

Transforming Information Into Knowledge: How Computational Methods Reshape Art History

Sabine Lang <sabine_dot_lang254_at_gmail_dot_com>, Interdisciplinary Center for Scientific Computing, Heidelberg Collaboratory for Image Processing, Heidelberg University

Björn Ommer <ommer_at_uni-heidelberg_dot_de>, Interdisciplinary Center for Scientific Computing, Heidelberg Collaboratory for Image Processing, Heidelberg University

Abstract

Current research in computer vision highlights the potential of using computational methods to analyze and access large datasets of real images and videos, performing tasks such as object detection or finding visual similarities. This essay describes how the application of these computational methods to digital art data transforms information inherent in images into new knowledge. We support this claim by presenting various research examples in the field of digital art history, which utilize computational methods for art analysis. We argue that in order to create new knowledge, we must involve transformative processes — from analog to digital data and digital to computational methods. Traditional methods used in art history to access datasets, link or edit images provide suggestions and validations for current practices, but are not sufficient role models for the processing of digital data because they were developed under varying technological conditions and standards and within a different historical context. Aby Warburg's (1866-1929) *Mnemosyne Atlas*, which aimed to visualize visual continuities from antiquity to the Renaissance, is one method often cited in the context of digital humanities. We argue that the characteristics of digital data, for example being reproducible or modifiable, require innovative computer methods that do not have a direct analog counterpart. Our argument is based on the success of current projects and personal experience: the authors have expertise in computer vision and art history. They have created and applied computational methods to analyze art data and are thus able to identify shortcomings of traditional approaches and suggest possible solutions. Eventually this essay presents solutions, which demonstrate the great potential of computational methods for a data analysis: approaches enable an easy and explorative access to digital image collections through visualization techniques or an automatic object search. Computational methods establish links between thousands of images, thereby identifying the adaptation of specific motifs or styles by artists over time, or enable a conspicuous editing of images, thus providing new insights for art history unobtainable with analog methods.

Introduction

Much effort is spent to preserve cultural artifacts through digitization including artworks, written documents and sites of cultural heritage: the *Time Machine Project*¹, a collaborative effort between science and humanities researchers and institutions worldwide, uses digital data and innovative computer technologies to build an image of European history covering more than a thousand-year period. Through large-scale digitization of cultural artifacts the project aims to preserve and revive the past of cities such as Venice, Amsterdam, Paris, Budapest or Jerusalem. The *Time Machine* is exemplary for current research projects within the digital humanities which digitize text documents and images (i.e., artworks), thus creating large data repositories, and use innovative methods for analysis. These repositories are equivalent to large storage boxes of information, which bear witness to the past, visual preferences or relationships between artists and herald new knowledge. In order to transform information inherent in digital images into new knowledge, we require computational methods, which are able to process and access thousands of images and consider the specific characteristics of digital data such as being reproducible and modifiable. By using computer

algorithms and technologies to view and analyze digital images, scholars can establish new connections between images or gain new insights regarding the reach of a specific motif or style or the frequency of artistic exchange. In addition image details become visible through, for example, X-rays, gamma rays or other physicochemical processes.

The availability of image collections and computational methods impacts society and scientific research, because they enable new ways of communication, the creation of a global (scholarly) network and a simplified access to images for experts and non-experts, because current systems to access collections operate on visual inputs — examples are presented in section four. So far scholars use digital images to reinforce arguments in publications or presentations, to view image details or place geographically distant artworks next to each other for comparison. While art historians already work with digital images, methods to view and present them have remained similar to traditional approaches: PowerPoint presentations mostly show comparisons of image pairs, online databases are structured according to established data models, for example a hierarchical tree structure, and links between images are influenced by the expertise of the art historian, who prepares the presentation: a specialist in seventeen-century Dutch painting most likely identifies similarities and differences in form and content between the works of Jan Vermeer (1632-75), Frans Hals (c1582-1666) or Rembrandt (1606-1669). The same person would probably find it more challenging to enlarge this cohesive group and define links between seventeen-century Dutch artists and other European artists of the same period or even modern painters. Moreover links are not only influenced by the expertise of the scholar, but also by the availability of information and facts validated and promoted by the art historical canon. However, this approach to link images is not representative of the possibilities offered by digital data. This essay argues that we need genuine computational methods to evaluate digital data, because traditional analog methods do not make use of the specific characteristics and possibilities of digital data. To create computational methods, which utilize the specific characteristics, requires transformative processes, namely from analog to digital data and digital to computational methods. We speak of traditional methods to refer to approaches in art history, which have been suggested for the study of artworks, such as Heinrich Wölfflin's (1864-1945) *Formalism* or Erwin Panofsky's (1892-1968) *iconographical-iconological method*. While the former amplifies formal properties of an artwork to identify an epoch or culture and mainly disregards content, Panofsky's method studies form and content. He established three stages: (1) A description of form, (2) decoding of attributes and symbols and (3) the study of the historical context in which the artwork was created [Hatt/Klonk 2006]. German art historian Aby Warburg (1866-1929) developed his *Mnemosyne Atlas* as a method to illustrate visual continuities from antiquity to the Renaissance in paintings, drawings, sculptures and other artifacts. To visualize these visual trails he utilized wooden boards covered in black cloth on which he pinned his reproductions [Hristova 2016]. Although Warburg left the project unfinished with his death in 1929, the *Atlas* was continued and further studied by his colleagues Gertrud Bing (1892-1964) and Fritz Saxl (1890-1948) [Kalkstein 2019] and presently by the Warburg Institute and Cornell University Library, which recreated ten panels from the *Atlas* online offering additional information and suggested readings [Cornell University Library Home 2016b]. Although these traditional methods vary in their focus and were developed for different purposes, they still have some commonalities: methods have been applied over a long period of time and are still used today; they were defined by humanists or art historians; operate on smaller image collections and results are based on visual observations and knowledge of art historians. In comparison computational methods are relatively new; their efficiency and the validity of results are still being evaluated and thus must be viewed critically. Computational methods are mostly developed by computer scientists to solve tasks such as classification and recognition using large collections of real images. Only in recent years have these approaches been developed in collaboration with art historians and applied to art data. In contrast to traditional approaches, computational methods operate on a visual level and do not consider external knowledge about historical conditions for analysis. Therefore differences between traditional and computational methods mostly refer to the developer, purpose, number of images, longevity and inclusion of external information. The shortcomings of traditional methods relate to these differences and include the relatively small image collection, which is analyzed, and the fact that results are often influenced by the knowledge of the art historian.

Our argumentation is supported by personal experience and recent research projects, which demonstrate the potential of computer methods for art historical research. These computer methods assist for example with restoration work or virtual reconstructions of missing artifacts. On July 8, 2019, a research team from the *Rijksmuseum* in Amsterdam together with collaborators from museums and universities in the Netherlands and worldwide embarked on a large-scale

research and restoration project: using digital imaging techniques and other advanced methods from computer science and artificial intelligence, researchers investigate the original appearance and current state of Rembrandt's *Night Watch* (1642) to acquire new information about the painting. The *Operation Night Watch* studies questions relating to the original commission or Rembrandt's preferred materials and painting techniques to understand changes in appearance and determine a treatment plan for restoration. Ultimately the project aims to preserve Rembrandt's masterpiece in digital form. The current status of the research project can be followed live on the museum's website [Rijksmuseum 2019]. In 2016 a team consisting of collaborators from ING, Microsoft and scholars from TU Delft, The Mauritshuis and the Rembrandt House Museum unveiled a new painting by Rembrandt. The project entitled *The Next Rembrandt* analyzed a large collection of paintings by Rembrandt to create a next Rembrandt utilizing high-resolution 3D scans of paintings, deep learning algorithms to process and chart images and a 3D printer. By analyzing the data, scholars were able to study the preferred subject of Rembrandt, his use of color, light and shadow in images [The Next Rembrandt 2016]. Other research examples use digital data and computer vision algorithms to estimate the source of illumination in paintings to reason about painting practice in the Renaissance and Baroque [Stork and Johnson 2006] or virtual reality techniques to reconstruct a segment of the Berlin Wall and surrounding neighborhoods. The project *Virtuelle Mauer/ReConstructing the Wall* led by Tamiko Thiel and Teresa Reuter creates a virtual installation, which aims to understand how the Wall and the possibility of escape impacted everyday life for residents [Thiel and Reuter 2019]. Many more examples at the intersection of art and technology are provided on Google's *Arts and Culture Experiments* website. The website presents projects from various artists, computer scientists and other researchers, which unlock culture on the basis of large datasets: the project "From a picture to a thousand stories" links art and books based on a single image; in "Draw to Art" a doodle is used as a query to discover related works of art [Arts and Culture Experiments 2016]. All listed research projects demonstrate the great potential of computer technologies and how these are deployed to support art historical research and consequently produce new knowledge for the field — unattainable by traditional methods or analog data. Lastly authors have expertise in computer vision and art history; together they created and applied computational methods to art data. Authors were thus able to identify shortcomings of traditional approaches and create solutions: projects included the development of a search engine for object detection, which was applied to for example a dataset of exhibition photographs to study changing exhibition contexts [Lang and Ommer 2018b], or the study of artistic style through a computational style transfer [Sanakoyeu et al. 2018] [Kotovenko et al. 2019b].

The article is structured as follows: after presenting related works, which include (theoretical) discussions within the digital humanities about current digital practices and research done in computer vision on the topics of classification, object detection, finding visual similarities or image generation, the article describes the transformation of analog to digital data and digital to computational methods. It then elaborates on the specific characteristics of digital data, including its modification and reproduction, the latter of which propels a wider distribution of images. The main part of the article discusses the insufficiency of simply turning analog into digital methods and provides examples of genuine computational methods, which support this argument and were created and used to analyze art data. Examples focus on object detection to access large image collections, data visualization to establish links between images and image generation and manipulation. Based on this we reach the following conclusions: first we require digital data and computational methods to produce new knowledge. It is not sufficient to simply translate existing analog into digital methods, because they do not consider and make use of the specific properties of digital data, such as being reproducible, modifiable, or able to capture granular image properties. Traditional approaches are unable to display thousands of images in one space or offer multiple ways to access image collections, which are necessary to cater for scholars' individual research questions. In order to fully exploit the potential of digital data and produce new knowledge, we need genuine computational methods based on neural networks. Examples given in the text show that these methods enable an automatic data analysis, facilitate a data exploration, or generate entirely new perspectives of the art object. Visualization techniques for example group portraits according to social status and age of the depicted person given only visual user input. In addition interfaces for object retrieval enable to search for objects such as the 'medal' in a large collection of images (see Figure 5). Both examples highlight that computer technologies and results facilitate a large-scale and fine-grained analysis of data collections to create overarching links or study changes in form, content or image context. Other examples will show that generative models produce new perspectives of the art object or give insights into artistic style and defining style properties. We demonstrate that by changing visual properties such as

shape and appearance or transferring the Impressionist style of Claude Monet (1840-1926) or the Post-Impressionist style of Paul Cézanne (1839-1906) onto real photos using computer models, we are able to study the effect of style on content or how the atmosphere of the image changes for different styles. Second, based on given examples and personal experience, we conclude that it is difficult to make precise predictions about the future of digital art history but that in order to create computer methods, which are helpful for art historical research and exceed the capabilities of analog methods, we must reduce skepticism towards new technologies and emphasize the value of the work of the respective other discipline. To achieve this, scholars from computer vision and art history must work together and computational tools must be made more accessible.

Related Works

Digital humanities and in particular digital art history is a vibrant and dynamic research field. In recent years there has been a significant increase of theoretical and research-related works, which discuss and illustrate how digital data and methods impact art history. Scholars elaborate on current digital and computational practices [Bishop 2018], point to insufficiencies of existing approaches and research gaps [Lang and Ommer 2018c] and, for example, refer to the often loose and unspecific usage of terms such as beauty and inconsiderate nature of many works. In her article “Against Digital Art History” Bishop gives the example of a facial-recognition software which had to establish, given 120,000 portraits from the thirteenth to the twentieth centuries, whether the concept of “beauty” had changed over time. Unsurprisingly to everyone who is at least a little bit familiar with modernism, the software determined a decrease of ‘beauty’ in the twentieth century. Authors of this work missed however to discuss the reason of this decrease, the subjectivity of ‘beauty’ or how the selected dataset influenced the result [Bishop 2018]. Some comment on the fact that computer-based projects provide a quantitative analysis but miss a zoomed-in study [Kienle 2017]. In order to assist research and produce knowledge, which is reflective of new possibilities, scholars need tools for a quantitative and qualitative analysis [Bonfiglioli and Nanni 2015]. Johanna Drucker’s article “Is There a ‘Digital’ Art History?” is a reflective comment on digital art history and provides valuable reference points for this essay [Drucker 2013]. She elaborates that cultural theories of the 1980s impacted art history, because they questioned the properties of an object, i.e., its function, social role or intention and thus our understanding of it; she demands the same effect for digital methods. Famously she makes the distinction between a digitized and digital art history. The former refers to the simple use of digital image collections, the latter includes the use of computational methods for data mining, network analysis or textual study [Drucker 2013]. The following section gives an overview of main tasks for which computer vision developed computational methods, these include the classification of artworks, recognition of objects, finding visual similarities between images and image generation. Current approaches to solve these tasks are based on neural networks, which are modeled according to the human brain and consist of multiple layers. In contrast to older approaches, networks are able to process thousands of images in an unsupervised manner, meaning that they require no human intervention. At every layer the network extracts features from images, which provide essential information to solve the given task. In lower layers features refer to color, shape or textures; higher layers include more complex information about content and relations between semantic concepts such as eyes or mouth [Goodfellow et al. 2016] [Krogh 2008]. Individual sections not only include research examples from computer vision but also digital art history. In recent years the field has adopted these methods and applied them to art datasets, performing tasks generally done by art historians (i.e., classification or finding visual similarities). Eventually given examples highlight the potential and success of a comprehensive, computational image analysis.

Classification Computer vision research has suggested numerous approaches, which address the task of an automatic image classification of thousands of images using neural networks [Krizhevsky et al. 2012] [Simonyan and Zisserman 2015]. Most approaches utilize a subset of ImageNet, a database consisting of over fourteen million images from more than 20,000 categories including plants or animals. Since 2010 the community takes part in the annual ImageNet LargeScale Visual Recognition Challenge which evaluates the efficiency of algorithms for the tasks of object localization/detection, scene classification and parsing, given a number of images from specific categories [Simonyan and Zisserman 2015]. Depending on the task the winner is the team which achieves the minimum average error or highest accuracy over all test images [ImageNet Large Scale 2016a]. In 2016 the winner for the task of object detection achieved an impressive mean accuracy of 66.27 percent on most object categories [ImageNet Results 2016b].

Following works presented solutions to improve the classification accuracy through for example an attention mechanism [Wang et al. 2017]. The neural network perceives information about visual properties (i.e., color) or pixel arrangements, which suggest the presence of an object. The localization of the object a priori simplifies the classification task [Wang et al. 2017]. These methods demonstrate the ability of networks to perceive content and their high efficiency to organize data accordingly. Approaches were then applied to the classification of art data. Karayev et al. classified 80,000 Flickr photographs and 100,000 images — mostly paintings — from the WikiArt collection according to styles using a convolutional neural network — a certain class of neural networks, which includes the mathematical operation of convolution. For the latter 25 style labels such as Symbolism, Expressionism or Art Nouveau were used [Karayev et al. 2013]. Others focused on predicting style, genre and artist based on features extracted by neural networks and provided similarity measures between images [Saleh and Elgammal 2015]. Based on a style classification of 77,000 paintings and a visualization of images, Elgammal et al. correlated the visible groupings with Heinrich Wölfflin's five contrasting principles of art history [Elgammal et al. 2018b]: linear versus painterly, plane versus recession, closed versus open, multiplicity versus unity and absolute versus relative clarity [Hatt/Klonk 2006].

Object detection The fast and precise localization and recognition of objects in images has been a prevalent computer vision task [Girshick 2015] [Ren et al. 2015]. Redmon et al. for example presented a neural network which is capable to localize objects in images and to provide class labels such as human, dog, horse and probabilities in one go [Redmon et al. 2016]. Current research papers have exploited information about relations between objects to improve the accuracy and speed of detection. Relations are based for example on the object's appearance or geometry [Hu et al. 2018]. Detection approaches have been transferred to art data, aiming to find objects such as horses or ships in paintings. Research works have studied the *domain shift* problem, evaluating the performance of classifiers trained on real images (mainly from ImageNet) versus paintings [Crowley and Zisserman 2016]. The task of finding identical and similar objects, which display little shape variances in a dataset of medieval manuscripts, was studied by Takami et al. (2014). Moreover Schlecht et al. utilized a template-based detector to find gestures in manuscripts. The authors first collected a small set of labeled instances of a gesture type and then found additional examples showing shape variations using computer algorithms [Schlecht et al. 2011]. Others added curvature information of objects to improve detections [Monroy et al. 2011]. A discriminative model based on parts and randomly aggregated compositions was further employed to assist with object detection and scene classification [Eigenstetter et al. 2014]. While these research projects demonstrate the abilities of a computer-based detection, they also show that networks are often challenged by unknown, often pre-modern object categories or objects defamiliarized by style properties. This is mainly because detection networks were trained on real photos and therefore have never seen instances of swords, medieval clothing or objects deformed by Cubism [Lang and Ommer 2018c].

Visual similarity Aspects such as relationships between a master and a student, the stay of artists at foreign courts, the circulation of books or exhibitions manifest in artworks through similar colors, brushstrokes, objects or compositions. The identification of these similarities is a central task of art historians. The computer vision community has presented numerous approaches to find visual similarities between images [Sanakoyeu et al. 2018]. Bautista et al. grouped images in cliques of similar samples around an initial exemplar and utilized information about highly similar and distant samples for the task of finding visually similar images [Bautista et al. 2016]. Kim et al. proposed a neural network to establish correspondences between semantically similar images, which is invariant to varying attributes such as color or texture [Kim et al. 2019]. Other methods leverage information about geometry and context to improve semantic matching [Ufer et al. 2019]. Qualitative results and quantitative evaluations in respective papers show the impressive abilities of networks to retrieve similar images from large datasets. To assist art historical research, these methods have been applied to art data. Seguin et al. established links using neural networks, pre-trained on human-annotated clusters of similar paintings [Seguin et al. 2016]. An additional approach was suggested by Shen et al. to link visual patterns in artworks of the Brueghel family and in photos of buildings and landmarks. The method relies on randomly sampled regions from images which are compared to all images in the dataset to find similar regions. A correspondence between regions is valid if surrounding regions can be matched as well [Shen et al. 2019]. Saleh et al. calculated the similarity between paintings regarding content or style using Euclidean distances to determine artistic influences [Saleh et al. 2016].

Generative models In machine learning the advent of Generative Adversarial Networks (GANs), first introduced by Goodfellow et al. in 2014, triggered a new interest in generative models 2 to alter shape or appearance of objects or to translate an image into a new domain: a painting by Monet turns into a real photograph, a zebra into a horse or a winter into a summer landscape. Goodfellow presented a framework to estimate generative models based on an adversarial, competitive process of two models: a generative model, which approximates the data distribution of the training set, and a discriminative model; the latter estimates the probability of a data sample belonging to a training set or generator G [Goodfellow et al. 2014]. Generative models 2 have been used for the task of image generation and transfer, for example for appearance and shape generation [Esser et al. VU-Net 2018b]. The model proposed by Esser et al. allows to “synthesize different geometrical layouts or change the appearance of an object, either shape or appearance can be retained from a query image, whereas the other component can be freely altered or even imputed from other images” [Esser et al. VU-Net 2018b]. Lorenz et al. utilized generative models to disentangle shape and appearance of an object, which is necessary because of large intra-class variations such as pose, texture, color or clothing. The authors evaluate their approach on the tasks of pose prediction and image synthesis and provide remarkable and authentically looking results for image generation. Results demonstrate how the model can estimate shape or how appearances can be exchanged for parts [Lorenz et al. 2019]. Generative models are also used for the task of style transfer, which describes the stylization of a real photo in the style of an artist, such as Claude Monet (1840-1926), Paul Cézanne (1839-1906) or Picasso (1881-1973). The task is either based on one input image [Gatys et al. 2015] [Johnson et al. 2016] or a collection of paintings [Sanakoyeu et al. 2018], aiming to fully capture an artist’s style. Subsequent works focused on improving the stylization process and image results by separating style and content [Gatys et al. 2016], gaining control over the degree of stylization for different image regions [Huang and Belongie 2017], or speeding up the stylization process [Johnson et al. 2016]. Other approaches pay attention to transform content in a style-specific manner [Kotovenko et al. 2019b] or capture variations and particularities of a style [Kotovenko et al. 2019a]. Stylized images show that generative models are not only able to perceive what is depicted on images but also how. Models were able to learn the particularities of style and produce — if not authentic paintings by Monet or Picasso – at least compellingly looking images.

Provided examples for the tasks of classification, object detection, finding visual similarities or image generation indicate the great potential of computational approaches for art analysis as they are able to process thousands of images, thus increasing the scope of an evaluation, assist with art historical tasks, alter the art object or create new perspectives via style transfer. Notably in comparison to an automatic image classification or object and similarity detection, generative models are genuinely computational and do not have an analog counterpart.

Others Other works utilized computer-based methods to address additional concerns relevant to art history. “How can we explore patterns and relations between sets of photographs, designs, or video, which may number in hundreds of thousands, millions, or billions?” [Manovich 2012] — this question was answered by media theorist and digital humanist Lev Manovich and his lab, who developed new methods to analyze and visualize large data collections according to individual visual characteristics of images. These methods consist of two steps: First visual properties of images such as color, contrast or saturation are translated into numbers through image processing. These values are then taken to position all images in a two dimensional visualization. Manovich and lab used this method for example to visualize more than one million manga pages according to different image characteristics [Manovich 2012]. Li et al. automatically extracted brushstrokes in paintings by van Gogh for authentication and dating purposes. Therefore the authors had to define his signature brushstroke: an algorithm was utilized to identify edges which indicate a brushstroke. Additional algorithms were facilitated to automatically fill in gaps if a brushstroke was not fully enclosed [Li et al. 2012]. A fully enclosed brushstroke was extracted “by setting the edge pixels as background and the non-edge pixels as foreground” [Li et al. 2012]. A brushstroke was considered valid if it fulfilled certain conditions: for example the skeleton was not severely branched or the ratio of broadness to length lied within a prior defined range. Brushstrokes were analyzed to determine their average length, width or orientation [Li et al. 2012]. Algorithms were also utilized to segment individual strokes in drawings by mostly Picasso, Henry Matisse and German artist Egon Schiele in order to assist with their attribution or to determine between original and fake. The approach is based on the quantification and comparison of stroke characteristics, such as shape, tone variations and local shape characteristics, using a stroke segmentation algorithm. This algorithm first extracted a stroke network and subsequently untangled the network to get individual

strokes [Elgammal et al. 2018a].

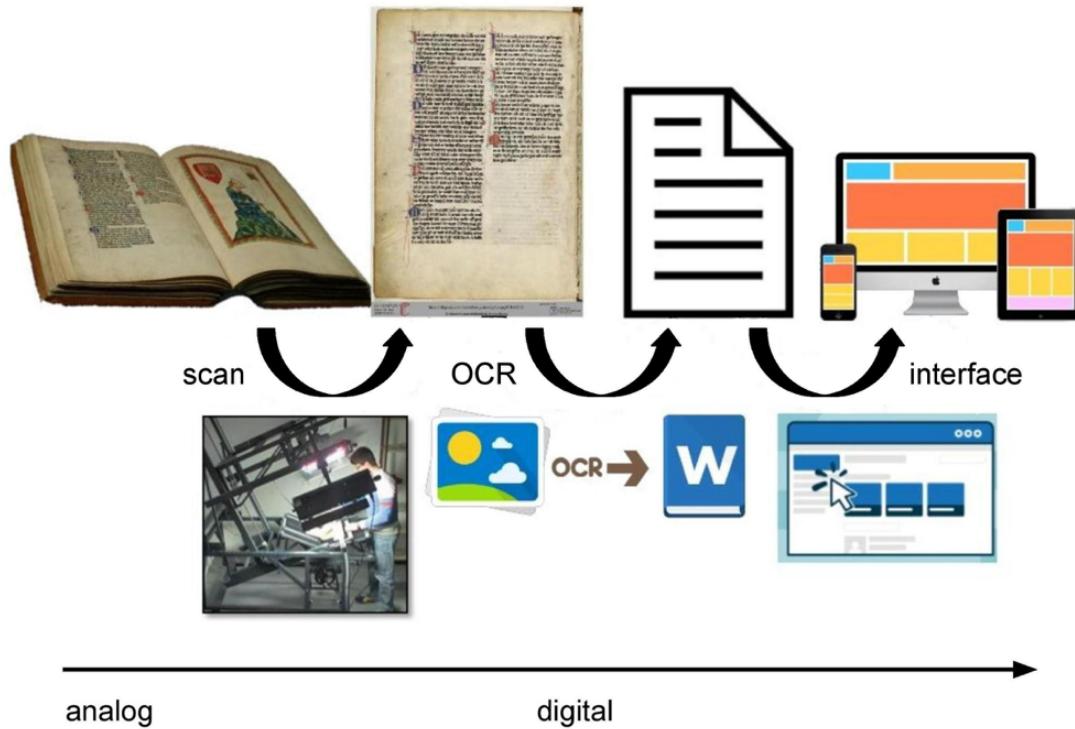


Figure 1. Shows the translation from analog to digital data using scanning tools or Optical Character Recognition (OCR). To make digital data accessible, collections are stored on interfaces © Copyright Computer Vision Group, Heidelberg University

3. Transformative processes: from analog to digital and digital to computational

Digitization describes the transformation of analog data into digital data through for example scanners or digital cameras: sensors implemented in devices convert a two-dimensional image into a numerical representation. Optical character recognition (OCR) is specifically used to translate a handwritten or typed document into a digital text document (see Figure 1). These techniques have produced large data repositories containing text, image, video or music files. Images include reproductions of artworks and other cultural artifacts which are gathered and published on museum websites and in online databases such as WikiArt and shared on social media platforms. The online availability of images has various consequences: images reach a wider audience, allow for an easy comparison of spatially distant artworks (because they belong to different collections or museums) and propel the development and use of methods which are able to process and analyze them thus shifting the focus from images to image processing techniques. In “Is There a ‘Digital’ Art History?” Drucker discusses this shift and explicitly distinguishes between a “digitized” and “digital” art history. “Digitized” refers to the building of digital image repositories and its subsequent use in research, while the latter describes the usage of computer-based methods for art analysis. Drucker lists the following methods as central to the digital humanities: “text ‘mark up,’ topic modeling, structured metadata, visualization of information, network analysis, discourse analysis, virtual modeling, simulation and aggregation of materials distributed across geographical locations” [Drucker 2013]. These methods still resemble traditional methods, because they for example rely on a model which structures the data, use pre-defined terminology to describe digital data or require manual annotations. Therefore these methods restrict the possibilities of digital data and essentially hinder the production of new knowledge. The current status of digital humanities and the fast progress in computer vision, especially in deep learning 3, require an extension of the process defined by Drucker. We argue that in order to exploit the potential of digital data and produce knowledge, an additional process must take place, namely from a digital to a computational art history. The latter demands the processing of digital data using automatic and unsupervised computational methods. These methods

require minimum or no human intervention and do not resemble existing, analog methods.

3.1 What are the specifics of digital data?

"We have to take into account the ways digital humanities more broadly have taken up computational techniques and then consider the specificity of visual art objects and their particular requirements and points of resistance" [Drucker 2013]. Drucker remarks on a process very common in the digital humanities: methods are developed first or existing approaches are simply taken and specificities and requirements of digital data are considered afterwards. However, this approach as stated does not consider characteristics of digital data during the phase of development and therefore restricts the potential of digital data and methods. Although we have transferred analog image processing into a digital space, we restrict ourselves to knowledge which could have been obtained with analog data and methods. Instead, we must identify special characteristics and requirements of digital data first and create corresponding methods afterwards. Examples of research projects in the last section of this article highlight the importance and validity of this thought process. But what are the specifics of digital data in comparison to analog data which should be considered and explicitly exploited by new methods?

Digitization through, for example scanning, describes the translation of an analog, continuous into a digital, discrete signal. Digital images are thus available as numeric representations. This means that certain aspects in images such as composition, spatial relations between objects and figures or color can be evaluated statistically and expressed in descriptive statistics. Through a statistical analysis we are able to summarize the content of one or multiple images in representative statistics. These provide factual and objective descriptions of digital artworks, which are impossible to create for analog images and with analog methods. By looking at these image statistics, art historians are able to study the popularity of a composition, the degree to which artists relied on compositional models or when a color originated and how frequent it was adapted by artists.

In addition digital images can be easily duplicated because they are stored as numerical representations. These digital copies are reproductions of existing artworks and not computer-generated images. In contrast to the analog image, the digital reproduction is transportable and easy to distribute. Digital image copies can be shared with colleagues or used for multiple projects, thus supporting collaborations and the founding of a global network of scholars. Digitization then supports the transformation of art history from a locally and nationally determined discipline, where each country favors different methods and topics, to an international and dynamic field. While the reproduction of images offers great potential for society and research, it also holds disadvantages: reproduction quality varies for images; color, illumination, sharpness or contrast are often misrepresented in digital images and falsify the appearance of the original artwork. These visual alterations must be considered in an interpretation or when developing computational methods. An erroneous color reproduction might establish a visual similarity of actually dissimilar images and a blurry image conceals details otherwise visible. Duplications of artworks also raise questions of originality, uniqueness and authenticity. German philosopher and art critic Walter Benjamin (1892-1940) famously addresses these issues in his essay "The Artwork in the Age of Mechanical Reproduction" first published in German in 1935:

Even the most perfect reproduction of a work of art is lacking in one element: its presence in time and space, its unique existence at the place where it happens to be. This unique existence of the work of art determined the history to which it was subject throughout the time of its existence. This includes the changes which it may have suffered in physical condition over the years as well as the various changes in its ownership. The traces of the first can be revealed only by chemical or physical analyzes which it is impossible to perform on a reproduction... [Benjamin 1969]

Therefore the digital artwork challenges the uniqueness of the original artwork and detaches it from any place or time restriction. According to Benjamin, the originality of an artwork and its existence at a particular time and place is the prerequisite of authenticity [Benjamin 1969]; for the viewer and in particular the scholar this attachment has consequences. The function and role as well as the physical effect of the artwork can never be fully comprehended through digital images. Instead, we are only able to guess and hypothesize, thus providing a first insight but never a full picture of the artwork in its original context. To do so we have to see it in person, if possible, in its actual environment:

13

14

15

16

the museum. While the removal of authenticity, uniqueness and original context are shortcomings of digitization, it allows to preserve art and cultural heritage in digital form. This is especially relevant for works which are exposed to weather conditions or pollution or which are made out of fragile material, examples include weavings, fresco paintings or stained glass windows. Street art is a current art form which faces — but also consciously plays with — these threats. As Benjamin mentions in his essay, reproduction techniques further blur physical damages such as cracks on the original artwork [Benjamin 1969]. The damaged artwork is thus restored in digital form. For art historians this aspect is ambiguous: on the one hand, it enables to study the unharmed artwork and to reason about its original appearance, on the other hand it gives a false impression of the artwork's current condition.

The representation of digital data as values makes it modifiable and editable through computer technologies. While this article focuses on the potentials of image editing and manipulation for art historical research, we feel propelled to express caution. The image observer must always be aware of the possibility of image manipulation and must challenge and critically evaluate the image content. We advocate for a transparent manipulation of digital art images, because the benefits for art historical research are too significant as following examples show. Editing programs such as AdobePhotoshop enable art historians to perform simple operations: they can crop details from images, alter the format and resize the image or change color or saturation. These alterations already offer new insights: a change of color emphasizes its importance for an artwork and illustrates how its perception and general meaning changes through the modification of color. By altering formal properties of images we produce different versions of the artwork and present new perspectives. Section four includes research examples which also produce unknown views of the object or artwork, however exceeding these early operations by far; computational methods not only change color, but alter textures, brushstrokes, content or the shape of objects.

Through editing programs scholars might also enlarge images to study brushstroke and texture of artworks or details, which would be overlooked or remain unnoticed otherwise. By enlarging image regions, art historians are able to study Jan van Eyck's *Arnolfini Portrait* (1434) in detail. The mirror depicted in the background was particularly interesting for art historians. Essays discussed the identity of the persons shown in the reflection or the frame depicting the ten stations from the Passion of Christ. The observer is further propelled to appreciate the haptic quality of the dog's fur, the material of the chandelier, the flowers seen through the window or the decorative details of the wooden bed [Postel 2017]. All details give proof of the mastery of van Eyck and can only be studied through the availability of the digital image by most scholars as they do not have access to the respective museum or are not allowed to see the physical painting up close (see Figure 2).

17

18

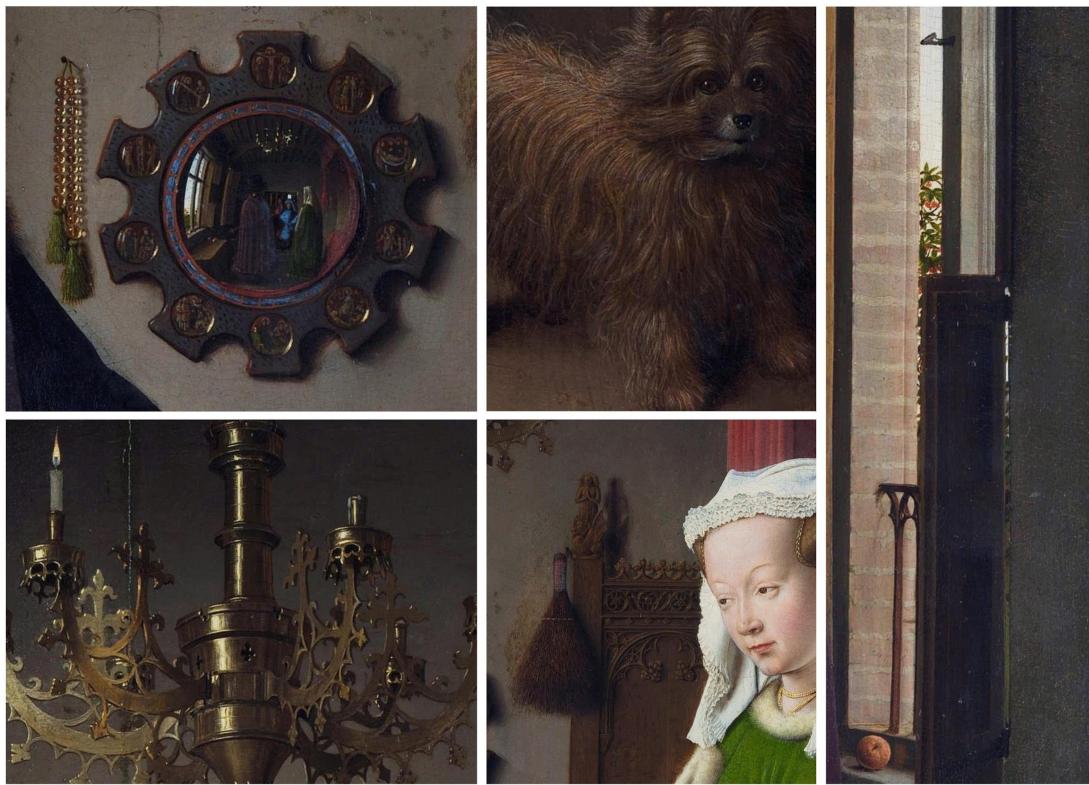


Figure 2. The digital reproduction of Jan van Eyck's *Arnolfini Portrait* (1434) allows scholars to zoom in on image details. © Copyright The National Gallery, London 2019

In contrast to digital images, analog images are restricted, often attached to a specific space and thus can only be enjoyed by a selected group of people who have access to specific spaces, museums or can afford entrance fees. The access to digital images is offered to much more people and simplified by computers and the internet. Seemingly audiences can immerse themselves in art and culture through digital images regardless of geographic location, nationality or social status. A study of the "World Bank" conducted in 2016 however has shown that while technologies have connected people on large scale, sixty percent of the world's population still have no access to the internet. This means that there is a growing digital divide between the rich and poor and rural and suburban regions and that advantages of digital technologies are mostly beneficial for the wealthy and educated [Elliott 2016]. Thus large parts of the population are still not able to enjoy art through digital images. Other factors further limit the access to images, this includes fees to image collections, copyright or political regulations. While art has become more inclusive through the wider and simplified access to digital art images, limitations and exclusions of specific demographics are still present.

Essentially the numerical representation of digital images facilitates their duplication, modification, availability or accessibility — all these are the specific characteristics of digital data. Aforementioned operations such as the changing of color, saturation or cropping of details are possible because of these characteristics, but only present a small fraction of current possibilities. The availability of digital images and their (numerical) representation also support the employment of computational methods for art analysis and their subsequent processing through neural networks. Computational methods by far exceed previous operations and capabilities and are able to produce even more knowledge. The specific characteristics of digital data and the presence of automatic methods in computer vision highlight that it is not enough to simply simulate analog methods in order to produce new knowledge for the field of art history but that computer scientists in collaboration with art historians must create new approaches.

3.2 Why it is not enough to simply translate analog into computational methods

The previous section has emphasized that analog and digital data are characterized by different qualities, the latter for example is reproducible and modifiable. These differences suggest that although both formats still carry the same information, methods to extract information and produce new knowledge must differ. Analog images are often visualized

in pairs; artworks are placed next to each other on a table or pinned on a board to compare visual properties. Heinrich Wölfflin's method to study artworks was essentially comparative: pairs of images, which embodied different sets of visual principles (linear versus painterly, plane versus recession, closed versus open, multiplicity versus unity, absolute versus relative clarity), were used to emphasize stylistic and historical shifts between the Renaissance and Baroque period [Hatt/Klonk 2006]. Aby Warburg extended this comparative approach with the *Mnemosyne Atlas*; his wooden panels provided space to visualize multiple images. An analog visualization might be further enlarged, but at one point the physicality of analog images demands too much space and the visualization is hardly comprehensible and understandable to the observer anymore. If we adapt this visualization method for digital data, it means that images are viewed — although on a computer screen — in pairs or a limited number. However, the presentation in pairs neglects the fact that a digital space offers to display significantly more images: thousands of images might be viewed and compared simultaneously; the user can change perspectives, zoom in or obtain an overview of the data. Visualizations of digital data are much more interactive and adaptable to needs and requirements of scholars. The *t-SNE Map* project, created by Cyril Diagne, Nicolas Barradeau and Simon Doury for Google's Arts and Culture Experiments, uses machine learning algorithms to map thousands of artworks in an interactive 3D landscape according to visual similarities [Diagne et al. 2018]. The project enables a virtual tour through the digital collections currently included in the Google Arts and Cultureonline website. The platform currently hosts collections from 1,200 museums and cultural institutions worldwide, including the Museum of Modern Art (New York), Uffizi Gallery (Florence), The Art Institute Chicago or the Kunsthistorisches Museum in Vienna [Google Arts and Culture 2011]. Although the *t-SNE Map* project focuses on data visualization and exploration and neglects a thorough data analysis, it enhances the possibility of computer technologies to access large image collections and study artworks from various angles and at different scales. If we continue to use techniques which resemble analog approaches to visualize digital images, we do not embrace the fact that computer-based visualization techniques significantly enhance our viewing practices.

The digital presence and reproducibility of digital data also enables its localization in many data networks or the combination of multiple networks. “The point is that we could situate a work within the many networks from which it gains meaning and value” [Drucker 2013]. This affects the analysis of single images, since its function and role within many networks can be studied, revealing its ambiguous meaning and adaptiveness to new social or political contexts. 22

In the past analog data such as prints, photos or books were often stored in rigid systems, organized according to categories, such as genre, technique or topic, or specific data models thus limiting the retrieval of records, exchange of data between different systems or users. Data models define the organization and meaning of the respective data collection [West 2011]. There are various models which are used to store data in databases and other information systems: popular examples are the network-model, relational model, entity-set model or hierarchical model. The former three were used as a basis for the entity-relationship model (ER-model), introduced in a paper by Peter Pin-Shan Chen in 1976 [Chen 1976]. The ER-model “incorporates some of the important semantic information about the real world” [Chen 1976] and provides information about entities and their relationships. The model offers different views of the data because it combines the network model, relational model and entity-set model. The network model consists of divided entities and relationships, which are presented in a non-hierarchical structure [Chen 1976]. The relational model is based on relation theory, where the term relation refers to its usage in mathematics: “Given Sets S₁, S₂, ..., S_n..., R is a relation on these n sets if it is a set of n-tuples each of which has its first element from S₁, its second element from S₂, and so on” [Codd 1970]. These relations if straightforward are expressed in arrays of columns [Codd 1970]. Lastly the entity set model is concerned with the properties of entities. Every entity belongs to a different entity set such as *secretary*, *engineer* or *department*; to test whether or not an entity belongs to a distinct set, we have to look at the properties attached to an entity set. It is assumed that if an entity is in the set *secretary*, it shares properties with all other entities in the set *secretary* as well [Chen 1976]. Although the ER-model superimposes structure onto digital art data, the value of early data models to separate elements, model their relationship and define common attributes among a set of similar elements must be recognized. The hierarchical model organizes data in a tree structure. This means that records are classified from general to more specific categories, which are predefined by scholars. For art history categories include but are not restricted to epoch, style, artist and topics recurrently painted by artists. The selection of a data model depends on the following aspects: the database in which it is integrated, topics and format of the data, its subsequent function, or the discipline which uses the data. Essentially models should simplify the storage, 23

retrieval and exchange of the data. However because listed data models operate on a fixed, rigid structure and databases often facilitate different models, it is difficult to link various databases or share data [West 2011]. Moreover we also limit their points of access: then images are only accessible through predefined categories, which prohibit a free and individual exploration of the data. After digitization techniques had produced large image repositories, scholars had to promptly decide on a way to store the data meaningfully. Consequently scholars simply applied traditional models to databases such as the hierarchical model and classified digital data accordingly. But in order for data to be accessible by scholars from different disciplines asking diverse research questions, databases have to be flexible, dynamic and open-ended. By assigning digital images to these rigid structures, we limit their flexibility and usability for different, personal visual networks. Consequently (digital) artworks should not be organized according to existing data models, which would describe and predefine them according to terms of the past, but according to flexible and adaptive models.

New forms of analysis become possible through the digital format of the data. Methods are manifold, varied, adaptive to individual research questions and often easier to use than analog methods. The latter is due to the fact that a computer-based analysis is often unsupervised, which means that the processing of images requires minimum human intervention and no descriptive image labels (i.e., genre or style). Computational methods are usable by a general audience, because they do not require expert knowledge to establish for example similarities between images; the link is suggested automatically by the computer, which is capable of processing thousands of images in a short amount of time. There are manifold and varied computational methods which combined provide a multilevel view of the data and make complex interpretations. In contrast to analog methods, computer-based approaches offer a statistical evaluation and visual analysis, the latter is performed by computer vision which transforms images into visual representations (and not numbers). Analog methods are not capable of a statistical analysis, because image information is not stored as numbers, which are easy to read for the computer but not for the human observer. For digital images, we might analyze and measure color distributions, saturation or even the image composition to answer questions such as what colors were preferably used in a given time period or how artworks are composed in general. The study of visual properties with computer technologies reveals overarching or unusual visual patterns. All this information is inherent in digital images and obtainable by computational methods.

Eventually, most analog methods were developed by art historians or adapted from other disciplines: in the early days of digital humanities, traditional methods from art history or computer vision were simply adapted and applied to digital art data without considering their specific needs [Lang and Ommer 2018c]. Aby Warburg's *Mnemosyne Atlas* and corresponding panels are models for early approaches: similar to the computer screen, Warburg's panels displayed numerous images (instead of pairs) and included hyperlinks and a non-linear, dynamic and open-ended structure [Kalkstein 2019]. However Warburg's space was restricted by the physical dimension of the wooden panels, which measured approximately 150 x 200 cm [Cornell University Library 2016a], and the number of links between images and hyperlinks to external sources were limited to the information and photographic reproductions available to Warburg. While digital approaches resemble Warburg's *Atlas* and his way of grouping images, computer technologies exceed Warburg's abilities and encourage a display of complex image structures and numerous possibilities to access image collections. At present art historians and computer scientists consider these new capabilities and jointly develop genuine computational methods for art historical research. Thus we must encourage collaborations between disciplines to ensure that new methods truly assist art historical research and reduce skepticism towards computational methods. An increased interest is not least visible in the considerable number of digital humanities conferences and journals worldwide; the Digital Humanities Conference, which takes place annually and is organized by the Alliance of Digital Humanities Organizations (ADHO), receives contributions from scholars working in the humanities and sciences, thereby demonstrating an increasing intersection of both disciplines [ADHO 2019]. However reactions of art historians towards computational methods and a collaboration have been ambiguous: while most use digital images for presentations to visualize research hypotheses or show image details, or online databases to gather images and information, comparably few actually engage in a collaboration with computer scientists and use computational methods for their research. While digital humanities as a subject has been increasingly established at universities worldwide, both art historians and computer scientists are still skeptic towards a general collaboration and the usability and validity of methods and results. A survey among art historians and computer scientists conducted in 2014 amplified these observations. It "...intended to shed light on field members' knowledge of the capabilities and applications of computer

technology, attitudes and perceptions about the use of it, and reactions to the meaning of this type pf digitization of the humanities" [Spratt and Elgammal 2014]. The questionnaire was completed by fifty-nine art historians and sixteen computer scientists and allowed two responses, namely approvals and disapprovals. While we are unable to discuss the survey in its entirety, we nevertheless want to highlight some findings. There was an overall consensus regarding the digitization of humanities: the majority of participants from both fields answered positively when asked about possible collaborations. However more than fifty percent of art historians did not view the implementation of AI technologies in humanities as a positive shift; in comparison, the majority of computer scientists did. The survey also brought forth that most art historians were unaware of the current possibilities of computer technologies for a visual analysis and only knew about the most basic applications, including a physical assessment of art or image retrieval based on descriptive labels [Spratt and Elgammal 2014]. "Why would I entrust my life's work to a computer just so I could be made irrelevant" [Spratt and Elgammal 2014], was one response by an art historian. Although the survey was conducted five years ago and the field of digital humanities has massively enlarged and evolved since then, the fear of being replaceable and a general skepticism is still perceptible. A solution to further reduce skepticism and encourage collaborations might be to be more transparent with each other, so that the other discipline, its intentions and methods do not remain a black box. Another solution could be to already approach students of art history and computer science and inform them about the potentials of a collaboration. So far students in art history are less affected by computational tools as access has not been widely available. Their use of digital methods and images has been limited to simple visualization techniques and online databases from which they obtain images and information, sources include WikiArt, Wikipedia, museum collections or university catalogues. Digital images have been mostly used for presentations and term papers and altered through basic editing programs, which allowed them to change color, saturation or size. In turn students of computer science utilize art data as another source for the classification or detection task or to increase the robustness of algorithms. However we must emphasize that a collaboration promises much more: it produces new methods to visualize data or to perform an automatic data analysis, which "draws upon multiple forms of visual analysis that together make complex interpretations of a given object or group of objects" [Spratt and Elgammal 2014]. Consequently methods enable new questions relating to the popularity or the spatial and temporal spread of a motif. In turn computer science develops robust algorithms and produces methods, which are not only capable to perceive what is depicted but also how. Lastly a collaboration enhances and creates new work fields for art historians and computer scientists.

Previous sections emphasized that it is not sufficient to simply translate analog into digital methods, because they are not reflective of characteristics and possibilities offered by digital data. It is only when we develop new computational methods to process digital data that we are capable of producing new knowledge, which is not possible with analog data and respective methods and which in turn affects not only art historians but also earlier generations of scholars, namely students.

3.3. Computational methods to access, link and edit digital images

Based on research projects such as the *Time Machine* or Google's Arts and Culture Experiments and previous remarks we argue that computational methods are especially beneficial to access, link and edit digital images. This section discusses these aspects in more detail and concludes that especially operations of accessing and linking are closely related.

"While access and availability of the art historical corpus have been facilitated by the development of online repositories, tools for... computational processing have been slow to emerge" [Drucker 2013]. Only in recent years have scholars spent increasing efforts to create sufficient tools and methods to access digital image collections. However most are still inspired by analog methods. We have already noted that many databases use data models, such as the network-model, relational model, entity-set model or hierarchical model, which structure collections according to predefined concepts but hinder individual access and a free data exploration. In contrast, computational methods independent of traditional database structures, provide access to large data collections on individual terms. Such methods, which allow easy access and a free exploration, have already been hypothesized at the middle of the twentieth century. In 1945 Vannevar Bush (1890-1974), American engineer and head of the US Office of Scientific Research and Development during World War Two, envisioned a "future device for individual use, which is a sort of mechanized private file and library... A memex

26

27

28

is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his (the user's) memory" [Bush 2015]. At the end of World War Two Bush already demanded what is at the forefront of digital humanities' research today, namely to build storage spaces for digital data and methods, which retrieve records fast and according to individual demands. Although Bush remained vague about how these devices should look — supposedly because necessary technologies were missing — , his request that repositories should be consulted with exceeding flexibility supports the current urge to be able to access image collections independent of predefined terms. As we will see in section four current computational methods for object detection, often implemented in interfaces, operate on purely visual information. These methods enable a search for objects, compositions or other arbitrary image regions in large data collections and do not require tedious and error-prone annotations to access data. Thus users formulate search queries based on visual input and do not need to provide text queries, which require knowledge about art historical terms. Instead of searching for examples of a "still life" users can simply mark an arrangement of fruits, cut flowers or vases to find similar instances. Access to repositories is not only granted through interfaces but also through data visualizations. The digital space is capable of displaying thousands of images at the same time, allowing a direct comparison of image content or style — this has been exemplified by the *t-SNE Map* project of Google's Arts and Culture Experiments [Diagne et al. 2018]. Through visualizations art historians can recognize patterns such as groups of artworks sharing a similar motif or formal property and identify gaps, which indicate that certain visual features were not popular at a certain time.

There have been art historians in the past, who explicitly focused on how art collections can be accessed through visualizations to be comprehensible for the viewer. Hamburg-born art historian and cultural scientist Aby Warburg and his *Mnemosyne Atlas* are often referred to as a starting point for current practices in digital art history. Digital Media professor Stefka Hristova compared present methods of data visualization with Warburg's method to examine "shifts and continuities that have shaped informational aesthetics [...] and data-driven narratives" [Hristova 2016].

29

Warburg accessed image repositories by establishing visual lines of continuity, mainly linking Western antiquity with the Renaissance. He coined the term "Bilderfahrzeuge" to describe the migration of images, objects and other visual features throughout time 4. The term and concept was materialized through the *Mnemosyne Atlas*: the scholar pinned reproductions of artworks on wooden panels covered with black cloth to group images along a common theme [Hristova 2016]. These images, which consisted of a relatively fixed number of 2000 reproductions, were often regrouped or completely removed emphasizing a great dynamic of the *Atlas* [Warnke 2000]. Warburg was interested in two main concepts, namely the expression of human emotion (Pathosformel) and the afterlife of antiquity (Nachleben der Antike) [Impett and Süssstrunk 2016]. Both concepts were traced from antiquity to the Renaissance through their depiction in paintings, drawings, sculptures and objects. He visualized this continuum by merging images and presenting them on one panel. In this way, the *Atlas* not only operated across genres and media but demonstrated an efficient way of information processing. The *Atlas* enabled Warburg to access a substantial part of the repository of art by linking images based on specific properties content. He focused on "images of great, symbolic, intellectual, and emotional power" from the art and culture of Western antiquity through the Renaissance, and up through his own day" [Kalkstein 2019]. Although Warburg was interested in non-Western cultures — in 1923 he gave a talk on images from the Pueblo Indian region — his *Atlas* focused on Western art and specific periods which are indicative of Western art history such as the Renaissance [Hvattum and Hermansen 2004]. This restriction might be due to the fact that he selected most reproductions from his own library or photographic collection, which both focused on Western art [Kalkstein 2019]: in 1909 Warburg wrote to his assistant Wilhelm Waetzoldt (1880-1945): "I have now got a library for cultural sciences (with a focus on Italy) of ca. 9,000 volumes, to which 600 are added every year, several thousand photographs and a few hundred slides. The aim: a new methodology of cultural science based on the 'reading' of the pictorial work. Area: Europe in the fifteenth century" [The Warburg Institute 2018]. Although his library was comprehensive, it focused on specific geographic regions and times, namely Europe and the Renaissance, and thus mainly neglected images of non-Western cultures. This is further indicated in his introduction to the *Mnemosyne Atlas* written in the late 1920s; the text presents Warburg's thoughts on "social memory, the origin of artistic expression and the psychological drama driving the history of European culture from classical antiquity onwards" [Warburg and Rampley 2009]. Although Warburg's *Atlas* utilized a restricted and mostly consistent number of images, it remains an important attempt and role model for current

30

visualization techniques and research attempts: he not only viewed the single image but placed it within a larger artistic network to make sense of visual continuities, to identify and understand similarities between artworks, which testify to artistic relationships and adaptation processes. The latter referring to the circulation of distinct motifs, compositions, styles and other visual properties over time and space.

The process of linking to find analogies is a central human behavior and studied in detail by neuroscience. To make sense of new experiences and impressions, our brain also uses associations: Neuroscientist Moshe Bar emphasizes the importance of analogies for associations and consequently to make predictions of the future. Bar proposes a framework which studies how the human brain "...generates proactive predictions that facilitate our interactions with the environment..." [Bar 2009]. He proposes that incoming information is not analyzed isolated in order to be understood, but that links between new and existing, already familiar information are used to make sense of the unknown. Like Warburg, Bar's theory is based on the presence of analogies and the search for similar information stored in our memory. He famously declared that when seeing a new object, we should not ask "What is this?" but rather "What is this like?" [Bar 2009]. Essentially the concept of linking is central for computer science: section four illustrates how computational methods based on neural networks find visual correspondences automatically; but even earlier data models are interested in links. The ER-model for example models entities and their relationships, both of which are expressed in a given data collection [Chen 1976]. The model considers various possible relationships, namely one-to-one, many-to-one or many-to-many relationships. The example *Secretary (entity) works for engineer (entity)* indicates one-to-many relationships: While every engineer can have only one secretary, a secretary can work for multiple engineers. This diagram is further enriched by adding descriptive attributes to entities such as name, color, or height [Teorey et al. 1986].

Both methods proposed by Warburg and Bush established links manually and based on knowledge which the scholar had prior to his research. Warburg's *Atlas* was mainly used to visualize his arguments and hypotheses. Thus most visual links were already known or at least suspected and therefore produced little new knowledge. In contrast, computational methods establish links automatically and independently of the art historian. Doing so, these methods hold the potential to identify links between thousands of artworks, which have not been known by art historians before. In addition new approaches overcome spatial limitations which characterize previous visualization techniques such as Warburg's *Mnemosyne Atlas*. Thus, art historians are able to establish comprehensive relations between artists, determine influences and the adaptation of motifs or styles by different artists. Accessing and exploring data collections through linking means that the artwork is not studied in isolation but always seen in comparison to others; by searching for predecessors or successors, scholars aim to make sense of short-term and long-lasting phenomena in the history of art. Where and how did a motif such as the skull evolve? In which context was it presented? Can we find predecessors for the Cubist style? These and similar questions can be answered by establishing visual links. The idea of linking is extended by, for example, hyperlinks, where connections to other data even of different media can be established simply by clicking — the internet consists of such hyperlinks.

Warburg's *Mnemosyne Atlas* has shown how the process of linking allows us to open up image collections of art, thereby creating visual trails and lines of continuity. The idea of linking was taken up by Bush for his *memex* device which was intended to operate on associations and therefore resembled the workings of the human mind [Bush 2015]. What has been performed by Warburg to a limited extent and only hypothesized by Bush has become reality through computational methods: thousands of art images can be visualized in one space, thereby revealing visual patterns common in essentially distant artworks. Consequently past works can only provide suggestions, guidelines and validations for current practices in the digital humanities, because the physicality of photographic reproductions and space limitations do not resemble current conditions.

Previous sections have provided more information about how computational methods are beneficial to access and link images; this last paragraph focuses on the aspect of editing. Eventually the combination of digital data and computational methods enables an editing of images. The format and reproducibility of digital images allow the modification of images without actually altering the physical artwork. One can change formal features such as color, saturation, brightness, contrast with easily available editing programs such as AdobePhotoshop. Art historians are thus able to improve the quality of reproductions or study the importance of color, etc., for an artwork. Through image editing,

31

32

33

34

scholars might also crop digital images to compare details or reconstruct the hanging of artworks in a digital space to study viewing practices or exhibition contexts. More complex computational methods such as style transfer or image synthesis, which are introduced in the following section, describe a more conspicuous editing of digital images. Recent works by Gatys et al. and Sanakoyeu et al. have used computational models to stylize a real photo in the style of a specific artist, this process is referred to as “style transfer” in computer vision [Gatys et al. 2015] [Sanakoyeu et al. 2018]. Visual features of a style such as color, object shape, texture and brushstroke are learnt by a neural network and transferred onto the real image. Image stylization thus holds the potential to generate new perspectives and visualize how an artist would have painted a modern scene. Generative models which are used for style transfer also offer the potential to generate unseen views of an object by for example synthesizing two images. Through image generation, lost, damaged or unfinished artworks can be reconstructed, producing new knowledge for art history. Both examples indicate that generative models enable an editing of images unimaginable with analog data or analog methods.

Previous sections have identified the potentials of using computer technologies for accessing, linking and editing images. The discussion has highlighted that methods to access collections or link images based on visual similarities have already been suggested by art history and studied by other fields such as neuroscience. While these attempts provide valuable guidelines for current computational methods, examples given in the next section emphasize that models based on neural networks exceed previous attempts by far.

35

4. Examples of how computer models assist with data analysis

Computational methods based on neural networks can access, link and edit digital images in an innovative way, thus enabling a large-scale analysis of art datasets, which is adaptive to individual research questions, or offering new perspectives of the art object. These methods include search engines for object retrieval or visualization techniques which spot overarching links between images based on visual similarities — often unknown to art historians — or perform a comprehensive study of image context. Doing so they process thousands of images in a short amount of time and do not rely on descriptive image labels but operate on a purely visual basis. By allowing to alter style properties or the shape of the digital object, crop details or recreate missing parts virtually, methods for image generation offer new perspectives of the art object. Thus art historians can study for example the impact of these changes to the meaning of the art object. The following section provides concrete examples which demonstrate these potentials in practice and highlight that computational methods significantly impact our existing research culture: “Techniques of computational analysis can be used to reveal features of art historical artifacts in novel ways. These allow us to rethink the identity, purpose, use, and substance of objects” [Drucker 2013]. Computational methods truly alter the discipline of art history and exemplify the process from a digitized to a computational art history, thereby producing new knowledge unattainable with analog methods. In comparison to analog methods which only consider a small set of images, computational methods are capable to process thousands of images in a short amount of time and produce an accurate image representation. Since the advance of deep learning in circa 2015 [LeCun et al. 2015] 3, convolutional neural networks, which resemble the human brain, are preferably used in state-of-the-art computer vision to perform tasks such as classification, object detection [Simonyan and Zisserman 2015] [Crowley and Zisserman 2016] or finding visual similarities or patterns in large data collections [Seguin et al. 2016] [Shen et al. 2019]. While the human eye views images one-by-one, pairwise or in small groups, the computer is able to perceive thousands of images at the same time. The following tasks employ convolutional networks to perform a large-scale analysis and answer questions about form and content of the artwork or artistic style.

36

If we want to access data through a visual search, we have to have a clear understanding of what to search for. However, if one is unable to define a specific search query yet, computer-based approaches still give access to large data collections: Lev Manovich has demonstrated how the visualization of thousands of images in one space can lead to new insights. To automatically measure a number of visual characteristics of manga images and compare them respectively, Manovich et al. use digital image processing. For every image, algorithms calculate, for example, the portion of certain colors, saturation, contrast or texture and use these values to create a 2D visualization of images. Visualizations then show images arranged according to these values: images with similar values are displayed closely in the feature space, whereas dissimilar images are more distant. Consequently depending on the selected feature,

37

visualizations of images differ and present various views of the data [Manovich 2012]. In his essay “How to compare one million images?” Manovich emphasizes that a manual study of a handful of images only enhances drastic changes between images. A computational analysis, based on calculated values, considers millions of images and also recognizes subtle changes in form and content. Manovich and his digital lab demonstrate that by visualizing data in a digital space, scholars are given an overview of the structure of a dataset to identify popular but also uncommon visual characteristics in images running through a considerable part of art [Manovich 2012]. Although we are not using a specific image as a starting point for comparison or formulate a visual search query, we are able to access a large dataset simply by visualizing it.

Another visualization example shows a grouping of digital images according to a user-defined similarity. In Figure 3 images were sorted according to “age” and “social status.” Given some instances which display these characteristics, the neural network learnt to cluster the remaining images accordingly. Indeed the visualization shows that painted portraits display a variance in age and social status; this grouping is visible along the vertical and horizontal line. This way, art historians might study conventions of portraiture throughout time for certain age and social groups or identify uncommon depictions.

38



Figure 3. Visualization techniques enable to display large image collections and group them according to individual, user-defined similarity dimensions such as “age” and “social status” © Copyright Computer Vision Group, Heidelberg University

The search for recurring motifs and compositions in order to identify relations between artists or the recurring adaptation of motifs and styles are core tasks of art historians and chores which cannot be performed manually for large datasets. Much effort has been spent on developing computational methods for an object and content based search in digital datasets [Takami et al. 2014] [Crowley and Zisserman 2016] [Seguin et al. 2016]. Results include interactive interfaces and search engines which allow users to access individual collections and search for specific objects, compositions or other image regions to study their form and content. These retrieval systems are used to study medieval manuscripts [Yarlagadda et al. 2013], prints, architectural drawings [Arnold et al. 2013], collections of paintings or photographs of exhibitions [Lang and Ommer 2018b]. The Computer Vision Group at Heidelberg University developed an interactive retrieval system (see Figure 4) to perform a multiple object search in large image collections. In contrast to other tools, the search is performed on visual input and does not require text annotations, which are error-prone and tedious to obtain [Lang and Ommer 2018a]. The system thus enables a search for previously unknown or not indexed object categories. As Manovich stated “...we will not be able to give names to all of the variations of textures, compositions, lines, and shapes used even in a single chapter of Abara [a Japanese manga], let alone one million manga pages. We can proceed with traditional approaches so long as we limit ourselves to discussing manga iconography and other distinct visual elements which have standardized shapes and meanings” [Manovich 2012]. However art historians also

39

aim to find less standardized, uncommon elements in artworks, which have not been considered or defined by scholars before. As Manovich suggests, to find these uncommon traits is one great value of computational methods.

The screenshot displays a user interface for a visual search engine. At the top, there are navigation links for 'Collections', 'Resources', and 'User'. Below this is a header bar with a search input field labeled 'New Collection', a 'Scaling' slider, and a 'Reset' button. The main area is a grid of image thumbnails, each representing a collection. The collections shown are:

- Portraits Picasso**: msimon, 7/12/18, 9:14 PM; 330 images. Buttons: Show collection, Edit.
- Ukiyo Images**: msimon, 7/17/18, 8:30 PM; 1335 images. Buttons: Show collection, Edit.
- Test**: msimon, 8/27/18, 11:46 AM; 33 images. Buttons: Show collection, Edit.
- Sachsenspiegel**: compvis, 8/29/18, 6:19 PM; 810 images. Buttons: Show collection, Edit.
- Os-Gemeos**: compvis, 10/18/18, 4:30 PM; 844 images. Buttons: Show collection, Edit.
- Faille**: compvis, 10/26/18, 1:09 PM; 964 images. Buttons: Show collection, Edit.

Below the grid, there are two more partially visible thumbnails.

Figure 4. The Computer Vision Group developed a visual search engine for objects and parts in artistic images © Copyright Computer Vision Group, Heidelberg University

On the basis of individually user-defined queries (up to five) the system utilizes neural networks to find identical and similar objects and parts in images (see Figure 5), which might not have been annotated beforehand. Since users can find multiple objects in artworks, the system is able to perform an iconographic analysis of artworks: we might search for specific subjects which require a certain arrangement of objects, such as the Adoration of Christ or Christ's Crucifixion, and find variations of it. The system encourages users to give feedback and to judge detection results, pick favorites or rearrange search results. The search for user-selected objects and motifs in large datasets demonstrates the significance to develop explicitly computer-based methods for digital images. These methods do not require annotations but are unsupervised meaning that they do not require labels to find objects or visual patterns in data; input images are given without any corresponding output variables.

40

By finding identical or similar objects in large image collections given a user-defined query, art historians are able to study form developments or changes in context over a great period of time or produce new image connections. The search engine thus produces new knowledge, because neural networks process thousands of images and transform images into visual representations; consequently unknown links are build without human intervention. In addition users can create multiple trails consisting of the same images and are able to enlarge existing links by adding more images. The process of linking digitized images based on similarities resembles the manual approach of Warburg, however significantly exceeding his capabilities.

41

Although we have emphasized that online repositories are also selections and not available to everyone, search engines benefit from the availability of large, digital image collections. On this basis, engines significantly enhance research in art history, for example the study of adaptation processes or the tightness of artistic networks. The individual digital object can be placed within many visual trails or networks, emphasizing its multiple functions and changing meaning in different contexts. Moreover the identification of links is not restricted to art historians, who know about visual patterns in art history, but performable by a wider audience.

42

Collection: Collection Paintings

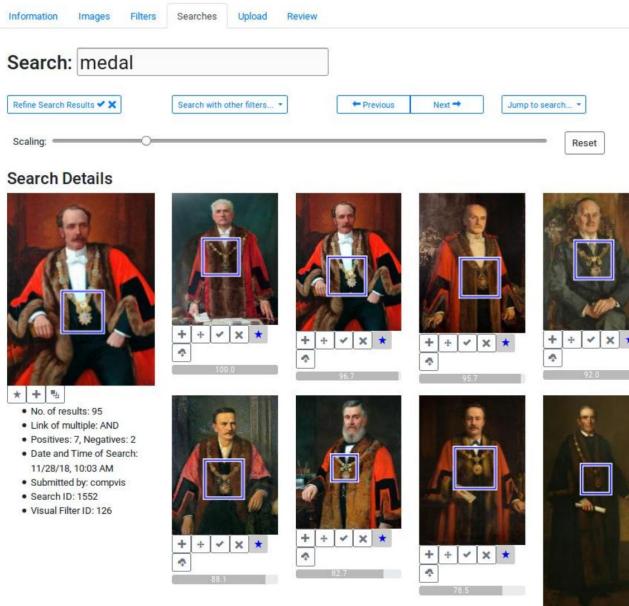


Figure 5. The illustration shows results for the search for a “medal” (query seen on the left) in a collection of paintings. Results are obtained by the search engine for objects of the Computer Vision Group at Heidelberg University © Copyright Computer Vision Group, Heidelberg University

The ability to edit images without actually altering the physical artwork is one advantage of digital data. While image editing programs such as Adobe Photoshop enable to cut details, change saturation or color very easily, thus providing knowledge about the impact of these changes for the presence and meaning of the art object, generative models first introduced by Goodfellow et al. perform a more compelling and drastic editing of images [Goodfellow et al. 2014]. In fact we can generate new perspectives of images or entirely new images. In order to reproduce the data and its specific properties, neural networks in generative models have to produce a data representation, which is a close approximation of the real data distribution. In contrast to discriminative models, which aim to classify data given certain features to predict a label, generative algorithms predict features on the basis of given labels: “GANs construct images, first generating arrangements of scene labels, filling in object labels from the scene labels, and then placing detail texture according to these sets of labels. In other words, GAN generator nets have learned how to plausibly compose objects together in scenes and to texture the object parts, following the instructions encoded in the vector” [Hertzmann 2019]. Eventually generative models obtain a detailed understanding of the data and perceive information about what is depicted and in which manner.



Figure 6. A two-story house with surrounding garden might be perceived and painted differently by various persons, thus producing diverse perspectives of one scene. © Zhou et al. (2018)

Generative models are capable to imitate the gaze of a person, for example the view of an artist, which is expressed in his or her paintings. The so-called “style transfer” describes the stylization of a real photo in the style of an artist such as Claude Monet (1840-1926) or Vincent van Gogh (1853-1890). Gatys et al. utilized a neural network for stylization: a real photo and a style example are given to the network; during processing the networks learns to retain the content of the photo but gradually approximating the style of the style example [Gatys et al. 2015]. This iterative process has been discarded by recent approaches for style transfer, which utilize GANs to improve image quality and style representation [Sanakoyeu et al. 2018]. However for both approaches a neural network learns the specifics of a style based on an individual style sample [Gatys et al. 2016] or a collection of images [Sanakoyeu et al. 2018] and transfers these learnt features to the real photo, thus generating a new image in the style of a specific artist. Research papers have argued that in order to generate a convincing stylization, it is better to use a collection of images since one example cannot fully comprehend and represent an artist’s style [Sanakoyeu et al. 2018] [Kotovenko et al. 2019b]. For example a landscape painting of Monet might accentuate different style features, such as color, brushstroke or shape, than a portrait in the same style. In order to replicate artistic style convincingly, networks have to consider these variations. Other works also focused on speeding up the stylization process [Johnson et al. 2016] or aimed to have more control over style transfer, determining the degree of stylization, select target regions, or combine multiple styles as input for one image [Huang and Belongie 2017]. Figure 6 shows the frontal view of a two-story house with surrounding garden. When given the task to paint the respective scene, an artist changes the color of the house and accentuates the contours of the transom windows; another artist highlights the flora in the foreground and blurs the contours of the house so that only a silhouette can be seen.



Figure 7. The figure shows stylization results for Paul Cézanne, Ernst Ludwig Kirchner, Picasso and Claude Monet using a collection of style images. © Kotovenko et al. (2019b)

Resulting images are representative of the subjective, very personal views of artists and provide information about personal and historical preferences and how the appearance of the scene differs for varying styles. Figure 7 shows stylizations of the same content image for artists Paul Cézanne (1839-1906), Ernst Ludwig Kirchner (1880-1938), Picasso (1881-1973) and Claude Monet (1840-1926). The stylized image in the style of Cézanne and generated by [Kotovenko et al. 2019b] is enlarged in Figure 8. The example gives an impression of how the artist would have painted the house; it shows characteristics which are prominent in his works from the late 1870s on onwards: black outlines, earthy tones and monochromatic image regions. By comparing all stylizations, we can study how style affects content and the appearance of the image. Kirchner's preference for bold, vibrant colors and vertical shapes creates a threatening impression. The house appears to be on fire and slowly dissolving; this visual effect is highlighted by the deformed windows. In contrast, the pure and pastel colors and pointillist brushstroke of Monet creates a spring-like and certainly lighter atmosphere. The outlines of the house are less accentuated and blend with the background. The style transfer performed by Kotovenko et al. specifically focuses on how style affects image content and alters fine image details such as the human figure or other objects in an artist-specific manner (see Figure 9) [Kotovenko et al. 2019b]. Style transfer does not remain a simple application for computer vision but has great relevance for art historical research. Art historians thus obtain new knowledge by asking: *What is style? Which properties define an artist's style? How does style affect content? How important are color, shape or brushstroke for an object? How does contemporary taste manifest itself in formal attributes?* Style transfer thus triggers a renewed style discussion which is built on broad observations about artistic style as well as small visual details magnified through stylizations.



Figure 8. The task of style transfer illustrates how artist Paul Cézanne (1839-1906) would have painted a scene. © Kotovenko et al. (2019b)

“Techniques of computational analysis can be used to reveal features of art historical artifacts in novel ways. These allow us to rethink the identity, purpose, use, and substance of objects” [Drucker 2013]. Style transfer presented a method which generates a new image by transferring formal properties, learnt from style examples, to a real photo and thus creating multiple views on one scene. Generative models then possess the ability to edit images in a novel way. Computational methods can go a step further and synthesize two images, altering the appearance of one object by adapting formal characteristics of the other [Isola et al. 2018] [Esser et al. VU-Net 2018b]. Figure 10 illustrates the synthesis of a person’s appearance with the shape of another: “to synthesize different geometrical layouts or change the appearance of an object, either shape or appearance can be retained from a query image, whereas the other component can be freely altered or even imputed from other images” [Esser et al. VU-Net 2018b]. To do so underlying algorithms have first learnt an exact representation of shape and appearance of an object in respective images. In the same manner, less complex objects such as shoes, bags or other clothing items can be manipulated. Computer models thus create images of different shoe types, displaying diverse colors or materials. Eventually these generated images which are still static can be brought to life [Esser et al. VU-Net 2018b]: by animating the object, the method allows an additional and alternative view, producing new knowledge about the artifact: how do different angles or motion impact the perception of the object? Moreover generative methods can be used to learn the behavior or control movements of a person. An activity such jumping rope is shown in a target video. The video data is used to train neural networks which then represent the appearance of a person and respective motion sequences. Based on this learning process the movement is transferred to another person (see Figure 11). In consequence, the process describes a synthesis of appearance and movement, instead of appearance and static pose as previously performed by Esser et al. (2018a). Similar to style transfer, image synthesis produces new knowledge about the artifact by altering its shape or appearance

with computational methods. It creates unseen views, highlights the specific characteristics of an object or style and presents the object in motion, thus giving new insights into the structure of an object or qualities of natural movement.



Figure 9. Style transfer models not only alter the entire image content but are capable of explicitly editing fine image details such as the human figure in an artistic-specific manner. The first row shows real image details and the second displays the same details but stylized. To evaluate the quality of stylizations, the last column holds patches from real artworks by Vincent van Gogh. © Kotovenko et al. (2019b)



Figure 10. Generative models enable the synthesis of a person's appearance with the shape of another. © Esser et al. (2018b)

5. Conclusion

This essay focused on digital data and computational methods and how a combination of both transforms information inherent in digital data into new knowledge. In order to do so we must involve transformative processes, namely from analog to digital data, from analog to digital methods and eventually from digital to computational methods. We argued that the specific characteristics of digital data, being for example reproducible or modifiable, require new computational methods different to analog methods. While analog methods such as Warburg's *Atlas* or Bush's suggested *memex* device provide valuable impulses, they can only give guidelines, suggestions and validations for computer-based approaches, because they were established under different conditions and mostly required knowledge about the expected outcome. Therefore, it is not enough to simply simulate known approaches but to develop new methods which are capable of performing these tasks with minimum or no human intervention. By considering the specifics of digital data, new computational methods offer a quantitative as well as a qualitative analysis of image repositories. Neural networks, which form the basis of current methods, automatically create links between thousands of images, allow to trace the adaptation of motifs or styles over a considerable amount of time and in various media, reveal analogies between images previously undiscovered or find local trends only shared by few artists. Manovich's visualizations of thousands of manga images or the search engine for objects and compositions in large datasets by the Computer Vision Group demonstrate how computational methods enable a simplified and explorative access to collections to study these visual patterns. This article also introduced generative models and their ability to significantly edit and alter images on the basis of a detailed data understanding. Models capture and transfer artistic styles, thus rendering a real photo in the style of an artist, generate new views or entirely new objects by synthesizing shape and appearance. Art

historians are thus propelled to study the specific characteristics of styles or how a change in shape impacts the appearance and meaning of the (art) object. Given examples demonstrate the great potentials, when art historians and computer scientists work together.

It is difficult to predict how the future of digital humanities will look like or where changes might lead. Future works however must address and reduce the skepticism towards computational methods and emphasize the general value of the work of the other discipline. Any efforts in this direction are beneficial and mark a positive change. This can be done by further encouraging collaborations and future research projects should aim to make image processing more transparent and machine learning understandable. So far computer scientists have relatively little insight into what happens during image processing in neural networks. In recent years research papers studied which regions are seen by the network when producing an image caption [Xu et al. 2015] or produced a representation of the image content during style transfer, illustrating how much content is preserved for a specific style by the computer model [Sanakoyeu et al. 2018]. More efforts must be spent to study the behavior of neural networks during image processing so that art historians are able to understand errors and give valuable feedback to computer scientists. Information or even hypotheses about why an image was classified incorrectly or why the network established a link between images can lead to an improvement of computational methods. Computer scientists are able to make corrections to models and assure that computational methods improve equally to the increasing amount and demands of digital data and become even more efficient for art historical research. We must further introduce seminars to the regular art history curriculum at universities, which teach on the usage of computational methods and provide an understanding of basic computational operations. This excludes the subject of digital humanities already taught at some universities. In addition we must communicate to students from computer science, how they can benefit from a collaboration: art data evokes new challenges due to different styles and new object categories and thus supports the development of robust algorithms. Style transfer for example not only requires an understanding of what is shown in the image but also in which manner. By teaching about digital possibilities and a cooperation, we not only reduce skepticism but create a new generation of researchers, who value a collaboration. Until then, we must further propel collaborations between researchers from computer vision and art history and make methods more accessible; so far only a limited number of people — mostly academics — have access to methods and tools. Not least the survey conducted by Emily L. Spratt and Ahmed Elgammal in 2014 indicated that we are on the right path. Authors concluded that art historians generally “wished they knew more about the movement of artificial intelligence into the humanities — or the attention to art history by computer scientists, of which half the group surveyed indicated that they were interested to get direct feedback from art historians as they developed new applications for computer vision technology” [Spratt and Elgammal 2014].



Figure 11. Can computational models learn the behavior or control movements of a person? The image shows persons in an activity priorly unknown to them. © Esser et al. (2018a)

Appendix

- The *Time Machine Project* originated out of the Venice Time Machine, which was initiated by Frédéric Kaplan and his team at the École Polytechnique Fédérale de Lausanne. It aims to sketch the history of Venice and its evolution over a period of thousand years. See the project page for more information: <https://timemachine.eu/> (accessed March 23, 2019).
- Generative models based on neural networks are used to learn explicit data distributions of training sets without any supervision. Data distributions consist of mathematical functions which show all values of the data. Based on learnt distribution, models generate new data points which show slight variations to the original data. By learning the particularities of the data and producing new data points, models are able to alter for example the shape or appearance of objects presented in images or transfer artistic style to real photos. The two most common generative models are Variational Autoencoders (VAE) and Generative Adversarial Networks (GANs). The VAE approach can learn a complex data distribution such as images without any human supervision using neural networks [Pandey 2018]. The autoencoder encodes “an input image to a much smaller dimensional representation which can store... information about the input data distribution” [Pandey 2018] and on this basis learns to generate similar looking images. A full mathematical explanation of VAEs is given in the tutorial by Doersch (2016). In contrast the GAN framework consists of two adversarial components, namely a generator and a discriminator. While the former learns to capture the data distribution, the latter must estimate if an image sample came from the generator or real data distribution. GANs were first introduced by Goodfellow et al. (2014).
- Deep learning is a subcategory of machine learning. Deep learning methods allow to learn representations of large datasets with different degrees of abstraction. These methods achieve state-of-the-art results in tasks such as classification, speech recognition or visual object recognition. (Convolutional) neural networks are large computing systems used in deep learning, which are inspired by biological neural networks and consist of multiple stages. At every stage the network extracts intermediate features from images: while low-level features encapsulate information about color, shape or edges, deeper stages contain high level information about semantically abstract concepts such as nose, eyes or faces and geometrical relations. Essentially, these features are descriptions of an image and are required to

understand the image and to extract information relevant for solving a given task (i.e., object recognition) [LeCun et al. 2015].

- The research project *Bilderfahrzeuge: Aby Warburg's Legacy and the Future of Iconology* explores the migration of images, objects , etc., on the basis of Warburg's theories. It studies the transfer of image concepts and also focuses on the image itself. The project theme is material images but also includes linguistic images and studies the access, treatment and use of images in different disciplines such as literature, the humanities and science. More information can be found on the website: <https://bilderfahrzeuge.hypotheses.org/> (accessed February 10, 2019).

Works Cited

ADHO 2019 Alliance of Digital Humanities Organizations 2019. Available from: <http://www.adho.org/>. (Accessed November 21, 2019).

Arnold and Tilton 2019 Arnold, T. and Tilton, L. "Distant viewing: analyzing large visual corpora," *Digital Scholarship in the Humanities*, fqz013, Oxford University Press, Oxford, 0:0 (2019).

Arnold et al. 2013 Arnold, M., Bell, P. and Ommer, B. "Automated Learning of Self-Similarity and Informative Structures in Architecture," *Scientific Computing & Cultural Heritage* (2013).

Arts and Culture Experiments 2016 Arts and Culture Experiments, Google. Available from: <https://experiments.withgoogle.com/collection/arts-culture>. (Accessed November 15, 2019).

Bar 2009 Bar, M. "The Proactive Brain: Memory for Predictions," *Philosophical Transaction of the Royal Society*, The Royal Society Publishing, London, 364 (2009): 1235-1243.

Bautista et al. 2016 Bautista, M. A., Sanakoyeu, S., Tikhoncheva, E. and Ommer, B. "Cliquecnn: Deep unsupervised exemplar learning." In *Advances in Neural Information Processing Systems*, NIPS, Barcelona, December 2016, 3846-3854.

Benjamin 1969 Benjamin, W. "The Artwork in the Age of Mechanical Reproduction." In *Illuminations: Essays and Reflections*, ed. Hannah Arendt. trans. Harry Zohn, New York, Schocken Books, (1969): 52-78.

Bishop 2018 Bishop, C. "Against Digital Art History," *International Journal for Digital Art History* 3 (2018): 123-133.

Bonfiglioli and Nanni 2015 Bonfiglioli, R. and Nanni, F. "From close to distant and back: how to read with the help of machines." In *International Conference on the History and Philosophy of Computing*, Pisa, October 2015, 87-100.

Bush 2015 Bush, V. "As We May Think", *The Atlantic*. Available from: [//www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/](http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/). (Accessed March 7, 2019).

Chen 1976 Chen, P. P.-S. "The Entity-Relationship Model – Toward a Unified View of Data", *ACM Transactions on Database Systems*, Association for Computing Machinery, New York, 1:1 (March 1976): 9-36.

Codd 1970 Codd, E.F. "A relational model of data for large shared data banks", *Communications of the ACM*, Association for Computing Machinery, New York, 13:6 (1970): 377-387.

Cornell University Library 2016a Cornell University Library 2016, "Mnemosyne. Meandering through Aby Warburg's Atlas," Cornell University. Available from: <https://warburg.library.cornell.edu/about>. (Accessed November 18, 2019).

Cornell University Library Home 2016b Cornell University Library 2016, "Mnemosyne. Meandering through Aby Warburg's Atlas," Cornell University. Available from: <https://warburg.library.cornell.edu/>. (Accessed November 20, 2019).

Crowley and Zisserman 2016 Crowley, E.J. and Zisserman, A. "The art of detection." In *Proceedings of the European Conference on Computer Vision*, ECCV, Amsterdam, October 2016, 721-737.

Diagne et al. 2018 Diagne, C., Barradeau, N., Doury, S. "t-SNE Map, Google's Arts and Culture Experiments." Available from: <https://experiments.withgoogle.com/t-sne-map>. (Accessed December 3, 2019).

Doersch 2016 Doersch, C. "Tutorial on variational autoencoders" arXiv preprint arXiv:1606.05908 (2016).

Drucker 2013 Drucker, J. "Is There a 'Digital' Art History?," *Visual Resources*, 29 (2013): 5-13.

Eigenstetter et al. 2014 Eigenstetter, A., Takami, M. and Ommer, B. "Randomized Max-Margin Compositions for Visual Recognition." In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, CVPR, Columbus, Ohio,

- Elgammal et al. 2018a** Elgammal, A., Kang, Y. and Leeuw, M.D. "Picasso, Matisse, or a Fake? Automated Analysis of Drawings at the Stroke Level for Attribution and Authentication." In *Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, February 2018.
- Elgammal et al. 2018b** Elgammal, A., Liu, B., Kim, D., Elhoseiny, M. and Mazzone, M. "The shape of art history in the eyes of the machine." In *Thirty-Second AAAI Conference on Artificial Intelligence*, New Orleans, April 2018.
- Elliott 2016** Elliott, L. "Spread of internet has not conquered 'digital divide' between rich and poor-report", *The Guardian*. Available from: <https://www.theguardian.com/technology/2016/jan/13/internet-not-conquered-digital-divide-rich-poor-world-bank-report>. (Accessed November 18, 2019).
- Esser et al. 2018a** Esser, P., Haux, J., Milbich, T. and Ommer, B. "Towards Learning a Realistic Rendering of Human Behavior." In *Proceedings of the European Conference on Computer Vision*, ECCV, workshops, Munich, September 2018, 409-425.
- Esser et al. VU-Net 2018b** Esser, P., Sutter, E. and Ommer, B. "A Variational U-Net for Conditional Appearance and Shape Generation." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Salt Lake City, Utah, June 2018, 8857-8866.
- Gatys et al. 2015** Gatys, L.A., Ecker, A.S. and Bethge, M. "A neural algorithm of artistic style" arXiv preprint arXiv:1508.06576 (2015).
- Gatys et al. 2016** Gatys, L.A., Ecker, A.S. and Bethge, M. "Image style transfer using convolutional neural networks." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Las Vegas, Nevada, June, July 2016, 2414-2423.
- Girshick 2015** Girshick, R. "Fast r-cnn." In *Proceedings of the IEEE international conference on computer vision*, ICCV, Las Condes, Chile, December 2015, 1440-1448.
- Gogolla 1994** Gogolla, M. "An extended entity-relationship model: fundamentals and pragmatics." Vol. 767. Springer, Heidelberg (1994).
- Goodfellow et al. 2014** Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. "Generative adversarial nets." In *Advances in Neural Information Processing Systems*, NIPS, Montreal, December 2014, 2672-2680.
- Goodfellow et al. 2016** Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*. MIT Press, Cambridge, London (2016).
- Google Arts and Culture 2011** Google Arts and Culture. Available from: <https://artsandculture.google.com/>. (Accessed December 1, 2019).
- Hatt/Klonk 2006** Hatt, M. and Klonk, C. *Art History. A Critical Introduction to its Methods*. Manchester University Press, Manchester (2006).
- Hertzmann 2019** Hertzmann, A. "Aesthetics of Neural Network Art" arXiv preprint arXiv:1903.05696 (March 2019).
- Hristova 2016** Hristova, S. "Images as data: cultural analytics and Aby Warburg's Mnemosyne," *International Journal for Digital Art History*, 2 (2016).
- Hu et al. 2018** Hu, H., Gu, J., Zhang, Z., Dai, J., Wei, Y. "Relation networks for object detection." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Salt Lake City, June 2018, 3588-3597.
- Huang and Belongie 2017** Huang, X. and Belongie, S.J. "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization." In *Proceedings of the International Conference on Computer Vision*, ICCV, Venice, February 2017, 1510-1519.
- Hvattum and Hermansen 2004** Hvattum, M. and Hermansen, C. *Tracing Modernity: Manifestations of the Modern in Architecture and the City*. Routledge, New York, London (2004).
- ImageNet Large Scale 2016a** Imagenet, *Large Scale Visual Recognition Challenge 2016* (ILSVRC2016), UNC Vision Lab. Available from: <http://image-net.org/challenges/LSVRC/2016/index#introduction>. (Accessed November 28, 2019).
- ImageNet Results 2016b** Imagenet, *Large Scale Visual Recognition Challenge 2016* (ILSVRC2016), UNC Vision Lab. Available from: <http://image-net.org/challenges/LSVRC/2016/results>. (Accessed November 28, 2019).
- Impett and Süsstrunk 2016** Impett, L., and Süsstrunk, S. "Pose and pathosformel in Aby Warburg's bilderatlas." In

- Isola et al. 2018** Isola, P., Zhu, J.Y., Zhou, T. and Efros, A. A. "Image-to-Image Translation with Conditional Adversarial Networks." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Salt Lake City, Utah, June 2018, 1125-1134.
- Johnson 2012** Johnson, C. *Memory, Metaphor, and Aby Warburg's Atlas of Images*. Cornell University Press, Ithaca (2012).
- Johnson et al. 2016** Johnson, J., Alahi, A. and Fei-Fei, L. "Perceptual losses for real-time style transfer and super-resolution." In *Proceedings of the European Conference on Computer Vision*, ECCV, Amsterdam, October 2016, 694-711.
- Kalkstein 2019** Kalkstein, M. "Aby Warburg's Mnemosyne Atlas: On Photography, Archives, and the Afterlife of Images," *Rutgers Art Review: The Journal of Graduate Research in Art History*, Rutgers School of Arts and Sciences, New Brunswick, 35 (2019): 50-73.
- Karayev et al. 2013** Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A. and Winnemoeller, H. "Recognizing image style" arXiv preprint arXiv:1311.3715 (2013).
- Kienle 2017** Kienle, M. "Digital Art History: Beyond the Digitized Slide Library. An Interview with Johanna Drucker and Miriam Posner," *Art@ s Bulletin*, 6.3:9 (2017).
- Kim et al. 2019** Kim, S., Min, D., Jeong, S., Kim, S., Jeon, S. and Sohn, K. "Semantic Attribute Matching Networks." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Long Beach, June 2019, Diego, 12339-12348.
- Kotovenko et al. 2019a** Kotovenko, D., Sanakoyeu, A., Lang, S. and Ommer, B. "Content and style disentanglement for artistic style transfer." In *Proceedings of the IEEE International Conference on Computer Vision*, ICCV, Seoul, Korea, October/November 2019, 4422-4431.
- Kotovenko et al. 2019b** Kotovenko, D., Sanakoyeu, A., Lang, S. and Ommer, B. "A Content Transformation Block for Image Style Transfer." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Long Beach, California, June 2019, 10032-10041.
- Krizhevsky et al. 2012** Krizhevsky, A., Sutskever, I., and Hinton, G.E. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, NIPS, Lake Tahoe, December 2012, 1097-1105.
- Krogh 2008** Krogh, A. "What are artificial neural networks?" *Nature Biotechnology*, Nature Publishing Group, London, 26 (February 2008): 195-197.
- Lang and Ommer 2018a** Lang, S. and Ommer, B. "Attesting Similarity: Supporting the Organization and Study of Art Image Collections with Computer Vision," *Digital Scholarship in the Humanities*, Oxford University Press, Oxford, 33:4 (2018): 845-856.
- Lang and Ommer 2018b** Lang, S. and Ommer, B. "Reconstructing Histories: Analyzing Exhibition Photographs with Computational Methods," *Arts, Computational Aesthetics*, 7:64 (2018).
- Lang and Ommer 2018c** Lang, S. and Ommer, B. "Reflecting on How Artworks Are Processed and Analyzed by Computer Vision." In *Proceedings of the European Conference on Computer Vision*, ECCV, workshops, Munich, September 2018, 647-652.
- LeCun et al. 2015** LeCun, Y., Bengio, Y. and Hinton, G. "Deep Learning," *Nature*, Nature Publishing Group, London, 521 (May 2015): 436-444.
- Li et al. 2012** Li, J., Yao, L., Hendriks, E. and Wang, J.Z. "Rhythmic brushstrokes distinguish van Gogh from his contemporaries: findings via automated brushstroke extraction." In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34.6 (2012), 1159-1176.
- Lorenz et al. 2019** Lorenz, D., Bereska, L., Milbich, T., and Ommer, B. "Unsupervised Part-Based Disentangling of Object Shape and Appearance." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Long Beach, June 2019, 10955-10964.
- Manovich 2012** Manovich, L. "How to compare one million images?" *Understanding Digital Humanities*, Palgrave Macmillan, London (2012): 249-278.
- Monroy et al. 2011** Monroy, A., Eigenstetter, A. and Ommer, B. "Beyond Straight Lines – Object Detection Using

Curvature.” In *Proceedings of the International Conference on Image Processing*, Brussels, September 2011, 3561-3564.

Pandey 2018 Pandey, P. “Deep Generative Models,” *Towards Data Science*. Available from: <https://towardsdatascience.com/deep-generative-models-25ab2821afd3>. (Accessed March 20, 2019).

Postel 2017 Postel, J.P. *Der Fall Arnolfini. Auf Spurensuche in einem Gemälde von Jan van Eyck*. Verlag Freies Geistesleben, Stuttgart (2017).

Redmon et al. 2016 Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. “You only look once: Unified, real-time object detection.” In *Proceedings of the IEEE conference on computer vision and pattern recognition*, CVPR, Las Vegas, June/July 2016, 779-788.

Ren et al. 2015 Ren, S., He, K., Girshick, R. and Sun, J. “Faster R-CNN: Towards real-time object detection with region proposal networks.” In *Proceedings of the Advances in Neural-Information Systems*, NIPS, Montreal, December 2015, 91-99.

Rijksmuseum 2019 Rijksmuseum. *Operation Night Watch to start at the Rijksmuseum*. Available from: <https://www.rijksmuseum.nl/en/press/press-releases/operation-night-watch-to-start-at-the-rijksmuseum>. (Accessed November 26, 2019).

Saleh and Elgammal 2015 Saleh, B. and Elgammal, A. “Large-scale Classification of Fine-Art Paintings: Learning the Right Metric on the Right Feature,” *International Journal for Digital Art History*, 2 (2015).

Saleh et al. 2016 Saleh, B., Abe, K., Arora, R.S. and Elgammal, A. “Toward automated discovery of artistic influence,” *Multimedia Tools and Applications*, Springer, Heidelberg, 75:7 (2016): 3565-3591.

Sanakoyeu et al. 2018 Sanakoyeu, A., Kotovenko, D., Lang, S. and Ommer, B. “A Style-Aware Content Loss for Real-time HD Style Transfer.” In *Proceedings of the European Conference on Computer Vision*, ECCV, Munich, September 2018, 698-714.

Schlecht et al. 2011 Schlecht, J., Carqué, B. and Ommer, B. “Detecting Gestures in Medieval Images.” In *Proceedings of the International Conference on Image Processing*, ICIP, Brussels, September 2011, 1285 – 1288.

Seguin et al. 2016 Seguin, B., Striolo, C. and Kaplan, F. “Visual link retrieval in a database of paintings.” In *Proceedings of the European Conference on Computer Vision*, Workshop, ECCV, Zurich, September 2016, 753-767.

Shen et al. 2019 Shen, X., Efros, A. A. and Aubry, M. “Discovering Visual Patterns in Art Collections with Spatially-consistent Feature Learning,” arXiv preprint arXiv:1903.02678 (March 2019).

Simonyan and Zisserman 2015 Simonyan, K. and Zisserman, A. “Very deep convolutional networks for large-scale image recognition.” In *International Conference on Learning Representations*, ICLR, San Diego, May 2015, 1-14.

Spratt and Elgammal 2014 Spratt, E. L. and Elgammal, A. “The Digital Humanities Unveiled: Perceptions Held by Art Historians and Computer Scientists about Computer Vision Technology” arXiv preprint arXiv:1411.6714 (2014).

Stork and Johnson 2006 Stork, D. G., and Johnson, M. K. “Estimating the location of illuminants in realist master paintings Computer image analysis addresses a debate in art history of the Baroque.” In *International Conference on Pattern Recognition*, ICPR, Hongkong, August 2006, 255-258.

Takami et al. 2014 Takami, M., Bell, P. and Ommer, B. “An Approach to Large Scale Interactive Retrieval of Cultural Heritage.” In *Eurographics Workshop on Graphics and Cultural Heritage*, Darmstadt, October 2014, 87-95.

Teorey et al. 1986 Teorey, T. J., Dongqing, Y., and Fry, J. P. “A logical design methodology for relational databases using the extended entity-relationship model,” *ACM Computing Surveys (CSUR)*, Association for Computing Machinery, New York, 18:2 (1986): 197-222.

The Next Rembrandt 2016 *The Next Rembrandt*, ING, Microsoft, TU Delft, Mauritshuis, The Rembrandt House Museum. Available from: <https://www.nextrembrandt.com/>. (Accessed November 26, 2019).

The Warburg Institute 2018 The Warburg Institute, The Library of Aby Warburg, School of Advanced Study University of London. Available from: <https://warburg.sas.ac.uk/library/library-aby-warburg> (Accessed November 19, 2019).

Thiel and Reuter 2019 Thiel, T. and Reuter, T. *Virtuelle Mauer/ReConstructing the Wall*. Available from: <http://www.virtuelle-mauer-berlin.de/index.htm>. (Accessed November 26, 2019).

Ufer et al. 2019 Ufer, N., Lui, K.T., Schwarz, K., Warkentin, P. and Ommer, B. “Weakly Supervised Learning of Dense Semantic Correspondences and Segmentation.” In *German Conference on Pattern Recognition*, GCPR, Dortmund,

- Wang et al. 2017** Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., Wang, X. and Tang, X. "Residual attention network for image classification." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, CVPR, Honolulu, July 2017, 3156-3164.
- Warburg and Rampley 2009** Warburg, A. and Rampley, M. "The absorption of the expressive values of the past," *Art in Translation*, Taylor & Francis, London, 1:2 (2009): 273-283.
- Warnke 2000** Warnke, M. "Aby Warburg. Der Bilderatlas Mnemosyne." In H. Bredekamp, M. Diers, K.W. Forster, N. Mann, S. Settis, and M. Warnke (eds), *Aby Warburg. Gesammelte Schriften*, Vol. 2,1, Berlin (2000).
- West 2011** West, M. "Developing high quality data models." Elsevier, Amsterdam (2011).
- Xu et al. 2015** Xu, K. et al. "Show, attend and tell: Neural image caption generation with visual attention." In *International conference on machine learning*, ICML, Lille, July 2015, 2048-2057.
- Yarlagadda et al. 2013** Yarlagadda, P., Monroy, A., Carque, B. and Ommer, B. "Towards a Computer-Based Understanding of Medieval Images." In H. G. Bock, W. Jaeger, M. J. Winckler (eds), *Scientific Computing and Cultural Heritage*, Springer, Heidelberg (2013): 89-97.
- Zhou et al. 2018** Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., and Torralba, A. "Places: A 10 million image database for scene recognition," *IEEE transactions on pattern analysis and machine intelligence*, PAMI, 40:6 (2018): 1452-1464.