Article

Digital Excavation of Mediatized Urban Heritage: Automated Recognition of Buildings in Image Sources

Tino Mager * and Carola Hein

Faculty of Architecture and the Built Environment, TU Delft, 2628 BL Delft, The Netherlands;
E-Mails: b.t.mager@tudelft.nl (T.M.), c.m.hein@tudelft.nl (C.H.)

* Corresponding author

Abstract
Digital technologies provide novel ways of visualizing cities and buildings. They also facilitate new methods of analyzing the built environment, ranging from artificial intelligence (AI) to crowdsourced citizen participation. Digital representations of cities have become so refined that they challenge our perception of the real. However, computers have not yet become able to detect and analyze the visible features of built structures depicted in photographs or other media. Recent scientific advances mean that it is possible for this new field of computer vision to serve as a critical aid to research. Neural networks now meet the challenge of identifying and analyzing building elements, buildings and urban landscapes. The development and refinement of these technologies requires more attention, simultaneously, investigation is needed in regard to the use and meaning of these methods for historical research. For example, the use of AI raises questions about the ways in which computer-based image recognition reproduces biases of contemporary practice. It also invites reflection on how mixed methods, integrating quantitative and qualitative approaches, can be established and used in research in the humanities. Finally, it opens new perspectives on the role of crowdsourcing in both knowledge dissemination and shared research. Attempts to analyze historical big data with the latest methods of deep learning, to involve many people—laymen and experts—in research via crowdsourcing and to deal with partly unknown visual material have provided a better understanding of what is possible. The article presents findings from the ongoing research project ArchiMediaL, which is at the forefront of the analysis of historical mediatizations of the built environment. It demonstrates how the combination of crowdsourcing, historical big data and deep learning simultaneously raises questions and provides solutions in the field of architectural and urban planning history.

Keywords
artificial intelligence; automated image content recognition; big data; computer vision; crowdsourcing; image repositories; urban heritage

1. Introduction

Comprehensive digitization and the dissemination of visual material via the internet have contributed to a scholarly concentration on the image, which in the 1990s became the focal point of several reorientations within art history (Alloa, 2015). They have also led to a visual turn in the history of architecture and urban planning. More and more digital visual data are produced, but the generation of new images is not yet balanced with the capacity of computers to read this data. While facial recognition has been rapidly advancing, 'building recognition' does not

even exist. Reading urban and architectural images remains a task for humans. Given the number of available images, relatively few are ever analyzed. As the amount of available data increases constantly through digitization, humanists need to critically reflect on their reading and interpretation of digital sources and strategize how best to explore the formerly unavailable heterogeneous and interconnected visual datasets. Preliminary reflection on research questions and methodologies is necessary to avoid the shortcomings of archaeologists of the past, many of whom were initially driven by the desire to literally find gold and other treasure. As they focused on excavating palaces and temples, they created a skewed understanding of the past and, as a result, today's knowledge of working-class housing in ancient cities or in suburban developments is more limited than the knowledge of the buildings of the elite. Similarly, urban planners need to critically evaluate the data that they are working with and the plans that they assess.

Urban and architectural historians need to go beyond their traditional, usually limited visual material—archival documents, physical collections or books. As they dig into a big new set of imagery—electronic repositories, crowdsourcing or web-scale datasets—they need to refine their theories and methods. When dealing with huge and unfamiliar data sets, questions will arise that go beyond the traditional hermeneutic reading of text and images. They must understand code as a cultural practice and learn to see qualitative data as the result of abstract 'technocratic' sorting that relies on established interpretation systems. Innovation in computer technology, both in crowdsourcing and in AI creates opportunities and challenges for urban and architectural history, notably the recognition of visuals in vast archives. Crowdsourcing metadata for historical images is an urban planning issue closely related to issues of communication, mediatization and urban futures.

This article explores ArchiMediaL's research into the development of image recognition tools. It explains how the project uses crowdsourcing and AI technology in combination and what this means for humanities-based archival research as a foundation for design. The combination of these different technologies allows for diverse approaches. Crowdsourcing can be both socially motivated and technologically important. It helps produce a kind of swarm intelligence studying, understanding and shaping cities based on their collective nature (Rossi, 1982, pp. 5, 24, 86). Following a more general reflection on the role of visuals in architecture and planning, the article explores the mediatization of visuals and the integrative research conducted in the context of the project ArchiMediaL.

2. Preliminary Reflections on the Nature of Visuals in Urban Planning

Visuals play an essential role in architecture and urban planning: prospective drafts and plans convey what is yet to be built, photographs show places of the past and capture lost sites for posterity. Since the 19th century, photography has facilitated the realistic depiction of the real world including the built environment; since the end of the 20th century digital tools have brought new possibilities to the conception and representation of future-oriented design. Computer-aided design (CAD) has become a standard in architectural and urban design practice and digitally generated images have become a part of the real world. Despite many decades of research in computer technology and enormous progress, it has only recently become possible to take the opposite path: to recognize and analyze the real world and its images with the help of computers. But, researchers, librarians and archivists still must spend countless hours identifying buildings in the huge amount of analog and digital data available. Automated recognition of buildings in historical photographs has yet to become reality, but developing this capacity can open up previously unseen image material that is important for research.

Visual information plays an essential role in the study of urban planning and its history. Urban form and architecture start as ideas that must be translated into matter. Both general urban concepts and urban plans emerge as thoughts based on logical considerations. The design of classical Chinese capitals could be sketched easily as a square with three gateways on each side and a system of crossing roads and regular patterns of gated neighborhoods. Grid cities were also a widespread practice in the Greek and Roman world. Similarly, simple vernacular buildings are often the result of century-old practices without the need to communicate the idea in the form of construction plans. Charles-Dominique-Joseph Eisen's depiction of the primitive hut—which became famous as the frontispiece of Marc-Antoine Laugier's (1755) treatise *Essai sur L'architecture*—shows a romantic idea of early architecture derived directly from nature. A similar type of construction was certainly the reality for many huts and shelters, but any type of building going beyond that required a comprehensible communication of its organization and planning. For thousands of years pictorial representations of urban plans, of buildings, facades and room layouts have been used as tools for this purpose. They not only facilitate the recording of complex thoughts, but also allow for their precise exchange. Visual representations of cities even date from the 7th millennium BC (Rochberg, 2014, p. 14) and early building floor plans have been preserved from the 3rd millennium BC. Like today's visualizations, the ancient ones also served the purpose of spatial representation and essentially pursued two communication goals: the prospective communication of space still to be built and the retrospective recording of already existing space.

Technical drawings as well as town plans are the product of experts. Although there is a need for research, it can be assumed that innovations in visualization technology ultimately contribute to the realization of new types of cities and buildings. To give just one example,

the Renaissance saw a wealth of innovation in the visualization of buildings, such as the central perspective or the veduta, and completely new forms of urban design and architecture inspired by antiquity, but enriched by new possibilities, came into being. The graphic representation of spatial ideas has been further refined over time and some designs have become autonomous works in their own right. Without ever being built, architectural and urban drawings have influenced the built environment and developed into an element of collective imagination and inspiration. Most notable are the well-known neoclassical visions of Étienne-Louis Boullée and Claude-Nicolas Ledoux, which anticipated later aesthetic developments and were even carried out centuries after their conception (Aire du Jura, Arlay, France). The power of urban and architectural visions captured on paper is also exemplified through the works of 20th century designers. Visions for cities and buildings that were never built had a great influence on architecture and urban planning, for example Antonio Sant'Elia's primarily graphic work, Frank Lloyd Wright's plans for Broadacre City, Le Corbusier's Plan Voisin, or the urban and architectural fantasies of Archigram and Superstudio.

Technical innovations in the representation of the world have always been put at the service of architecture. Photography has been used since its beginnings to document the built environment and it has helped disseminate both architectural heritage and the latest architectural fashions. Noël Paymal Lerebours' *Excursions Daguerriennes* (1840) and the *Missions Héliographiques* initiated by Prosper Mérimé in 1851 are impressive examples of the early use of photography in the recording of historical monuments (Figure 1). They can be counted among the earliest sources that reliably represent the built environment. In the 20th century, Walter Gropius was a pioneer in the use of photography as a messenger for architectural ideals, he even had pictures retouched to make the buildings on them appear more progressive (Eckstein, 1994, p. 29). Numerous publications exist on architecture and photography. Some focus on major works, others on vernacular architecture. Their focus is on the analysis of the aesthetics of the image, the format or the content (Colomina, 1987; Lichtenstein, 2018). They are not used as tools for computer analysis. To what extent novel forms of representation have influenced the development of the built form requires further research, but some innovations of the late 20th century clearly show this impact. CAD applications that had been developed by French and American car companies since the 1960s became affordable and truly three-dimensional in the 1980s (Riccobono, 2014, pp. 35–37). They were relatively late to be introduced to architectural and planning offices (Corser, 2012, p. 13) and left obvious traces in the built environment from the late 1990s onwards. Peter Cook and Colin Fournier's Kunsthaus in Graz or Frank Gehry's Guggenheim Museum in Bilbao are two examples of buildings that were only made possible by advanced CAD applications.

In the meantime, CAD has gone mainstream as the operation of programs has become more intuitive and people have become accustomed to using input devices. For urban planning, a broad variety of digital tools has similarly been created and extended from technological instruments to means of inclusive and participatory planning (e.g., the AvaLinn mobile app and the 3D visualization tool Earth Autopsy). At present, it is the complete immersion in virtuality that promises new possibilities for the design and visualization of what is to be created. Virtual reality (VR) is on the threshold of becoming a ground-breaking environment for urban design
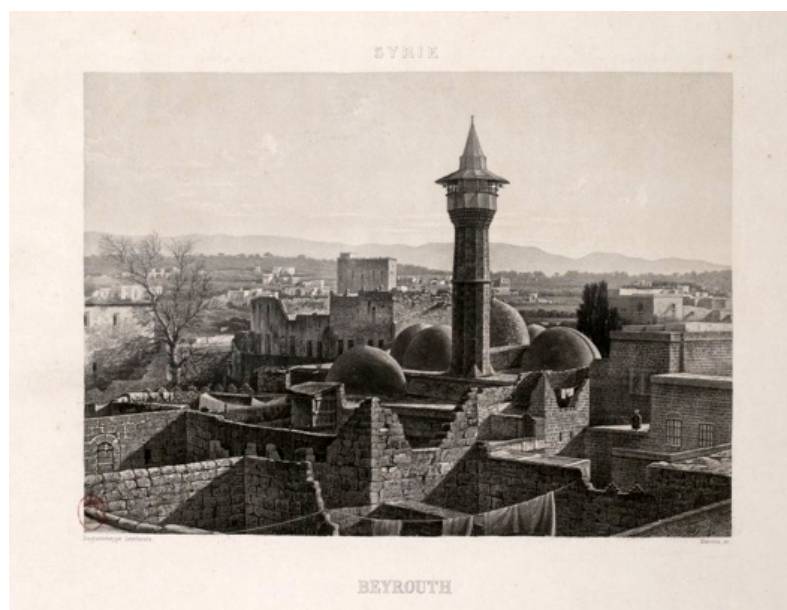
Figure 1. Cover of *Excursions Daguerriennes*, Beirut, 1840. Since daguerreotypes are not replicable, Lerebours had them lithographed so that they could be printed in series. Source: Lerebours, 1840.

and architecture as VR labs at leading research institutions demonstrate (e.g., the VR-Lab at TU Delft Faculty of Architecture and the Built Environment). In the VR-lab, designers become part of the world to be created, shape and move construction elements with their own hands and freely change the scale of the virtual world—at the same time, this almost playfully created digital model is the executable file for the production of real-world matter whereby any abstraction in the design process disappears. This also applies to the presentation of buildings and cities that have not yet been built: they are conveyed in such realistic renderings that they appear like photographs of already existing places, if one does not stumble upon the excessive perfection of the world depicted.

While the visual communication of architecture and visual planning tools are at the center of professional and academic attention regarding digital approaches, the analysis of mediatizations is largely carried out by conventional means in a painstakingly slow process that requires researchers with extensive background knowledge to carefully examine and annotate depictions of urban and architectural form. As libraries and repositories around the world fill in metadata on these images, they do so in different languages and styles, often making it impossible to find the same image if the metadata is not identical. In the process, millions of drawings, plans and photographs are lost to research. They are or will be made available digitally, but their content cannot be specified without great effort beyond the annotations made during digitization, insofar as there are no automated processes for this. Even as archives around the world are digitized, the knowledge that they include is not made available through metadata. For the history of architecture and urban planning, a challenge is to make these mediatizations usable for historical analysis and, in doing so, to also focus on those objects that have so far been insufficiently considered. Furthermore, the expensive digitization of visuals needs to be contextualized by urban and architectural historians. The databases that are currently digitized and that serve as the foundation for research represent only a small fraction of the built environment. Databases reflect particular collection strategies. A database on colonial architecture, for example, will only include that type of structure. A computer trained on these sources may assume that all architecture is colonial, which is certainly not the case. Computerized practices therefore need specialists from the humanities with sufficient computer knowledge to recognize opportunities and challenges and to translate and to apply them in the history of architecture and urban planning.

3. Deciphering Mediatizations

As mentioned above, digitization can facilitate access to a huge stock of visual material. Furthermore, digital cataloging is essential to inform about the existence and (virtual) location of the material globally and effectively.

Digital catalogs such as Europeana, the German Digital Library or the Digital Public Library of America offer access to millions of digitized materials, including historical architectural and urban images, and specific repositories such as Colonial Architecture & Town Planning (colonialarchitecture.eu) contain enormous amounts of visual resources. Their digital availability—unless legally constrained—is a blessing for research, as these sources can be viewed and analyzed without much effort (and if required, the digitized object can still be physically inspected in the corresponding archive). Search filters enable researchers to easily find photos of certain objects and places or the work of particular illustrators or photographers. But what exactly is found via the search terms?

The names of objects, places and persons as well as the keywords are ultimately metatags, which have been assigned to the images at some point by somebody—and, ideally, systematically and correctly. They thus represent a historical knowledge that was attributed to the image and its content on the basis of external sources or through the specific expertise of a beholder. The keywords associated with the images can only be updated very slowly if this is done by hand and can therefore hardly extend to new terms that are important for current research. Any new findings regarding the image content can only in certain cases be included in the metadata. Far more consequential, however, is the fact that faulty, incorrect or missing meta tags mean that the image material may not be found at all. This is particularly consequential for the many images whose content is largely unknown or not recognized by collectors: Millions of images end up in repositories without any means of searching their content (Figure 2). Unfortunately, this primarily concerns precisely those buildings and areas that have so far received less attention in research (Löffler, Hein, & Mager, 2018). Longitudinal analysis of the built environment is one of the most important sources of knowledge and inspiration for today's planning. Since illustrations in turn are among the most important sources for architecture and urban history, it is of greatest relevance to make accessible those source materials that have not yet been consulted. Researchers and experts can take on this task but can only access a tiny part of the media concerned. In view of the abundance of material, even larger teams would not be able to make any significant stock available. Here, the very technologies that are also driving the production of buildings and cities promise to help: Digital technologies—in this case computer vision and open linked data.

The digital tools for structural analysis, design and planning developed over the last decades have so far been virtually unparalleled by tools that transform the visual analysis of the existing environment and its transformation not only into digital data, but also into comprehensive information. Only a few years ago, computer vision made a great leap forward and promised to enable the recognition and classification of real-world el-

Figure 2. Unknown building on a photo from 1895 in the Colonial Architecture & Town Planning Repository. Source: "Woning, vermoedelijk in Manado," 1895.

ements for a variety of tasks. AlexNet, a convolutional neural network designed in 2012, can be regarded as a ground-breaking step in this development (Krizhevsky & Sutskever, 2012). Since then, the superhuman performance of AI in certain areas of image recognition—as well as other fields—can no longer be doubted. Today, algorithms are fairly reliable in their ability to recognize the face of smartphone owners and they recognize streets, road users and cancer cells and can outperform humans in many areas. Therefore, they also appear promising in the analysis of historical images of the built environment of the past. In a joint venture between four European universities, ArchiMediaL investigates the possibilities of using current information technologies to open up previously unexplored architectural and urban image material for research by developing strategies for automatic image content recognition. The participating architects, architectural and planning historians as well as deep learning and linked open data specialists form a multidisciplinary team that operates at the interface between quantitative and qualitative methods and explores their integration.

4. Integrative Research

The research project started out by exploring the automatic recognition of buildings in historic images by AI. Analogous to automatic facial recognition, buildings are to be recognized and identified. The input objects are historical images of buildings whose contents are localized by a specially designed and trained artificial neural network. The localization allows unique identification. The recognition can be realized for a specific area by providing the computer with many images of already identified, i.e., localized buildings. This training of the network enables the computer to recognize buildings in historical images unknown to it. First, however, the training data

set with several hundred historical images with identified buildings must be created.

Despite the goal of facilitating the opening up of large numbers of images on lesser known or little explored topics, the study must start in an area that is known to the researchers involved and that is also well investigated by architectural and urban history, since the performance and reliability of the algorithm can only be tested if the topic to which it is applied is well known. In this case that meant starting with several hundred thousand photographs of Amsterdam—a city that has been thoroughly researched by urban historians and is easily accessible for the project team—which could be obtained from the city's image archive (Beeldbank). The more than $400,000$ images, covering the period from the mid-19th to the end of the 20th century, contain daguerreotypes, black and white as well as color images, and have very different resolutions. They are mostly annotated, although in varying degrees of completeness—some annotations even include addresses or neighborhoods. This is also a good basis for creating a high-quality training dataset. In addition, the building stock of the Amsterdam city center has changed relatively little during the period covered by the images and is at the same time well documented. These conditions are favorable for the development of algorithms for the automated recognition of buildings in historic images.

The geolocation of buildings can be determined using a reference system that contains images of the building facades and their locations. This information can be obtained, for example, from online map services that feature facade images like Google Street View or Mapillary. Using the dense network of geolocated and oriented 360° images of Mapillary, covering most of Amsterdam's streets, it was possible to extract the location of today's facades and thus create a visual reference for the historic images (Figure 3). In order to enable an automated recog-
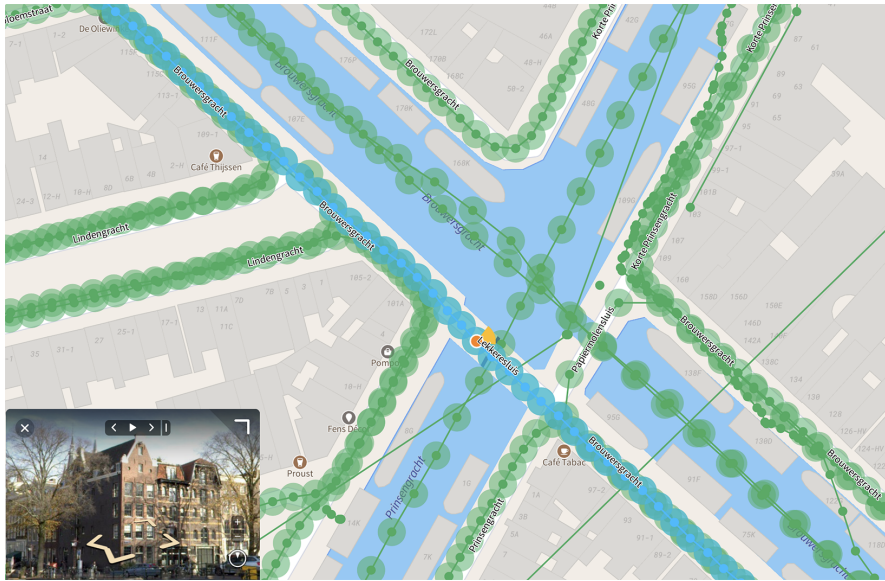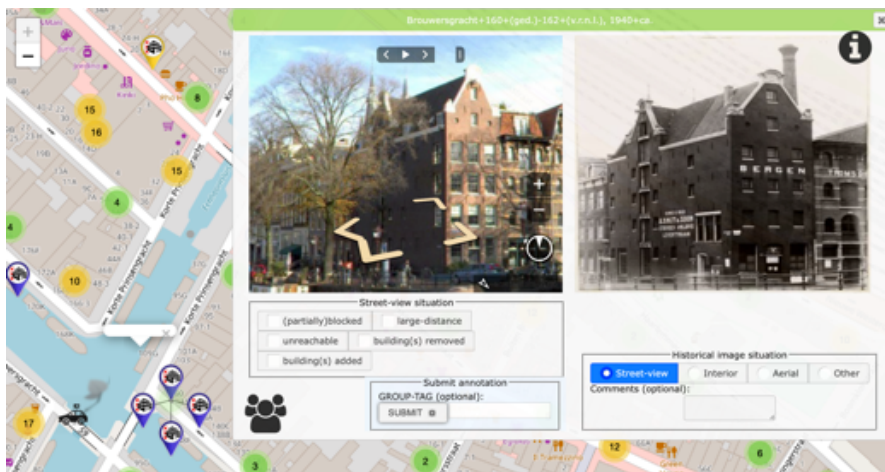
Figure 3. Amsterdam, Brouwersgracht 160 on Mapillary. The green dots in the map indicate the positions of 360° images. Source: Mapillary (2016).

nition through geolocation of the buildings in the historical images via deep learning, several hundred of these images have to be matched with the corresponding facade images from Mapillary. This teaches the algorithm to match a geolocated image with a building in a historic photograph. After training, the algorithm can then also locate buildings that were not included in the images of the training set. A large and well compiled training set increases the recognition performance.

In order to build the training dataset and to be able to evaluate the performance of the algorithm, it is necessary to exactly match buildings from the historical images with buildings in the Mapillary images. This task must be first performed by human beings through crowdsourcing and must meet ergonomic requirements. However, it must also be possible to perform it in such a way that it connects well with digital information processing. The er-

gonomic requirements refer to a simple and fatigue-free method of data entry. In addition, it is necessary to involve many people, as performing more than a few dozen matches quickly becomes a tedious and tiring task due to its repetitive nature, even if the crowdsourcing tool provides good ergonomics. In this type of crowdsourcing, the challenge is to design the matching process in such a way that it is both playful and intuitive and can be carried out without data entry errors. In order to meet these requirements, an online tool was developed that enables users to determine the scene corresponding to the historical photo by simple navigation in the virtual street space. Horizontal rotations and movements as well as zooming can be used to easily find a largely similar image section and register it as a match by a click (Figure 4). It is possible to add comments and report on possible problems— e.g., participants can indicate whether a place is inacces-

Figure 4. A historical picture of the Brouwersgracht 160 in Amsterdam from the 1940s and the corresponding street view scene in the submission form of the crowdsourcing tool. Source: ArchiMediaL, 2020.

sible or is of a different type than an outdoor street scene. They can also indicate whether buildings have been removed or added or whether their visibility is obstructed.

The administrators receive an illustrated list with the matches and can easily check their validity and also the performance of the contributors. A special login syntax makes it possible to distinguish between different groups of contributors while keeping their identity anonymous. This allows the performance of street scene recognition between e.g., architecture students, IT specialists or historians to be analyzed and compared. Targeted expansion of the user groups could allow statements as to whether local users can grasp the historical situations of their city more quickly and better than users who are unfamiliar with a city or who come from an area with a completely different urban development. This investigation is currently ongoing as a priority was the creation of the complete first training set. To date, more than 1500 matches have been made. An initial analysis shows that approximately two thirds of the assignments are valid and indicate the current location of the historic building. In addition, a smartphone app is being developed that allows users to compare historical images from their own collection with today's scenery in Mapillary and to submit their images. The resulting geolocation of the historical images, which even includes the orientation of the facade, supplements existing repositories with precisely located images from private collections. These images will help to complement historical image collections and contribute to a more complete picture of the past of specific locations.

5. Automated Image Content Recognition

The data set generated by crowdsourcing makes it possible to train a convolutional neural network in such a way that it is ultimately able to recognize buildings on the remaining historical Beeldbank images that were not part of the training data set. Despite the progress made in pattern recognition in recent years, and especially in the recognition of buildings (e.g., Amato, Falchi, & Gennaro, 2015; Andrianaivo & Palma, 2019; Gada & Mehta, 2017), this is an unprecedented task that first requires basic research. Even when it comes to on site recognition of buildings, based on current appearance, "the literature on how to develop effective neural networks to detect architectural features is still limited, as well as the availability of architecture-related datasets" (Andrianaivo & Palma, 2019, p. 77). However, AI-based recognition of buildings in historical images poses an additional challenge because the buildings have changed over time and are visually not identical to images from a georeferenced reference set (Mapillary). Moreover, it is a new challenge to let AI recognize objects in images from different domains. While a smartphone only has to recognize a single face that always appears frontal and is taken with the same camera, historical pictures of buildings are very different, as they show the object to be recognized from different angles and at different focal lengths; they may be blurred or sharp, over- or underexposed, black and white or in the tones of old color films. AI requires large training data sets to learn new tasks. To assure the success of the project, we opted to pair the AI task with the crowdsourcing mentioned above. This allowed ArchiMediaL to both use and study the knowledge of human observers of the built environment and those of the computer.

For this specific purpose, we designed an age-invariant feature learning convolutional neural network model with an attention aggregation module (for details see Wang & Li, 2019). Buildings can be clearly identified by their address or geolocation. Nevertheless, a location can contain different buildings at different times, and these may themselves be subject to changes such as partial demolition, extensions, additions or renovations. Therefore, the algorithm should ideally be robust against minor changes—and also against partial obstructions in the image, such as trees or cars—but still able to reliably identify particular buildings.

The data resulting from crowdsourcing provides a basis for evaluating the performance of the algorithm. The validity of the image content recognition can be observed through an expert view of the result. This only works for a limited number of findings—a much larger number of exactly matching images is required to provide a more reliable evaluation. The necessary validation and test data sets are also generated by the crowdsourcing method. Errors and distortions, which are also found in human thinking, are thus sometimes transferred to systems of AI (Leavy, 2018). This also refers to the canon of architectural history and urban or planning principles where colonial or gendered perspectives of the built environment may prevail. AI can help open up undiscovered areas of the documented past, but it is limited by the way it is trained. In this case the recognizability of the image content is limited to the areas covered by Mapillary. This means that backyards and private areas, for example, are only included in exceptional cases. Also, recognizability will be limited to those buildings that are still preserved or their adjacent buildings, otherwise there are no visual matches that can be recognized. The performance is still being evaluated and the data set has not yet been published.

6. Mixing Methods

While digital applications for design and visualization purposes are widely used in architecture and urban planning, architectural and planning history are humanities and social science disciplines that operate with different methods. So far, quantitative approaches to the visual analysis of historical representations of buildings constitute new territory. Although there have been numerous approaches to virtually reconstruct historical situations of cities (e.g., the Time Machine network) and progress has been made in 3D scanning and the printing of buildings and parts of buildings (see, e.g., "3D printer used

to reproduce Mauritshuis," 2017), these technologies do not focus on the interpretation of the digitally recorded imagery or structures and thus hardly facilitate the investigation of humanities research questions. The first conference on digitalization in art history, "Computers and Their Potential Application in Museums," took place at New York's Metropolitan Museum of Art in 1968. But even half a century later, digital approaches in the humanities dealing with imagery and space seem to have hardly been explored in depth. Only very few research projects, such as Urban Panorama (North Carolina State University) or Composito (University of Heidelberg), include automated image analysis. One reason for this is certainly the only recently achieved progress in the field of computer vision and the competitive demand from science and industry for these possibilities. Another reason is that efforts at methodological reflection and innovation in the humanities have so far been largely neglected (Hahn et al., 2020).

A central challenge here is the clash between qualitative and quantitative research. Established methods such as source criticism, discourse analysis, hermeneutics and morphological studies have provided a meticulous picture of certain urban planning phenomena and a high standard of analysis. Against this background it seems as if AI-powered automation and big data belong to a different world and may hardly be able to contribute to the intellectual task of pictorial-spatial analysis. But the quantitative approaches should not be seen as competing with the established methods, but rather as extending the existing possibilities regarding access to and handling of source material (Mager & Hein, 2019). The automated recognition of buildings in images can help to identify a large amount of image content, making a wealth of images searchable and easier to navigate as sources for research. It can also contribute to improving the availability of sources for forms of building and settlement that have hardly been considered up to now. As a result, previously less well researched areas such as informal settlements or vernacular architecture can be studied more comprehensively. Ultimately, the automated recognition of buildings in images offers the long-term possibility of identifying all existing photographic sources for a building or location and thus also contributes to creating a more solid basis for answering architectural historical research questions. Which insights can be gained by quantitative approaches will remain speculation until this field is duly investigated. New quantitative directions in the historical sciences reveal astonishing insights (Spinney, 2019). They can also lead the spatial sciences to innovate. This not only refers to the accessibility of source material, but also, for example, to the global distribution of architectural forms and structures or the worldwide analysis of the distribution of different concepts of land-use and zoning (Moosavi, 2017). Until recently, it was not possible to analyze the distribution of architectural or urban forms on the basis of millions of objects and places. The possibility of consulting (historic) big data may well provide opportunities to ask and pursue new research questions. Such questions need to be carefully framed by humanities scholars in light of existing biases—colonial, gendered, or other—and their potential transfer into the digital realm. But since AI uses data to create algorithms, possible bias results from this data even if the bias is not readily apparent. Biases that are present in the language and images used as training are taken over and continued (ALGB-WG, 2017; Koene, 2017). Therefore, in addition to careful programming, data quality is of great importance, and humanities scholars are called upon to think carefully about what can be considered ideal and universally valid. AI and large data could indeed be transformative in the sense that they allow phenomena to be compared on a global scale and over a long period of time. While humanities research addresses complex issues with a limited number of sources and case studies, new technologies can help to analyze a much larger number of sources and also make it easier to analyze visual information. We regard these possibilities as new tools that can lead to new and further-reaching questions, not as research instruments that provide their own interpretation.

In order to be able to take this direction at all, important basic research is needed. This concerns methodological issues as well as the practical experience of research teams consisting of humanities scholars and IT specialists. ArchiMediaL's approach here is to formulate research problems in such a way that they represent a challenge for all disciplines involved and that no science appears as a mere support. This approach begins with building understanding of how research is conducted in other scientific fields, what is interesting and what is possible, what experiences exist with other (auxiliary) sciences or what publication strategies prevail. The attempt to analyze historical big data with the latest methods of deep learning, to involve many people—laypeople and experts—in research via crowdsourcing and to deal with partly unknown visual material has provided further understanding of the new possibilities and the results they can generate. By locating buildings in the more than 400, 000 historical photographs from the Beeldbank archive, they become discoverable and thus accessible for research. In addition, the automated building recognition developed can also be applied to other locations and in the future serve to identify less well researched areas. The basic research carried out represents a breakthrough in the field of computer vision and offers methodological incentives for historical architectural and urban research.

7. Outlook and Conclusion

The findings of the ArchiMediaL project open up new perspectives for planning history in diverse areas. Researchers, politicians and planners can explore 4D reconstructions of the past (e.g., in their respective websites, the HistStadt4D research project or various local Time Machine projects) to increase historical under-

standing, to enrich tourist experiences, or to facilitate planning decisions. A look at the number of images collected in specific areas of the city provides a first step into research. Why are more materials available for some locations than for others? How is this spread over time? How can we complement the available data through crowdsourced intervention? The historical data available from ArchiMediaL can be used to advance community engagement and serve as a hub to collect local stories. Such stories could be complemented by visuals contributed by individual citizens that are not in official archives, but that are needed to complement the existing data with more vernacular elements. Local stories could help create new leisure and tourist locations and open up new themes and directions.

Digital reconstructions can also provide advanced understanding of processes in the past and the role of planning therein. For example, available photos can be connected to specific moments in time when urban plans were made or implemented. As a result, we may be able to study moments of transition and temporalities, scales and perspectives of planning intervention. Assessment of past crises and the intervention of public actors through policy or urban interventions, for example, related to public health events (pest, cholera, yellow fever), environmental disasters (earthquakes, floods, windstorms), or changes in the energy landscape (petroleum revolution), can provide insight into contemporary challenges from global diseases to climate change and the (re-)introduction of sustainable energy sources. Modeling of the past would complement contemporary tools that are aimed at designing the future. Smart city discussions project the future, but they usually do not acknowledge where the current environment comes from. A better understanding of the planning processes that have created our cities can help identify path dependencies and critical junctures. It may allow us to combine spatial and social data about the past to model neighborhoods and entire regions and the intersection of spatial, social and cultural developments.

New research questions can be framed through the availability of such data. Using the ArchiMediaL tool can raise numerous questions. For example, scholars could examine bubbles where more or less data is available raising questions such as: How does the availability of pictorial data from the past overlap with the architectural quality of the building stock, or the socioeconomic composition of its citizens? In the case of Amsterdam, many datasets with spatial information are available in digital form, including ones based on the age of buildings, the number of breeding birds in green areas, climate information (heat, drought, flooding), postwar monumental wall art and land value, to name but a few. The crossing of this data with the visual sources localized within the work of the project allows for the framing of new research questions that investigate the connection between architectural and urban form with phenomena such as property value or gentrification. The expansion to other cities and areas will make it possible to formulate new findings on the basis of a high number of correlations and thus to make more general statements than those that emerge from individual case studies.

The application of AI for historical research is not a mere information technology task. As with any mixed methods approach (Creswell & Plano Clark, 2018), it requires the meeting and communication of different disciplines and profound expertise in the humanities. Interpretation of the past needs careful framing of the available data to achieve meaningful findings: Such a step can only be made through transdisciplinary collaboration among humanities scholars, computer scientists, historians and designers. Moreover, this project has required people to contribute their knowledge, both to create the training dataset and to eventually evaluate the performance of the algorithm. Crowdsourcing can offer an important opportunity for participation—important not only when it comes to identifying past worlds, but also when it comes to involving people in research, integrating their point of view and ultimately awakening their interest in questions of urban history and urban development.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

References

3D printer used to reproduce Mauritshuis. (2017, January 16). *TU Delft News*. Retrieved from https://www.tudelft.nl/en/2017/bk/3d-printer-used-to-reproduce-mauritshuis

ALGB-WG. (2017). P7003—Algorithmic Bias Considerations. *Algorithmic Bias Working Group*. Retrieved from https://standards.ieee.org/project/7003.html

Alloa, E. (2015). Iconic turn: A plea for three turns of the screw. *Culture, Theory and Critique*, *57*(2), 228–250. https://doi.org/10.1080/14735784.2015.1068127

Amato, G., Falchi, F., & Gennaro, C., (2015). Fast image classification for monument recognition. *Journal on Computing and Cultural Heritage*, *8*(4), 1–25. https://doi.org/10.1145/2724727

Andrianaivo, L. N., & Palma, V. (2019). Architecture recognition by means of convolutional neural networks. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *LXII 2*(w15), 77–84. https://doi.org/10.5194/isprs-archives-XLII-2-W15-77-2019

ArchiMediaL. (2020). Image matching crowdsourcing tool, Brouwersgracht 160, Amsterdam. *ArchiMediaL*. Retrieved from http://archimedial.eu/beeldbank/marker-clustering-geojson2.php

Colomina, B. (1987). Le Corbusier and photography. *Assemblage*, *October*(4), 6–23.

Corser, R. (2012). Rereadings: Looking back at recent discourse in emerging technologies. In R. Corser (Ed.), *Fabricating architecture: Selected readings in digital design and manufacturing* (pp. 11–19). Hudson, NY: Princeton Architectural Press.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*. Los Angeles, CA: Sage.

Eckstein, K. (1994). Inszenierung einer Utopie: Zur Selbstdarstellung des Bauhauses in den zwanziger Jahren [Staging a utopia: On the self-portrayal of the Bauhaus in the twenties]. In A. Haus (Ed.), *Bauhausideen. Bibliographie und Beiträge zur Rezeption des Bauhausgedankens* [Bauhaus ideas. Bibliography and contributions on the reception of the Bauhaus idea] (pp. 15–29). Berlin: Reimer.

Gada, S., & Mehta, V. (2017). Monument recognition using deep neural networks. In *2017 IEEE international conference on computational intelligence and computing research* (pp. 1–6). Piscataway, NY: IEEE. https://doi.org/10.1109/ICCIC.2017.8524224

Koene, A. (2017). Algorithmic bias: Addressing growing concerns. *IEEE Technology and Society Magazine*, *36*(2), 31–32.

Krizhevsky, A., & Sutskever, I. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, *25*(2), 1097–1105.

Laugier, M. A. (1755). *Essai sur l'architecture* [An essay on architecture]. Paris: Duchesne.

Leavy, S. (2018). Gender bias in artificial intelligence: The need for diversity and gender theory in machine learning. In E. Abraham, E. Di Nitto, & R. Mirandola (Eds.), *GE '18: Proceedings of the 1st international workshop on gender equality in software engineering* (pp. 14–16). New York, NY: Association for Computing Machinery.

Lerebours, N.-M. P. (1840). *Excursions daguerriennes: Vues et monuments les plus remarquables du globe* [Daguerreotype excursions, the world's most remarkable views and monuments; lithographies]. Estampes et Photographie (ark:/12148/btv1b8458373v). Bibliothèque nationale de France, Paris, France. Retrieved from https://gallica.bnf.fr/ark:/12148/btv1b8458373v/f249.item.r=Excursions%20Daguerriennes,%20Beyrouth.planchecontact

Lichtenstein, T. (2018). *Image building: How photography transforms architecture*. New York, NY: Parrish Art Museum.

Löffler, B., Hein, C., & Mager, T. (2018). Searching for Meiji-Tôkyô: Heterogeneous visual media in times of global urban history, digitalization, and deep learning. *Global Urban History*. https://globalurbanhistory.com/2018/03/20/searching-for-meiji-tokyo-heterogeneous-visual-media-and-the-turn-to-global-urban-history-digitalization-and-deep-learning

Mager, T., & Hein, C. (2019). Mathematics and/as humanities: Linking humanistic historical to quantitative approaches. In L. D'Acci (Ed.), *The mathematics of urban morphology* (pp. 523–528). Basel: Birkhäuser.

Mapillary. (2016). Brouwersgracht 160, Amsterdam. *Mapillary*. Retrieved from https://www.mapillary.com/app/?lat=52.38043128108052&lng=4.8881926757478595&z=17&focus=map&pKey=8UePEDeTX9kO_SxQxtH6Qg&x=0.7220591253875135&y=0.49880793345495045&zoom=1

Moosavi, V. (2017). Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet. *Computer Science*. Retrieved from https://arxiv.org/abs/1709.02939

Hahn, W., Hein, C., Löffler, B., Mager, T., Meyer-Sickendiek, B., Prescott, A., & Schneider, B. (2020). *Mixing methods. An agenda for the humanities in the digital age*. Manuscript in preparation.

Riccobono, A. (2014). *Architectural design in the digital era. Identifying computer influences and new expressive trends in current architecture*. Palermo: University of Palermo.

Rochberg, F. (2014). The expression of terrestrial and celestial order in ancient Mesopotamia. In R. J. A. Talbert (Ed.), *Ancient perspectives: Maps and their place in Mesopotamia, Egypt, Greece, and Rome* (pp. 9–46). Chicago, IL: University of Chicago Press.

Rossi, A. (1982). *The architecture of the city*. Cambridge, MA: MIT Press.

Spinney, L. (2019, November 12). History as a giant data set: How analysing the past could help save the future. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2019/nov/12/history-as-a-giant-data-set-how-analysing-the-past-could-help-save-the-future

Wang, Z., & Li, J. (2019). *Attention-aware age-agnostic visual place recognition*. Paper presented at the ICCV workshop on e-heritage 2019, Seoul, South Korea. Retrieved from https://arxiv.org/abs/1909.05163

Woning, vermoedelijk in Manado. (1895). [Photograph]. KITLV Collection (Shelf location 124934). Leiden University Libraries, Leiden, the Netherlands. Retrieved from http://colonialarchitecture.eu/obj?sq=id%3Auuid%3A1754e787-e353-482d-993e-f0149822752c

About the Authors

Tino Mager studied media technology in Leipzig and art history and communication science in Berlin, Barcelona and Tokyo. Since 2017 he has been a Post-Doc at the Chair of History of Architecture and Urban Planning at TU Delft and is working on the development of methods for the use of AI in architectural historical research. He has been awarded grants from, among others, the Leibniz Association, the Andrew W. Mellon Foundation, the University of Queensland, and the Dutch Research Council.

Carola Hein is Professor of the History of Architecture and Urban Planning at Delft University of Technology. She has published widely and received a Guggenheim and an Alexander von Humboldt fellowship as well as other major grants. Her books include *Adaptive Strategies for Water Heritage* (2019), *The Routledge Planning History Handbook* (2017), *Uzō Nishiyama, Reflections on Urban, Regional and National Space* (2017), *Port Cities* (2011), *The Capital of Europe* (2004), *Rebuilding Urban Japan after 1945* (2003), *Cities, Autonomy and Decentralisation in Japan* (2006), and *Hauptstadt Berlin 1957–58* (1991).