibliography	31
A	Additional plots	36

Chapter 1

Introduction

In the field of urban history, archives of city maps represent one of the richest sources of information, as they provide a historical record of the city's morphology at a given time. Prior to the XVIth century, these maps commonly represented cities by roughly drawing the main buildings within their walls in perspective or axonometry [18]. The function of these maps was mainly to illustrate the city's architecture while displaying the urban space organisation. However, in the XVIth century, the purpose of maps started to shift. Beginning with the famous map of Imola drawn by Leonardo Da Vinci in 1502 [34], the popularity of planar representations, with a high level of details in the urban fabric, started to rise [18]. This style was predominantly used for military cartography, as the function of these maps was primarily practical. At the end of the XVIIth century, the need for precise representation of public space significantly grew, and thus, the techniques developed by the military spread to a larger audience. All over Europe, cities were mapped with planar geometry, and their ease of creation led to an explosion in their production during the XVIIIth century [18].

With all these maps, urban historians have a wealth of information to study. In particular, the abundance of maps created with planar geometry makes it possible to compare a city's topology through time effectively. Such a comparative analysis is called *map regression* [27]. This technique consists of compiling a sequence of maps of different time periods depicting the same geographical region and comparing a set of predefined features. For instance, this analysis can highlight the expansion of a city in certain areas or can help to observe the densification of a neighbourhood. Map regression is therefore a tool that enables urban historians to better understand the development and transformations of a city [27].

However, performing map regression is not a straightforward process. While the planar geometry grants a common ground for comparison, maps may still differ in their scale or orientation. For this reason, the corpus of studied maps is typically first realigned on today's topology. To do so, the maps are first scanned and then manually georeferenced with the help of a geographical information system software [1]. This preprocessing step can be time-consuming, especially

when dealing with a large corpus of maps. Furthermore, while the realignment alleviates the comparison, map regression remains a tedious task. Indeed, while it is relatively simple to observe the similarities and differences between two realigned maps that are put side by side or superimposed, the task is significantly more difficult when dealing with corpora of hundreds of maps of diverse styles. It thus appears that while map regression is a great analysis tool, performing it manually can turn out to be too time-consuming and may even be impossible for large scale studies.

If map regression appears to be too difficult to perform manually when dealing with large amounts of data, it is logical to investigate computational methods that would alleviate the workload and allow analysis of larger map corpora. This work thus proposes a two-step pipeline to automatically realign map scans to enable computational map analyses. More specifically, this project is divided into three main steps. First, a semantic segmentation step is applied to the scanned maps, which allows us to remove unnecessary information and to only extract the road networks. Then, these networks are automatically realigned on today's topology using local feature detection and matching. Finally, we propose a small set of computational methods that can be used to analyse the resulting digital corpus.

The first step of this pipeline, semantic segmentation, takes advantage of the latest innovation in Computer Vision. Indeed, while this task used to yield poor and non-generic outcomes, the recent advances in Deep Learning, and more specifically with Convolutional Neural Networks (CNN) [28], allowed to obtain flexible and accurate results [40]. In 2021, Petitpierre et al. studied different CNN architectures to segment historical city maps and achieved state-of-the-art for this task [33]. Since then, however, new network architectures for semantic segmentation have been developed and the performance of this task keeps growing, especially with the rise of Transformers [24]. Therefore, as a first research question, this work will explore whether modern Deep Learning models can surpass the current state-of-the-art for historical city map semantic segmentation. The results of our work showed that OCRNet, a recent Transformer model, brings a significant increase in the mean IoU score [45].

In the second step, we performed the automatic realignment of these segmented maps on the current city's topology. As the maps differ in their scale and orientation, this task consists of finding the best homography matrix that maps the historical road network to the current one. A common method to find the homography between two images consists of first finding and describing keypoints. Similar keypoints are then matched, and the resulting pairs can be used to compute the perspective transformation that warps one image onto another one [22]. The second research question of this work therefore seeks to assess whether feature detection and matching can be efficiently used to automatically realign historical maps. In particular, we compared the performances of two algorithms, SIFT and Superpoint+Superglue on a manually georeferenced dataset of Parisian maps. Our results showed that Superpoint+Superglue correctly realigns 20% more maps than SIFT, but with a 16.48m higher average error. We also developed a fine-tuning method based on SIFT to reduce the error of a first coarse realignment. This approach improved the performance of Superpoint+Superglue by 4.85m. Finally, we selected 20 additional

maps of five different cities to qualitatively evaluate the genericity of our realignment methods. Achieving a similar accuracy as for the Parisian dataset, we believe our methods to be able to realign maps of any European city.

For the final step of our study, we used the methods described above to create a dataset of realigned Parisian road networks, spanning 189 years of the XVIIIth, XIXth, and XXth centuries. While this work has no intention of providing deep and precise historical analysis, as a proof-of-concept, we present computational methods that provide entry points to answer three research questions related to the evolution of the city of Paris. First, did the transformations of the city of Paris allow for a shortening of the travel distances inside the city? To answer this question, we proposed a measure representing the length of the shortest paths linking five arbitrary locations. Second, did the Haussmannian transformations increase the number of roads at intersections ? We used the average number of connections at each crossing to provide insight into this question. Third, in what time frame did the Parisian road network develop the most ? To determine a candidate period, we plotted the evolution of the number of intersections, as well as the length of the road network, as a function of the distance from Notre-Dame.

Chapter 2

Previous work

2.1 Semantic segmentation of historical maps

With the traditional techniques for semantic segmentation, each situation requires a specific and handcrafted method. In the case of historical maps, approaches taking advantage of the colour [8, 21, 25, 35] and of the texture [4, 42] of the maps were developed. However, while these methods find success on their specific datasets, they do not generalise well to maps of different styles. It is only with the advances in Deep Learning that the performance and genericity of semantic segmentation tasks improved significantly. In particular, CNNs, which were developed to solve tasks on images, allowed training neural networks for semantic segmentation. Petitpierre et al. extensively studied this tool to segment historical city maps [33]. As their work also provides an annotated dataset that serves as a benchmark for this task, we use it as the starting point for this step.

In their work, Petitpierre et al. proposed a division of city maps into five classes that were used to label 330 map patches of Paris. This dataset, which was published in open access, contains a training set and a testing set. The authors trained and compared the performance of different CNN architectures using this dataset. Among the UNet [37], PSPNet [48] and UperNet [43] that they assessed, they found that the UNet coupled with a ResNet [19] encoder, implemented in the PyTorch [30] version of dhSegment [2], was not only the most effective architecture on the test set, but was also the one with the fastest training time. This model achieved a 63.63% mean IoU in the five class case, and a 89.05% mean IoU in the three class case, where three of the five classes are merged into one. Not only did their method surpass the performance of previous work, but it was more generic as it worked well on diverse styles of maps.

2.2 Map realignment

There are numerous methods for realigning historical maps on today's topology. Indeed, since the scale of the maps, the type of landscape they represent, and the elements they depict heavily influence the realignment process, research based on different kinds of maps tends to approach this task from completely different angles with distinct techniques.

Duan et al. introduced an algorithm in which, for each point of the vectored map, all the possible shifts allowing the vector to be aligned on the underlying map were computed [14]. Then, the most frequent shift in neighbouring points was chosen and used to translate the vectors. In later research, Duan et al. exploited the technology of reinforcement learning. In their approach, the vectored road network was considered as an agent that could move on the raster historical map [15]. This agent was rewarded by being aligned on a consistent colour. Wu et al. introduced a technique that determines the translation between two road networks by identifying the one that minimises the least-square cost. While these methods performed well in their scope, a strong limitation was that they required the two maps to be already approximately realigned, and these algorithms can only be used for fine-tuning. Moreover, a possible unwanted behaviour of the two former methods is that they allow local deformations of the road networks. Costes and Perret proposed an approach that does not have these weaknesses and that makes use of hidden Markov models [10]. Their algorithm seeks to match two spatial networks using their underlying topology.

Numerous other map realignment methods also rely on matching features from two target maps. Chen et al. came up with a solution that matches the road intersections [5]. However, their method also assumes an approximate pre-realignment and may be unfit for historical maps as the intersections may not be the same as today. Another method is to directly match the road networks. Cléry et al. matched the roads through an iterative process that refines the precision at each iteration [6], while Yu et al. used a novel matching algorithm based on the Fast Fourier Transform to match the roads [44]. In the scope of the Jadis project, Petitpierre matches features computed from the geometry of the blocks [31]. A slightly different approach consists of letting an algorithm find the most descriptive features to match. This technique was used by Luft and Schiewe to match large scale maps, where features were automatically detected on semantically segmented water bodies [23]. Zambanini took advantage of the SIFT algorithm [22] to describe and match local features to automatically realign aerial photos of the second World War [46].

2.3 Computational map analysis

One of the common approaches for computationally analysing an urban map is to consider its road network as a graph, with the intersections acting as the nodes, and the roads as the edges. Such a mathematical representation allows to compute diverse metrics that can be significant for urban historians. Crucitti et al. studied the concept of centrality to analyse networks of urban

streets [11]. This metric was extensively used in different studies to analyse the evolution of road networks [3, 20, 47]. In order to measure the expansion of a city, Barthelemy et al. also used more basic graph metrics, such as the evolution and variation of the number of nodes or edges over time [3].

The book *Paris Haussmann* [26] introduced numerous metrics to characterise urban fabrics. Amongst them, the concept of grid density, computed as the number of intersections per km^2 , indicates the fineness of the urban grid. Different measures of pedestrian accessibility allowed to identify and to visualise the most vibrant neighbourhoods. The average surface area of blocks per km^2 was also computed to "get a sense of the city's porousness" [26].

Chapter 3

Methods

3.1 Semantic segmentation of historical maps

In the first step of our proposed pipeline, the historical maps undergo a semantic segmentation process. The goal of this step was to remove unnecessary information from the maps, such as the frame or stylistic elements. In order to develop a solution that remains efficient with diverse map styles, we chose to perform the semantic segmentation with a neural network [33]. The chosen architecture was an OCRNet [45], which was trained on a corpus of Parisian map patches [32].

3.1.1 Dataset and data augmentation

The dataset used to train and test the neural network in this work was created by Petitpierre et al. [32]. It contains 330 map patches of 1000x1000 pixels that were randomly extracted from a corpus of 1'500 maps representing the city of Paris. The majority of these maps were published between 1800 and 1950 and their scale varies between 1:25'000 and 1:2'000 [32]. The map patches were manually labelled using an annotation ontology that contains five classes : frame, road network, block, water, and non-built. As the differentiation between block, water, and non-built may not be of interest for all tasks, Petitpierre et al. also proposed to merge them into a single class, called *blocks*, that comprises "all cartographic elements that do not belong to the road network" [32]. The dataset was divided into a training set containing 300 annotated map patches, and a testing set containing 30 annotated map patches.

To avoid overfitting, the training samples undergo a series of transformations at each epoch. The pipeline of transformations is as follows:

1. Normalizing the colours. The mean and standard deviation values used in this step were computed over all the patches.

2. Scaling with a random ratio comprised between 0.75 and 1.25.
3. Randomly cropping to a size of 512x512 pixels.
4. Flipping horizontally with a probability of 0.25.
5. Sequentially applying photometric distortions. The map patch received a series of distortions in the BGR and HSV colour spaces. Each of them was applied with a probability of 0.5. They comprised adding random brightness, contrast, saturation, or hue.

3.1.2 Network architecture and configuration

The model of deep network we chose is an Object-Contextual Representations network (OCRNet) [45] with a High-Resolution network (HRNet) [41] backbone. Previous deep models for semantic segmentation exploited the context of a given pixel to predict its class. This context was defined as the properties of a set of surrounding pixels. OCRNet’s primary driving theory is that, to describe a given pixel, neighbouring ones may not be the most appropriate. Their idea was that pixels belonging to the same class would constitute a better context. To implement this solution, OCRNet was divided into three main steps. First, a coarse segmentation is performed by a deep network (in our case, an HRNet). Second, an estimated representation of each class is computed by aggregating the representation of their pixels, as predicted by the backbone. Last, the representation of each pixel is augmented with the context of its class, called the *object-context representation*. This data is then used to train a Transformer encoder-decoder network that predicts the final segmentation.

We used the implementation of OCRNet by OpenMMLab [7]. This network was trained with a batch size of 4 and was initially configured with a learning rate of 0.001, exponentially decaying at a rate of 0.9. The model was optimised with the Stochastic Gradient Descent algorithm with a momentum of 0.9 and an L2 penalty was applied to the weights with a factor of 0.0005. The network was first trained for the five class case, and it was initialised with weights pre-trained on the Cityscapes dataset [9]. The resulting weights from this task were then used as initialising weights for the three class case. For both networks, we computed the IoU score for each class across the 30 samples of the test set and compared our results to those of Petitpierre et al. [33].

3.2 Map realignment

In our approach to realigning historical city maps onto today’s topology, we first extracted the road networks using the semantic segmentation step and skeletonized them [39] to remove the roads’ width. The current road network, called *the anchor*, was obtained from OpenStreetMap and was also skeletonized. Then, we used an algorithm to describe local features in both road

networks. These features were then matched and the pairs were used to determine the homography matrix that maps the historical road network on its anchor. Lastly, we fine-tuned this realignment by performing this step a second time, but by defining the keypoints ourselves on the road intersections and by forcing them to match close keypoints.

A quantitative evaluation of the realignments produced by our method was performed on a manually georeferenced corpus of Parisian maps. A qualitative assessment was also performed on 5 different cities to evaluate the genericity of our approach.

3.2.1 Local feature detection, description, and matching

We tested two different algorithms to detect and describe local features on each map. The first one was the Scale-Invariant Feature Transform (SIFT) [22]. SIFT first finds locations of interest by identifying the extrema after having applied a difference-of-Gaussian function. For each of these points, a model is fitted to determine its scale. Then, keypoints are selected from these candidate locations by measures of their stability. An orientation is also assigned to each keypoint based on the local image gradient directions. Finally, descriptors of keypoints are computed using the histogram of the image gradient at the selected scale around the keypoint. SIFT descriptors can then be matched between the two images using a nearest neighbour algorithm. Since many of the obtained matches may be incorrect, the ones whose second closest neighbour is not far enough are discarded. The idea is that, "for false matches, there will likely be a number of other false matches within similar distances due to the high dimensionality of the feature space" [22]. The main advantage of these descriptors is that they are invariant to translation, scale and rotation. Initial tests showed that correctly calibrating two hyperparameters was crucial to obtaining good results. First, it was important to apply a Gaussian blur to the road networks to enhance the quality of the feature detection. Second, while SIFT is scale invariant, we still noticed that the results of the algorithm were highly sensitive to the size of the image. Therefore, we also shrank the road networks, and a simple grid search allowed us to find the optimal configuration for these hyperparameters.

The second tested algorithm was a combination of Superpoint [12], for local feature detection and description, and Superglue [38] for feature matching. Superpoint uses a fully-convolutional encoder-decoder deep model to both find and describe keypoints. This network was trained on a synthetic dataset of 2D shapes and performs well on any kind of image. Superglue is a graph neural network that is trained to match two sets of local features. The downside of this matcher is that the available model was trained on images of outdoor scenery. Training Superglue on our own dataset of road networks would have been too demanding for the scope of this project, since the training code was not released to the public. However, as the complexity of our road networks is lower than that of outdoor scenery, we believed that these pre-trained weights were still appropriate for our task. Using these two algorithms requires setting two hyperparameters. First, the Non Maximum Suppression (NMS) [36] radius needs to be defined for the Superpoints. A small radius means that multiple Superpoints can coexist close to each other, while with a

greater radius, they may be merged into a single Superpoint. Then, the size of all input images has to be set to the same value. However, as the algorithm performs badly for large images [38], and as we were limited by the memory of our GPU, we had to limit the largest side to 2000px. Again, these hyperparameters were optimised by a grid search.

For both methods, the final matched pairs were finally used to determine the homography matrix that warps the historical road network on the anchor. To this end, we used OpenCV's `findHomography` function [29], with the RANSAC [16] algorithm to filter out erroneous matches.

3.2.2 Homography fine-tuning

Even in the case where our methods of local feature detection, description, and matching perform well, there is still room to enhance the realignments. Indeed, since keypoints can cover a relatively large area, two of them may match the approximate same location, but be centred on different points of the map. This may result in a first homography yielding a slightly shifted realignment. To resolve this potential issue, we applied SIFT to find a second homography, analogously as described in section 3.2. However, instead of letting the algorithm detect the points of interest itself, we created keypoints ourselves at each road intersection. Then, we employed the first coarse realignment to force keypoints of the historical maps to match close keypoints on the anchor. A grid search was used to find the optimal keypoint size and maximal distance to find matches. We finally used these new pairs to find a second homography matrix that, presumably, corrected the potential shifts in the realignment.

3.2.3 Quantitative and qualitative evaluation

To evaluate the correctness of the realignments produced by our methods, we first randomly selected fifty maps from a corpus of Parisian maps of the Bibliothèque Nationale de France available on Gallica [17]. Then, we manually georeferenced them using a custom software. This georeferencing was made by selecting five homologous points on each map. For consistency, we chose the approximate same landmarks as homologous points for each map. The selected landmarks were the following:

- Place de Vendôme : one of the corner of the place
- Places des Vosges / Place Royale : one of the corner of the place
- Paris Observatory : one of the corner of the building
- Pont Neuf : northern end of the bridge
- Notre-Dame de Paris : northern west corner of the building. When the monument itself was not precisely represented, one of the neighbouring bridges (Pont de l'Archevêché or

Pont Saint-Louis) was chosen.

To determine whether a homography matrix correctly realigned one of the maps, we used it to map the positions of these five historical landmarks on the anchor. For each of them, we computed their euclidean distance to the actual location that was manually annotated. If the average of these distances was below 75 meters, we considered the map to be correctly realigned. Above this threshold, the homography matrix was considered to be incorrect and the map imprecisely realigned.

Since this quantitative evaluation was uniquely performed on Parisian maps, we carried out a qualitative assessment to verify that our methods were not overfit to the selected corpus and remained generic. To this end, we selected four additional maps for five cities : Toulouse, Lyon, Strasbourg, Montpellier, and Rome. We applied the semantic segmentation and realignment steps to each of these maps and simply visually evaluated the results to determine if they were correctly realigned or not.

3.3 Computational map analysis

We created a dataset of 346 Parisian maps using our two-step pipeline. Then, as a proof-of-concept that this data can be used to study urban history, we asked three basic historical research questions for which we provided a computational analysis. Since these questions study the evolution of different features over time, we grouped the maps of our dataset into ten time periods that have an equal amount of data. While this division means that the time periods do not have the same length, it ensures that there is enough data in each group to obtain significant results.

3.3.1 Dataset

To create the dataset, we first downloaded a corpus of 750 Parisian maps from Gallica [17]. All maps were manually inspected to remove those that did not represent the full city. We then used our methods to extract the road networks and to automatically realign them. We finally checked all outputs to discard the ones with unsatisfactory semantic segmentation or realignment. In the end, we obtained a dataset of 346 realigned Parisian road networks. The publication dates of the maps in this corpus are comprised between 1760 and 1949. As it can be observed in Figure 3.1, most of these maps were published during the XIXth century. This time frame is ideal as it allows us to observe the topology of Paris before, during, and after the Haussmannian transformations. However, because of the scarcity of maps published in the XVIIIth and the XXth centuries, we had to discard the first and last divisions of the dataset, as they spanned too many years to be meaningful. A small subset of our dataset can be visualised in Figure 3.2.

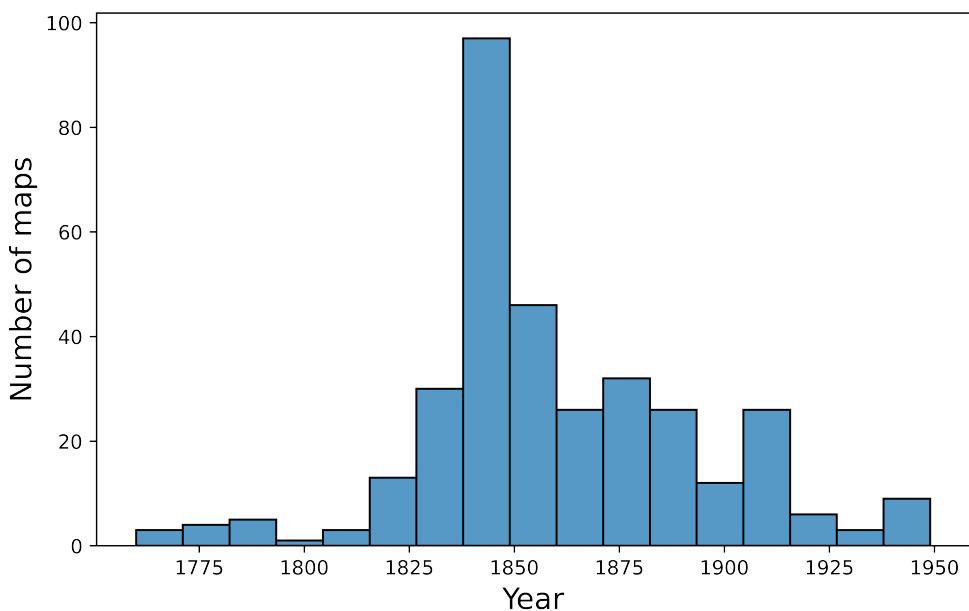


Figure 3.1: Distribution of the publication dates of the 346 maps of Paris constituting our dataset for the computational analysis.

3.3.2 Travel distances

Our first historical research question was "Did the transformations of the city of Paris allow for a shortening of the travel distances inside the city ?". Our hypothesis is that, over more than a century, the transformations of the urban fabric were planned to speed up the navigation within the city.

To conduct this analysis, we converted each road network skeleton into a bi-directional weighted graph, with edges representing roads, and nodes representing intersections. The weight of an edge is defined as the euclidean distance between the two intersections it connects. We removed intersections with only one road shorter than 30m connected, since these dead-ends were probably the product of segmentation noise. We then picked five arbitrary locations inside Paris. The first location was selected at the centre of the city, near Notre-Dame. The four others were selected at approximately 1400 metres north-east, north-west, south-east, and south-west of the cathedral. For each graph, we selected the nodes that were the closest to these locations. If no node was found at less than 50m from any of these locations, the graph was discarded. Then, we used Dijkstra's algorithm [13] to find the ten shortest paths that link each pair of locations. Finally, the distances of each of these paths were added up and this value was used as our travel distance metric for a given map.

However, this metric may not be accurate for every map as it is heavily reliant on the quality of the semantic segmentation. Indeed, if our OCRNet cuts a road in two, the shortest path may be wrongly extended. Moreover, a map drawn with fewer details may have not represented short



Figure 3.2: Representation of a sample of the dataset we created, where 30 randomly selected road networks are superimposed.

roads that would have reduced the length of the shortest path. To resolve these issues, for each time frame, we only considered the graphs with a travel distance metric below the median. We finally used a Mann-Kendall test with a significance level of 0.10 on these values to evaluate the general trend of this metric over time.

3.3.3 Number of roads at intersections

For our second historical research question, "Did the Hausmannian transformations increase the number of roads at intersections", we hypothesised that there was a higher average number of connections at intersections after the Hausmannian works. To perform this analysis, we transformed the road networks into graphs, as explained in Figure 3.3.1. By the definition of these graphs, the nodes' degrees represent the number of roads connected to them. Therefore, we computed the mean node's degree of each graph in an area of 1km around Notre-Dame, and averaged this value for each time period. Finally, we also used a Mann-Kendall test with a significance level of 0.10 to evaluate the trend over time.

3.3.4 Rythm of expansion of the Parisian road network

Our third historical research question asked during which time period the Parisian road network developed and expanded the most. Our method was directly taken from the study of Barthelemy et al. [3]. First, we selected a reference point at the city centre, in our case Notre-Dame. Then, we counted the number of intersections inside circles of radius r centred on this reference point, where r ranges from 200m to 5900m with a 300m interval. Plotting these numbers of intersections, as a function of the distance from the city centre for each time frame allows us to visualise the expansion rate of the city of Paris over time. Furthermore, by subtracting these values from those of the following time frame, and plotting these differences as a function of r , we can observe which time frame had the most development. We then performed a similar analysis, where instead of counting the number of intersections, we summed the length of the road network. We plotted these values, as well as their differences, as described above, to obtain a second measure of the expansion of the Parisian road network.

Chapter 4

Results and discussion

4.1 Semantic segmentation of historical maps

The training phase of the OCRNet for the semantic segmentation with five classes ended after 15'000 iterations and it reached its best performance on the test set after the 10'800th iteration. Examples of the segmentation produced by this network on map patches from the test set can be observed in Figure 4.1. These examples show that, while the detection of the road network is rather clean, the model has trouble differentiating the blocks, water, and non-built, which is also the reason these three classes are merged to train the next model.

The training phase for the three class case was then initialised with the weights previously obtained for the five classes. This greatly helped the network to converge faster. We interrupted it after 4500 iterations and it attained its best performance on the test set after the 1800th iteration. Examples of the results obtained with this network can be observed in Figure 4.2.

Table 4.1 provides the detailed measurements of the performance of both networks. The metric reported is the Intersection over Union of each class, as well as its average. These values are compared with those of the UNet obtained by Petitpierre et al. [33]. However, it is worth noting that their IoU for the segmentation of the frame is not comparable with ours since this class was pre-segmented in their work. Therefore, to have a common ground for comparison, we recomputed the mean IoU without taking the frame into account. The analysis of this table reveals that our OCRNet outperforms the UNet in multiple areas. Most notably, we observe that the mean IoU of the OCRNet is approximately 12% higher than that of the UNet for the five class case. For three classes, while the difference is not as large, there is still a 3% increase in mean IoU compared to the UNet. We can therefore assess that our deep network significantly improves the quality of the semantic segmentation for historical city maps. The table also indicates that the OCRNet for the three class case performs the best on the road networks. Since we are looking to extract this class for the next step, we will use this network in the rest of this project.

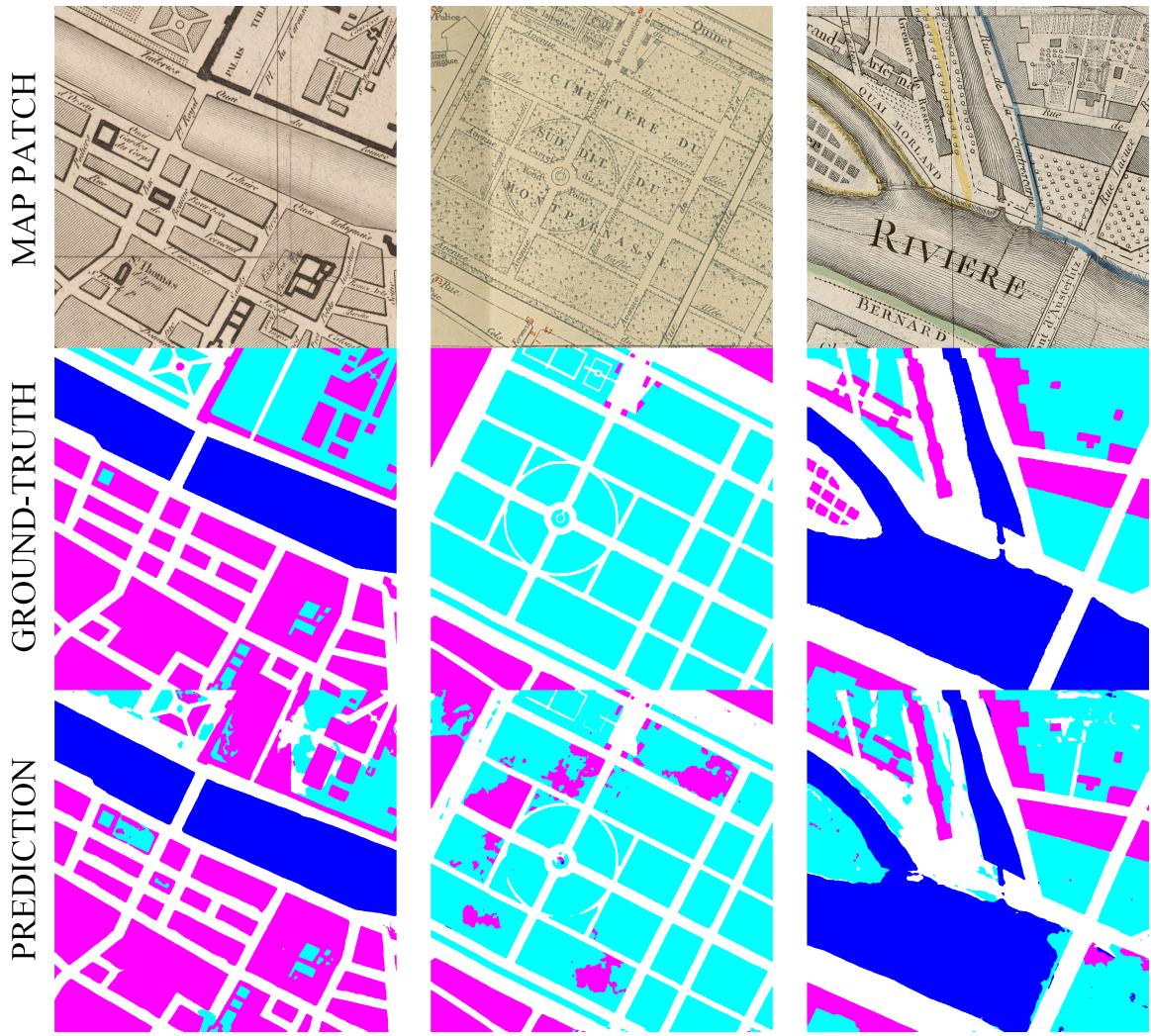


Figure 4.1: Comparisons of map patches from the test set, with their ground-truth, and their OCRNet prediction. This figure shows the five class case, with the road network (white), blocks (magenta), non-built (cyan), water (blue), and frame (black, non-present here).

4.2 Map realignment

4.2.1 SIFT

With the optimal hyperparameters configuration (Gaussian kernel of 17x17px and a scaling factor of 0.25), SIFT was able to correctly realign 29 out of our 50 maps. Figure 4.3 displays the density estimation of the distribution of euclidean distances between each pair of homologous points from our dataset after being realigned by this method. We can observe two distinct peaks in this plot : one in the magnitude of the tens of metres, and the other one around the thousands of metres. Such a plot shows that, when the realignment is considered correct, its precision is relatively high, with an average error of 11.91 meters, but when it is not, the homography



Figure 4.2: Comparisons of map patches from the test set, with their ground-truth, and their OCRNet prediction. This figure shows the three class case, with the road network (white), blocks (magenta), and frame (black).

matrix yields a completely wrong result. Despite trying to further develop this method, we were not able to obtain better realignments for these failing cases. While inaccurate semantic segmentation of the maps is a factor that causes some of the unsuccessful instances, we were not able to determine the reason why SIFT was unable to perform well for the others. Analysis of the dataset's metadata showed no correlation of success related to the scale or publication date of the maps. Therefore, our conclusion is that our realignment method based on SIFT is bound to fail in approximately 40% of the cases. Nevertheless, this method still yields accurate results for the remaining 60%.

Class	OCRNet 5 classes	UNet 5 classes	OCRNet 3 classes	UNet 3 classes
Mean w/o frame	0.6624	0.5426	0.8697	0.8380
Frame	0.9194	/	0.9207	/
Blocks	0.7459	0.5657	0.9210	0.9181
Road Network	0.7993	0.7132	0.8184	0.7580
Water	0.5175	0.4682	/	/
Non-built	0.5870	0.4235	/	/

Table 4.1: Detailed report of the IoU metric of our OCRNet on the test set and its comparison with the results of Petitpierre et al. [32]

4.2.2 Superpoint+Superglue

This method successfully realigned 39 out of 50 of our maps when the optimal hyperparameters configuration (NMS radius of 5px and the maximal side of the image set to 1400px) was used. The density of the distribution of euclidean distances between each pair of homologous points, after being realigned by this method, is displayed in Figure 4.3. Similarly as for the SIFT method, this plot shows two peaks. This indicates that, again, the realignment of the maps that were considered imprecise are completely wrong. Additionally, we notice that these two peaks are shifted to the right in comparison to SIFT’s. Indeed, the average error of this method for the correctly realigned maps is of 28.39m, which is 16.48m more than for SIFT. Therefore, it appears that Superpoint+Superglue yields less precise realignments than SIFT. This imprecision can be explained by the smaller size of the images that can be used by this method. However, this approach allowed us to find a correct realignment for 11 more maps than SIFT. Since the next step of the pipeline seeks to correct the realignment inaccuracies, and since we opted for a method that did not discard too much data, we decided to use Superpoint+Superglue for the coarse realignment to fine-tune in subsection 4.2.2, as well as for the creation of our dataset described in subsection 3.3.1.

4.2.3 Homography fine-tuning

With the optimal hyperparameters configuration (keypoints of 19px of radius that look for matches in a radius of 175px), we were able to reduce the average realignment error obtained with Superpoint+Superglue by 4.85m, bringing it down to 23.54m. Although this result demonstrates the effectiveness of the fine-tuning method, it is unfortunate that the improvement is not more evident. The strategies outlined in section 2.1 that make use of a first coarse realignment may be interesting to test in future studies.

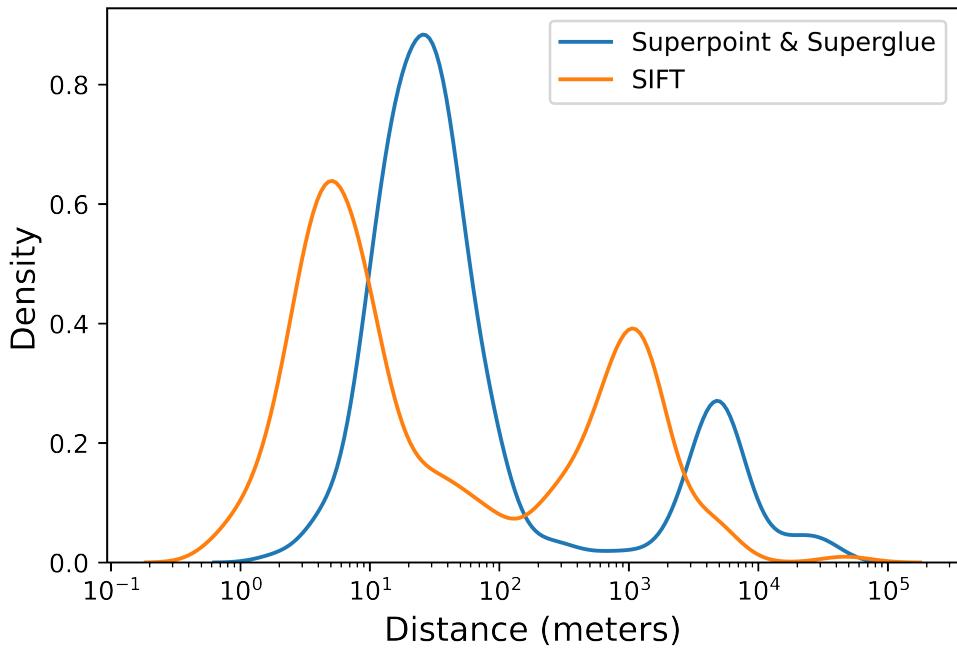


Figure 4.3: Kernel density estimations of the distributions of Euclidean distances of each pair of homologous points. Results obtained by our methods based on SIFT and on Superpoint+Superglue are displayed.

4.2.4 Qualitative assessment

After extracting the road networks from our 20 test maps, executing the Superpoint+Superglue automatic realignment on each of them on their respective anchor, and fine-tuning them with SIFT, the number of maps that we visually evaluated to be appropriately realigned is shown in Table 4.2. All correctly realigned maps can be observed in Figure A.3. As these results show a similar accuracy to the one obtained on the dataset of Paris, we can conclude that our automatic realignment is generic enough to be applied to other European cities with an accuracy close to 75%.

City	Correct realignments
Rome	4/4
Strasbourg	3/4
Montpellier	3/4
Lyon	1/4
Toulouse	3/4
Total	14/20

Table 4.2: Number of maps per city that we visually assessed were correctly realigned by the Superpoint+Superglue method.

4.3 Computational map analysis

4.3.1 Travel distances

The results of the analysis of travel distances in Paris can be observed in Figure 4.4. A first visual inspection of the evolution of the mean travel distances seems to reveal that there is indeed a decrease across time frames. The mean distance required to perform the 10 trips between 1840 and 1844 is 14.68km, whereas it is only 13.91km between 1889 and 1907. Moreover, the Mann-Kendall test indicates that there is indeed a decreasing trend, with a p -value of 0.06. However, further work is recommended to improve this analysis. For instance, the shortest paths of each time frame could be superimposed to discover which new roads allowed shorter trips. Such an analysis may reveal whether the decreasing travel distance is due, for instance, to the construction of bridges or to the opening of new road axes.

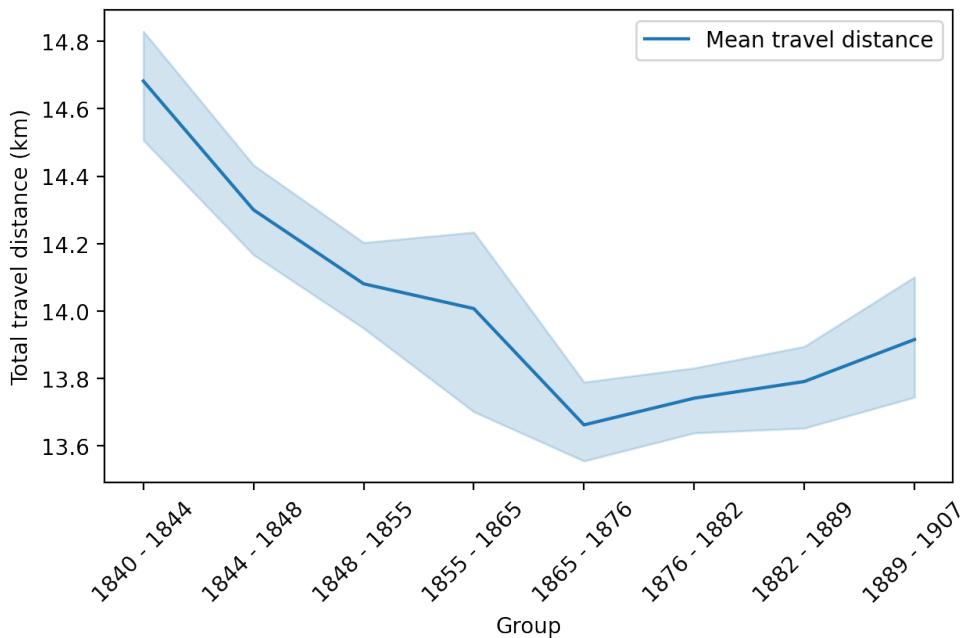


Figure 4.4: Mean travel distance for each time frame, with the semitransparent area depicting the standard deviation.

4.3.2 Number of roads at intersections

Figure 4.5 shows the results of the analysis of the number of connections at intersections. This plot immediately reveals that there is no augmentation of this value, even less during the period of the Haussmannian transformations. Notably, since the y-axis spans a short range of values, all the variations shown by this graph are minimal. The Mann-Kendall test reveals that there is no trend, with a p -value of 0.11. Therefore, this analysis completely refutes our initial hypothesis,

and it rather shows that the overall number of connections at intersections remained stable.

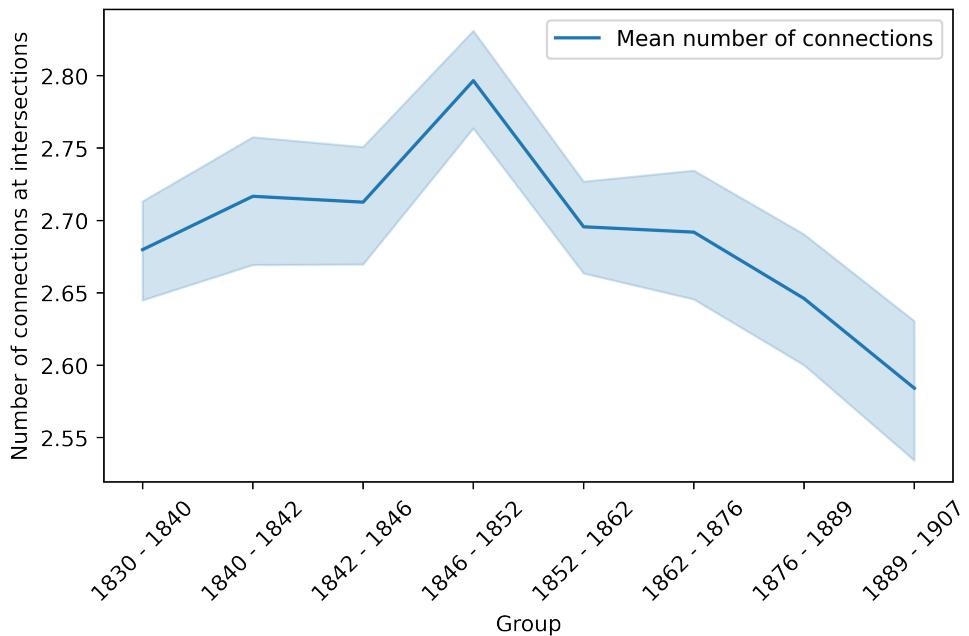


Figure 4.5: Mean number of connections at intersections for each time frame, with the semitransparent area depicting the standard deviation.

4.3.3 Rythm of expansion of the Parisian road network

The analysis of the expansion of the number of intersections can be observed in Figure 4.6. The above plot demonstrates that across the time period covered by our dataset, the number of intersections up to 2 km from Notre-Dame hardly changed. This observation may be due to the city centre's advanced development prior to 1830. There was no space to construct new roads because the majority of the region was already populated. Then, we notice that for earlier time periods until 1846, the number of intersections plateaus at around 4km from the city centre. It only starts to rise during the time period from 1846 to 1852, with more than 1000 new intersections overall, as shown in the below plot. The pace of creation of new road intersections speeds up from this moment, with the exception of the period from 1877 to 1889. The latest time period is also the one with the greatest expansion of road intersections.

Figure 4.7 shows the analysis of the expansion of the length of the Parisian road network. As it can be expected, these results are similar to the ones of Figure 4.6. Indeed, the length of the road network does not increase over the whole time period in the 2km area around Notre-Dame. We notice a difference for the 1842-1846 period, which shows a relative great augmentation of road network length compared to the number of new intersections. Such an observation indicates that longer roads were built during that period. Conversely, the last time frame, which had the greater expansion of intersections, shows a relatively small augmentation in road network length.

This result may indicate that fewer main and long road axes were built during that time.

Overall, these results indicate that the Parisian road network developed the most in the second part of the XIXth century, and kept expanding at a great pace during the beginning of the XXth century. It is interesting however to observe a pause in this expansion in the period from 1877 to 1889. It is necessary to conduct additional historical research to identify any potential causes of this pause.

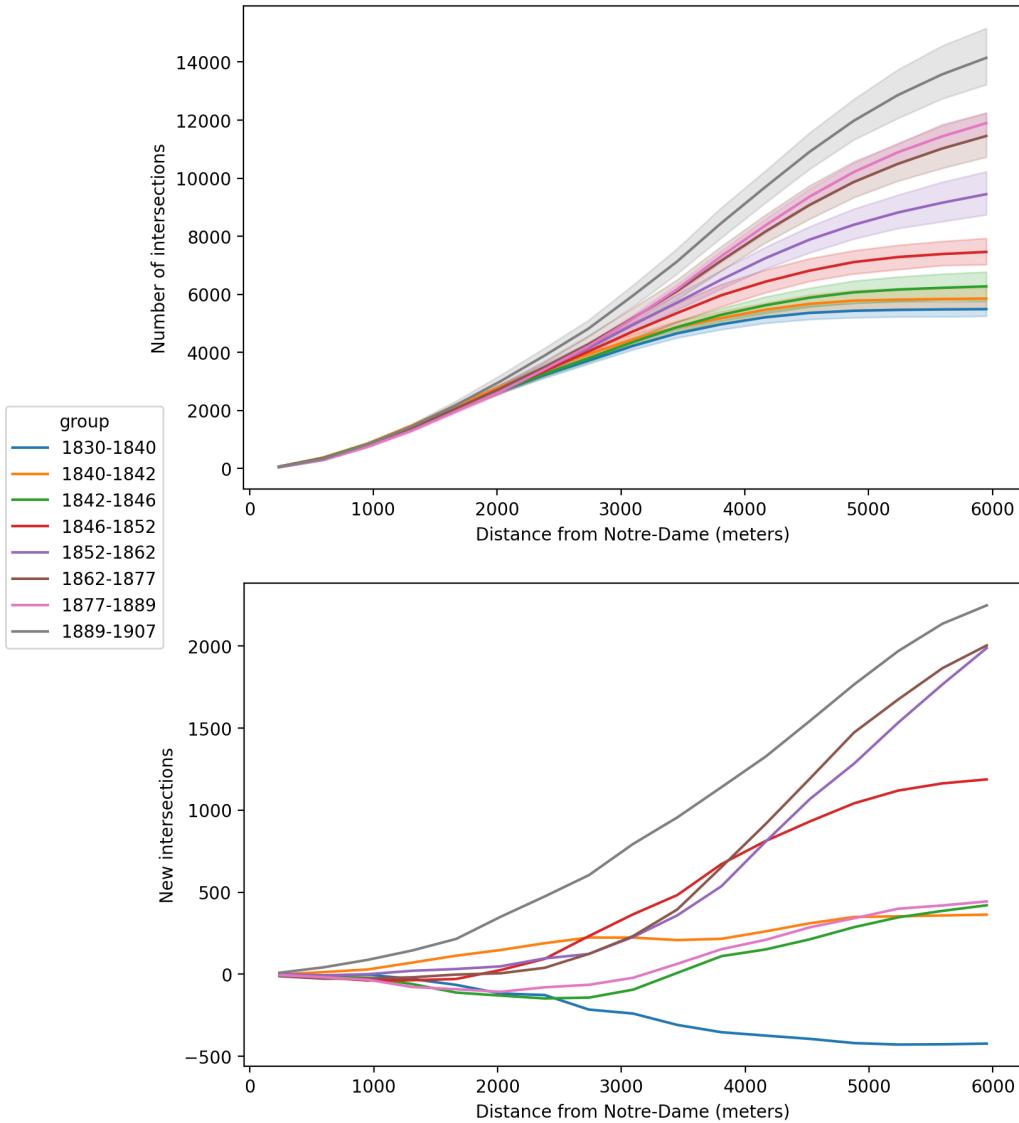


Figure 4.6: Above : number of intersections in Paris' road networks represented as a function of the distance from Notre-Dame. Below : number of new intersections created from the previous time period.

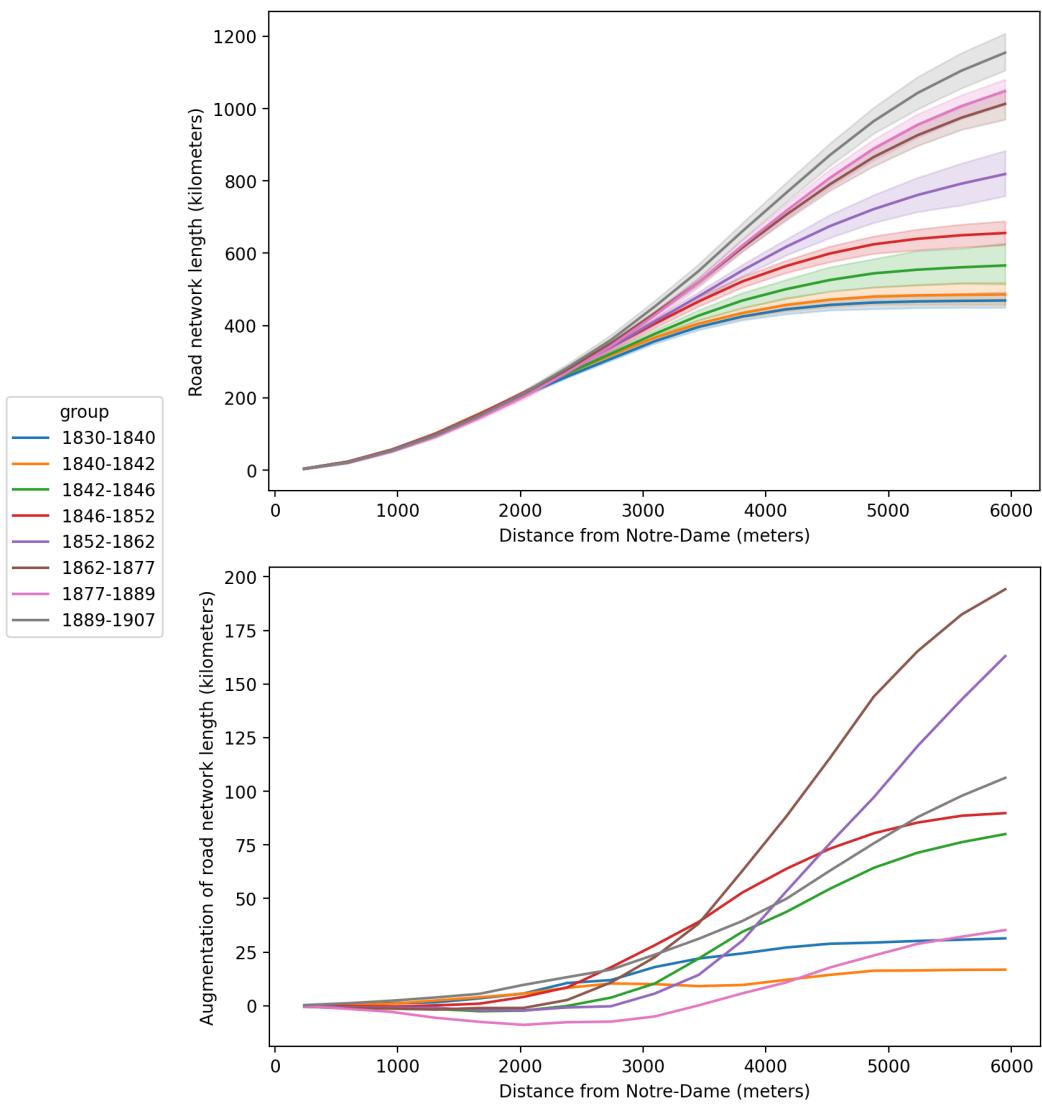


Figure 4.7: Above : average length of the Parisian road network represented as a function of the distance from Notre-Dame. Below : augmentation of the length of new roads created from the previous time period.

Chapter 5

Conclusion

In this work, we proposed a two-step pipeline to process historical city maps with the goal of enabling computational map analysis. In the first phase, we extracted the road networks using semantic segmentation. Our method was based on OCRNet, a deep neural network that we trained with a pre-existing dataset of map patches representing the city of Paris [32]. Our approach allowed an increase in the state-of-the-art mean IoU results by approximately 12%. Furthermore, we obtained a maximal IoU score of 81.84% for the road networks. This result allowed us to extract this class with good precision, and therefore to obtain an accurate skeleton of the road network for any historical city map.

In the second step, this skeleton was used to georeference the map by realigning it on its anchor. Our method was based on local feature detection and matching, with the idea that homologous points in historical and current road networks could be paired to find the homography that allows warping the historical map on today's topology. We employed two algorithms: SIFT and Superpoint+Superglue. Both techniques worked well on a subset of maps from our evaluation dataset, but completely failed on the others. However, while Superpoint+Superglue showed less precise realignments than SIFT, it succeeded on 78% of the test maps, compared to the 58% success rate of SIFT. With a final fine-tuning step, Superpoint+Superglue achieved an average precision of 24 metres for the correctly realigned maps. We therefore chose it as our preferred method, trading some accuracy to discard as few maps as possible.

Finally, we used our pipeline to create a dataset of 346 realigned road networks of Paris. We also proposed simple research questions to showcase how computational methods could be used on this type of data to study urban history. We first showed a method taking full advantage of converting the road networks into graphs, where we computed the shortest paths between different locations within the city. This metric allowed us to demonstrate that the travel distances within the city decreased over time. Then we also used the graphs to count the average number of connections at each intersection. This analysis refuted our hypothesis that this value increased with the Haussmannian works. Our last method compared the growth of different metrics in

an area expanding from the city centre to the whole city. This approach revealed that, within the time frame represented by our dataset, the periods that had the greatest development were the second half of the XIXth century, as well as the beginning of the XXth century. With these methods, we provided entry points for deeper computational urban historical analyses.

Reflecting on this work, we identify numerous limitations and areas that could be improved through additional research. In particular, the automatic map realignment methods we developed are not completely satisfactory. First, even our best approach fails to realign approximately one quarter of the maps. Moreover, the maps need to be first sorted by the city they represent, and need to have a scale close to their anchor. Due to these limitations, manual work is still required throughout the process and our method is not applicable to any kind of city maps. We anticipate that future research in this area will result in a solution that is even more general, allowing any type of map to be input without any prior knowledge.

Bibliography

- [1] Kurt Anschuetz, Richard Wilshusen, and Scheick Cherie. “An Archaeology of Landscapes: Perspectives and Directions”. In: *Journal of Archaeological Research* 9.2 (2001), pp. 157–211.
- [2] Sofia Ares Oliveira, Benoit Seguin, and Frederic Kaplan. “dhSegment: A generic deep-learning approach for document segmentation”. In: *Frontiers in Handwriting Recognition (ICFHR), 2018 16th International Conference on*. IEEE. 2018, pp. 7–12.
- [3] Marc Barthelemy, Patricia Bordin, Henri Berestycki, and Gribaudi Maurizio. “Self-organization versus top-down planning in the evolution of a city (vol 3, 2153, 2013)”. In: *Scientific Reports* 3 (Nov. 2013). DOI: 10.1038/srep02415.
- [4] Regine Bruegelmann. “Recognition of hatched cartographic patterns”. In: *18th Congress, International Society for Photogrammetry and Remote Sensing, Vienna, Austria 1996*. Ed.: K. Kraus. Pt. B3. S. 82-87. (*International archives of photogrammetry and remote sensing*. 31,B3.) 1996.
- [5] Ching-Chien Chen, Craig A. Knoblock, Cyrus Shahabi, Yao-Yi Chiang, and Snehal Thakkar. “Automatically and Accurately Conflating Orthoimagery and Street Maps”. In: *Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems*. GIS ’04. Washington DC, USA: Association for Computing Machinery, 2004, pp. 47–56. ISBN: 1581139799. DOI: 10.1145/1032222.1032231. URL: <https://doi.org/10.1145/1032222.1032231>.
- [6] I. Cléri, M. Pierrot-Deseilligny, and B. Vallet. “Automatic Georeferencing of a Heritage of old analog aerial Photographs”. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3 (2014), pp. 33–40. DOI: 10.5194/isprsaannals-II-3-33-2014. URL: <https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/II-3/33/2014/>.
- [7] MM Segmentation Contributors. *MM Segmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark*. <https://github.com/open-mmlab/mmsegmentation>. 2020.
- [8] Amândio Cordeiro and Pedro Pina. *Colour Map Object Separation*. Jan. 2006.
- [9] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. “The Cityscapes Dataset for Semantic Urban Scene Understanding”. In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.

- [10] Benoît Costes and Julien Perret. “A hidden Markov model for matching spatial networks”. In: *Journal of spatial information science* 18 (2019), pp. 57–89.
- [11] Crucitti, Paolo and Latora, Vito and Porta, Sergio. “Centrality in networks of urban streets”. In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 16.1 (2006), p. 015113. DOI: 10.1063/1.2150162. eprint: <https://doi.org/10.1063/1.2150162>. URL: <https://doi.org/10.1063/1.2150162>.
- [12] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. “SuperPoint: Self-Supervised Interest Point Detection and Description”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. June 2018.
- [13] Edsger W Dijkstra. “A note on two problems in connexion with graphs”. In: *Numerische mathematik* 1.1 (1959), pp. 269–271.
- [14] Weiwei Duan, Yao-Yi Chiang, Craig A. Knoblock, Vinil Jain, Dan Feldman, Johannes H. Uhl, and Stefan Leyk. “Automatic Alignment of Geographic Features in Contemporary Vector Data and Historical Maps”. In: *Proceedings of the 1st Workshop on Artificial Intelligence and Deep Learning for Geographic Knowledge Discovery*. GeoAI ’17. Los Angeles, California: Association for Computing Machinery, 2017, pp. 45–54. ISBN: 9781450354981. DOI: 10.1145/3149808.3149816. URL: <https://doi.org/10.1145/3149808.3149816>.
- [15] Weiwei Duan, Yao-Yi Chiang, Stefan Leyk, Johannes H. Uhl, and Craig A. Knoblock. “Automatic alignment of contemporary vector data and georeferenced historical maps using reinforcement learning”. In: *International Journal of Geographical Information Science* 34.4 (2020), pp. 824–849. DOI: 10.1080/13658816.2019.1698742. eprint: <https://doi.org/10.1080/13658816.2019.1698742>. URL: <https://doi.org/10.1080/13658816.2019.1698742>.
- [16] Martin A. Fischler and Robert C. Bolles. “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”. In: *Commun. ACM* 24.6 (June 1981), pp. 381–395. ISSN: 0001-0782. DOI: 10.1145/358669.358692. URL: <https://doi.org/10.1145/358669.358692>.
- [17] Gallica. [Online; accessed 14-August-2022]. URL: <https://gallica.bnf.fr/accueil/fr/content/accueil-fr?mode=desktop>.
- [18] Bernard Gauthiez. *The history of urban mapping in XVIe-XVIIIté centuries, appearance and development of the representation modes*. July 2006.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. *Deep Residual Learning for Image Recognition*. 2015. DOI: 10.48550/ARXIV.1512.03385. URL: <https://arxiv.org/abs/1512.03385>.
- [20] Alec Kirkley, Hugo Barbosa, Marc Barthelemy, and Gourab Ghoshal. “From the betweenness centrality in street networks to structural invariants in random planar graphs”. In: *Nature Communications* 9.1 (June 2018). DOI: 10.1038/s41467-018-04978-z. URL: <https://doi.org/10.1038%2Fs41467-018-04978-z>.

- [21] Stefan Leyk and Ruedi Boesch. “Colors of the past: Color image segmentation in historical topographic maps based on homogeneity”. In: *GeoInformatica* 14 (Jan. 2010), pp. 1–21. DOI: 10.1007/s10707-008-0074-z.
- [22] David G. Lowe. “Sift-the scale invariant feature transform”. In: *Int. J* 2.91-110 (2004), p. 2.
- [23] Luft, Jonas and Schiewe, Jochen. “Automatic content-based georeferencing of historical topographic maps”. In: *Transactions in GIS* 25 (2021), pp. 2888–2906. DOI: 10.1111/tgis.12794. URL: <https://onlinelibrary.wiley.com/doi/10.1111/tgis.12794>.
- [24] Yujian Mo, Yan Wu, Xinneng Yang, Feilin Liu, and Yujun Liao. “Review the state-of-the-art technologies of semantic segmentation based on deep learning”. In: *Neurocomputing* 493 (2022), pp. 626–646. ISSN: 0925-2312. DOI: <https://doi.org/10.1016/j.neucom.2022.01.005>. URL: <https://www.sciencedirect.com/science/article/pii/S0925231222000054>.
- [25] Sebastian Muhs, Hendrik Herold, Gotthard Meinel, Dirk Burghardt, and Odette Kretschmer. “Automatic delineation of built-up area at urban block level from topographic maps”. In: *Computers, Environment and Urban Systems* 58 (July 2016), pp. 71–84. DOI: 10.1016/j.compenvurbsys.2016.04.001.
- [26] Umberto Napolitano, Benoît Jallon, and Franck Boutré. *Paris Hausmann*. Park Books, 2017.
- [27] National Library of Scotland and Archaeology Scotland. *Step by step map analysis*. [Online; accessed 10-August-2022]. 2013. URL: <https://digital.nls.uk/mapping-history/learn-about-maps/step-by-step/>.
- [28] Keiron O’Shea and Ryan Nash. “An introduction to convolutional neural networks”. In: *arXiv preprint arXiv:1511.08458* (2015).
- [29] OpenCV. *Feature Matching + Homography to find Objects*. [Online; accessed 13-August-2022]. URL: https://docs.opencv.org/3.4/d1/de0/tutorial_py_feature_homography.html.
- [30] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. “PyTorch: An Imperative Style, High-Performance Deep Learning Library”. In: *Advances in Neural Information Processing Systems* 32. Curran Associates, Inc., 2019, pp. 8024–8035. URL: <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>.
- [31] Rémi Petitpierre. *BnF-jadis/projet: v1.0-beta*. Version beta. May 2022. DOI: 10.5281/zenodo.6594483. URL: <https://doi.org/10.5281/zenodo.6594483>.
- [32] Rémi Petitpierre. *Historical City Maps Semantic Segmentation Dataset*. <https://zenodo.org/record/5497934>. 2021. DOI: 10.5281/zenodo.5497934.
- [33] Rémi Petitpierre, Frédéric Kaplan, and Isabella di Lenardo. “Generic Semantic Segmentation of Historical Maps”. In: *Proceedings of the Workshop on Computational Humanities Research (CHR 2021)* (Nov. 17–19, 2021). Amsterdam, NL: CEUR, 2021.

- [34] Branko Puceković. "Leonardo da Vinci and his contributions to cartography". In: *Kartografija i geoinformacije (Cartography and Geoinformation)* 12.20 (2013), pp. 34–52.
- [35] Romain Raveaux, Jean-Christophe Burie, and Jean-Marc Ogier. "Object Extraction from Colour Cadastral Maps". In: (2008), pp. 506–514. DOI: 10.1109/DAS.2008.9.
- [36] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. *You Only Look Once: Unified, Real-Time Object Detection*. 2015. DOI: 10.48550/ARXIV.1506.02640. URL: <https://arxiv.org/abs/1506.02640>.
- [37] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. DOI: 10.48550/ARXIV.1505.04597. URL: <https://arxiv.org/abs/1505.04597>.
- [38] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. "SuperGlue: Learning Feature Matching with Graph Neural Networks". In: *CVPR*. 2020. URL: <https://arxiv.org/abs/1911.11763>.
- [39] Scikit-Image. *Skeletonize*. [Online; accessed 13-August-2022]. URL: https://scikit-image.org/docs/stable/auto_examples/edges/plot_skeleton.html.
- [40] Farhana Sultana, Abu Sufian, and Paramartha Dutta. "Evolution of image segmentation using deep convolutional neural network: a survey". In: *Knowledge-Based Systems* 201 (2020), p. 106062.
- [41] Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, and Bin Xiao. *Deep High-Resolution Representation Learning for Visual Recognition*. 2019. DOI: 10.48550/ARXIV.1908.07919. URL: <https://arxiv.org/abs/1908.07919>.
- [42] Jianhua Wu, Pengjie Wei, Xiaofang Yuan, Zhigang Shu, Yao-Yi Chiang, Zhongliang Fu, and Min Deng. "A New Gabor Filter-Based Method for Automatic Recognition of Hatched Residential Areas". In: *IEEE Access* 7 (2019), pp. 40649–40662. DOI: 10.1109/ACCESS.2019.2907114.
- [43] Tete Xiao, Yingcheng Liu, Bolei Zhou, Yuning Jiang, and Jian Sun. *Unified Perceptual Parsing for Scene Understanding*. 2018. DOI: 10.48550/ARXIV.1807.10221. URL: <https://arxiv.org/abs/1807.10221>.
- [44] Zhanwu Yu, V. Prinet, Chunhong Pan, and PinXiang Chen. "A novel two-steps strategy for automatic GIS-image registration". In: *2004 International Conference on Image Processing, 2004. ICIP '04*. Vol. 3. 2004, 1711–1714 Vol. 3. DOI: 10.1109/ICIP.2004.1421402.
- [45] Yuhui Yuan, Xiaokang Chen, Xilin Chen, and Jingdong Wang. *Segmentation Transformer: Object-Contextual Representations for Semantic Segmentation*. 2019. DOI: 10.48550/ARXIV.1909.11065. URL: <https://arxiv.org/abs/1909.11065>.
- [46] Sebastian Zambanini. "Feature-based groupwise registration of historical aerial images to present-day ortho-photo maps". In: *Pattern Recognition* 90 (2019), pp. 66–77. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2019.01.024>. URL: <https://www.sciencedirect.com/science/article/pii/S0031320319300342>.

- [47] Yuanyuan Zhang, Xuesong Wang, Peng Zeng, and Xiaohong Chen. "Centrality Characteristics of Road Network Patterns of Traffic Analysis Zones". In: *Transportation Research Record: Journal of the Transportation Research Board* 2256 (Dec. 2011), pp. 16–24. DOI: 10.3141/2256-03.
- [48] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. *Pyramid Scene Parsing Network*. 2016. DOI: 10.48550/ARXIV.1612.01105. URL: <https://arxiv.org/abs/1612.01105>.

Appendix A

Additional plots

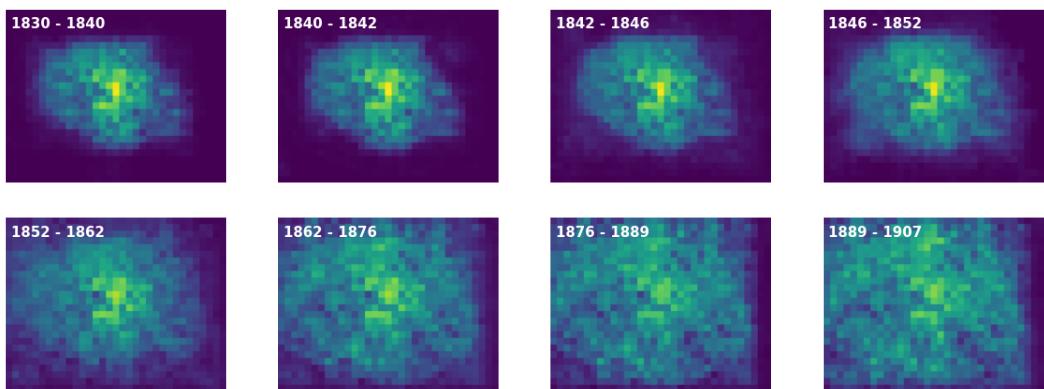


Figure A.1: Abandoned analysis : visualisation of the expansion of the density of the Parisian road network

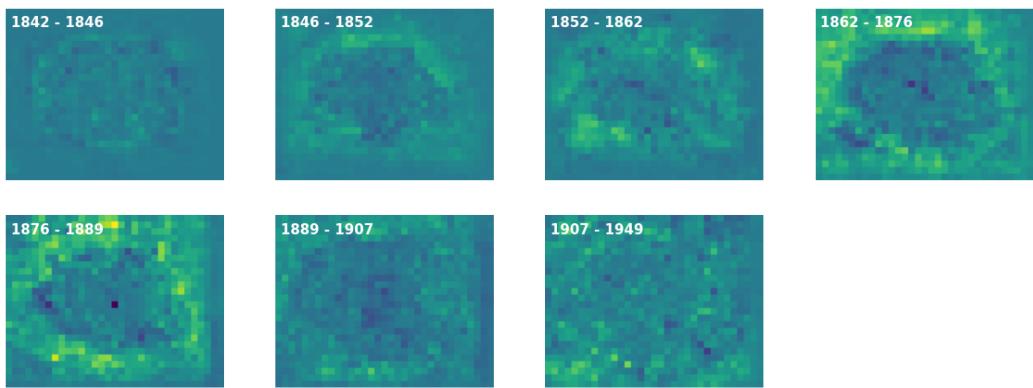


Figure A.2: Abandoned analysis : visualisation of the modifications in the density of the Parisian road network for each time period.

ROME



STRASBOURG



MONTPELLIER



LYON



TOULOUSE

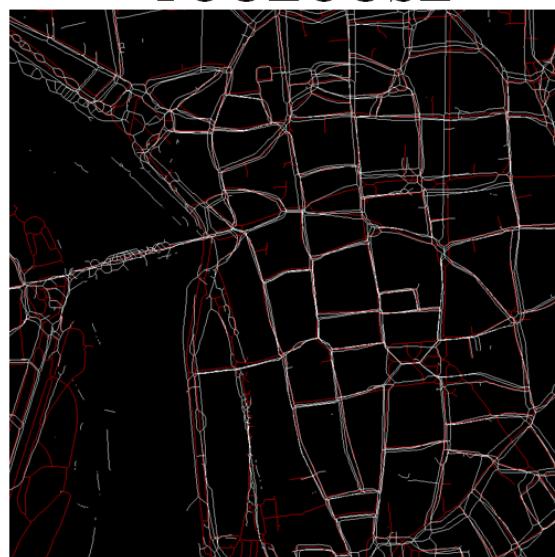


Figure A.3: For each city of our test dataset, the road networks of the maps that were correctly realigned (in white) are stacked on top of their anchor (in red).