Visual Content Analysis and Linked Data for Automatic Enrichment of Architecture-Related Images

Mager, Tino; Khademi, Seyran; Siebes, Ronald; Hein, Carola; de Boer, Victor; van Gemert, Jan

DOI
10.1007/978-3-030-15200-0_18

Publication date
2019

Document Version
Final published version

Published in
Digital Cultural Heritage

Citation (APA)
Mager, T., Khademi, S., Siebes, R., Hein, C., de Boer, V., & van Gemert, J. (2019). Visual Content Analysis and Linked Data for Automatic Enrichment of Architecture-Related Images. In H. Kremers (Ed.), *Digital Cultural Heritage* (pp. 279-293). Springer. https://doi.org/10.1007/978-3-030-15200-0_18

Visual Content Analysis and Linked Data for Automatic Enrichment of Architecture-Related Images

Tino Mager, Seyran Khademi, Ronald Siebes, Carola Hein, Victor de Boer and Jan van Gemert

Abstract Built form dominates the urban space where most people live and work and provides a visual reflection of the local, regional and global esthetical, social, cultural, technological and economic factors and values. Street-view images and historical photo archives are therefore an invaluable source for sociological or historical study; however, they often lack metadata to start any comparative analysis. Date and location are two basic annotations often missing from historical images. Depending on the research question other annotations might be useful, that either could be visually derived (e.g. the number or age of cars, the fashion people wear, the amount of street decay) or extracted from other data sources (e.g. crime statistics for the neighborhood where the picture was taken). Recent advances in automatic visual analysis and the increasing amount of linked open data triggered the research described in this paper. We provide an overview of the current status of automated image analysis and linked data technology and present a case study and methodology to automatically enrich a large database of historical images of buildings in the city of Amsterdam.

Keywords Architectural history · Architectural heritage · Computer vision · Linked open data · Automated image analysis

---

T. Mager (✉) · C. Hein
Faculty of Architecture and the Built Environment, Department of Architecture, TU Delft, Julianalaan 134, 2628 BL Delft, The Netherlands
e-mail: b.t.mager@tudelft.nl

S. Khademi · V. de Boer
Faculty of Sciences, Department of Computer Science, VU Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

R. Siebes · J. van Gemert
Faculty of Electrical Engineering, Mathematics and Computer Science, Department of Intelligent Systems, TU Delft, Mekelweg 4, 2628 CD Delft, The Netherlands

H. Kremers (ed.), *Digital Cultural Heritage*,
https://doi.org/10.1007/978-3-030-15200-0_18

1 Image Use in Architectural History

The creation of and the research in architecture is closely related to images. From architectural drawing as a cornerstone of design to paintings or photographs, visual representations allow the investigation of not only existent, but also unbuilt, changed or destroyed buildings. Thereby, the images facilitate select insights and eventually allow for contemplating on timely or spatially unreachable objects. In this regard, especially photography plays an important role due to its low grade of abstraction and a relatively trustworthy representation of specific situations.[1] It might be by mere chance that *View from the Window at Le Gras*, the world's first photograph, depicts a building.[2] But the then new technology was soon taken into account for conserving the views of historical buildings. After individual approaches, the Missions Héliographiques was a first large-scale institutional approach: in 1851 Prosper Mérimeée, France's Inspector General of Historical Monuments, initiated the project to photograph France's monuments on a country wide scale. This acknowledgement of photography's possibilities to provide reliable depictions of buildings belongs to an early stage of the modern chapter of the discipline of architectural history, but it has remained its validity ever since.

Much later, in the second half of the 20th century, the pictorial turn established images not only as scientific sources, but also as an object of investigation in its own right. Today in architectural history, research on image content as well as on issues of creation or fake concerning image content is accompanied by considerations on how the digital availability of images changes and challenges the character of research and the notion of pictorality. At the border to art history and computer science, a broad variety of considerations can be observed. At this point, we argue that the definition of the difference between chemical or digital photography as well as questions of originality and materiality might allow for interesting excursions into art theory but will be of lesser relevance than a much more game-changing aspect of digitality: the autonomous processing of image content by machines.

The recent leaps forward in computer science and machine learning, particularly deep learning, have shown astonishing results in applying non-human intelligence to a broad range of purposes. Even if rather trivial tasks for humans, such as driving a car, seem to remain relatively difficult for machines, they outsmart us on more specific tasks like learning chess from scratch and beating the chess champion within a day, or learning a language. Moreover, computers are pretty good at detecting faces—just check e.g. the Fotos app on your phone—and can do so in the millions, that means to an extent much further than human beings would be able to recognize their fellows. This brings us back to architecture: if computers can recognize individual human beings by looking at their face, they should be able to detect specific buildings by looking at their visual representations. In principle, this is the aim of ArchiMediaL,

[1]The authors are aware of the selective and sectional character of photographic images as well as the possibility of manipulative use of photographs and their falsification. However, at this point the important thing is photography's ability to capture a situation closely to reality.

[2]Nicéphore Niépce, *View from the Window at Le Gras*, heliographic image, 1826 or 1827.

an interdisciplinary research project conducted by architectural historians and computer scientists from TU Delft, VU Amsterdam, HafenCity University Hamburg and University Duisburg-Essen. Its general aim is to research computational approaches for the automatic detection of built forms and architectural elements in visual representations. So what's the benefit?

Along with digital photography and mass digitalisation in the recent years, the availability of visual representations of architecture increased exponentially. Unfortunately, in many cases this availability remains theoretical and does not come along with accessibility. To be found and investigated, images of architecture need to be annotated with information concerning the depicted object, otherwise they are unavailable for (re)search. Moreover, images of the same object can hold divergent annotations in different repositories, not to mention mistakes or different languages or spellings [1]. Here, a recognition and identification of the visual content—the depicted object—could overcome the shortcomings of text-based search. If such a vision-based search could be conducted across databases and archives, it would link their repositories and thereby enhance the search for visual sources (once the related legal questions are solved, a serious issue not to be addressed here). However, the interlinking of the repositories requires the advanced approach of linked open data.

2 Linked Open Data

Physical and digital objects and events can be described in various ways and for various purposes. An encyclopedia is one example where natural language is used to describe noteworthy things for historical and educational purposes. A number-plate on a car contains a sequence of numbers and letters used to uniquely identify a vehicle. Humans have a high tolerance to syntactic variations and sloppiness used to describe or identify things by relying on redundancy and common sense to deduce the intended 'meaning' of a sequence of characters. Computers are less flexible and rely on explicit schemas and rigid grammar. For more than two decades, just after the World Wide Web became part of our daily lives, researchers in the domain of Computational Logic came up with the idea to use the Web infrastructure as a vehicle to collaboratively describe and link 'things' in a formal way and create a standard to define an event or object by reusing descriptions made and shared by others. This research domain is called the *Semantic Web* [2] and the resulting formal (in various levels of computational complexity) and interconnected descriptions *Linked Data* [3]. When this data is publicly available, we call it *Linked Open Data (LOD)*, and the whole collection of Linked Open Data sets is called the *Linked Open Data cloud*. Since structured language is grounded in formal logic, computers can automatically reason over this data. For example, if one wants to describe an archaeological finding of a Delftware ceramic plate at a location in Alkmaar, a Dutch city, the place is already been described by the *Geonames Ontology* [4] and the type of artifact, 'Delftware', by the *Art & Architecture Thesaurus* available as Linked Data [5]. Since the computer 'understands' the geographical property of sub-regions, and the hierarchical property of historic artifact classifications, the archaeological finding would be automatically included when searching for: ‘pottery findings in The Netherlands’. The use of Linked Data to represent both data and metadata allows for a flexible data model that is easily updated and extended. Linked Data and the Resource Description Framework (RDF) [6] are the web-standard for publishing heterogeneous data on the Web, making it possible to build interconnected knowledge graphs, enriching original digital collections.

Digital Humanities is one of the earliest adopters and leading contributors to the Linked Open Data cloud, ranging from cultural heritage [7], archaeology [8], history [9] to politics [10] and census data [11]. Architecture is another domain where relevant Linked Data became available. For example, The Netherlands’ Cadastre, Land Registry and Mapping Agency—in short Kadaster—collects and registers administrative and spatial data on property and the rights involved. This also goes for ships, aircraft and telecom networks. Linked Data has been used extensively for knowledge representation and data publishing in the digital Cultural Heritage domain [12]. Europeana [7], the European cultural heritage aggregator is using Linked Open Data to publish information from European museums, archives and libraries on the Web of Data using the Europeana Data Model (EDM), which is defined in RDF and reuses well-known vocabularies such as SKOS and Dublin Core.

Previous Digital History projects that use linked data include *Dutch Ships and Sailors*, where heterogeneous maritime historical datasets are integrated [13], and *BiographyNet* which focused on biographical data [14]. By adhering to Linked Data standards, the project results can easily be integrated in (international) research infrastructures such as Europeana or CLARIAH [15], who also base their data publishing on Linked Data principles.

Apart from inter-linking images by content across repositories, an automatic detection of buildings supports a shift of focus within architectural history. Even though architectural historians began to use images relatively early, the discipline did not sufficiently reconsider its own methodological principles. Architectural historians for a long time mainly looked at the masterpieces of architectural history. As they focused on palaces and temples, they created a skewed understanding of the past. Today, despite the availability of millions of visual representations from around the globe, the consideration to the new possibilities that have arisen is by far insufficient. As a result, our knowledge of working class housing in ancient cities or on suburban developments is more limited than knowledge on the buildings of the elite. If today the flood of images can only be considered fully with the help of computer technology, the new possibilities also entail the obligation to question the established perspectives and to change them in the direction of a holistic view of architecture based on refined theoretical and methodological foundations. It is therefore crucial to critically reflect on the reading and interpretation of these sources and to develop global and balanced approaches to explore the previously unavailable heterogeneous and interconnected datasets. Here again, an automated recognition of image content would help researchers to gain systematic visual access to the rather unknown world of non-capital-a architecture.

3 Visual Content Recognition

Unlike humans, computers are good at memorizing huge amounts of seemingly unrelated images; however, deducing complex semantic relations for these images is among the more challenging tasks in computer vision. There are recent advances on data-driven approaches that impressively improved the later capability of the visual intelligent systems. This could be a benchmark for computer scientists to interact more meaningfully with other disciplines.

In this part, we aim to spot and sketch the joint research directions for visual data processing and humanities researches in urban planning and history of architecture within the scope of the three-year project, with the focus on the state of the art computer vision techniques. We aim for mutual understanding between the two fields that enables true interdisciplinary research. In the following, example areas are highlighted where computer vision can offer an aid in architectural history research, which is further elaborated in our case study at the end of this paper.

3.1 Image Retrieval

There has always been a demand from scholars to locate a research object such as an image in a collection of available objects, for instance in a repository of images. Content Based Image Retrieval (CBIR) is the task of finding a query image in the gallery of a reference image database. This is a timely problem of image based search engines that can be a challenging task once the query image does not have an exact match (same object representation) in the gallery. The most explored research area in CBIR is the representation learning, i.e. to effectively learn the descriptor of the query and the gallery images in order to maximize the distance metric between un-matched objects, and the similarity between the same objects [16, 17]. Another active area of research is distance metric learning that studies the best measure of the similarity and dissimilarity between objects. We refer to the combination of these topics as “matching problem”. Unlike the classification and segmentation tasks that deep learning techniques are successful at, the matching problem is not yet impressively excelled by deep nets. This is mostly attributed to the lack of available data to train a network in an end-to-end manner, which is known to be the most crucial setting to achieve the competing results. Given this, there are lots of research opportunities in this area.

4 Location Prediction

Where is the location of a depicted scene in a query image? Firstly, this is to help curators to automatically annotate a given object in the repository. Secondly, the attributes that a computer learns for this task may reveal the discriminative features over location and invariants over time. The latter can be a subject for further analysis by humanities researchers, e.g. to study the dynamics of the appearance in buildings over time or in a specific area. Hence, the research question for humanities is to study the causality of these discrepancies, where the related question for the computer scientist is how to visualize the learned discriminative attributes [18, 19]. Consistently, is it possible to combine these two functionalities to serve both the curators and the humanities researchers? In the following we introduce a case study on Amsterdam and the available and potential databases for the location prediction task and try to address this question.

4.1 Data Visualization

There are vast amounts of digital records in art and heritage that are well-documented, but in general they are far from being sufficiently analysed. Currently, there is a variety of advanced search tools that help humanities researchers to access these resources neatly, surpassing the hassle of old-school library search. In general scale, Google search platforms are among the most popular, that cover wide ranges of publicly available data. Recently, CLARIAH, a specific platform for humanities and art researchers has been developed to facilitate search queries and to unify different databases in the Netherlands, which is anticipated to be scaled at the international level [20]. Despite the improving situation, important questions remain: Where is the bottleneck in information retrieval? Why is it still difficult to analyse all these data? Is there anywhere that pattern recognition and computer visualisation can help researchers to better process these valuable resources? We believe human vision is most powerful to analyse incredible amounts of information in an amazingly efficient way and computer visualisation could be an auxiliary tool to extend this capability. Therefore, data visualisation is one of the most promising computer-aided means that researchers would benefit from. The idea behind the data visualisation is to ease the interaction with data where more relevant information are clustered together.

4.2 Object Detection and Recognition

Moreover, time consuming and relatively tiring tasks like identifying objects of interest for specific research could be conducted by machines. For example, the identification of buildings with brick facades or thatched roofs in a city or region is rather trivial for a computer as the data is available in various street view applications. Specific details of facades could also be investigated in regard to their correlation to other non-architectural phenomena, as long as this information is available in a digital format. These are but a few examples for the benefits of an automatic object recognition, e.g., in buildings and architectural forms. The nature of the task is not trivial and has not been solved in the computer vision community; however, there is a good chance to soon reach a point where the above mentioned aspects become common tasks for trained machines. This will tremendously change the way architectural historians are conducting research. It is rather obvious that the possibility to investigate phenomena of millions of objects and on a global scale will open up new fields of research within the discipline. Be it global networks of oil-related architecture, architecture of power or poverty, informal urbanism or correlations of architecture and economy, politics or culture—a shift to quantity calls for novel theories and a reconsideration of methods. Concerning the latter, we regard it as essential to not only ask computer scientists for help, but to create research projects that understand interdisciplinarity as a type of research that tackles issues for everyone involved. We believe that jointly formulated, common research questions will facilitate a collaboration that provides motivation for the members of all fields involved. This requires the ability to compromise but also results in new views on a disciplines' objects and methods. That aspect should be taken into consideration, when the rising intelligence of machines is regarded as a thread to humanities and humanities' values. The term "intelligence" has been well-disputed and it remains a matter of perspective if algorithms can be considered intelligent or not. The same goes for the term "learning", that is being contested in regard to its application to machines [21]. Machines—or to be more precise: networks—learn basically by adapting their processing to the valuation of their output in regard of a certain input. Much like e.g. a child learns to distinguish animals or a language. Experience results in knowledge. If there's a preference to use the term "training" in regard to deep learning, we should be careful to reserve a specific term to humans just because it seems inadequate for a machine to learn. This could lead to dangerous underestimations. The historian Yuval Harari points out that we are at a changing point in history and that we are coming closer to reconsidering humanism itself as we have no reason to claim several human features, such as consciousness, as exclusive [22]. However, as we have to adapt to new technologies and methods, a core ability of the humanities researcher will remain uncontested, at least for the foreseeable future: critical reflection. Therefore, we should be open for critically adapting new technologies, and need to establish literacy in fields not yet regarded as closely related disciplines.

In the case of ArchiMediaL, the close cooperation quickly revealed that some of the architectural historians' wishes are far less realistic than first assumed. On the other hand, we discovered several possibilities that are rather unrealistic for classical approaches, and thereby slowly brought our naturally divergent research interests insofar together, that there is something for every involved field—architectural history, deep learning and linked data. The aim to use computer technology to improve access to visual representations of e.g. pre-colonial or vernacular architecture turned out to be limited by the data available. Deep learning requires quite large training sets with several images per depicted object and the databases available to us just did not provide this variety.[3] Therefore we decided to step back and to use a well known and

[3]E.g. the TU Delft repository *Colonial Architecture and Town Planning*. URL: http://colonialarchitecture.eu [access: 12 April 2018].

visually well represented environment as a playground to test transdisciplinary cooperation and its opportunities. This does not mean that the tools and knowledge that are under development and already developed are specific to the beeldbank project, but it is a more reliable database to start with due to its completeness. Once the tools have been validated with the beeldbank repository, they will be adjusted and used for other visual historical databases.

4.3 The Beeldbank Case Study

People who are very familiar with a neighborhood often are able to know where a historic picture from some street in that area is taken by recognizing buildings and other visual features, even if some structures are demolished and the businesses, fashion and cars have changed. One of the first challenges that we are dealing with in the ArchiMediaL project is to automatically match the content of historical images with recent “Street View” images (SVIs). SVIs come with various metadata, like a timestamp and geo-data. By visual matching historical images with SVIs successfully, the geo-coordinates from the SVIs will be inherited by the matched images. In order to start this challenging endeavour, a lot of data is required. As mentioned earlier, the artificial neural networks for visual content analysis need to be trained, and they often only perform well when a large amount of training and evaluation data is available. To achieve this goal, we need to find an urban area (e.g. a city) for which both a rich source of SVIs and historical images are available and where we can find people who are familiar with the area and who can be incentivised to manually create training and evaluation data.

For this reason we selected Amsterdam in The Netherlands, since: (1) the city archives provide access to beeldbank [23], a large publicly available online database of historical images of streets and buildings, (2) the area of the city is widely covered by various Street View approaches (e.g. Google Street View and Mapillary) and (3) the authors of this article are geographically closely connected to organisations that can help with finding people for manual annotation and curation. To find the locations of the buildings in the historic photographs, a two-stage approach is pursued: firstly, linking and comparing any available metadata accompanied with the respective historical image to narrow down the geographical area where the image is taken. Secondly, applying automatic visual content analysis on the sets above to find the exact matches.

4.4 Annotating Historical Images with Linked Open Data

If we are able to derive the exact geo-location and the timestamp of a historical image, we will need to transform it into URI’s to match the Linked Open Data used by e.g. the Dutch National SDI (PDOK)—a central facility for unlocking geo-datasets of

Fig. 1 Historical image from the Beeldbank database depicting a street corner in the centre of Amsterdam [24]

national importance, like cadastral information. After that we can easily look up all relevant information about buildings, for example the intended function, its floor-plan, age, address, neighbourhoods, etc. Every image in the beeldbank database has meta-data containing information about the type of figure (e.g. drawing, map, photo, author, date, title, description).

Figure 1 contains a historical image from the beeldbank dataset depicting a street corner in the centre of Amsterdam.

A part of the available metadata is being transformed into Linked Data format by the City Archive of Amsterdam, for example meta-data for Fig. 1 is available via their endpoint [25]. The location depicted in the image can be derived by humans reading the description "*De zijgevel van Brouwersgracht 137 en links Brouwersgracht 208-212*". For a Dutch speaker it is easy to conclude that the building on the right side in the image is on the "Brouwersgracht" street, number 137. And a simple lookup with the whole title in Google Maps [26] matches this building, however that is not always the case with other titles. Luckily, for most titles from beeldbank our heuristic worked and we got a geo-coordinate 'close' to the buildings depicted in the image:

https://www.google.com/maps/place/Brouwersgracht+137,+Amsterdam

Geo-coordinate: 52.3810575,4.8846276

This is good enough for this first stage, since we at least narrowed down the search space.

Since we now have the geo-coordinates we can prepare for the second step by gathering the StreetView images around the found coordinates. The Mapillary API

[27] provides various ways to gather street-view images, for example all images within a given radius from a geo-coordinate, a bounding box, the *n* nearest images for a geo-coordinate etc.

This example fetches the 100 images in a radius closest to the given coordinate 52.3810575,4.8846276:

https://a.mapillary.com/v3/images/?closeto=4.8846276,52.3810575&radius=100&per_page=100&client_id=<CLIENT-ID>

(*note that in order to try this link, one has to register for a Mapillary account*)

Which results in:


```
type:                "FeatureCollection"
features:
  0:
    type:            "Feature"
    properties:
      ca:            0
      camera_make:   "Trimble"
      camera_model:  "Trimble TMX"
      captured_at:   "2016-08-15T08:44:53.268Z"
      key:           "yyHfvTjH7HEd_kRbBRxr2A"
      pano:          true
      user_key:      "jAve28suG8-ezvTWVkNHag"
      username:      "amsterdam"
    geometry:
      type:          "Point"
      coordinates:
        0:           4.886246514513118
        1:           52.38147442251949
  1:
    type:            "Feature"
    properties:
      ca:            0
      camera_make:   "Trimble"
      camera_model:  "Trimble TMX"
      captured_at:   "2016-08-15T08:44:54.648Z"
      key:           "qHC3Ycz7oi2jjbwj8-s9jA"
      pano:          true
      user_key:      "jAve28suG8-ezvTWVkNHag"
      username:      "amsterdam"
```


The images can be fetched by taking the *key* property and compose a URL via the Image Source URL pattern. For example, the yyHfvTjH7HEd_kRbBRxr2A key results in (Fig. 2).

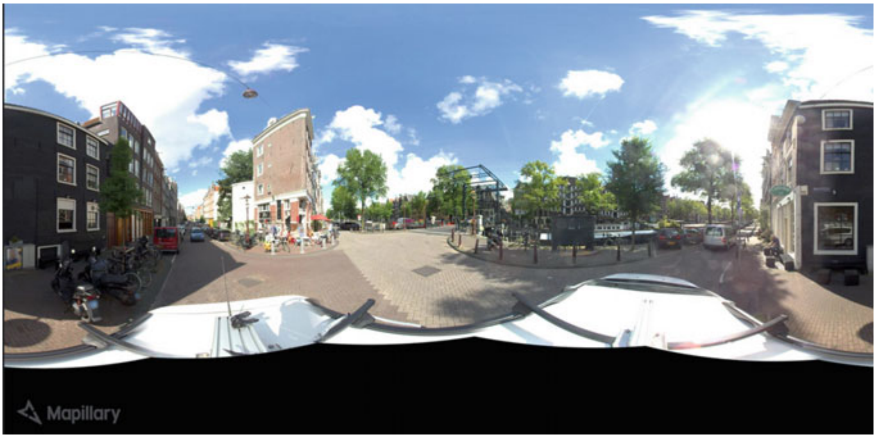

Fig. 2 Street view image retrieved via the Mapillary API for a given radius and geo-coordinate [28]

4.5 Beeldbank Data Visualization

The beeldbank data repository contains 300k+ images of Amsterdam including indoor and outdoor scenes, objects and people. The representation manifest through photography, sketches, paintings, maps etc. The first step to understand and evaluate a database is to visualise the content. For this purpose, we used a network pre-trained on Pittsburgh google street view dataset [R. Arandjelovic 2012] for representing the images in the beeldbank dataset. A small subset of the data repository is clustered and plotted in Fig. 3 to visualize the objects in the dataset. This will give a tangible representation of the beeldbank images for further investigation.

4.6 Image Matching of Historical and Street View Images

The aim of the historical image matching is to find a common built form across images in the beeldbank and the SVIs collected for that registered image, i.e. cross domain image matching task. The source domain is the SVIs which contains label (geo-tagged) information and the target domain is the beeldbank data repository containing coarse annotation (mostly up to the street level). The challenging computer vision task is to train a neural network for the image matching task on the domain with given labels (SVIs) and then transfer the knowledge to the domain without much annotated data, which is the beeldbank repository in our setup. To elevate the domain disparity, transfer learning algorithms [29] are required, which is the subject of research for this task. A further step is to align images from the two domains to create a 4D map of Amsterdam streets with historical images on top of the available SVIs. The procedure for collecting the SVIs are explained in the following:

Fig. 3 Visualization of a subset of the beeldbank repository

4.7 Revealing Discriminative Features in the Amsterdam Neighborhoods

The distinguishing research element here in relation to the conventional geo- localization task (image matching) is to not only look at the accuracy or precision of the algorithm but to understand what features are learnt by a computer to perform the localization task. What are the visual clues that a computer learns to discriminate the location of a scene? [30] Can we find visual attributes that represents regional features in a city like Amsterdam?

We proposed the classification task as a starting point to approach this problem, i.e., given a query image, the algorithm assigns one regional district to the object as a prediction of the location. The districts (classes) are comprehensive and disjoint for the whole city of Amsterdam. The research question to be answered here is if this problem statement for localization serves better for the humanities researchers to understand the underlying features of changes across the city? Does classification as a form of supervised data clustering discover higher level of discriminative semantic features? We aim to experimentally justify this hypothesis with visualising the learned features at the higher level of a deep network trained on the Amsterdam SVIs. The training data are collected to be street view images labeled by the name of neighborhoods in Amsterdam. The classification task is to assign the name of the neighborhood to the query image. The research area is to visualise the learnt attributes for this task and observe the visual discriminative elements in different regions in Amsterdam. This would serve as an attempt to make interpretable classifiers for humanities researchers that can interact with the machine.

4.8 Prospect

The automatic detection of buildings and architectural details in visual representations is helpful to make a large amount of non annotated images available for search and research. The case study on Amsterdam explains basic issues and promising approaches to achieve the recognition of architectural image content. Once established, the challenge will be to apply the knowledge and the technique to sets of images from lesser investigated contexts, e.g. vernacular architecture, informal settlements or pre-colonial buildings. Here, the approach outlined in this paper can provide a valuable method to (1) get better access to images of these lesser investigated fields of architecture and to (2) provide information on images of these architectures that does not contain meta-tags or annotations. On the long term this will allow for a more global and balanced writing of architectural history. Moreover, the analysis of discriminative features of neighborhoods will contribute to a better understanding of the correlation between built form and non-architectural factors. This can be achieved by combining e.g. data on local income, housing prices, types of businesses etc. with changing features of local facades e.g. state of maintenance or the amount of satellite antennas. We expect, that the results brought to architectural history by quantitative methods will reveal several unknown aspects within the field and thereby facilitate novel qualitative investigations of built form.

References

1. Löffler, Beate; Carola Hein; Tino Mager. „Searching for Meiji-Tokyo: Heterogeneous Visual Media and the Turn to Global Urban History, Digitalization, and Deep Learning“, in: *Global Urban History*, 20 March 2018. URL: https://globalurbanhistory.com/2018/03/20/searching-for-meiji-tokyo-heterogeneous-visual-media-and-the-turn-to-global-urban-history-digitalization-and-deep-learning/ [access: 12 April 2018].
2. Berners-Lee, Tim, James Hendler, and Ora Lassila. “The semantic web.” Scientific American 284.5 (2001): 34–43.
3. Bauer, Florian, and Martin Kaltenböck. “Linked open data: The essentials.” Edition mono/monochrom, Vienna (2011).
4. Wick M, Vatant B, Christophe B. Geonames ontology. URL http://www.geonames.org/ontology. 2015. Here: http://www.geonames.org/2759899/about.rdf [access: 20 March 2018].

5. The Getty Research Institute: *Art & Architecture Thesaurus® Online*. URL: http://www.getty.edu/research/tools/vocabularies/aat/index.html Here: http://vocab.getty.edu/page/aat/300140808 [access: 20 March 2018].
6. https://www.w3.org/RDF/.
7. Isaac, Antoine, and Bernhard Haslhofer. "Europeana linked open data–data. europeana. eu." Semantic Web 4.3 (2013): 291–297.
8. Ronzino, Paola, Achille Felicetti, and Sara Di Giorgio. "Integrating Archaeological Datasets: the ARIADNE Portal." AI* CH@ AI* IA. 2016.
9. Meroño-Peñuela, Albert, Ashkan Ashkpour, Marieke Van Erp, Kees Mandemakers, Leen Breure, Andrea Scharnhorst, Stefan Schlobach, and Frank Van Harmelen. "Semantic technologies for historical research: A survey." Semantic Web 6, no. 6 (2015): 539–564.
10. Van Aggelen, Astrid; Hollink, Laura; Kemman, Max; Kleppe, Martijn; Beunders, Henri, The debates of the European Parliament as Linked Open Data, Semantic Web, vol. Preprint, no. Preprint, pp. 1–10, 2016. https://doi.org/10.3233/sw-160227.
11. Shadbolt, Nigel, Kieron O'Hara, Tim Berners-Lee, Nicholas Gibbins, Hugh Glaser, and Wendy Hall. "Linked open government data: Lessons from data. gov. uk." IEEE Intelligent Systems 27, no. 3 (2012): 16–24.
12. Schreiber, Guus, et al. "Semantic annotation and search of cultural-heritage collections: The MultimediaN ECulture demonstrator." Web Semantics: Science, Services and Agents on the World Wide Web 6.4 (2008): 243–249.
13. Victor de Boer, Jur Leinenga, Matthias van Rossum and Rik Hoekstra. Dutch Ships and Sailors Linked Data Cloud. In Proceedings of the International Semantic Web Conference (ISWC 2014), 19–23 October, Riva del Garda, Italy, 2014 (p. 229–244).
14. Niels Ockeloen, Antske Fokkens, Serge ter Braake, Piek Vossen, Victor de Boer, Guus Schreiber, and Susan Legene. BiographyNet: Managing Provenance at multiple levels and from different perspectives. 3rd International Workshop on Linked Science 2013—Supporting Reproducibility, Scientific Investigations and Experiments.
15. http://www.clariah.nl [access: 15 March 2018].
16. Relja Arandjelovic, Petr Gronat, Akihiko Torii, Tomas Pajdla, and Josef Sivic. NetVLAD: CNN architecture for weakly supervised place recog- nition. CoRR, abs/1511.07247, 2015.
17. Albert Gordo, Jon Almazan, Jerome Revaud, and Diane Larlus. Deep image retrieval: Learning global representations for image search. CoRR, abs/1604.01325, 2016.
18. Matthew D. Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. CoRR, abs/1311.2901, 2013.
19. Jason Yosinski, Jeff Clune, Anh Mai Nguyen, Thomas J. Fuchs, and Hod Lipson. Understanding neural networks through deep visualization. CoRR, abs/1506.06579, 2015.
20. The Common Lab Research Infrastructure for the Arts and Humanities (CLARIAH), URL: https://www.clariah.nl/en/.
21. Wolfgang Coy and Gertraud Koch in a public discussion at *Embedded Digitalities* in Basel, 5 April 2018.
22. Harari, Yuval Noah. *Homo Deus – A Brief History of Tomorrow*. London: Vintage, 2017. pp. 136, 137–40, 361–2, 462.
23. The beeldbank archive of historical images from Amsterdam, The Netherlands. https://beeldbank.amsterdam.nl/.
24. https://beeldbank.amsterdam.nl/beeldbank/weergave/record/?id=B00000013972.
25. https://data.adamlink.nl/saa/Beeldbank/table?subject=http%3A%2F%2Fbeeldbank.amsterdam.nl%2Fafbeelding%2FB00000013972.
26. https://www.google.com/maps/search/De+zijgevel+van+Brouwersgracht+137+en+links+Brouwersgracht+208-212/@52.3815462,4.88439,17z/data=!3m1!4b1.
27. https://www.mapillary.com/developer/api-documentation/.

28. https://d1cuyjsrcm0gby.cloudfront.net/yyHfvTjH7HEd_kRbBRxr2A/thumb-2048.jpg.
29. Mei Wang, Weihong Deng."Deep Visual Domain Adaptation: A Survey", School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China., arXiv:1802.03601v3 [cs.CV] 24 Apr 2018.
30. Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros. "What makes Paris look like Paris?" ACM Trans. Graph., 31(4):101:1–101:9, July 2012.