



HAL
open science

Segmentation of historical maps without annotated data

Aurélie Lemaitre, Jean Camillerapp

► **To cite this version:**

Aurélie Lemaitre, Jean Camillerapp. Segmentation of historical maps without annotated data. 6th International Workshop on Historical Document Imaging and Processing (HIP'21), Sep 2021, Lausanne, France. pp.19-24, 10.1145/3476887.3476909 . hal-03374571

HAL Id: hal-03374571

<https://inria.hal.science/hal-03374571>

Submitted on 12 Oct 2021

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.

Segmentation of historical maps without annotated data

AURÉLIE LEMAITRE, Univ Rennes, CNRS, IRISA, France

JEAN CAMILLERAPP, Univ Rennes, CNRS, IRISA, France

This paper presents the method which we submitted to the competition of Historical Map Segmentation, in ICDAR'21. The goal is to segment document images of Paris maps from the beginning of the 20th century: delineate the content of the map and locate the graticule line intersections, which give indications on the geographic coordinates of the map. Our contribution is entirely led by a rule-based method. Thus, it does not require a training phase, nor annotated data. We extract the line segments, in the image, at various resolutions. Then, a grammatical description combines those elements in a logical way to filter the relevant map contents. This is an original approach, at the era of deep learning techniques, but that is very convenient for old non-annotated documents. We validated this work during MapSeg competition. It obtained the third position for map segmentation and the second position for graduate line detection, with a localization score of 89.2%.

CCS Concepts: • **Applied computing** → **Document analysis**.

Additional Key Words and Phrases: document analysis, historical maps, rule based system, line segment extraction, multiresolution analysis

ACM Reference Format:

Aurélie Lemaitre and Jean Camillerapp. 2021. Segmentation of historical maps without annotated data. In *The 6th International Workshop on Historical Document Imaging and Processing (HIP '21), September 5–6, 2021, Lausanne, Switzerland*. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3476887.3476909>

1 INTRODUCTION

In this paper, we present the method that took part of MapSeg competition at ICDAR'21 [2]. This competition focuses on the segmentation of historical city maps, more precisely maps of Paris, from the beginning of the 20th century (Figure 1).



Fig. 1. Extract from a map of Paris, 1925

City maps are of great interest for geographers. They contain many-detailed information, with an accurate representation. However, the interpretation of historical images of maps is a difficult challenge [4]. The large density of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

Manuscript submitted to ACM

information in the documents makes their interpretation difficult. For this vectorization process, geographers are used to long and tedious manual transcription.

Our work focuses on two tasks that are the key elements of the pipeline process of automatic map recognition. The first one is *map content segmentation*, which consists in the separation of meta-data like titles, legends, from core map content. On Figure 6, this corresponds to the detection of the yellow frame.

The second work is the detection of *graticule line intersection*. This work aims to map historical coordinate system into a reference coordinate system. Graticule lines indicate the North/South/East/West major coordinates in the map. Their intersections are very useful to geo-reference the map image. On Figure 7, the graticule lines are in orange, and their intersections are in red.

In the era of deep learning, well-trained neural networks could solve those tasks. Nevertheless, this may require many annotated data. That is why we propose a concurrent approach, which does not require any annotated data. Our approach is entirely led by a rule-based system. The principle is to extract line segments in the maps, at different levels of resolution, and then to combine those visual clues using logical rules.

This method has been validated during MapSeg ICDAR'21 competition, and obtained the 2nd place in the graticule line task.

Our main contributions are:

- the extraction of line segments, at different resolution levels, as visual clues;
- a rule based approach for map segmentation, that does not require annotated data;
- the validation during the MapSeg competition of ICDAR'21.

2 RELATED WORK ON HISTORICAL MAPS

2.1 Historical map segmentation

The segmentation of historical maps is a difficult task. Thus, it must deal with the difficulties related to historical documents: degradation, bad quality of printing. However, map vectorization is also constrained by the complexity of the map contents. Thus, general techniques for document analysis and graphics recognition cannot be directly applied to map processing [5]. Neural networks have been used for a long time in the domain. However, their applications remain limited to extracting the information from text.

Recent convergence of advancements in the domains of training deep neural networks significantly improved results, inspired by computer vision applications [8]. Thus, map digitization requires visual object and pattern recognition, such as the identification of edges (parcel and building lines), symbols, text, and patterns. Neural networks can be trained for such tasks, especially when it comes to vision classification problems [4].

For example, convolution neural networks (CNN) have shown a high level of performance for boundary detection [11]. Some more recent work propose to combine deep learning with mathematical morphology tools [3]. He *et al.* [7] note that exploiting multi-scale representations is critical to improve edge detection for objects at different scales. They propose a Bi-Directional Cascade Network (BDCN) architecture, where an individual layer is supervised by labeled edges at its specific scale, rather than directly applying the same supervision to different layers. The main limit of using neural networks for historical map vectorization is the constraint of the limited amount of training data available [8].

2.2 Approaches proposed for MapSeg competition

Baloun *et al.* [1] proposed the winning method for map content segmentation task. Their method uses a Fully Convolutional Network, trained on 26 historical map annotated images. This system provides interesting results but requires an additional post-processing to produce consistent results. Thus, the authors use an OTSU binarization to detect small components. Then, they combine the predicted FCN with the connected components, using morphological operations. CNN team [2] also provide some morphological operations. They obtain the map by dilatation from a rectangle in the center of the document.

Concerning the graticule line intersections, the UWB team [2] uses a succession of dedicated processes to detect graticule lines. Their work uses Hough accumulator and does not require any learning. The L3IRIS team [2] uses a deep learning approach, based on a Unet model. However, their results are not satisfactory, and they need to post process them with Hough accumulator.

The approaches proposed by the other competitors show well the interest of non-learning-based methods. However, the proposed approaches are dedicated for each specific task.

In a previous work, we have shown the interest of a rule-based system for document analysis [9]. We have shown the interest to combine several points of view, several resolutions of a document, to improve recognition. We propose to apply this method for the MapSeg competition.

3 RULED BASED APPROACH FOR MAP SEGMENTATION

Our system works with visual clues that are detected in the images. Then, a rule-based method enables to compute the searched elements, using the visual clues. For the implementation, we used an existing line segment extractor, and we described specific rules, dedicated to the problematic of map segmentation. We first describe the visual clues, before explaining the rules.

3.1 Visual clues

Our system uses a line segment extractor. The particularity of our approach is to combine line segment extraction at different resolution levels.

3.1.1 Kalman filtering. We apply an efficient line segment extractor [10], based on a Kalman filtering. The main idea is to follow a prediction/verification principle. A line segment is initialized as a set of touching black pixels, with a regular slope and a regular thickness. At each step of the analysis, the algorithm predicts the position of the following pixels, and adjusts the parameters of the model for the remaining analysis. All the details are given in [10].

The main properties of this extractor are:

- following the slope and the curvature,
- dealing with white spaces and following the dotted lines,
- crossing noisy regions, like intersections with other black pixels.

The Figure 2 illustrates the ability of the line segment extractor to detect segment in noisy regions.

3.1.2 Available clues. One of the originality of this contribution is to apply the line segment extractor at several resolution levels of the image. Indeed, in the map image, the nature of line segments is different depending on their usage. As can be seen in the Figure 1, the document contains:

- triple rulings outside of the map,

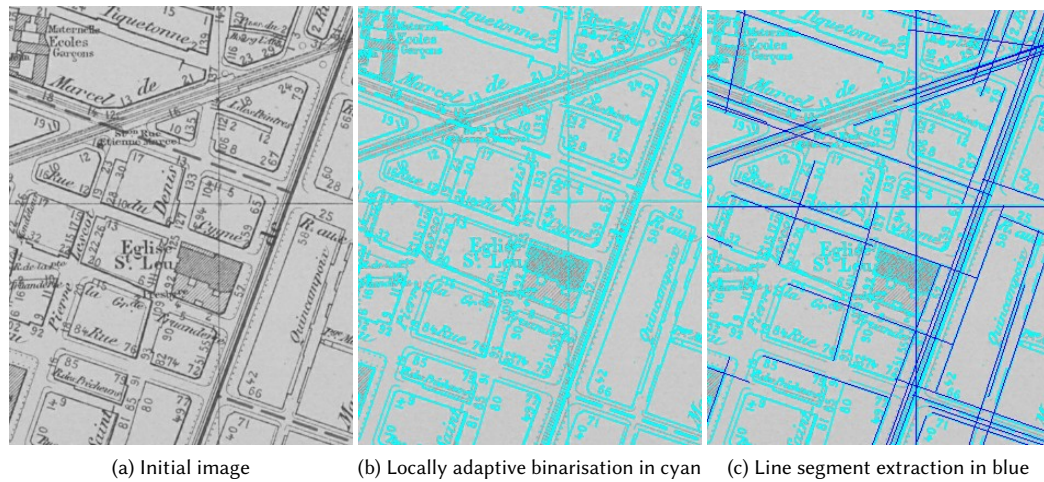


Fig. 2. Interest of Kalman filtering for line detection in noisy environment

- simple thick rulings for the border of the map and legend regions,
- simple thin rulings for the graticule lines,
- many other rulings for roads, railroads, buildings...
- noise, such as paper folding, which could be detected as a line segment.

Thus, we extract the line segments at three levels of resolution:

- at low resolution: image dimensions divided by 16 (Figure 3), the line segments provide clues for map outer frame detection;
- at medium resolution: image dimensions divided by 4 (Figure 4), the line segments provide clues for border and legend position;
- at high resolution: image dimensions divided by 2 (Figure 5), the line segments provide clues for graticule lines.



Fig. 3. Line segment extraction at low resolution: visual clues for map outer frame detection (181 detected segments)



Fig. 4. Line segment extraction at medium resolution: visual clues for borders and legend (2295 detected segments)



Fig. 5. Line segment extraction at high resolution: visual clues for graticule lines detection (6424 detected segments)

The extracted line segments are only clues for the final result. Thus, even if the line segment detector is efficient, the high density of the content, and the quality of printing of ancient document sometimes causes over-segmentation of line segments. Moreover, the line segment extractor detects many candidates on each image (several thousands). Consequently, we propose to use a rule-based system to smartly combine those clues, and build the expected lines.

3.2 Map content area segmentation

Our first grammatical rule aims to detect the outer border. We use the low-resolution extraction (Figure 3), that provides visual clues on the presence of the triple rulings outside of the map region. Then, a grammatical rule describes the frame like a geometric shape of rectangle. As a result, we obtain the green contour on Figure 6.

The final goal is to detect the precise content area segmentation. The line segment at medium resolution provide good clues (Figure 4). Using this resolution, we first build a rectangle shape, parallel to the outer frame. Then we consider that, in any corner of that rectangle, there may be a legend. Consequently, we describe the contour with the following rules.

```

mapContour ::=
  AT(insideFrame) &&
  rectangleShape R &&
  AT(topLeftCorner R) &&
  possibleLegend L1 &&
  AT(topRightCorner R) &&
  possibleLegend L2 &&
  AT(bottomRightCorner R) &&
  possibleLegend L3 &&
  AT(bottomLeftCorner R) &&
  possibleLegend L4.

```

The AT operator defines the position of the analysis, relatively to the position of the rectangle shape R. The rule possibleLegend can succeed with a small rectangle in the corner, or with nothing. This ensure the possibility to find from zero to four legends in each map. Then we compute the final contour, it is the yellow line in Figure 6.

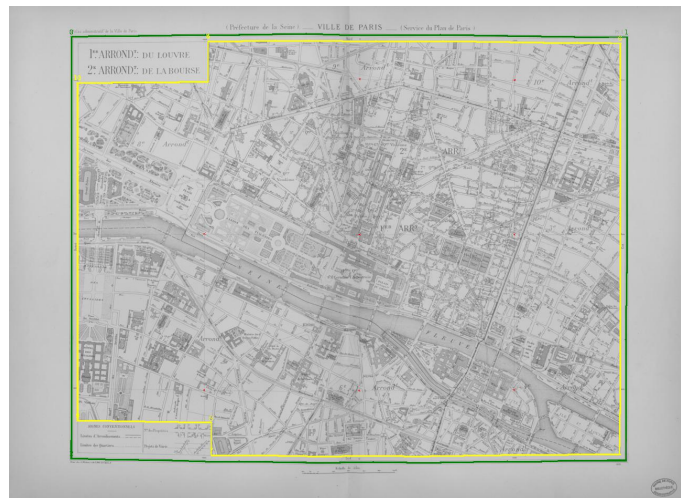


Fig. 6. Map area segmentation: in green the outer border, in yellow the final contour of map content area.

3.3 Graticule line intersection

The graticule line detection is carried out as a continuation of the previous step. Thus, the analysis is located inside of the detected contour. The analysis uses the clues detected at high-resolution level (Figure 5). The rule-based grammar has a huge importance, as there are many candidates: line segments are also detected on streets, houses, rivers, railroads... Figure 5 shows this problematic. Nevertheless, some simple but strong constraints make it possible to restrict the analysis.

The grammatical rules describe a map as made of a major cross, with two perpendicular rulings. We explore the segments in descending order of size: the analysis tries to build a long enough cross using the longest segment. Then the grid is composed of additional parallel rulings. This is expressed in our system by the following rule:

```

gridPattern Frame Grid ::=
  AT(middleFrame Frame) &&
  FIND(firstCross Horiz1 Verti1) COND(usingLongestLine)&&
  horizontalParallelGrid Horiz1 &&
  verticalParallelGrid Verti1 &&
  checkGridRegularity.

```

At the end of the rule, the predicate `checkGridRegularity` checks that the graticule line found respects a regular horizontal and vertical spacing. In particular, it eliminates false positives and allows the detection of lines that may have been missed.

The rules `horizontalParallelGrid` `verticalParallelGrid` imply the recognition of independent `graticuleLines`. We describe a graticule line as a succession of aligned line segments, in the same axis, that is overall long enough.

```

graticuleLine ::=
  firstLineSegment S &&
  AT(beforeSegment S)&&
  alignedBeginOfGraticuleLine B &&
  AT(afterSegment S) &&
  alignedEndOfGraticuleLine E &&
  checkLongEnough (S B E).

```

We show the results in Figure 7: the first cross is in purple, the remaining of the grid is in orange. We compute the intersections, in red on Figure 7.

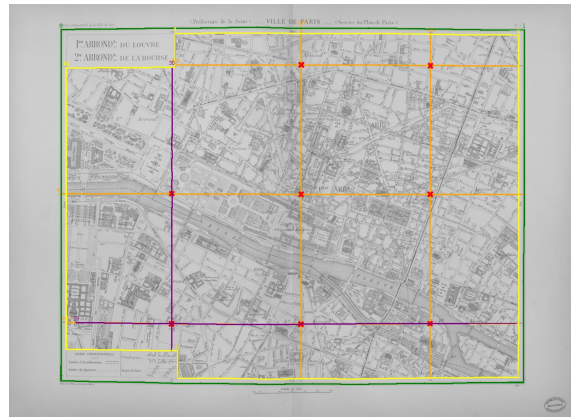


Fig. 7. Result of graticule line detection: in purple the reference cross, used as a base to detect the rest of the grid in orange. In red: intersection of graticule lines.

4 IMPLEMENTATION

Our grammatical rule system is implemented using the existing DMOS method [9], based on logical programming. This method uses the specific EPF formalism. EPF enables to describe both the physical and logical content of the document. Once a document has been described in EPF language, it is compiled and DMOS method automatically produces the

associate analyser for document recognition. DMOS method has been used for more than 15 years, applied to many kinds of documents, and validated at a large scale. DMOS method is now available for research purpose [6].

Concerning map segmentation, we used EPF language to describe the organisation of maps. The rules presented in sections 3.2 and 3.3 are expressed in EPF language. Then, after a compilation step, DMOS method produced our analyser. Using DMOS method is a real advantage for designing a new recognition system, as it benefits from a complete existing analysis system. The only thing to do is to describe the special rules for map segmentation.

The second interest of using DMOS method is that the system well masters the complexity, using the backtracking of logical programming. Thus, the system tries all possible combinations of segments to find the first one that will check all constraints. This is done automatically.

5 EXPERIMENTATION

We validated this work by the participation of MapSeg competition of ICDAR'21. This competition proposed three tasks. In the competition, the task to segment the map content area is called *task 2*. The task of graticule line detection is called *task 3*. We did not work on the first task proposed by the competition.

5.1 MapSeg competition database

The documents for evaluation are provided by MapSeg competition at ICDAR'21. All the data and metrics are available online ¹. The maps are taken from 9 atlases of the city of Paris, France, from 1894 to 1937. The dataset is made of 26 images for train, 6 images for validation. The competition is led on 95 test images. The input are JPEG RGB images which are quite large (about 10 000 x 10 000 pixels). Annotations were made manually, by the organizers of the competition, for both tasks.

5.2 Segment map content area

The goal is to provide a mask of the content area: white pixels inside of the mask, and black pixels outside. We produced masks using the yellow borders presented on all our results (Figures 6, 7, 11). Figure 8b presents a part of a generated final mask.

5.2.1 Metrics. The quality of the detected contours are evaluated using the Hausdorff distance between the ground-truth shape and the predicted one. More specifically, the organizers chose to use a "Hausdorff 95" variant which discards the 5 percentiles of higher values (assumed to be outliers) to produce a more stable measure. Finally, the average of the measures for all individual map images produces the global score. The resulting measure is a float value between 0 and a large value. A lower value is better.

5.2.2 Competition results. The table 1 presents the competition results for content area segmentation [2]. Our system obtains the third place, with an average Hausdorff distance of 112, on the 95 test images of the competition. The methods proposed by the two first candidates, UWB and CMM are described in section 2.2.

5.2.3 Discussion. As our method is explainable, we know clearly what its strengths and weaknesses are. First, we must mention that, as the Kalman filtering is able to follow a local curvature, our system properly detects the border of the document, even if the paper was folded during digitization. For example, Figure 8 shows good results of detection, and the associated generated mask. This ensures that our method can cope with wonky scans.

¹<https://icdar21-mapseg.github.io/>

Table 1. Final competition scores for content area segmentation. Scores range from 0 (best) to arbitrarily large values.

Rank	Method	Hausdorff 95 distance
1	UWB	19
2	CMM	85
3	IRISA (ours)	112
4	L3IRIS	126



(a) In yellow: the recognized border



(b) Upper part of the generated mask

Fig. 8. Success of detection: following the curvature of the document

Concerning the drawbacks of our method, the main limit is that it totally relies on the presence of line segments. Thus, our method is not able to build the border of the map when the content goes out of the frame (Figure 9). Nor can it detect the legends when a line segment does not delimit them (Figure 10). Those two key points should be improved if we want to lower the final error score for this task. Note that this occurs frequently in the competition set, whereas it was not the case in the training set.

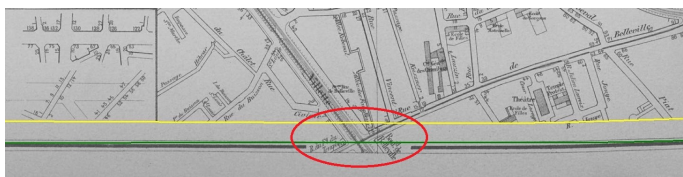


Fig. 9. Limit of detection: the content is out of the border

5.3 Locate graticule lines intersection

The goal of so-called *task 3* is to provide a list of points that are the intersection of graticule lines. The red crosses in figures 7 and 11 illustrate our results.

5.3.1 Metrics. The metric evaluates both the detection and the localization of the intersections. It is detailed in [2]. Each predicted point is associated to an expected one, if it is located within 50 pixels. Then, the metric take the area under a



Fig. 10. Limit of detection: the legend is not found as not delimited by a line segment.

curve " F_β score vs threshold" (with $\beta = 0.5$) as a performance indicator. Thus, such indicator blends two elements: point detection and spatial accuracy. The average of the measures for all individual map images produces a global indicator with a confidence measure. The resulting measure is a float value between 0 and 100%. A higher value is better.

Overall, this metric quantifies the accuracy of each detected point: 50 pixels of distance represents a 100% error, whereas an error of 5 pixels costs 10% in the final score, for this point. Concerning detection, this metric does not penalize missed points very much. On the other hand, false positives (or points located more than 50 pixels away) are very penalizing.

5.3.2 Competition results. The table 2 presents the competition results for graticule line intersection location [2]. Our system obtains the second place, with an average score of 89.2%. The method proposed by the first candidate, UWB, is described in section 2.2.

Table 2. Final competition scores for graticule line intersection. Scores range 0 (worst) to 100% (best).

Rank	Method	Detection score
1	UWB	92.5%
2	IRISA (ours)	89.2%
3	CMM	86.6%
4	L3IRIS	73.6%

5.3.3 Discussion. One of the main interest of our system is that it can handle a high level of combinatorial complexity. The rules enable to select the better candidates for the global construction of the grid, even if there are many candidate line segments. This is possible with of DMOS method, relying on logical programming.

Our system is able to deal with slope. The Figure 11 presents an example with inclined graticule lines. As both the line segment detectors and the rules for the extension of segments follow the slope, the graticule lines are correctly extracted.

The remaining errors are due to a high density of line segments. The figure 12 present an example of document with horizontal fortification and railroad. Unfortunately, those drawings are horizontal, in a reasonable zone to find a graticule line. The extracted visual clues shows that, even for a human, finding the position of a graticule line is not obvious. This is the main reason of our low scores, as railroad often occurs in the competition database, whereas they were not present in the test set.



Fig. 11. Success of detection: inclined graticule lines.



(a) Initial image: in green, the position where a graticule line could be found, taking into account the rest of the grid



(b) Visual clues: candidate line segments in blue. The position of graticule lines is not obvious.



(c) Result of graticule line recognition: the pink horizontal line is expected, but the orange one, just below, is found instead, on horizontal fortification

Fig. 12. Limit of detection: misplaced horizontal ruling

6 CONCLUSION

In this paper, we have presented our method for map segmentation. We focused on two tasks, evaluated in MapSeg competition of ICDAR'21. The first task is map content segmentation, and the second is graticule line detection.

We proposed a consistent approach for those two tasks. Our method first extracts line segment clues, at different resolution images. Then, an efficient rule based system selects the best candidates to provide the final segmentation. The main advantage of this approach is that it does not require any annotated data. Indeed, annotation is often costly and tedious for ancient documents.

We have specifically designed the rules for the MapSeg competition tasks, using the training dataset. However these rules are very simple, and could be used for any other similar corpus of map segmentation. In the era of deep-learning supremacy, it is interesting to show that some traditional methods can still perform well. This is the interest of a competition such as MapSeg. During the competition, our method obtained the third place in the map content segmentation task, and the second place for the localization of graticule line intersection.

In a future work, we plan to enrich the input visual clues of our method, in order to complete the use of line segments. For example, using DMOS method, we could easily use as input results of deep-neural networks, and combine it with logical rules.

REFERENCES

- [1] Josef Baloun, Pavel Král, and Ladislav Lenc. 2021. ChronSeg: Novel Dataset for Segmentation of Handwritten Historical Chronicles. 314–322.
- [2] Joseph Chazalon, Edwin Carlinet, Yizi Che, Julien Perret, Bertrand Duménieu, Clément Mallet, and Thierry Géraud. 2021. ICDAR 2021 Competition on Historical Map Segmentation. In *ICDAR 2021*.
- [3] Yizi Chen, Edwin Carlinet, Joseph Chazalon, Clément Mallet, Bertrand Duménieu, and Julien Perret. 2021. Combining Deep Learning and Mathematical Morphology for Historical Map Segmentation. In *IAPR International Conference on Discrete Geometry and Mathematical Morphology (DGMM)*. Springer, Uppsala, Sweden.
- [4] Yao-Yi Chiang, Weiwei Duan, Stefan Leyk, Johannes H. Uhl, and Craig A. Knoblock. 2020. *Historical Map Applications and Processing Technologies*. Springer International Publishing, Cham, 9–36.
- [5] Yao-Yi Chiang, Stefan Leyk, and Craig Knoblock. 2014. A Survey of Digital Map Processing Techniques. *Comput. Surveys* 47 (07 2014).
- [6] Bertrand B. Couasnon and Aurélie Lemaitre. 2017. DMOS, It's your turn !. In *1st ICDAR-OST*. Kyoto, Japan. <https://hal.inria.fr/hal-01659131>
- [7] Jianzhong He, Shiliang Zhang, Ming Yang, Yanhu Shan, and Tiejun Huang. 2020. BDCN: Bi-Directional Cascade Network for Perceptual Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020), 1–1.
- [8] Jelena Ignjatić, Bojana Bajic, Aleksandar Rikalovic, and Dubravko Culibrk. 2018. Deep Learning for Historical Cadastral Maps Digitization: Overview, Challenges and Potential.
- [9] Aurélie Lemaitre, Jean Camillerapp, and Bertrand Couasnon. 2008. Multiresolution cooperation makes easier document structure recognition. *IJDAR* 11, 2 (2008), 97–109.
- [10] Charles Queguiner, Jean Camillerapp, and Ivan Leplumey. 1995. Kalman Filter Contributions Towards Document Segmentation. In *ICDAR*. Montreal, 765–769.
- [11] Saining Xie and Zhuowen Tu. 2015. Holistically-Nested Edge Detection. *CoRR* abs/1504.06375 (2015).