

Towards a Classification of Neoclassical Objects in Interior Scenes

Simon Donig, Maria Christoforaki, Bernhard Bermeitinger und Siegfried Handschuh

Digitale Forschungsinstrumente sind disruptive Technologien mit dem Potential, grundlegend die Art und Weise zu verändern, wie wir in den historischen Bildwissenschaften, in der Architektur- und Designgeschichte oder den Material Culture Studies forschen, ja welche Vorstellungen wir von unseren Disziplinen entwerfen. Künstliche Intelligenz und Verfahren maschinellen Lernens werden in den bildbezogenen Digitalen Geisteswissenschaften zum Zweck der Bildanalyse sowie der Klassifizierung visueller Merkmale genutzt. Waren Methoden im Bereich der Bildanalyse lange Zeit von niedrigschwelligen Methoden wie der Bildung von Clustern auf der Grundlage von Farbverteilungen oder lokalen Kontrasten für diesen Bereich bestimmend, zielen Forschungsinstrumente nunmehr darauf ab, Bildmerkmale zu identifizieren, zu klassifizieren oder gar semantisch zu segmentieren. Sie bieten also, anders ausgedrückt, erstmals informatisch halbautomatischen Zugang zur dem Bild eigenen Ebene der Repräsentation. Das Neoclassica-Projekt hat zum Ziel, Forschenden einen Zugang zu Instrumenten und Methoden zu bieten, die eine computergestützte Analyse materieller Kultur, genauer struktureller Merkmale und ästhetischer Formen klassizistischer Artefakte bieten. Wir möchten damit sowohl mehr hypothesen-getriebene, als auch explorative Zugänge zum Medium Bild unterstützen. Gegenwärtig konzentrieren wir uns dabei auf Raumkunst, insbesondere Mobiliar und Innenräume, aber auch Architektur und deren jeweilige bildliche Darstellung. In diesem Beitrag stellen wir unseren Zugang zur Analyse von Zimmerbildern vor, angefangen mit der Klassifizierung in Darstellungen von Einzelartefakten bis hin zur Identifizierung von Objekten in komplexen zeitgenössischen Interieurdarstellungen. Abschließend zeigen wir Anwendungsmöglichkeiten der Technologie und Perspektiven der weiteren Forschung auf.

Introduction

The introduction of digital instruments in the history of design, art and architecture like so many disruptive technologies are about to fundamentally change the way we conduct and even envision our very disciplines.¹ Lately, Artificial Intel-

1 This transformation has been hailed as a “Computational Turn” (David Berry, “The Computational Turn: Thinking about the Digital Humanities”, in: *Culture Machine* 12 (2011), pp. 1–22, URL: <http://sro.sussex.ac.uk/49813/>, accessed 2018-12-14) and has been early on stated by art historians such as Hubertus Kohle, “Ordnung und Umbruch in der Wissenschaft der Kunstgeschichte: Der Einfluss des Digitalen”, in: *Reibungspunkte: Ordnung und Umbruch in Architektur und Kunst*, ed. by Hanns Hubach, Petersberg 2008, pp. 141–144. Recently espe-

ligence and Machine Learning techniques are used in the Visual Digital Humanities for the purpose of image analysis and the classification of features alike.² They reach beyond a low-level image processing (e.g. by grouping images by colour distribution or contrast) seeking to identify, classify, and even semantically segment visual features in imagery.³ In other words, they provide for the first time computational access to the representational dimension of images.

The Neoclassica project was conceived to harness these technologies in order to provide scholars with such new instruments and methods for the computer-assisted analysis of material culture, specifically, structural features and aesthetic forms of neoclassical artefacts. We aspire to support both hypothesis-driven and exploratory approaches in the Bildwissenschaften, Cultural Studies and Humanities alike. We currently concentrate on Raumkunst (in particular furniture and interior spaces) as well as architecture and their respective visual representations.

Here, we present our path towards a visual analysis of interiors starting from the initial classification of single artefact images and moving to artefact identification in complex depictions of period interiors. Finally, we present our contemplations on the use of this methodology and possible extensions of it.

The Classical Movement and the Interior

Antiquarianism played an important role in shaping the rise of European modernity. The ascent of classical antiquity to a cultural paradigm during the *Sattelzeit*⁴ was profoundly interwoven with processes of fundamental social, economic and cultural transformation, such as growing urbanisation, the evolution of a consumer-culture, new conceptions of privacy and property, and even the human self. The Classical movement was remarkable in its spread both geographically and socially as well as the intensifying processes of interaction and exchange across territories. Patrons returning from their *grand tour* formed learned clubs

cially the impact of computer vision on the Visual Digital Humanities has been discussed by Simon Doing et al.: "Der Ferne Blick. Bildkorpora und Computer Vision in den Geistes- und Kulturwissenschaften: Stand – Visionen – Implikationen", in: *DHd 2018: Kritik der digitalen Vernunft*, ed. by Georg Vogeler, Köln 2018, pp. 86–89, URL: <http://dhd2018.uni-koeln.de/wp-content/uploads/boa-DHd2018-web-ISBN.pdf>, accessed 2018-12-14.

2 Peter Bell and Björn Ommer, "Training Argus, Ansätze zum automatischen Sehen in der Kunstgeschichte", in: *Kunstchronik* 68/8 (2015), pp. 414–420; Björn Ommer and Peter Bell, "Digital Connoisseur? How Computer Vision Supports Art History", in: *Il metodo Del conoscere – approcci, limiti, prospettive Connoisseurship nel XXI secolo*, ed. by Stefan Albl and Alina Aggujaro, Roma 2016, pp. 187–200.

3 Sabine Lang and Björn Ommer: "Attesting Similarity: Supporting the Organization and Study of Art Image Collections with Computer Vision", in: *Digital Scholarship in the Humanities* 33/4 (2018), pp. 845–56. DOI: <https://doi.org/10.1093/llc/fqy006>, accessed 2018-12-14.

4 For a recent assessment of the scholarly productivity of Reinhart Koselleck's concept see Daniel Fulda, "Sattelzeit. Karriere und Problematik eines kulturwissenschaftlichen Zentralbegriffs", in: *Sattelzeit: Historiographiegeschichtliche Revisionen*, ed. by Elisabeth Décultot and id., Berlin 2016, pp. 1–16.

of Dilettanti, maintained vast letter networks, or even transformed their private collections into public institutions like the Boston Athenaeum.⁵

The medial revolution around 1800 spread knowledge of the classical *goût* far beyond the most wealthy and powerful in their respective societies. Publications such as the *Journal de la mode et du goût* (appeared under different titles 1786–1793), *Ackermann's Repository of Arts, Literature ...* (app. 1809–1829) or the *Journal des Luxus und der Moden* (app. 1797–1812) contributed to the formation of consumers' tastes. Specialised craftsmen travelled across vast territories spreading new techniques and technologies in carpentry or metal work, learning both from their peers and through the study of the antiquities, later forming their own workshops (often maintained for multiple generations). Some of those had a fundamental impact on the reception of the Classical tradition itself such as the Roentgen's workshop in Neuwied or the Jacob's workshop in Paris,⁶ and contributed significantly to the creation of almost global markets for consumer goods such as furniture and bronzes.

The common reference to classical antiquity, as well as aesthetic divergence due to local traditions, social requirements and financial limitations or even political conditions, provides in our eyes a field particularly well suited for the application of computational methods. The widespread reception of classical antiquity, as well as technological improvements such as printmaking techniques in the form of stipple engraving, mezzotint, and aquatint,⁷ produced a vast wealth of contemporary visuals.

Additionally, these visuals provide a specific kind of evidence, due to a transformation of the very practice of seeing and depicting. During this period the modern notion of the interior "with significance as a physical, three-dimensional space, as well as an image, whether it be a two-dimensional representation such as a painting, a print in a portfolio of decoration, or a flat backdrop that could conjure up an interior as a theatrical scene" emerged.⁸ As a genre, the interior deviated from the emblematic and allegorical image-language of the baroque, employing a more 'naturalistic' and 'reconstructive' approach, now supposed to mirror the personality of its inhabitants.⁹ Christiane Lukatis has pointed out that the interior started out as a primarily private form of memory that hence depicted

5 Jason M. Kelly, *The Society of Dilettanti*, New Haven, CT 2009; Bruce Redford, *Dilettanti: The Antique and the Antique in Eighteenth-Century England*, Los Angeles, CA 2008; Hina Hirayama, "With Éclat" – the Boston Athenæum and the Origin of the Museum of Fine Arts, Boston, Boston, MA 2013.

6 Wolfram Koeppe, *Extravagant Inventions: The Princely Furniture of the Roentgens*, New York, NY 2012; for the Jacobs see the older but still authoritative biographies by Hector Lefuel, *Georges Jacob, ébéniste du XVIIIe siècle*, Paris, 1923 and id., *François-Honoré-Georges Jacob-Desmalter Ébéniste de Napoléon Ier et de Louis XVIII.*, Paris 1926.

7 Jeremy Aynsley and Charlotte Grant, ed., *Imagined Interiors: Representing the Domestic Interior since the Renaissance*, ed. by, London 2006, p. 100.

8 Charles Rice, *The Emergence of the Interior – Architecture, Modernity, Domesticity*, London 2007, p. 2.

9 Rainer Schoch, "Repräsentation und Innerlichkeit. Zur Bedeutung des Interieurs im 19. Jahrhundert", in: *Mein blauer Salon: Zimmerbilder der Biedermeierzeit*, ed. by Christiane Lukatis, Nürnberg 1995, pp. 11–16, here p. 14.

very personal and rarely representational settings. Following its private function, interior images are usually executed as drawings rather than oil paintings.¹⁰

The genre thus ranged from the private *lieux de souvenir* to the “painted inventory”,¹¹ as well as, from the idealised milieu¹² as a portrait to a strict documentary function.¹³

Interior images may hence fulfil multiple functions of representation. Depending on the research interest, they may hold evidentiary value, informing about domestic culture, perceptions of humans of their environment or programmatic and aesthetic conceptions held.¹⁴ They can also inform about the study of processes of stylistic transformation and the application of iconographic programs. The huge digitisation efforts in recent years have given us unprecedented access to sources, facilitating the employment of methods and instruments from the Digital Humanities that directly benefit such research in multiple ways.

The Neoclassica Framework

The *Neoclassica* framework evolved in the above context and is composed of a top-down human created knowledge codification harnessing the domain-expert knowledge and a bottom-up data-driven approach to knowledge extraction from sources in multiple modes, exploiting the powers of computationally processing large amounts of data.

The Neoclassica ontology

Domain expert knowledge has to be formalized so that it can be automatically processed by a computer. In *Neoclassica*, we do this by means of a multilingual research-oriented formal ontology.¹⁵ Additionally, we strive to reflect the specific

10 Christiane Lukatis, “Zimmerbilder – Entwicklung und Charakter des Genres”, in: *Mein blauer Salon: Zimmerbilder der Biedermeierzeit*, ed. by ead., Nürnberg 1995, pp. 17–26, here pp. 19, 21.

11 Ibid.

12 This observation holds also for other genres such as the social caricature or the political portrait. For instance, in George Cruikshanks temperance inspired series of etchings called *The Bottle* from 1847 (George Cruikshank and Charles Mackay, *The Bottle. In Eight Plates [...]*, London 1847, URL: <https://www.bl.uk/collection-items/the-bottle--a-series-of-temperance-themed-illustrations-by-george-cruikshank-with-poetry-by-charles-mackay>, accessed 2018-12-14) the gradual social and moral demise of the protagonist is represented through the gradual loss of the bourgeois interior surrounding the family; or the 1808 portrait of Charles-Maurice de Talleyrand-Périgord by François Gérard depicts him in a relatively private setting yet with furniture in the Louis XVI style. In the context of Talleyrand having recently resigned as a foreign minister to protest Napoléons policy towards Russia, this choice of an ancien régime style at the heyday of the Empire style in a painting that was publicly exhibited after its completion in 1808 can hardly be called accidental. (For visual references see Fig. 10.4, Farbteil.)

13 Lukatis, “Zimmerbilder”, p. 24.

14 Peter Burke, *Eyewitnessing: The Uses of Images as Historical Evidence. Picturing History*, London 2001, pp. 81–87.

15 Simon Doing et al., “Neoclassica – A Multilingual Domain Ontology. Representing Material Culture from the Era of Classicism in the Semantic Web”, in: *Computational History and Data-Driven Humanities. CHDDH 2016*, ed. by Bojan Bozic et al., Cham 2016 (IFIP Advances in

chronological period semantics. Thus, concepts and terms used to describe artefacts and their structure are based on period sources such as drawing, model, and pattern books or other treatises reflecting the conceptual control over the production of artefacts. The ontology contains concepts regarding furnishing and furniture, their components and aesthetic features as well as the relationships between them.¹⁶ We chiefly include elements common to both architecture and furniture since our initial development focuses primarily on *Raumkunst*. In the context of this paper, the ontology provides for the annotation of a training set for the data-driven approach.

The bottom-up approach

The *Neoclassica* bottom-up approach consists of the application of Artificial Intelligence techniques specifically Deep Learning for knowledge discovery. This technique allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction.¹⁷ A *Deep Neural Network* is a specific architecture implementing *Deep Learning*. In this paper, the terms *Deep Learning*, *Deep Neural Network*, and *Neural Network* are used as synonyms. Deep Learning like all machine learning techniques requires an annotated training set used by the algorithm to learn particular features and a disjoint test set (i.e. not sharing any common items) used for the evaluation of the results. In our case, we have compiled the *Neoclassica Open Corpus* annotated according to the *Neoclassica* ontology.

At the time of writing the corpus comprises both modern photographs and digitised period imagery. Both depict either single objects in one image or multiple objects. The latter can be divided in modern photographs from exhibition spaces and period rooms, historic photographs of interiors, period images taken either from pattern-books (both printed and distributed or unique workshop designs such as the Bellangé album),¹⁸ or solitaire images (e.g. drawings, paintings) as well as thematically cohesive albums (e.g. the Wittelsbacher Album),¹⁹ digitised period books, all depicting interiors.²⁰

Information and Communication Technology 482), pp. 41–53, DOI: 10.1007/978-3-319-46224-0_5, accessed 2019-09-23.

16 The ontology builds on CIDOC-CRM for its core concepts, since CIDOC-CRM is the basis of the ISO standard 21127:2014 for the interchange of cultural heritage information (CIDOC Documentation Standards Working Group, *What Is the CIDOC CRM?*, URL: <http://www.cidoc-crm.org/>, accessed 2019-10-23; Martin Doerr, “The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata”, in: *AI Magazine* 24/3 (2003), pp. 75–92).

17 Yann LeCun et al., “Deep Learning”, in: *Nature* 521 (2015), pp. 436–444.

18 Sylvain Cordier, “The Bellangé Album and New Discoveries in French Nineteenth-Century Decorative Arts”, in: *Metropolitan Museum Journal* 47/1 (2012), pp. 119–147, DOI: 10.1086/670144, accessed 2018-12-14.

19 *Das Wittelsbacher Album*, ed. Hans Ottomeyer, München 1979; Thomas Langenholt, *Das Wittelsbacher Album: das Interieur als kunsthistorisches Dokument am Beispiel der Münchner Residenz im ersten Drittel des 19. Jahrhunderts*, Norderstedt 2002.

20 The imagery was sourced from institutions such as the Metropolitan Museum of Art, New

A particular challenge of any classification experiment in the humanities lies in the fact that the bodies of images available for training are rather small. Most Deep Learning approaches, however, are trained on millions of images. We thus applied pre-training, a common method for improving the overall performance of a neural network. The assumption is that the classifier learns specific features of a basic nature (like round shapes and specific edges) that are exploited in the unknown dataset. Most available pre-training sets such as *ImageNet*,²¹ the one we used, consist, however, of modern-day photographs. Thus, we decided to train the algorithm on a specific *ImageNet* subset containing mainly photos of modern furniture objects like tables, chairs, and cabinets.

Before each experiment, we randomly split the corpus: 80% were used to train the algorithm and 20% to evaluate the results. Additionally, we also used control corpora with newly sourced independent images for some experiments.

Single Object Detection

In order to reach our goal of identifying and classifying specific furniture forms in complex interior scenes, we applied a stepwise approach. Our first experiment layouts followed a resource efficient approach in image classification through deep learning, by concentrating on single-class image classification, focussing on photographs.

For this kind of experiments, we used *Convolutional Neural Networks* (CNNs).²² In our first experiments, we employed a corpus that was collated from open data of the *Metropolitan Museum of Art* (MET), New York. The first experiment²³ was conducted using 2167 images belonging to 42 classes and when evaluated, gave

York, the Victoria & Albert Museum, London, the Wallace Collection, London, the Cooper-Hewitt Smithsonian, New York, as well as the Hermitage, St. Petersburg. All sources are available under permissive licenses.

- 21 Jia Deng et al., "Imagenet: A Large-Scale Hierarchical Image Database", in: *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, URL: <http://ieeexplore.ieee.org/abstract/document/5206848>, accessed 2018-12-14.
- 22 Convolutional networks (alternatively, convolutional neural networks, or CNNs), are a specialized kind of neural network employing the namesake mathematical operation (Ian Goodfellow et al., *Deep Learning*, Boston, MA 2016). We used two layouts, VGG-16, (Karen Simonyan and Andrew Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, Ithaca, NY 2015, URL: <https://arxiv.org/abs/1409.1556>, accessed 2018-12-14), for the first experiment and VGG-19 (Olga Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge", in: *International Journal of Computer Vision* 115/3 (2015), pp. 211–252, DOI: <https://doi.org/10.1007/s11263-015-0816-y>, accessed 2018-12-14) for the second. See Alex Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks", in: *Proceedings of the 25th International Conference on Neural Information Processing Systems* 2012, Red Hook, NY 2012, pp. 1097–1105, URL: <http://dl.acm.org/citation.cfm?id=2999134.2999257>, accessed 2018-12-14, for a different implementation of VGG-19, also employed by us.
- 23 Bernhard Bermeitinger et al.: "Object Classification in Images of Neoclassical Furniture Using Deep Learning", in: *Computational History and Data-Driven Humanities. CHDDH 2016*, ed. by Bojan Bozic et al., Cham 2016 (IFIP Advances in Information and Communication Technology 482), pp. 109–112.

an average F1 score of 0.442 and 0.43 accuracy. The metrics for the experiment suggested that the composition of the corpus affected negatively the learning process. The images depicted not only whole single objects but also parts, pairs, or sets of objects. Each of these was assigned the label for the artefact as a whole, tainting the training of the classifier. We thus concluded that the common approach in machine learning – assuming that the quantity of data would level out variances in the source data – did not work out, since our data was not large enough.

We thus decided to clean the corpus to improve data quality by excluding all images that showed only part of an artefact, additionally by splitting images that contained multiple artefacts, and by covering neighbouring objects in images. Furthermore, not all classes possessed a sufficient number of images to create meaningful results thus we removed all classes with less than five instances. This led to a further reduction of the corpus size to 1246 images belonging to 30 classes. We ran a second experiment²⁴ which yielded an accuracy of 0.77 and an F1 score of 0.76. We later repeated the same experiment with a different layout resulting in an even better mean accuracy of 0.80 and a mean F1 score of 0.72.

In order to test whether period depictions can be classified correctly when applying a classifier trained chiefly with modern day photographs of artefacts from a museum context,²⁵ we collated a control corpus of period graphics depicting furniture from a compilation of Thomas Sheraton's furniture designs created in 1910.²⁶ We concentrated on four prime classes that were particular populous in the source material (chair, armchair, bed, and settee) and achieved a mean accuracy of 0.63. This led us to conclude that apparently, factors like mediality, technique, or the degree of abstraction of the depiction affect the classification process. As an illustration of the diversity of the corpus (see Fig. 9.1, Farbteil)²⁷ displays a sample of settees in various modalities and techniques. In the first row a correspondence of forms can be observed in a settee attributed to Samuel McIntire and its archetype in plate 35 of Thomas Sheraton's *The Cabinet-Maker and Upholsterer's Drawing-Book* (1793) (the second image). The same holds for the third row where a Greek Revival settee by William Hancock is depicted in direct correspondence to plate two of Thomas Sheraton's *Cabinet Encyclopedia* (London, 1805).²⁸ In order to further explore this hypothesis, we laid out another set of

24 Bernhard Bermeitinger et al., "Object Classification in Images of Neoclassical Artifacts Using Deep Learning", in: *Digital Humanities 2017: Conference Abstracts*, ed. by Rhian Lewis et al., Montréal 2017, URL: <https://dh2017.adho.org/abstracts/590/590.pdf>, accessed 2018-12-14.

25 The classifier was trained with photographs because there were not enough period depictions per class available to conduct a meaningful experiment.

26 The original designs for the furniture date from between 1770 and 1790. Thomas Sheraton, *The Furniture Designs*, ed. by J. Munro Bell, London 1910.

27 All images are in the public domain, taken from the Metropolitan Museum of Art, New York, ascension numbers from left to right, top to down: 26.207; NK2229 .S54 1793; 2007.368; 51.624.2; 48.164.1; NK2229 .S54 1793; 60.4.1.

28 Cf. Marshall B. Davidson and Elizabeth. Stillinger, *The American Wing at the Metropolitan Museum of Art*, New York, NY 1985, pp. 143, 159.

control experiments. We decided to retrain the classifier adding distinct subsets of images that reflected the different properties. We expanded the corpus to include 682 images from the Victoria and Albert Museum, 121 images from the *Wallace Collection*, 196 from a re-edition of Hepplewhite's *The cabinet maker and upholsterer's guide*.²⁹ We also included drawings corresponding to 24 images from the so-called *Bellangé Album*, a calf-leather bound folio attributed to the workshop of Pierre-Antoine Bellangé (1757–1827) and his son, Louis-Alexandre (1797–1861) as well as several other artists. We annotated the total of collected 2020 images by using the domain ontology, resulting in the use of 54 classes. We found that we needed at least eight annotated images per class, thus reducing the workable image set to 1674 images. We also employed a newer implementation framework of our CNN layout, which resulted in a median accuracy of 0.78 and F1 score of 0.77. (For an extended discussion of the analysis of the factors affecting the classification, see below in the discussion).

Transition to Interiors

Seeing only a marginal potential left for optimisation of the single artefact classification, we moved on to start experimenting with complex interior images. For this, we hence decided to use an RCNN³⁰ being the application of CNNs not to the image as a whole but to specific image regions. This required re-annotating all our images, defining regions by drawing polygons and associating them with labels corresponding to the *Neoclassica* ontology classes. We annotated all the images of the previously described corpus.

In order to test this approach, we chose to classify single objects and compare the results with our previous experiments.³¹ For this purpose, we used a subcorpus of 618 images representing 29 classes entirely from the MET. Of these images, 90% are photographs. This yielded a mean accuracy of 0.94. To better assess the impact of mediality, technique, the degree of abstraction of the depiction we used again the independent control corpora from the single image classification experiments, resulting in average mean precision of 0.60 (V&A), 0.69 (Wallace), 0.48 (Sheraton), 0.50 (Hepplewhite) and 0.64 (Bellangé). An overview of the experiments is provided by Table 1: Experiments conducted by the *Neoclassica* project with regard to single artefacts. There is a visible discrepancy between the results among the different corpora which can be distinguished between a) better results

29 George Hepplewhite, *The Cabinet Maker and Upholsterer's Guide; or, Repository of Designs for Every Article of Household Furniture*, London 1794, URL: <https://archive.org/details/cabinetmakerupho00ahepuoft>, accessed 2018-12-14.

30 Ross Girshick et al., "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation", in: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587. DOI: <https://doi.org/10.1109/CVPR.2014.81>, accessed 2018-12-14.

31 Simon Donig et al., "Bildanalyse durch Distant Viewing – Zur Identifizierung von klassizistischem Mobiliar in Interieurdarstellungen", in: *DHd 2018: Kritik der digitalen Vernunft*, ed. by Georg Vogeler, Köln 2018, pp. 130–138, URL: <http://dhd2018.uni-koeln.de/wp-content/uploads/boa-DHd2018-web-ISBN.pdf>, accessed 2018-12-14.

achieved among the MET and all the other corpora and b) the better results for the corpora that contain mainly digitised photographs (MET, V&A, Wallace) as well as those containing coloured drawings (Bellangé – cf. Fig. 9.1, Farbteil, second row on the right) as opposed to monochromatic period prints and graphics (Sheraton, Hepplewhite – see Fig. 9.1, Farbteil, 1st and 2nd row on the right). This might be attributed to the stylization of the prints.

Object Detection in Interiors

As at the time of the experiment, we did not yet have fully annotated interiors we compiled a test set including period interiors in multiple modalities and techniques from the *Cooper-Hewitt Smithsonian Design Museum*, the *Wittelsbacher Album*, the *Metropolitan Museum of Art*, New York, as well as photographs of period rooms³² and exhibition spaces, again from the MET. This lack of an annotated interior scene corpus allowed us only a qualitative evaluation of the experiment. This evaluation essentially showed that the identification and classification in period images worked mainly with the most populous classes (i.e. chairs).

We then proceeded to annotate interiors so that we could train with these annotations and concentrated on kinds of objects that were more likely to be found in interiors. We included classes that already formed part in the previous training processes as well as adding new classes not well represented in the previous training sets since the number of such artefacts is usually limited in a typical organisational context. For instance, a museum may have only a given number of chandeliers in its exhibition spaces and storage facilities. We thus augmented the set to include classes like chandeliers, mantelpieces, architectural glass (a term we coined to include over-mantle mirrors and pier glass for instance). These were sourced both from freely available retro-digitised print material published by museums and freshly annotated interior images cropped to the bounding box of particular artefacts as well as images from hitherto incomputable classes that we had so far retained. This led to overall 24 classes, including various kinds of seating furniture, tables, case furniture, lighting devices and architectural features. We trained a classifier³³ with both the annotated interiors and the anno-

32 There is a minor epistemological problem in using period rooms to train a classifier which we are very well aware of. As the period room essentially reimagine past in a particularly set framing they are chiefly used for training and testing a classifier in a noisier environment. Assuming they would provide for a considerable part of a corpus and would provide a basis for further inferencing (e.g. the distribution of artefact-classes in a particular type of room) this would pose an issue graver. (For a general characterisation of the subsequent modes of this kind of representation see Luc Noppen, “L’habitat mis en scène”, in: *Continuité* 51 (1991), pp. 19–25).

33 MaskRCNN structure, backbone feature extraction done with Inception V2 (MaskRCNN: Kaiming He et al., *Mask R-CNN*, Ithaca, NY 2018, URL: <http://arxiv.org/abs/1703.06870>, accessed 2018-12-14; Inception V2: Christian Szegedy et al., *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*, Ithaca, NY 2016, URL: <http://arxiv.org/abs/1602.07261>, accessed 2018-12-14).

Classifier	VGG-16 (CNN)	VGG-19 (CNN)	VGG-19 (CNN)	VGG-19 (CNN)	VGG-19 (CNN)	VGG-19 (CNN)	Faster-RCNN with ResNet101
Framework	Lasagne/Theano	Lasagne/Theano	Keras/Tensorflow	Keras/Tensorflow	Keras/Tensorflow	Keras/Tensorflow	Tensorflow
Corpus Characteristics							
Number of Images	2167 (overall)	1246 (overall)	1246 (overall)	1246 (MET), 64 (Sheraton)	1674 (overall)	618 (overall)	618 (MET), 371 (V&A), 61 (Wallace), 89 (Sheraton), 75 (Hepplewhite), 24 (Bellangé)
Number of Classes	42	30	30	4	54	29	29
Training	MET	MET	MET	MET	MET, Sheraton, V&A, Wallace C., Hepple-white, Bellangé (fused)	MET	MET
Test Corpus	MET	MET	MET	Sheraton	MET, Sheraton, V&A, Wallace C., Hepplewhite, Bellangé (fused)	MET	MET, V&A, Wallace, Sheraton, Hepplewhite, Bellangé
Evaluation Measures	Mean accuracy: 0.43 Mean F1 score: 0.44	Mean accuracy: 0.77 Mean F1 score: 0.76	Mean accuracy: 0.80 Mean F1 score: 0.72	0.63 Mean Top-1 accuracy: 0.63 Mean Top-2 accuracy: 0.78 Mean Top-3 accuracy: 0.84	Median accuracy: 0.78, Median F1 score: 0.77	Mean accuracy: 0.94	Mean accuracy of 0.94 0.60 mAP, 0.69 mAP, 0.48 mAP, 0.50 mAP, 0.64 mAP

Table 1: Experiments conducted by the Neoclassica project with regard to single artefacts.

tated single objects. We obtained precision of 0.53, recall of 0.51 and F1 score of 0.52. Although the results were not directly comparable there is a noticeable drop both in precision and F1 score as compared to the single image classification experiment. As was to be expected the most populous classes resulted in better classification scores as did some of the more prominently featured artefacts such as chandeliers. Since there is no established ground truth regarding this kind of data in order to explain the variance of the metrics between the two types of experiments, we provide a couple of considerations in the following paragraph.

Discussion

There are multiple observations to be made regarding factors of potential impact on the classification process.

Factors affecting the training of a classifier

As the most important factor for image classification experiments in the Humanities and Visual Cultural Studies we regard the size of corpora. Typical data sets in machine learning, comprise ten thousand to some million labelled images. Our data, in comparison, is extremely low scale (approaching the size of a single *ImageNet* class for the whole data set) and scattered through many classes.

While the limited size of corpora is induced by the domain or research project – there is for instance only a finite number of period interiors in existence after all – other factors are having an equal impact on the composition of the corpus. Since we don't digitise material ourselves we rely on digitisation programs of cultural institutions. This makes us highly dependent on their licencing policies. Such licensing may vary from permissive public domain licensing as is the case with the *Metropolitan Museum of Art* to relatively restrictive licensing that does not permit a redistribution like the CC-BY-NC-ND applied by some institutions or even the requirement to erase data downloaded for scholarly purpose after a certain period of time. Last but not least, a considerable number of institutions especially in Europe do not offer any kind of data for public or scholarly use at all. It is not by accident that no French artefact in our data set stems from a French institution. The same holds for the German institutional landscape.³⁴

Additionally, the corpus is affected by the documentation practices of the holding institutions. First the size of the corpus is limited by the available quality of the images resulting from retro digitisation, often favouring mass over quality (e.g. low resolution, low dynamic range scanning of 35mm film material in museums, digitisation of printed material instead of the source images). Second, the creation of images used for documentation of artefacts in an institution's collections management system involves decisions regarding the range and variety

³⁴ This makes initiatives like the recent Berlin call-to-action Cultural Heritage for the Future of Europe (URL: <http://www.europanostra.org/our-work/campaigns/berlin-call-action>, accessed 2019-05-20) all the more important.

of the visuals created: for the same object there may exist images taken from different angles (see Fig. 9.1, Farbteil, 4th row from above where a settee attributed to New York based *ébéniste* Duncan Phyfe is depicted from slightly different viewpoints, with and without cushioning). Furthermore, the same photographs may be attributed to different but visually similar objects (e.g. chairs belonging to a set). This might contribute either to overfitting³⁵ or, on the positive side, enable the classifier to learn a richer set of features for each class. The general availability of image data, however, is only one aspect. There are more subtle processes that equally affect the classification procedure and that relate to the composition of the data set regarding the kinds of objects it contains.

Foremost, this composition is affected by the collecting practices of the cultural institutions and their donors. Both have a tendency to concentrate on high-end, well-documented pieces. For instance, neoclassical chests of drawers are almost entirely absent from the material made accessible by the MET while they represent one of the most common household pieces in period time. The overall space of material culture that could be covered for training a classifier, is furthermore limited by two factors: systematic processes of loss (such as destruction or modification) especially for low end artefacts and effects of period production practices such the expansion of the production of chairs for middle-class households during the last quarter of the 18th century.³⁶

In addition, there is an observable impact on the data collection grounded in the geographic distribution of artefacts, as for instance our classifier was trained almost without artefacts of Russian, Baltic, Scandinavian or colonial origin as no such material has been available to us. These factors lead to an inherent imbalance of any corpus built, further enforced by the collecting practices of organisations and private collectors alike.

Classification process impact factors

The classification process is furthermore affected by the mediality, technique, degree of abstraction of the depiction as noted above. For instance, the photographs of the exhibition spaces of the MET give much better results than many period depictions of similar interiors, as can be seen in the juxtaposition of Figure 9.2, Farbteil and Figure 9.3, Farbteil. Figure 9.2 displays a period room, the so-called Richmond room, from the MET,³⁷ while Figure 9.3 shows a local artist's inter-

35 Overfitting may be defined as a “machine learning model becoming so specially tuned to its exact input data that it fails to generalize to other, similar data.” Field Cady, *The Data Science Handbook*, Hoboken, NJ 2017, p. 9.

36 Percy Macquoid, *A History of English Furniture, Vol. 4: The Age of Satinwood*, London 1908, p. 100.

37 Richmond Room, 1810-1811 (Richmond, Virginia); View towards southeast corner of room (68.137) as installed in galleries, including: chandelier (34.75.1), part of a suite of seating furniture (60.4.1-15) including four sidechairs, one armchair and a sofa, and carpet (52.59), <https://www.metmuseum.org/art/collection/search/3411> (accessed 2018-12-14), for more information see Amelia Peck, *Period Rooms in the Metropolitan Museum of Art*, New York, NY 1996.

pretation of a room in the Governor's residence in Hermannstadt (rom. Sibiu) in 1841.³⁸ In the case of the Richmond room the classifier proved very capable of identifying most artefacts correctly, such as a settee, chairs, armchairs, a gueridon table, a mantelpiece and a chandelier, in Figure 9.3, however, only some of the artefacts have been identified at all (correctly: gueridon, chandelier, incorrectly: chair as armchair, chair as stool, chair and trestle-table as a gueridon). This particular drawing displays a bad command of perspective, folding multiple planes into one image, thereby probably affecting the capability of the classifier. Perspective distortion was likely also the reason for the classifier being incapable to identify the armchair in the foreground of Figure 9.2.

The differences in the results between photographs and period depictions may be attributed to following factors: The pre-training process with photographs could affect the performance of the classifier when it comes to period drawings. The bulk of training data for interior scenes consists from single object images, mainly modern day photographs. This reinforces the classification bias towards this kind of visuals. Furthermore, deep learning frameworks are developed to be trained with photographs which might have an effect on the results when applied to period data. So far none of these effects has been independently verified, mostly because the period depictions are few. What can be said, however, is that images that display a higher degree of visual similarity to photographs – such as naturalistic oil paintings usually work better with the classifier.

The confidence rate for the classification of a fauteuil, for instance, dropped by more than thirty per cent between an oil-painting (see Fig. 9.4, right hand side, Farbteil) and its reproduction in a period print (see Fig. 9.4, left hand side, Farbteil); also, a false positive of a sofa-table has additionally been identified.³⁹

The relative consistency across photographs of different creators, processes and media in turn, may well be attributed to the comparative similarity of visual representations created by the technical process of documentary photography.

In a similar way, pattern books represent a group of images in their own right, following conventions of reproducibility introduced by technical ways of drawing. While the degree of misclassification with this group over all experiments was larger than for instance for photographs, we assume that this is due to the underrepresentation of this medium as compared to the other kinds of visuals in

38 For more details on the image see Charlotte Gere, *Nineteenth Century Interiors: an Album of Watercolors*, London 1992, p. 60, plate 16. The image was scraped from the Cooper Hewitt Smithsonian Design Museum, URL: <https://collection.cooperhewitt.org/objects/18708151>, accessed 2018-12-14.

39 The two images depict Charles Maurice de Talleyrand-Perigord, among other things foreign minister under Napoléon seated in a fauteuil next to a ceremonial bureau plat with a settee partially visible in the background. Image sources are <https://www.metmuseum.org/art/collection/search/395807> for the engraving by Auguste Boucher Desnoyers and <https://www.metmuseum.org/art/collection/search/441969> for the oil-painting by François Gérard respectively (both accessed 2018-12-14).

the training set. Further experiments are required to derive safe conclusions in that regard.

Additionally, the ‘shallow’ nature of the classification process is not capable of taking into consideration context knowledge. This leads to a number of misclassifications that could be regarded as false negatives if context knowledge was taken into consideration. For instance, in Figure 9.5, Farbteil, left hand side, an armchair was classified as a chair. While for the human observer an armchair is a special sub-category of chairs and this result would be regarded as being not incorrect, for the algorithm this result is considered as a misclassification. In a similar way, a *bergère* (an armchair with closed armrests) classified as a *fauteuil* (an armchair with open armrests) would be incorrect for the classifier (see Fig. 9.5, Farbteil, right hand side) while for a human it may be considered as sufficient.⁴⁰

One way of overcoming this limitation is to actually take into account the knowledge represented in the ontology for the classification process, where these kinds of relationships are properly represented. Another approach could be using textual corpora in conjunction.⁴¹

Furthermore, it has to be taken into consideration that all classifiers that we use, apply some kind of scaling and transformation on the image to be classified. For instance, the frameworks when we started our experiments require a fixed resolution between 224 x 224 and 33 x 331 pixels while more recent frameworks can work with arbitrary resolutions up to 1024 x 1024, 1365 x 1365 and even 1987 x 1987. This however still results in smaller details getting lost in complex interior scenes, even if the quality and resolution of the image per se would be sufficient for a classification. This may account for a lot of misses of annotated features in an image. In particular with regard to research related to aesthetic features and stylistic correspondence and transformation, an extension of the methodology to account both for the image as a whole and for individual features will be required.

Computational Analysis of Interiors

As a research instrument the identification and classification of both, neoclassical artefacts and aesthetic features in an interior provides researchers with two fundamental capabilities: first a (partial) semantization of the depiction (semantic segmentation) and second with a quantification of these segmentations, either by frequency or by correlation. This enables the researcher to identify typicalities or statistical anomalies occurring in the image corpus. For the researchers, this means that they gain an additional level of abstraction, while at the same time

40 Figure 9.5, Farbteil: An armchair by Seddon & Sons, ca. 1790, classified as a chair. A *bergère* by the Jacob-Desmalter workshop, ca. 1805, classified as a *fauteuil*. Both Victoria & Albert Museum, London. V&A Museum Numbers W.2:1-1968 & W.2C-1987.

41 Donig, Simon et al., “Vom Bild zum Text und wieder zurück”, in: *DHd 2019. Digital Humanities: multimedial & multimodal. Konferenzabstracts*, ed. by Patrick Sahle, Frankfurt 2019, pp. 227–232.

still being capable to transparently assess the observations in a qualitative way in the images themselves.

For instance, for a given larger corpus of visuals the method improves retrieval operations by being based on the actual semantic content of an image.⁴² Usually visuals are represented with relative sparse metadata probably supplemented by a textual description. The search can be improved by relying on additional extracted semantic information, e.g. in the case of Figure 9.4, *Farbteil*, this instrument could not only return a result upon a query for *Talleyrand*, but also when searching for *fauteuil* or *bureau plat*.

Additionally, images of interior spaces may be clustered by the presence and distribution of visual features belonging to specific classes. To provide but an example, images depicting particular types of rooms such as reception areas or bedrooms are likely to be clustered together. This process of clustering may correspond to different research interests.

For instance, a distribution of types of chair backs might inform about the 'typicality' of this particular construction feature, leading potentially to a long-tail-distribution indicating for more or less common types of chair backs in the set and the time span represented by this set for the given corpus. (This could be particularly useful when taken into account that for some regions such as the Baltics, Poland or the territories of the former German East, surviving specimen of furnishing *in situ* are relatively few and provenance information is often scarce, while the number of visual depictions is much broader and often comparably context-rich).

In the same vein, the choice of one or more aesthetic features, might inform researchers about trends in the desirability or fashionableness of this feature at a given time. The interpretation of these findings in turn obviously lies with the researcher.

In the artefact level the observation of changes in the presence and distribution of pieces of furniture over time may inform about the transformation of socio-economic conditions, social practices or aesthetic expectations towards interiors. For example, Percy Macquoid's observation of the increase of the number of chairs in the late quarter of the eighteenth-century England⁴³ pointing to the expansion of a consumer culture by the provision of more and inexpensively produced artefacts, could be visually confirmed by assessing the number of occurrences of this kind of artefacts in respective period images of middle class households. Totally unrelated, Charlotte Gere noted a discrepancy in the density of the furnishing in a series of depictions of the chamber of the Duchesse de Berry in the Tuileries from the end 1820s and later compared with the previous Napoleonic era. She attributed the presence of a large amount of seating furniture to a

⁴² Lang and Ommer, "Attesting Similarity", p. 846.

⁴³ Cf. above, footnote 36.

change in habits, with hosts expecting guests to interact more familiar and thus be seated instead of standing when attending a private or semi-private space.⁴⁴

The clustering of interior spaces by the presence and distribution of furniture may also lead to the establishment of prototypical rooms and study of their transformation over time. Obviously, the implication of what typicality or comparability actually means in a given research situation is entirely dependent on the choice of the corpus the classifier is confronted with and thus an integral part of the research question to be addressed.

Conclusion

The introduction of digital instruments in primarily qualitative disciplines is about to change how we conduct research. We presented our path towards a visual analysis of interiors using artificial intelligence and machine learning techniques specifically deep learning. We used convolutional neural networks to classify single object images and complex interior scenes. Our corpus stemmed from museums and collections as well period sources such as pattern and drawing books.

While the single object image classification could be described as extremely successful (leading up to a mean accuracy of 0.94), the complex interiors proved far more challenging. We attributed this to two factors affecting the classification process: The composition of the corpus and the mediality, technique, degree of abstraction of the source material, the pretraining process as well as external factors such as licensing or collection practices.

The computational analysis of interiors holds great promise for the research process since regions in visuals – corresponding e.g. to artefacts – can be semantically identified and retrieved. This introduces a method for both exploratory as well as hypothesis driven research.

Such semantical segmentations can be quantified, either by frequency or by correlation. This enables the researcher to identify typicalities or statistical anomalies occurring in the image corpus. For the researchers, this means that they gain an additional level of abstraction, while at the same time still being capable to transparently assess the observations in a qualitative way in the images themselves.

The transformation of aesthetic features, for instance, might inform researchers about trends in the desirability or fashionableness of this feature at a given time, or the distribution of specific artefacts in particular settings may point to the change of socio-economic conditions, social practices or aesthetic expectations towards interiors.

As we have also shown this methodology can be further improved. An important challenge remains the inclusion of contextual knowledge to image ana-

⁴⁴ Gere, *Nineteenth Century Interiors*, p. 24. A reproduction of the image may be found in: Mario Praz, *An Illustrated History of Interior Decoration: From Pompeii to Art Nouveau*, London, 2008, p. 200, plate 167, a reprint of the original (Milan 1964).

lysis. In particular, closing the semantic gap could help to cope with competing hypotheses about the nature of an artefact. For example, even for a human observer it is almost impossible to make a proper assessment if a closed *commode à vantaux* is indeed a *commode* or rather a *buffet bas*. Taking into account textual information would effectively resolve this situation. In a similar way the issue of misclassifications that relate to conceptual similarity will require combining more advanced algorithms with refined classification procedures, probably involving multiple classifiers as well as the *Neoclassica* ontology. A multimodal approach joining textual and visual data could open a new road towards a more effective way of solving these problems.

Bibliography

Sources

Cruikshank, George and Charles Mackay, *The Bottle. In Eight Plates [...]*, London 1847, URL: <https://www.bl.uk/collection-items/the-bottle-a-series-of-temperance-themed-illustrations-by-george-cruikshank-with-poetry-by-charles-mackay>, accessed 2018-12-14.

Das Wittelsbacher Album, ed. by Hans Ottomeyer, München 1979.

Hepplewhite, George, *The Cabinet Maker and Upholsterer's Guide; or, Repository of Designs for Every Article of Household Furniture*, London 1794, URL: <https://archive.org/details/cabinetmakerupho00ahepuoft>, accessed 2018-12-14. Sheraton, Thomas, *The Furniture Designs*, ed. by J. Munro Bell, London 1910.

Secondary Publications

Aynsley, Jeremy and Charlotte Grand, ed., *Imagined Interiors: Representing the Domestic Interior since the Renaissance*, London 2006.

Bell, Peter and Björn Ommer, "Training Argus, Ansätze zum automatischen Sehen in der Kunstgeschichte", in: *Kunstchronik* 68/8 (2015), pp. 414–420.

Bermeitinger, Bernhard, André Freitas, Simon Doing and Siegfried Handschuh: "Object Classification in Images of Neoclassical Furniture Using Deep Learning", in: *Computational History and Data-Driven Humanities. CHDDH 2016*, ed. by Bojan Bozic, Gavin Mende l-Gleason, Christophe Debruyne and Declan O'Sullivan, Cham 2016 (IFIP Advances in Information and Communication Technology 482), pp. 109–112.

—, Simon Donig, Maria Christoforaki, André Freitas and Siegfried Handschuh, "Object Classification in Images of Neoclassical Artifacts Using Deep Learning", in: *Digital Humanities 2017: Conference Abstracts*, ed. by Rhian Lewis, Cecily Raynor, Dominic Forest, Michael Sinatra and Stéfan Sinclair, Montréal 2017, URL: <https://dh2017.adho.org/abstracts/590/590.pdf>, accessed 2018-12-14. Berry, David, "The Computational Turn: Thinking about the Digital Humanities", in: *Culture Machine* 12 (2011), pp. 1–22. URL: <http://sro.sussex.ac.uk/49813/>, accessed 2018-12-14.

Burke, Peter, *Eyewitnessing: The Uses of Images as Historical Evidence. Picturing History*, London 2001.

Cady, Field, *The Data Science Handbook*, Hoboken, NJ 2017. CIDOC Documentation Standards Working Group, *What Is the CIDOC CRM?*, URL: <http://www.cidoc-crm.org>, accessed 2019-10-23.

- Cordier, Sylvain, "The Bellangé Album and New Discoveries in French Nineteenth-Century Decorative Arts", in: *Metropolitan Museum Journal* 47/1 (2012), pp. 119–147. DOI: 10.1086/670144, accessed 2018-12-14.
- Davidson, Marshall B. and Elizabeth Stillinger, *The American Wing at the Metropolitan Museum of Art*, New York, NY 1985.
- Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Li Fei-Fei, "Imagenet: A Large-Scale Hierarchical Image Database", in: *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, URL: <http://ieeexplore.ieee.org/abstract/document/5206848/>, accessed 2018-12-14.
- Doerr, Martin, "The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata", in: *AI Magazine* 24/3 (2003), pp. 75–92.
- Donig, Simon, Maria Christoforaki and Siegfried Handschuh, "Neoclassica – A Multilingual Domain Ontology. Representing Material Culture from the Era of Classicism in the Semantic Web", in: *Computational History and Data-Driven Humanities. CHDDH 2016*, ed. by Bojan Bozic, Gavin Mendel-Gleason, Christophe Debruyne and Declan O'Sullivan, Cham 2016 (IFIP Advances in Information and Communication Technology 482), pp. 41–53, DOI: 10.1007/978-3-319-46224-0_5, accessed 2019-09-23.
- , Bernhard Bermeitinger, Maria Christoforaki and Siegfried Handschuh, "Bildanalyse durch Distant Viewing – Zur Identifizierung von klassizistischem Mobiliar in Interieurdarstellungen", in: *DHd 2018: Kritik der digitalen Vernunft*, ed. by Georg Vogeler, Köln 2018, pp. 130–138, URL: <http://dhd2018.uni-koeln.de/wp-content/uploads/boa-DHd2018-web-ISBN.pdf>, accessed 2018-12-14.
- , Siegfried Handschuh, Canan Hastik, Hubertus Kohle, Björn Ommer and Malte Rehbein: "Der Ferne Blick. Bildkorpora und Computer Vision in den Geistes- und Kulturwissenschaften: Stand – Visionen – Implikationen", in: *DHd 2018: Kritik der digitalen Vernunft*, ed. by Georg Vogeler, Köln 2018, pp. 86–89, URL: <http://dhd2018.uni-koeln.de/wp-content/uploads/boa-DHd2018-web-ISBN.pdf>, accessed 2018-12-14.
- , Maria Christoforaki, Bernhard Bermeitinger and Siegfried Handschuh, "Vom Bild zum Text und wieder zurück", in: *DHd 2019. Digital Humanities: multimedial & multimodal. Konferenzabstracts*, ed. by Patrick Sahle, Frankfurt 2019, pp. 227–232.
- Fulda, Daniel: "Sattelzeit. Karriere und Problematik eines kulturwissenschaftlichen Zentralbegriffs", in: *Sattelzeit: Historiographiegeschichtliche Revisionen*, ed. by Elisabeth Décultot and id., Berlin 2016, pp. 1–16.
- Gere, Charlotte, *Nineteenth Century Interiors: an Album of Watercolors*, London 1992.
- Girshick, Ross, Jeff Donahue, Trevor Darrell and Jitendra Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation", in: *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587. DOI: <https://doi.org/10.1109/CVPR.2014.81>, accessed 2018-12-14.
- Goodfellow, Ian, Yoshua Bengio and Aaron Courville, *Deep Learning*, Boston, MA 2016.
- He, Kaiming, Georgia Gkioxari, Piotr Dollár and Ross Girshick, *Mask R-CNN*, Ithaca, NY 2018, URL: <http://arxiv.org/abs/1703.06870>, accessed 2018-12-14.
- Hirayama, Hina, "With Éclat" – *the Boston Athenæum and the Origin of the Museum of Fine Arts, Boston*, Boston, MA 2013.
- Kelly, Jason M., *The Society of Dilettanti*, New Haven, CT 2009.
- Koepe, Wolfram, *Extravagant Inventions: The Princely Furniture of the Roentgens*, New York, NY 2012.

- Kohle, Hubertus, "Ordnung und Umbruch in der Wissenschaft der Kunstgeschichte: Der Einfluss des Digitalen", in: *Reibungspunkte: Ordnung und Umbruch in Architektur und Kunst*, ed. by Hanns Hubach, Petersberg 2008, pp. 141–144.
- Krizhevsky, Alex, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", in: *Proceedings of the 25th International Conference on Neural Information Processing Systems 2012*, Red Hook, NY 2012, pp. 1097–1105, URL: <http://dl.acm.org/citation.cfm?id=2999134.2999257>, accessed 2018-12-14.
- Lang, Sabine and Björn Ommer: "Attesting Similarity: Supporting the Organization and Study of Art Image Collections with Computer Vision", in: *Digital Scholarship in the Humanities* 33/4 (2018), pp. 845–56. DOI: <https://doi.org/10.1093/lilc/fqy006>, accessed 2018-12-14.
- Langenholt, Thomas, *Das Wittelsbacher Album: das Interieur als kunsthistorisches Dokument am Beispiel der Münchner Residenz im ersten Drittel des 19. Jahrhunderts*, Norderstedt 2002.
- LeCun, Yann, Yoshua Bengio and Geoffrey Hinton, "Deep Learning", in: *Nature* 521 (2015), pp. 436–444.
- Lefuel, Hector, *Georges Jacob, ébéniste du XVIII^e siècle*, Paris, 1923.
- , *François-Honoré-Georges Jacob-Desmaltre Ébéniste de Napoléon I^{er} et de Louis XVIII.*, Paris 1926.
- Lukatis, Christiane, "Zimmerbilder – Entwicklung und Charakter des Genres", in: *Mein blauer Salon: Zimmerbilder der Biedermeierzeit*, ed. by ead., Nürnberg 1995, pp. 17–26.
- Macquoid, Percy, *A History of English Furniture, Vol. 4: The Age of Satinwood*. London 1908.
- Noppen, Luc, "L'habitat mis en scène", in: *Continuité* 51 (1991), pp. 19–25.
- Ommer, Björn and Peter Bell, "Digital Connoisseur? How Computer Vision Supports Art History", in: *Il Metodo Del Conoscitore – Approcci, Limiti, Prospettive Connoisseurship Nel XXI Secolo*, ed. by Stefan Albl and Alina Aggujaro, Roma 2016, pp. 187–200.
- Peck, Amelia, *Period Rooms in the Metropolitan Museum of Art*, New York, NY 1996.
- Praz, Mario, *An Illustrated History of Interior Decoration: From Pompeii to Art Nouveau*, London, 2008.
- Redford, Bruce, *Dilettanti: The Antic and the Antique in Eighteenth-Century England*, Los Angeles, CA 2008.
- Rice, Charles: *The Emergence of the Interior – Architecture, Modernity, Domesticity*, London 2007.
- Russakovsky, Olga, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge", in: *International Journal of Computer Vision* 115/3 (2015), pp. 211–252, DOI: <https://doi.org/10.1007/s11263-015-0816-y>, accessed 2018-12-14.
- Schoch, Rainer, "Repräsentation und Innerlichkeit. Zur Bedeutung des Interieurs im 19. Jahrhundert", in: *Mein blauer Salon: Zimmerbilder der Biedermeierzeit*, ed. by Christiane Lukatis, Nürnberg 1995, pp. 11–16.
- Simonyan, Karen and Andrew Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition*, Ithaca, NY 2014, URL: <https://arxiv.org/abs/1409.1556>, accessed 2018-12-14.
- Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke and Alex Alemi, *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning*, Ithaca, NY 2016, URL: <http://arxiv.org/abs/1602.07261>, accessed 2018-12-14.