

Fachartikel aus:

Hartmut Beyer / Thomas Mandl (Hg.): Bildähnlichkeit und Bildsuche: Geistes- und informationswissenschaftliche Zugänge zu historischem Material (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 8). 2026. DOI: [10.17175/sb008](https://doi.org/10.17175/sb008)

Titel:

Aiding Provenance Research. A Computer-Assisted Image Retrieval in Auction Catalogs

Autor*in:

Mathias Zinnen

Kontakt: mathias.zinnen@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg

GND: [1393911307](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-393911307) ORCID: [0000-0003-4366-5216](https://orcid.org/0000-0003-4366-5216)

Contribution (CRediT): [Data curation](#) | [Formal analysis](#) | [Investigation](#) | [Methodology](#) | [Software](#) | [Visualization](#) | [Writing – original draft](#) | [Writing – review & editing](#)

Autor*in:

Sabine Lang

Kontakt: sab.lang@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg

GND: [1100083812](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1100083812) ORCID: [0000-0003-2543-0085](https://orcid.org/0000-0003-2543-0085)

Contribution (CRediT): [Conceptualization](#) | [Formal analysis](#) | [Investigation](#) | [Visualization](#) | [Writing – original draft](#) | [Writing – review & editing](#)

Autor*in:

Andreas Maier

Kontakt: andreas.maier@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg

GND: [138422893](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-138422893) ORCID: [0000-0002-9550-5284](https://orcid.org/0000-0002-9550-5284)

Contribution (CRediT): [Supervision](#) | [Funding acquisition](#) | [Resources](#)

Autor*in:

Vincent Christlein

Kontakt: vincent.christlein@fau.de

Institution: Friedrich-Alexander-Universität Erlangen-Nürnberg

GND: [1107595517](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-1107595517) ORCID: [0000-0003-0455-3799](https://orcid.org/0000-0003-0455-3799)

Contribution (CRediT): [Supervision](#) | [funding acquisition](#) | [Resources](#) | [Writing – review & editing](#)

DOI des Beitrags:

[10.17175/sb008_005](https://doi.org/10.17175/sb008_005)

Nachweis im OPAC der Herzog August Bibliothek:

[196699317X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-196699317X)

Erstveröffentlichung:

30.06.2026

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

30.03.2026

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Auktion](#) | [Bildverarbeitung](#) | [Kunstgeschichte](#) | [Maschinelles Lernen](#) | [Provenienzforschung](#)

Empfohlene Zitierweise:

Mathias Zinnen / Sabine Lang / Andreas Maier / Vincent Christlein: Aiding Provenance Research. A Computer-Assisted Image Retrieval in Auction Catalogs. In: Hartmut Beyer / Thomas Mandl (Hg.): Bildähnlichkeit und Bildsuche: Geistes- und informationswissenschaftliche Zugänge zu historischem Material (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 8). Wolfenbüttel 2026. 30.06.2026. HTML / XML / PDF. DOI: [10.17175/sb008_005](https://doi.org/10.17175/sb008_005)

Mathias Zinnen / Sabine Lang / Andreas Maier / Vincent Christlein

Aiding Provenance Research. A Computer-Assisted Image Retrieval in Auction Catalogs

Abstract

Provenance research examines the origin of objects and aims to reconstruct their ownership history. For this purpose, researchers increasingly use online resources such as the database *German Sales*, which contains auction and sales catalogs, through a full text search. While this is very valuable, a text-based search has limitations due to missing or varying information. This motivated us to search for a different, image-based method which enables searches for identical and similar images in large data collections using neural networks. We provide qualitative and quantitative results showing the effectiveness of our method and discuss the technical conceptualization of image similarity underlying the presented algorithm.

Provenienzforschung untersucht die Herkunft von Objekten und deren Besitzgeschichte. Dafür verwenden Forscher*innen Online-Datenbanken wie *German Sales*, die Auktions- und Verkaufskataloge enthält, die im Volltext durchsuchbar sind. Obwohl dieser Zugang wertvoll ist, limitieren fehlende oder abweichende Informationen den Erfolg einer textbasierten Suche. Dies hat uns motiviert, eine bildbasierte Methode zu entwickeln, die Suchen nach identischen und ähnlichen Bildern in Datensammlungen ermöglicht und auf neuronalen Netzen basiert. Wir präsentieren qualitative und quantitative Ergebnisse, die die Wirksamkeit der Methode zeigen, und diskutieren die technische Konzeptualisierung der Bildähnlichkeit, die dem Algorithmus zugrunde liegt.

1. Introduction

Provenance research examines the origin and history of objects, ideally tracing them from their creation to the present day. It seeks information about previous owners and the circumstances surrounding the transfer of ownership, and is a core task in museums, libraries, archives, and other cultural institutions.¹ In the German context, it encompasses objects looted by the National Socialists, property from colonial contexts, the German Democratic Republic (GDR) and Soviet Occupation Zone, or general war-related losses. To find out about an object's provenance, researchers study the (physical) object, explore historical contexts related to persons and institutions, and utilize archival documents, literature and online-resources such as [German Sales](#).²

[1]

The *German Sales*-database provides access to circa 15,000 auction and sales catalogs primarily from German-speaking countries, making it an essential source for studying the art market and collecting practices in the 20th century and for provenance research. All catalogs in *German Sales* are open access and searchable via a text entry in the full text.³ Despite the significant utility of a full-text search, it has its limitations: Missing information, varying titles, dates, or attributions for the same object might lead to unreliable or incomplete results. The following example illustrates this. A painting by the German artist Ludwig von Zumbusch (1861–1927) was titled *Römische Ideallandschaft* in an auction catalog in 1931⁴, appeared as *Landschaft mit Birken und Pappeln* in 1933⁵, and eventually offered as *Einsames Land* in 2019⁶ (see Fig. 1).

[2]

¹ Cf. Zuschlag 2022, pp.11–12.

² Cf. German Lost Art Foundation et al. 2020, p. 6 and table of contents chapter 3.

³ Cf. German Sales 2025.

⁴ Cf. Helbing 1931.

⁵ Cf. Internationales Kunst- und Auktionshaus 1933.

⁶ Cf. Ketterer Kunst 2019.



Ludwig von Zumbusch, 1896

Munich 1931: *Römische Ideallandschaft*
(Helbing 1931)

Berlin 1933: *Landschaft mit Birken und Pappeln*
(Internat. Kunst- und Auktionshaus 1933)

Munich 2019: *Einsames Land* (Ketterer Kunst 2019)

Figure 1: A painting by Ludwig von Zumbusch was offered at three auctions, each time under a different title. [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: Ketterer Kunst 2019]

A search with the current title would have yielded no direct results. This motivated us to look for a different method. What if, instead of text, we use computational methods to perform an image-based search? To this end, we developed a process which automatically detects and crops images from auction catalogs, extracts quantifiable feature vectors, and stores them in a feature database. Additionally, we created an application to search for images in the processed auction catalogs. Users can upload arbitrary images to the application and are provided with the most similar images in the database according to the extracted feature vectors. This system can be used to detect different depictions of identical objects⁷ or to explore similar images. The capability to explore similar images positions our work within a tradition of using *computer vision* for the analysis of historical art works to complement traditional humanities methodologies with new data-driven approaches.⁸

[3]

In the following sections, we first introduce related works and subsequently the dataset at hand. We will then present our methodology for image search in auction catalogs, followed by experiments and case studies. As the central topic of this special issue is image similarity, we will discuss the technical conceptualization of image similarity underlying the presented algorithm and contrast it with a human perspective. We conclude this article by summarizing the main findings and point to future work.

[4]

2. Related Work

This work depends on two established computer vision techniques: object detection, which is required for the initial cropping of the images from the auction catalogs, and image retrieval, essential for computing image similarities based on extracted feature vectors. Conceptually, our method is anchored in recent and innovative approaches which apply digital methods to study the art market and support provenance research. We will first explore the technical underpinnings of our methodology before contextualizing our approach within the broader fields of digital art market analysis and provenance studies.

[5]

⁷ When we speak of objects in the following, we primarily refer to paintings, graphics or sculptures. However, since everyday objects etc. are also depicted in the auction catalogs, we use the more inclusive term ›object‹.

⁸ Cf. Crowley / Zisserman 2016; Ufer et al. 2021.

2.1 Object Detection

Object Detection is one of the fundamental tasks in computer vision and maintains its prominence as one of the most active research areas.⁹ It goes beyond image-level classification by requiring not only the classification but also the localization of objects within an image. Image-level classification, in contrast, involves assigning a label to an entire image, which generally poses less complexity. Classification benchmarks get their complexity from the amount of categories that need to be distinguished, as exemplified by the classic **ImageNet**¹⁰ benchmark, which required differentiation among 1,000, or in some versions, even 21,000 classes.¹¹ [6]

In detection, two-stage methods like the *Region-based Convolutional Neural Network (R-CNN)* family,¹² have dominated the field, offering strong accuracy at the cost of increased runtime. In contrast, one-stage methods like *You only look once (YOLO)*¹³ or *Single Shot MultiBox Detector (SSD)*¹⁴ have traditionally prioritized speed over precision, although the performance gap has narrowed with more recent architectures. Recently, transformer-based detection architectures, particularly those based on the **DEtection TRansformer (DETR)**,¹⁵ have become the dominant paradigm. However, one-stage and two-stage methods maintain their relevance. Notably, recent versions of YOLO continue to show competitive performances on many detection benchmarks and are very popular due to their easy usage, efficient training and inference times. For our purposes, the complexity and performance of modern detection benchmarks far exceed our demands. The task of detecting images in printed layouts is straightforward and can nearly be considered as a resolved problem within the field of computer vision. Consequently, we favor the **YOLO-v8**¹⁶ architecture over more powerful but slower modern two-stage or DETR-based architectures. [7]

While applying object detection to artworks is less common than in natural image analysis, it is recognized as a well-established practice within digital humanities. Pioneering work in object detection for artworks was conducted by Elliot J. Crowley and Andrew Zisserman. They trained detection algorithms on the photographic dataset **Pascal VOC**¹⁷ and employed transfer learning to apply the trained algorithms for identifying objects such as vehicles or animals in artworks.¹⁸ Another line of research focused on the detection of smell-related objects in artworks,¹⁹ a topic addressed with the introduction of the **ODOR** challenge and dataset. Special cases of object detection, such as the detection of depicted faces²⁰ and persons²¹ have a great relevance for art history, e.g. for studying portraiture, figure painting, or genre painting. Westlake et al. introduced a dataset of annotated persons depicted across a wide range of artistic styles.²² Closely related to that is the application of *pose estimation*, which typically requires an initial person detection stage. Pose estimation was used to analyze and cluster artworks based on body postures,²³ hand gestures,²⁴ and overall image composition,²⁵ or to recognize smell-related²⁶ and sensory²⁷ gestures. [8]

⁹ Cf. Zou et al. 2023.

¹⁰ Cf. Russakovsky 2015.

¹¹ Cf. Ridnik et al. 2021.

¹² Cf. Girshick et al. 2014; Girshick 2015; Ren et al. 2016.

¹³ Cf. Redmon et al. 2016.

¹⁴ Cf. Liu et al. 2016.

¹⁵ Cf. Carion 2020.

¹⁶ Cf. Jocher et al. 2023.

¹⁷ Cf. Everingham 2010.

¹⁸ Cf. Crowley / Zisserman 2014; Crowley / Zisserman 2015; Crowley / Zisserman 2016.

¹⁹ Cf. Zinnen et al. 2022a.

²⁰ Cf. Bengamra et al. 2021; Mermet et al. 2020.

²¹ Cf. Westlake et al. 2016.

²² Cf. Westlake et al. 2016.

²³ Cf. Impett / Moretti 2017; Impett / Süssstrunk 2016; Bell / Impett 2019; Impett 2020.

²⁴ Cf. Bernasconi et al. 2023.

²⁵ Cf. Madhu et al. 2020; Madhu et al. 2023.

²⁶ Cf. Zinnen et al. 2023.

²⁷ Cf. Zinnen et al. 2025.

A common challenge for object detection in artworks is the sparsity of annotated datasets. Gonthier et al. addressed this by employing a weakly supervised training approach, using image-level labels before assessing the method on instance-level labels.²⁸ Additionally, they shifted their focus from detecting modern categories to identifying objects with art historical relevance. This thread was picked up by Marinescu et al., who adapted COCO categories to be consistent with historical contexts,²⁹ and by Reshetnikov et al., who compiled a large dataset categorized by art historical themes.³⁰ [9]

Another approach to address the challenge of data sparsity is the use of style transfer to mimic artistic object representations.³¹ This technique leverages existing annotations from photographic data to obtain an adaptation to the artistic target domain using transfer learning. This strategy is commonly employed for various tasks, including pose estimation,³² emotion recognition,³³ or painting captioning,³⁴ and has also been successfully applied to object detection.³⁵ [10]

The issue of data sparsity, at least for person detection and pose estimation, has been significantly mitigated with the introduction of the Human-Art dataset.³⁶ This dataset features 50,000 images of artistic creations like sculptures, paintings, or cartoons annotated with the position of persons and pose estimation keypoints. [11]

2.2 Image Retrieval

The retrieval of similar depictions of identical objects from auction catalogs can be contextualized within several closely related research areas, with *Content-Based Image Retrieval (CBIR)* being the most prominent. CBIR, often used synonymously with ›Image Retrieval‹, can be defined as the problem of searching for relevant images in a database given a query image based on visual features.³⁷ [12]

The retrieval process can also be framed as a problem of *Near Duplicate Detection (NDD)* when considering different depictions of the same object as near duplicates. Primarily aimed at identifying digital copies or manipulated images, NDD differs from CBIR in its application focus. While CBIR is a relatively open task that accommodates a wide range of queries and image similarities, NDD specifically targets the duplicate nature of the query and target images.³⁸ Irrespective of whether it is labeled as NDD or CBIR, a retrieval system typically operates in two phases: A preparatory phase, where a database of quantifiable image representations is created, and an online phase, where this database is queried.³⁹ These queries are not limited to image inputs alone; for example, a shared image-language embedding space, as provided by *Contrastive Language-Image Pre-Training (CLIP)*,⁴⁰ enables cross-modal querying.⁴¹ However, in the context of our application, we are focusing on image queries. In this scenario, the online phase entails a feature extraction similar to that during the database creation. This process maps the query input into the same embedding space as the already processed image corpus, enabling the computation of distances between the query features and the previously extracted features to assess their similarity. As computing the feature distance across all stored feature vectors can become intractable for large image corpora, the embedding [13]

²⁸ Cf. Gonthier et al. 2018; Gonthier et al. 2022.

²⁹ Cf. Marinescu et al. 2020.

³⁰ Cf. Reshetnikov et al. 2023.

³¹ Cf. Sanakoyeu et al. 2018; Gatys et al. 2015; Gatys et al. 2016.

³² Cf. Madhu et al. 2023; Springstein et al. 2022.

³³ Cf. Patoliya et al. 2024.

³⁴ Cf. Lu et al. 2021.

³⁵ Cf. Jeon et al. 2020; Kadish et al. 2021; Smirnov / Eguizabal 2018.

³⁶ Cf. Ju et al. 2023.

³⁷ Cf. Chen et al. 2022.

³⁸ Cf. Thyagarajan / Kalaiarasi 2021.

³⁹ Cf. Zhou et al. 2017.

⁴⁰ Cf. Radford et al. 2021.

⁴¹ Cf. Garcia / Vogiatzis 2018.

space is often clustered or indexed to enable efficient querying. Chen et al. categorize methods based on their reliance on off-the-shelf models, originally trained for different tasks than retrieval, versus methods that incorporate an additional training stage dedicated to fine-tuning the models specifically for image retrieval.⁴² In our application, we employ the simpler approach by using pre-trained models to extract features. Moving forward, we plan to improve our system by integrating more elaborate training schemes specifically designed for retrieval tasks.

The process of fine-tuning models for retrieval is closely related to the field of metric learning. However, the term *metric learning* is often discussed more from a technical standpoint than from its practical application. Metric learning involves mapping data into an embedding space where similar data points are closer together and dissimilar data points are further apart, guided by a distance metric such as the *Euclidean norm*.⁴³ Metric learning can be used for many purposes such as classifying samples according to their nearest neighbors in the embedding space or in self-supervised pre-training. However, retrieval can be considered its most natural application as it typically entails the search for the closest query result in the feature space. This way, retrieval can be framed as an instance-level open-set classification, where the query has to be matched to its closest neighbors in the feature space. [14]

Independent of whether models are specifically trained for retrieval is the consideration of which features are used for the distance computation. Before *deep learning*, approaches typically relied on local features obtained with methods such as *Scale-Invariant Feature Transform (SIFT)*.⁴⁴ These local features can be aggregated into a global descriptor using methods like *Bag-of-visual-Words (BoW)*.⁴⁵ Such combinations of classical local features and their global aggregation have been applied in various contexts, including the recognition of CD-covers,⁴⁶ object retrieval⁴⁷ and product search.⁴⁸ [15]

The question of feature aggregation remains relevant in modern, deep-learning-based approaches: Even when employing neural networks, various feature representations can be considered, involving different scales⁴⁹ (intra-model) or different models⁵⁰ (inter-model). Effective fusion of different feature levels can lead to a more meaningful feature representation.⁵¹ For our application, a comparatively simple strategy is sufficient. We simply use the flattened output of the last layer of a pre-trained *Convolutional Neural Network (CNN)* to represent the image contents and compute similarities. Future research could explore different scales and more elaborate feature fusion, but also revisit classical feature extraction methods such as SIFT, and analyze their impact on the type of similarity encoded in the query results. [16]

In the fields of digital humanities and computational cultural heritage, research formerly centered around the visual retrieval of similar images or image parts to identify patterns across numerous artworks and uncover relations between them.⁵² Eventually, projects and works also focused on the development of interfaces to search for images or image parts.⁵³ Besides feature vector distances, works also explored the application of different metrics to capture diverse aspects of image similarity, for example color concepts,⁵⁴ image composition,⁵⁵ body posture,⁵⁶ or even symbolic meaning.⁵⁷ [17]

⁴² Cf. Chen et al. 2022.

⁴³ Cf. Musgrave et al. 2020.

⁴⁴ Cf. Lowe 2004.

⁴⁵ Cf. Csurka et al. 2004.

⁴⁶ Cf. Nistér / Stewénius 2006.

⁴⁷ Cf. Philbin et al. 2008.

⁴⁸ Cf. He 2012.

⁴⁹ For example Sun et al. 2015 or Tolias et al. 2015.

⁵⁰ Cf. Yokoo et al. 2020.

⁵¹ Cf. Chen et al. 2022.

⁵² Cf. Seguin et al. 2016; Ufer et al. 2021; Shen et al. 2019; Castellano et al. 2021.

⁵³ Cf. Ufer et al. 2021; Springstein et al. 2021; Offert / Bell 2024; PortApp 2025.

⁵⁴ Cf. Yelizaveta et al. 2005.

The wealth of works which apply computational methods for art analysis and understanding is reflected in several review papers. While Bengamra et al.⁵⁸ provide a summary of significant computer vision applications for art, they specifically focus on object detection. An overview of datasets and works on the task of recognizing and extracting patterns in visual arts using deep learning is given in a paper by Giovanna Castellano and Gennaro Vessio.⁵⁹ Another review provides an overview of how computational methods are used for classification, object detection, similarity retrieval or multimodal representations, among others.⁶⁰ Amalia Foka then presented past computer vision applications for art historical research and future possibilities.⁶¹ [18]

2.3 Digital Art Market & Provenance Studies

Computational methods have also been applied to art market research: Utilizing a subset of over 267,000 sale transactions from the [Getty Provenance Index](#) and complex network science, Schich et al. studied the history of the art market and collection dynamics to reveal social, temporal, and spatial networks.⁶² Fletcher et al. studied the art market in London between 1850 and 1914 on the basis of complementary datasets and visualizations.⁶³ Scheithauer et al. suggested a two-step pipeline to analyze the layout and content of auction sales catalogs utilizing object detection and text sequence labeling models.⁶⁴ A similar approach and goal to the one presented in this paper was announced in 2021. Then, the Fraunhofer Institute published a report which informed about a feasibility study. The study developed AI-methods for image search in auction catalogs, enabling a successful comparison between current and historical images.⁶⁵ A second feasibility study, which again highlighted the success of the methods, was published in the magazine *Museumskunde* in 2024.⁶⁶ In the humanities, the art market has been a research topic for many years. Numerous publications on the art market between 1901 and 1945 in Germany and other German-speaking countries focus on individual art dealers⁶⁷, auction houses⁶⁸, or the valuation and price development.⁶⁹ An overview of relevant literature is given in the bibliographies *German Sales 1901–1929*⁷⁰ and *German Sales 1930–1945*.⁷¹ [19]

The relatively new field of digital provenance research studies the impact of digitality on provenance research, focusing on chances and challenges, and uses digital methods for analyzing provenance data.⁷² Works report on the development of research databases⁷³, the presentation and communication of provenance information and research results online⁷⁴, the future of provenance research and digital infrastructures in Germany, especially focusing on tendencies and desiderata.⁷⁵ Special attention has [20]

⁵⁵ Cf. Madhu et al. 2023.

⁵⁶ Cf. Impett / Moretti 2017; Bell / Impett 2019.

⁵⁷ Cf. Sartini et al. 2023; Sartini / Gangemi 2021.

⁵⁸ Cf. Bengamra et al. 2024.

⁵⁹ Cf. Castellano / Vessio 2021.

⁶⁰ Cf. Cetinic / She 2022.

⁶¹ Cf. Foka 2021.

⁶² Cf. Schich et al. 2017.

⁶³ Cf. Fletcher et al. 2012.

⁶⁴ Cf. Scheithauer et al. 2024.

⁶⁵ Cf. Vicente-Garcia 2021.

⁶⁶ Cf. Vicente-Garcia 2024.

⁶⁷ Cf. Hoffmann / Kuhn 2016.

⁶⁸ Cf. Hopp 2012.

⁶⁹ Cf. Jeuthe 2014.

⁷⁰ Cf. Bommert 2019.

⁷¹ Cf. Bähr 2013.

⁷² Cf. Lang 2023a.

⁷³ Cf. Werner 2020.

⁷⁴ Cf. Haffner 2020; Haffner 2019.

⁷⁵ Cf. Hopp 2018.

been paid to aspects of incompleteness and vagueness in provenance research.⁷⁶ Rother et al. study the transformation of unstructured provenance records into *Linked Open Data* and how computer-based methods can be utilized for a comprehensive analysis of provenance records.⁷⁷

3. German Sales

The database *German Sales* was launched in 2013 and currently holds circa 15,000 digitized sales and auction catalogs mostly from German-speaking countries and bibliographic metadata. Various projects contributed to the development of *German Sales*. The initial project started in 2010 with a collaboration between the Getty Research Institute, Heidelberg University Library, and the Art Library in Berlin. Back then, the aim was to provide (online) access to auction catalogs held and preserved throughout Germany, Austria, and Switzerland for the period from 1930 to 1945.⁷⁸ Subsequent sub-projects focused on different time periods or locations and not only included sources relevant to the secondary but also primary market such as gallery catalogs or stock books.⁷⁹ All catalogs are published in open access and accessible through a full text search. The database itself offers different views on the data (see Fig. 2): Users can activate and deactivate the metadata block, overview, facsimile, or OCR full text depending on their preferences and interests. The catalogs are very heterogeneous in content and form: They vary in length, layout, and with regard to the information they contain. Most catalogs include a cover page, an introduction, and illustrations of the included lots. These illustrations vary in size, rotation and quality and are embedded within the text or printed on separate pages at the end of the catalogs. All catalogs contain, to varying extents, a list of the lots with information about the title, artist, date, measurements, technique, and (sometimes) a description (see Fig. 3).

[21]

The screenshot displays the user interface of the German Sales database. At the top, there are navigation tabs: 'Metadaten', 'Überblick', 'Faksimile', 'OCR-Volltext', and 'Vollansicht'. Below these, a search bar contains the text 'z.B. IV, 145, Seite' and a button labeled 'Seite: Tafel I'. The main content area is split into two columns. The left column, titled 'METADATEN', shows the following information:

- A. Müller-Fontana <Zürich> [Hrsg.]; A. Vogel <Genève> [Hrsg.]**
- Auktion in Zürich: hervorragende französische Möbel des XVIII. Jahrh. und der Empirezeit, Gemälde und Handzeichnungen, Schweizer Ansichten und Trachten, französische und englische Stiche ... ; Versteigerung in Zürich am 28. und 29. November 1934 im Zunfthaus Zur Meise – Zürich, 1934
- DOI Seite / Zitierlink: <https://doi.org/10.11588/digit.12619#0011>
- Startseite des Bandes
- Bibliographische Information
- arthistoricum.net: German Sales
- Rechte vorbehalten / freier Zugang
- Nutzung/Bestellung

 The right column, titled 'Überblick', shows a facsimile view of the auction catalog page, displaying a grid of numbered pages (8-26) and a 'Tafel II' section. The pages contain text and illustrations of auction lots.

Figure 2: *German Sales* offers different views of the data. Users can choose between the metadata, overview, facsimile, or OCR full text. [Screenshot: Mathias Zinnen / Sabine Lang 2024]

⁷⁶ Cf. Lang 2023b; Mariani 2022.

⁷⁷ Cf. Rother et al. 2023; Rother et al. 2022.

⁷⁸ Cf. Huemer 2014, pp. 273–278.

⁷⁹ Cf. German Sales: Project description 2004.



Figure 3: Exemplary pages from different catalogs showing their heterogeneity regarding, for example, the number of images per page, the variety of offered lots, color of the pages and placement within the catalogs. [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]

The lots themselves are also very heterogeneous and include paintings, drawings, prints, sculptures, books, ceramics, furniture, and other objects. Fig. 4 visualizes the types of offered objects and their frequency as a word cloud, created using the keywords provided for each catalog in the bibliographies of Astrid Bähr and Britta Bommert.⁸⁰ While we focus on paintings in this work, the heterogeneity of object types provides an excellent opportunity to extend our approach to different kinds of auction objects in the future. [22]



Figure 4: Word cloud visualizing the frequency of sales / object types found in the German Sales bibliographies. [Visualization: Mathias Zinnen / Sabine Lang 2024]

4. Methodology

In the following sections we present our methodology which enables searches for identical and similar images in a data set containing auction catalogs from German Sales. We first describe the necessary data preprocessing steps, then describe the feature-based retrieval and conclude with some remarks on the demo-app which allows users to test the described method. [23]

4.1 Data Preprocessing

Before we can store the feature representations of objects depicted in the auction catalogs, we must prepare the catalogs and crop the illustrations. First, we parse the PDF-files of the two bibliographies⁸¹ and convert the information into a structured, machine-readable format. Table 1 lists the extracted metadata fields describing the catalogs. [24]

⁸⁰ Cf. Bähr 2013; Bommert 2019.

⁸¹ Cf. Bähr 2013; Bommert 2019.

Field Name	Explanation
title	Title of the catalog in the bibliography.
location	Place where the auction took place.
year	Year of the auction. Exact dates are transformed to years.
types	List of object types offered in the auction as specified in the bibliography.
uri	Permanent link to the catalog entry in <i>German Sales</i> .
fn	File name of the downloaded catalog PDF used for further processing.

Table 1: List of metadata extracted from the auction catalog bibliographies (cf. Bähr 2013; Bommert 2019).

In a second step, we filter the catalogs using the metadata and process them further. Due to computational constraints, we initially limit our scope to a smaller number of catalogs, focusing on locations in Switzerland. Additionally, we narrow our selection to catalogs covering sales of ›Gemälde‹ as specified by the keywords in the bibliographies. This process results in a set of 86 catalogs, with an additional 25 catalogs which are included later for evaluation. We plan to extend the dataset in future work, eventually covering all catalogs in *German Sales*. [25]

Using the links provided in the bibliographies, we download the catalogs from the online services of the Heidelberg Library and convert each page to an image file using the `PyMuPDF`⁸² library. To minimize memory use, we store the pages as JPEG-files with 95 % export quality compression. Subsequently, we employ the `YOLO-v8x`⁸³ object detection algorithm to automatically detect images of objects on the pages, crop the detected images and save them separately. To train the detection algorithm, we manually label a set of 181 catalog pages with the position for 316 depicted objects. We split 33 for validation (containing 60 objects), and train the algorithm with the remaining 148 annotated pages (256 objects). Finally, the cropped object depictions are converted to grayscale and rotated to align with the longer edge, ensuring a consistent representation in terms of print color and orientation. Fig. 5 provides a complete overview of the preprocessing steps described above. [26]

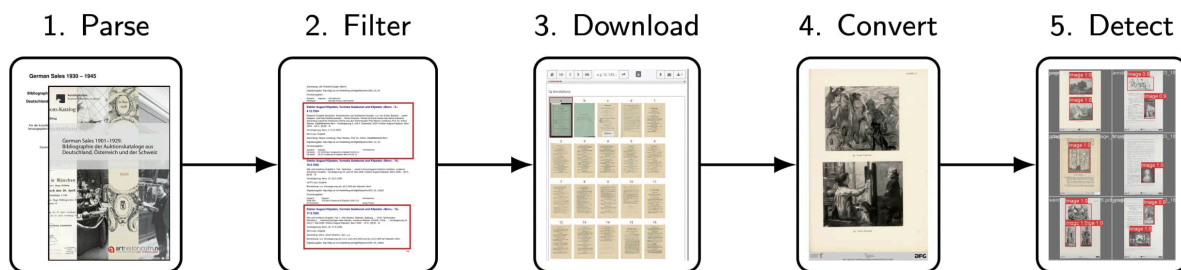
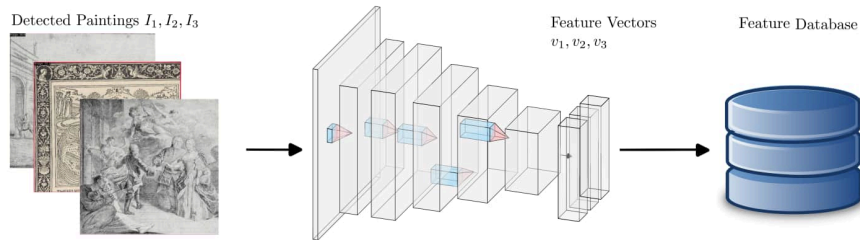


Figure 5: Complete pipeline of preprocessing steps applied to the auction catalogs prior to extracting features from the depicted objects. [Graphic: Mathias Zinnen / Sabine Lang 2024]

⁸² Cf. PyMuPDF 2024.⁸³ Cf. Jocher et al. 2023.

4.2 Feature-Based Retrieval

Preparatory Phase: Feature Database Creation



Retrieval Phase: Database Querying

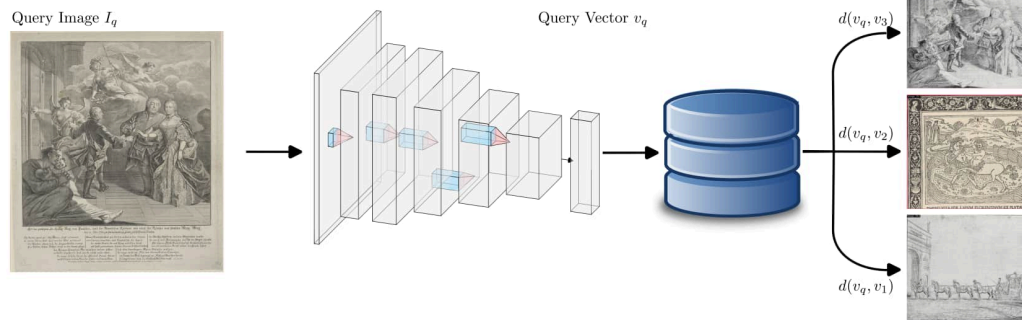


Figure 6: Process employed to perform an image-based search for identical or similar images in auction catalogs. [Graphic: Mathias Zinnen / Sabine Lang 2024]

Fig. 6 illustrates the process of feature extraction and query employed for the reverse image search in the auction catalogs. To prepare the system, the previously cropped images from the auction catalogs are used as an input to various feature extraction methods to compile a database of feature vectors as shown in the top row of Fig. 6. Specifically, we employ three different ResNet-50⁸⁴ feature extractors, each pre-trained for different tasks:

[27]

1. Classification of 1,000 classes in the ImageNet⁸⁵ dataset,
2. Detection of smell-related objects in a dataset of historical artworks (ODOR)⁸⁶,
3. Recognition of 17 pose estimation keypoints in the COCO⁸⁷ dataset.

The motivation behind selecting these three pre-training schemes is to evaluate whether the choice of extraction method can influence which images are found to be similar. The underlying hypothesis is that the embeddings extracted from the images can be related back to the task the extraction model was originally trained for. Accordingly, pose recognition embeddings would be expected to emphasize the body posture of depicted persons. ImageNet embeddings, on the other hand, have been shown to generalize towards various tasks and put a specific emphasis on texture versus shape.⁸⁸ The model trained for object detection in premodern artworks was selected to test whether the availability of artistic object representations in the training data supports the models in extracting more relevant embeddings for artwork similarity.

[28]

⁸⁴ Cf. He et al. 2016.

⁸⁵ Cf. Deng et al. 2009, model weights obtained from the mmpretrain framework (cf. OpenMMLab Pre-Training 2024).

⁸⁶ Cf. Zinnen et al. 2022b; Zinnen et al. 2025, models trained by the authors.

⁸⁷ Cf. Lin et al. 2014, models obtained from the mmpose framework (cf. OpenMMLab Pose 2024).

⁸⁸ Cf. Geirhos et al. 2018.

After deploying the application, users can upload arbitrary images to the system. These images are fed to the feature extractor, similarly to the process during the creation of the feature database. Afterwards, we compute vector distances $d(v_q, v_i)$ between the query vector v_q and the n precomputed vectors $v_i, i \in \{1, 2, \dots, n\}$ and return the k images l_1, l_2, \dots, l_k where the distance between the feature and query vectors are lowest (see bottom row of Fig. 6). [29]

Our approach simplifies the process of image retrieval by using the flattened final feature vector obtained just before the final classification layer of the ResNet-50 architecture, which typically has 2048 dimensions. Incorporating more elaborate strategies and vector representations that account for multiple scales might further enhance performance. This improvement is a potential direction for future work, although it is beyond the scope of this current study. [30]

For efficient querying, we implement **FAISS (Facebook AI Similarity Search)**⁸⁹, which performs initial clustering and hashing in the feature space to speed up the search process. The similarity between vectors is computed using the *Euclidean Distance* formula: [31]

$$d(v_q, v_r) = \sum_{i=1}^{2048} \sqrt{(q_i - r_i)^2}, \quad [32]$$

where q_i and r_i denote the i -th element of the feature vectors v_q and v_r extracted from the query images l_q and l_r , respectively. [33]

4.3 A User-Interface for Image Search in Auction Catalogs

In order to test our method and eventually enable its use by provenance researchers, we develop a user-interface with the open source package **Gradio**.⁹⁰ Fig. 7 shows a screenshot of the interface and an exemplary search. Users first select a query image either by simply uploading it or by using the drag-and-drop functionality. They can then select one of the three models discussed above to determine the feature embedding used to compute image similarities. Before users initiate the search, they can adjust the number of results on the right hand side of the interface. After the search is started, the most similar images to the query are computed using Euclidean Distance (see paragraph 31–33 for more details). Eventually, the search results are displayed underneath the query with additional metadata, including the DOIs of the respective catalogs. [34]

⁸⁹ Cf. Johnson et al. 2019.

⁹⁰ Cf. Abid et al. 2019.

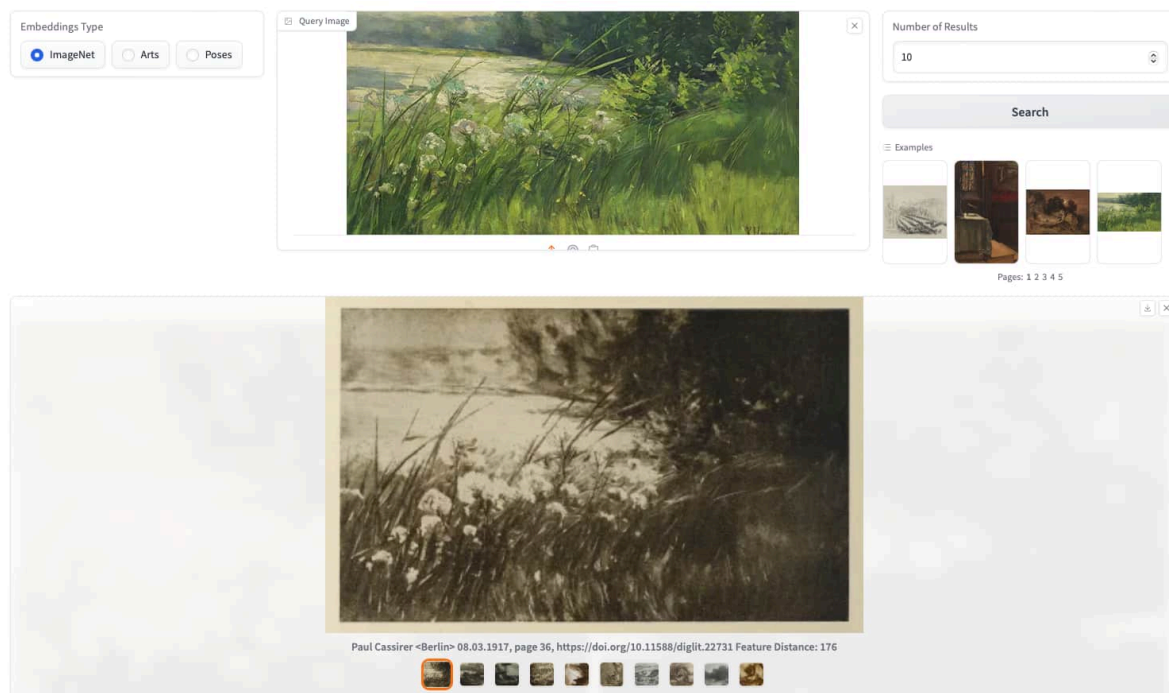


Figure 7: Easy-to-use interface developed by using the open source software Gradio. [Screenshot: Mathias Zinnen / Sabine Lang 2024]

5. Experiments

We quantitatively measure the performance of the proposed retrieval system for images of objects using common metrics defined below. Furthermore, we assess the performance of the preparatory image detection step. [35]

5.1 Detecting Images in Auction Catalogs

To measure the performance of the image detection step, we apply *mean average precision* as defined in the COCO challenge (*COCO mAP*).⁹¹ COCO mAP is the most widely used evaluation metric for object detection algorithms to date and realizes a trade-off between precision and recall of detected objects by averaging over multiple confidence and overlap thresholds.⁹² Using the training and validation splits defined in Section 4.1, we train the detection algorithm for 50 epochs. We then evaluate the model on the 33 unseen validation pages and achieve a COCO mAP of 98.6 %. Exemplary predictions from the validation split are illustrated in Fig. 8. [36]

⁹¹ Cf. Lin et al. 2014.

⁹² A detailed definition is beyond the scope of this work and can be found at Common Objects in Context 2024.

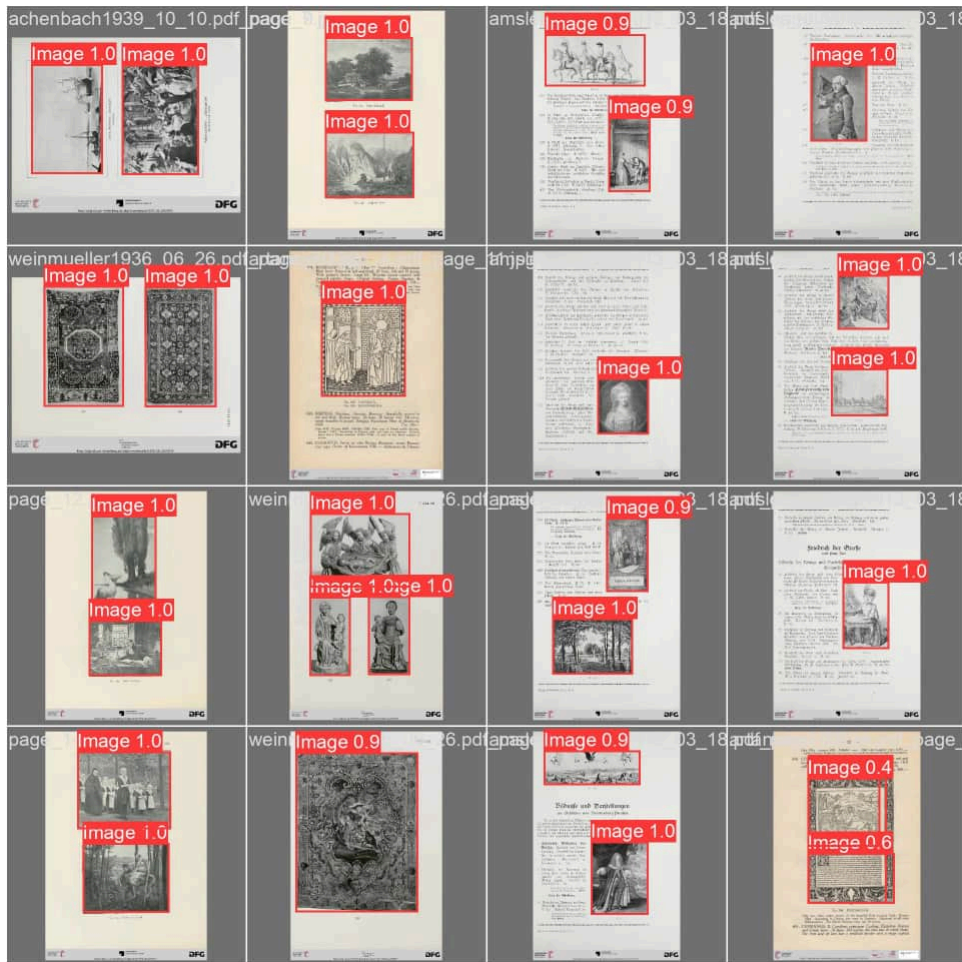


Figure 8: Exemplary predictions from the image detection stage: most images are reliably detected. However, the algorithm has problems differentiating between different object types / elements as evidenced by detected carpets, sculptures, and ornated text (bottom right). [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]

The detected depictions of objects other than paintings could likely be corrected by integrating another training stage specifically targeted at recognizing objects which are not paintings. However, in our use case we can tolerate these false predictions as they will not produce visual features similar to any query image. Thus, we do not expect any negative impact apart from a slightly increased memory usage. Generally, the examples visually confirm the strong performance of the detection system, highlighting that the detection of images in auction catalogs is an easy-to-solve problem for modern detection algorithms.

[37]

5.2 Retrieval

5.2.1 Metrics

We measure the performance of our retrieval system using *top-1 accuracy*, *top-5 accuracy*, and *mean average precision for retrieval (retrieval mAP)*. *Top-k accuracy* reflects the percentage of evaluation queries where a target image ranks among the first *k* suggestions. For this study, we particularly focus on top-1 and top-5 accuracies, which, in our case, measure how often we find a depiction of the query image in the auction

[38]

catalogs as the first result or among the first five results, respectively. Retrieval mAP is a standard metric in retrieval tasks that balances precision and recall by averaging the precision achieved at various recall levels r , similar to COCO mAP⁹³.

Precision at r is the proportion of target images correctly identified in the first r retrieval results. Conversely, recall is the ratio of all target images successfully retrieved by the algorithm. As the number of relevant predictions r increases, the precision typically decreases while the recall increases. This is because an expansion of the result set increases the probability that target images are included. Consequently, retrieval mAP is computed as the average of precisions for multiple values of considered retrieval results r . [39]

Instead of considering all possible ranks k , we only consider up to 50 retrieval results (mAP@50). Accordingly, we compute the average precision (AP) for each query artwork q as follows: [40]

$$AP_q = \frac{1}{R} \sum_{r=1}^{50} P_q(r) * rel_q(r), \quad [41]$$

where: [42]

- R is the number of relevant matches in the corpus,
- $P_q(r)$ is the precision at r (i.e. the fraction of correctly retrieved images at rank r)
- $rel_q(r)$ is an indicator function that returns 1 if the retrieved image at rank r is a depiction of the same image and 0 otherwise.⁹⁴

The mean average retrieval precision (mAP) is then computed as the mean of AP over all Q evaluation queries: [43]

$$\frac{1}{Q} \sum_{q=1}^Q AP_q \quad [44]$$

5.2.2 Results

We evaluate our method with $Q = 18$ images. As queries, we select already available digital images of artworks; these are not included in our corpus of auction catalogs. The distribution of corresponding images R in the corpus for the evaluation images is as follows: 15 images of artworks have one matching target image, two images have two matching target images, and one has three. To enable a convenient reproduction of the evaluation, we include all evaluation images as examples in the app. [45]

Table 2 presents top-1 and top-5 accuracies, along with the retrieval mAP for the three pre-training schemes detailed in section 4.1. All models share the same ResNet-50 architecture but vary in their pre-training. No specific fine-tuning was applied to the models for the retrieval task. [46]

⁹³ Cf. Lin et al. 2014.

⁹⁴ Cf. Zhou et al. 2017, p. 14.

Training Dataset	Top-1 Acc.	Top-5 Acc.	mAP
ImageNet	72.2	88.9	73.3
Arts	77.8	83.3	72.2
POSES	33.3	38.7	29.6

Table 2: The table shows the metrics for our method where the extraction network was trained on the *ImageNet*, *Arts* or *Poses* datasets. While for the *top-1 accuracy* we achieve best results with the Arts model, the ImageNet model scores the best *top-5 accuracy* and *mAP*.

Comparing the ImageNet and Arts pre-trained models, we do not see strong differences. While Arts pre-training shows slightly better performance in top-1 accuracy, ImageNet pre-training yields higher top-5 accuracy and retrieval mAP. From the two models evaluated, we cannot conclude that pre-training within the target domain (premodern paintings vs. photographs) increases the retrieval performance. Instead, these results suggest that the features learned from ImageNet classification are sufficiently generic to capture similarity in artistic representations of reality. Qualitative examples which illustrate successful queries are discussed in section 6. [47]

The decline in performance for the model pre-trained for pose estimation is striking. We hypothesize that the features learned for the body posture estimation are too specialized towards their original application to effectively capture relevant aspects of artwork similarity such as image composition. They also fail to recognize similarities between landscapes or inanimate objects, such as clouds, trees, houses, and abstract pictorial elements. To validate this hypothesis, we conducted another evaluation with only six images featuring at least one person very prominently in the image and reported the results in Table 3. This setting led to an increased performance for all models, particularly for the Arts model, which achieved perfect top-1 accuracy. The pose estimation model also displayed significant improvement, especially when compared to the ImageNet model. This indicates that models trained for pose recognition indeed capture artwork similarity better when persons are depicted. [48]

Model	Top-1 Acc.	Top-5 Acc.	Retrieval mAP
ImageNet	85.7	88.1	85.7
Arts	100	100	100
POSES	71.4	71.4	66.3

Table 3: Evaluation of the models' performances when the query images are restricted to images of artworks with at least one person depicted. We see an increased performance in all models with the Arts-pre-trained models even achieving perfect retrieval metrics. Specifically the model trained for pose estimation performs considerably better and partly closes the gap compared to the two other pre-training schemes.

Two artworks could not be retrieved by any method (see Fig. 9). This drastically lowers the evaluation metrics. To identify the cause, we experimented with different image rotations and compression applied to the query image. However, neither of these measures could resolve the issue. We also confirmed that the images were correctly extracted during the preparatory image detection step (see section 5.1) to eliminate the possibility that it might have failed. [49]



Figure 9: These two artworks drastically lower the quantitative metrics: Adrian Ludwig Richter's *Hirten am Feuer*, (c. 1861) on the left and Heinrich Bürkel's *Der Kochelsee mit den Häusern von Schlehdorf* (c. 1863 / 1867) on the right. [Graphic: Mathias Zinnen / Sabine Lang 2024, image sources: Ketterer Kunst 2018a; Ketterer Kunst 2018b]

Further investigation revealed that the images cropped from the auction catalogs (see Fig. 10) were of poor quality, characterized by blurriness and noise. However, as illustrated by the two cropped images of artworks in the second row of Fig. 10, poor image quality alone is not a determining factor for the model not being able to find the images. Although they also have a bad quality, the model is able to detect the works in the dataset. [50]

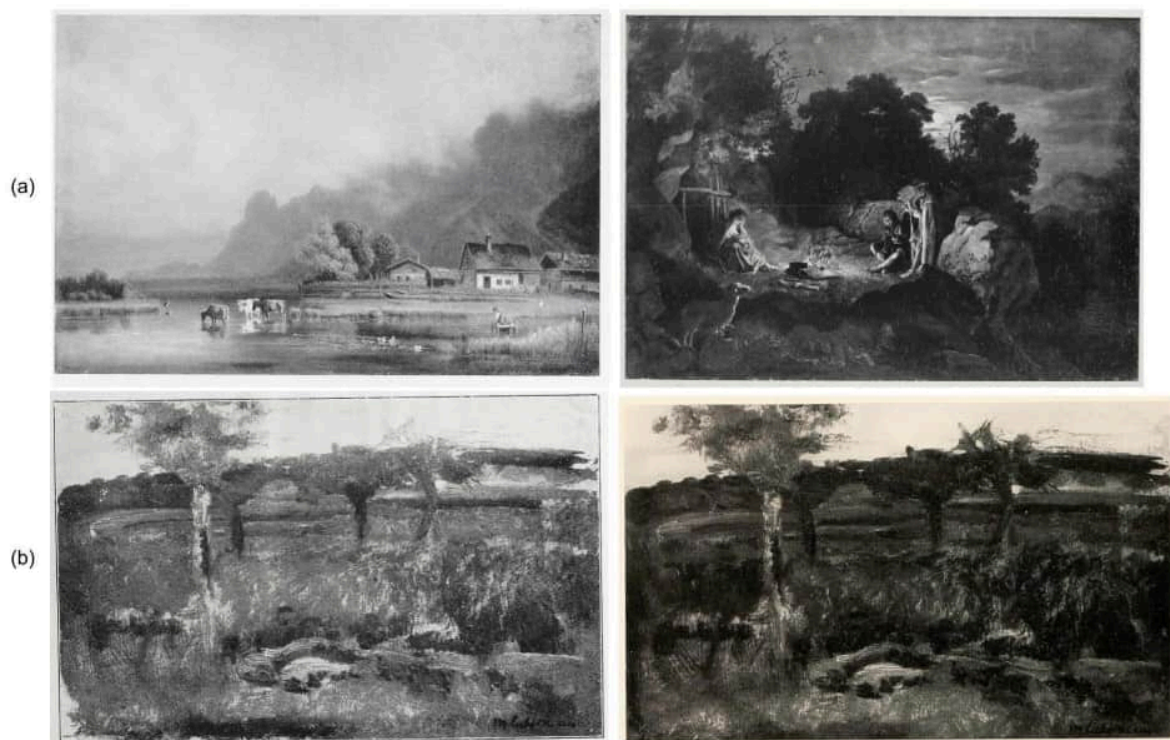


Figure 10: Crops of the target images of artworks as detected in the auction catalogs. Row (a) shows the images which could not be found by any of the extraction methods. The second row (b) shows detections of an artwork by Max Liebermann with a similarly bad image quality which were found (note also the difference in quality and color between the two identical images). [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]

In the following, we present two case studies to illustrate how the presented method can be successfully used to search for identical and similar images. This way, the method not only aids provenance researchers with reconstructing an object's origin but also facilitates the study of visual patterns. [51]

6. Case Studies

6.1 Case Study: Retrieving Identical Images

In this section we want to demonstrate the potential and performance of our method to search for identical images in auction catalogs, thus assisting with the reconstruction of an object's provenance. To perform the search, we use the demo-app which we have introduced in section 4.3. In June 2024, the Munich-based auction house Neumeister offered a work by Johann Sperl (1840–1914) entitled *Sommerlust*.⁹⁵ Can we find out more about the painting's provenance using our method? To this end, we perform a search in our collected data set using the app. We upload the query image, set the number of results to ten and initiate the search. The results appear after a few seconds underneath the query image (Fig. 11). The first three results are identical to the query image, suggesting that the painting was offered in three auctions. The first result refers to an auction which took place on April 24, 1928, in the Kunsthaus Lempertz in Cologne.⁹⁶ The painting was listed as lot 32, titled *Kinder auf der Wiese*⁹⁷ and displayed on panel 13.⁹⁸ Notably, the name differs from the current title, highlighting the issue of a text-based search. Two years later, on November 14, 1930, Sperl's painting was included in an auction at Paul Cassirer's *Kunstsalon* in Berlin.⁹⁹ The respective auction catalog shows the work on page 88. Back then, however, it was entitled *Kind auf der Wiese* – a slight variation of the 1928-title.¹⁰⁰ The last result points to an auction at Hugo Helbing's gallery on March 26, 1927 – the earliest auction date in our list of results.¹⁰¹ The corresponding catalog lists the work as *Sommerlust* (lot 106) on page 22¹⁰² and includes a reproduction of the painting on panel 16.¹⁰³ The example of Johann Sperl demonstrates the efficiency of our method and is indeed confirmed by the provenance information given on the website of Neumeister which mentions all three provenances.¹⁰⁴ [52]

If we look at the calculated feature distance, we notice that the first result (Lempertz) had a distance of 257, the second (Cassirer) a distance of 259 and the last (Helbing) a calculated distance of 448 to the query image. All images show the same content to the query; why do they have a different distance then? Here is a possible explanation: The coloration of the images vary slightly as well as the rotation (Helbing). These deviations from the query image might cause these observed distances, especially to the Helbing example.¹⁰⁵ [53]

⁹⁵ Cf. Neumeister 2024.

⁹⁶ Cf. Kunsthaus Lempertz 1928a.

⁹⁷ Cf. Kunsthaus Lempertz 1928b.

⁹⁸ Cf. Kunsthaus Lempertz 1928c.

⁹⁹ Cf. Kunstsalon Paul Cassirer 1930a.

¹⁰⁰ Cf. Kunstsalon Paul Cassirer 1930b.

¹⁰¹ Cf. Auktionshaus Hugo Helbing 1927a.

¹⁰² Cf. Auktionshaus Hugo Helbing 1927b.

¹⁰³ Cf. Auktionshaus Hugo Helbing 1927c.

¹⁰⁴ Cf. Neumeister 2024.

¹⁰⁵ It is unclear whether the coloration of the pages is present in the originals or a result of the digitization process and only visible in the digital reproductions.

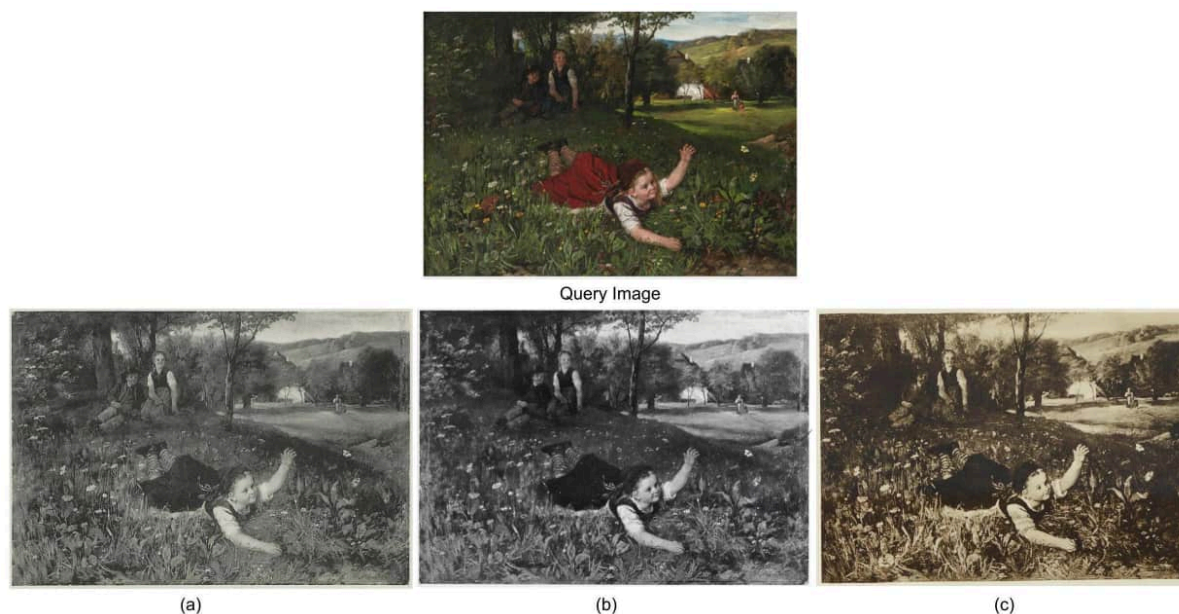


Figure 11: Search results for Johann Sperl's *Sommerlust*: The work appeared in a catalog in (a) 1928, (b) 1930 and (c) 1927. [Graphic: Mathias Zinnen / Sabine Lang 2024]

6.2 Case Study: Retrieving Similar Images

For many visual disciplines such as art history it is not only relevant to find identical but also similar images. Finding these similarities offers insights into reception processes over time and space, prevailing taste, artistic networks and focuses of auction houses and collectors. Can we use our method to address the following questions: [54]

1. In which contexts does the motif of the boat appear?
2. What can be said about the drawing style of Max Liebermann (1847–1935)?

6.2.1 The Motif of the Boat

In order to answer the first question, we choose Andreas Achenbach's (1815–1910) painting *Fischerboot auf stürmischer See* (1895) as a query image and initiate the search (Fig. 12). The first result is identical to the query, therefore we disregard this image in our analysis. [55]

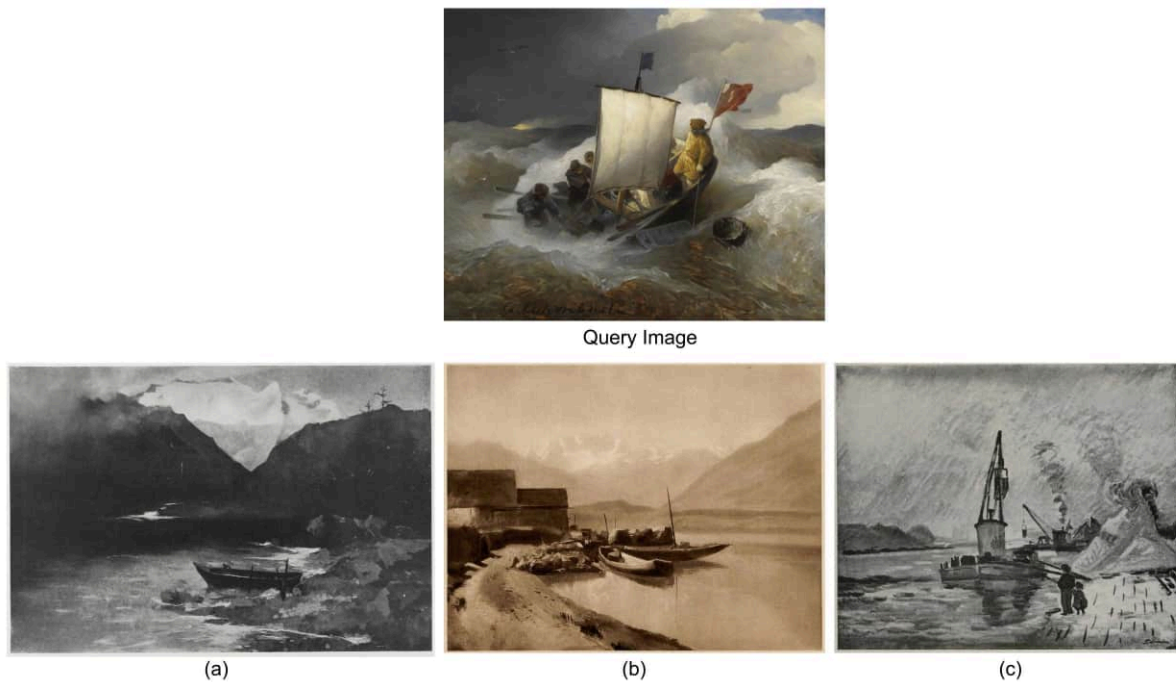


Figure 12: The first search results (excluding the first / identical) for Andreas Achenbach's *Fischerboot auf stürmischer See*: The painting was offered at auctions in (a) 1936, (b) 1931 and (c) 1928. [Graphic: Mathias Zinnen / Sabine Lang 2024]

The second image (a) shows a boat close to the shore, embedded within a mountain view. It was painted by Max Buri (1868–1915) and offered as *Brienzersee* (1894) in an auction hosted by the Gallery Neupert in Zurich on April 4, 1936.¹⁰⁶ Similar to the query, the boat is shown at a central position; however, the stormy water and atmosphere is replaced by a calmness and tranquility. In addition, Buri's painting is devoid of any human life. The third result shows a similar visual pattern: A boat floats on a calm lake at a central position, surrounded by a deserted landscape, mountain range, small houses and shore. The image was painted by Otto Frölicher (1840–1890) and offered as *Barken* in an auction catalog of G. & L. Bollag in Zurich in 1931.¹⁰⁷ These first results suggest a content similarity, possibly influenced by the motifs of the boat and / on water. The last result discussed in this case study stems from a catalog of the *Kunstsalon Dr. Störi* in Zurich; the respective auction took place in March 1928.¹⁰⁸ The image shows a painting by Guillaumin Armand (1841–1928) and is entitled *Der Kran*.¹⁰⁹ We see a crane vessel in the middle ground, another one is visible in the background. The right side of the image displays piles of sand and two standing figures who turn their back towards the viewer. Compared to the first results the image conveys a sense of motion and liveliness, this is accentuated by the painting style (the reproduction suggests visible brushstrokes and different color regions). In addition to the content similarity, the last example thus suggests a similarity based on the mood and effect of the image. This concise study suggests that the motif of the boat mainly appears within a landscape setting which shows very few signs of human life. If we look at their time of creation and the life dates of the artists, we can observe that all paintings were created in the late 19th century and early 20th century (Armand), thus suggesting a preference for the motif during that time.

[56]

¹⁰⁶ Cf. Galerie Neupert.

¹⁰⁷ Cf. G. & L. Bollag 1931.

¹⁰⁸ Cf. Kunstsalon Dr. Störi 1928a.

¹⁰⁹ Cf. Kunstsalon Dr. Störi 1928b.

6.2.2 The style of Max Liebermann

Max Liebermann is one of the most important artists of the 19th and 20th centuries; he is known for his elaborate oil paintings as well as for his delicate drawings and sketches. We utilize the suggested method to gain more insight into Liebermann's drawings by looking at similar images (see question two). We take his chalk drawing *Einholung Bismarcks in Berlin* (1890) as a query. Interestingly, all first results stem from an auction held at the *Kunstsalon* of Paul Cassirer in March 1925 (Fig. 13).¹¹⁰ The title of the catalog already indicates that the auction only included drawings by Liebermann (namely 316 works).¹¹¹ While this result does not enable us to study similar drawings by other artists, it is interesting, because it suggests a strong individual artistic style. According to the results Liebermann is most similar to himself. Further analysis might study the drawings in more detail and look at the motifs, composition or textures he used.

[57]

Both case studies demonstrate the potential of our method to search for identical and similar images in a large dataset, thus not only assisting provenance researchers but any discipline interested in visual patterns over time and space. The examples of Achenbach and Liebermann highlighted that by searching for similar images, we can address a variety of research questions which go beyond the provenance of objects.

[58]



Figure 13: Search results for Liebermann's drawing *Einholung Bismarcks in Berlin* (1890). All results appear in an auction catalog published by the *Kunstsalon Paul Cassirer* in March 1925. [Graphic: Mathias Zinnen / Sabine Lang 2024]

7. Similarity

The previous sections described how our method allows to search for identical and similar images based on a given query. Thus, the term *similar* featured often. Although many scholars emphasize the blurriness of the term, *similarity* is a widely discussed concept and has been addressed within art history¹¹² and media studies, among others.¹¹³ Dorothee Kimmich offers a comprehensive introduction to the concept of similarity, particularly within modernity.¹¹⁴ In her book, Kimmich highlights the general unpopularity of vague terms such as similarity: Philosophy proclaims its unusability and cultural studies consider it historically outdated.¹¹⁵ Critics argue that since everything can be similar to everything else in some way, no new knowledge is

[59]

¹¹⁰We disregard the first result, because it is identical to the query image.

¹¹¹Cf. Kunstsalon Paul Cassirer 1925.

¹¹²Cf. Gaier et al. 2012.

¹¹³Cf. Winkler 2021.

¹¹⁴Cf. Kimmich 2017.

¹¹⁵Cf. Kimmich 2017, p. 15.

produced.¹¹⁶ Depending on which criteria are picked, things can either be similar or dissimilar.¹¹⁷ Today, similarity is seen as a mental and subjective concept that enables and organizes recognition.¹¹⁸ While the concept of similarity is often rejected and criticized, its vagueness and blurriness might also provide a chance to explore and discuss its meaning within a specific (thematic) context, free from prior assumptions or guiding definitions. This paper makes use of this degree of freedom by looking at *similarity* in relation to the method employed. Related terms such as mimesis, imitation / mimicry, similitudo, or iconicity will not be discussed, as it would go beyond the scope of this paper.¹¹⁹ After providing some general remarks, we focus on two aspects of similarity: quantification and the process of abstraction.

7.1 General Remarks on Similarity

Comparative processes are integral to determining similarity.¹²⁰ Humans, for example, compare paintings to reach a conclusion about their similarity. This process is guided by various criteria, which are manifold, subjective and often difficult to grasp. These criteria might include specific motifs, the image composition, color, forms, texture or artist, time period and location and thus encompass both internal and external criteria of the image. Our proposed method suggests a similarity based on image-inherent, visual criteria. This similarity might be determined by global image structures, numerous objects and their arrangement, or singular objects. Following this approach, we understand similarity as determined by criteria inherent and visible in an image. Thus, external criteria described for example by metadata, such as artist, title, year, or technique, are disregarded.

[60]

While our method focuses on internal criteria, it still allows us to explore different similarity dimensions: By selecting a suitable pre-training method, the computer scientist can try to influence the type of similarity underlying the retrieval process. For example, we can assume that an extraction network trained to recognize body postures will put more weight on this aspect of similarity. Therefore, the process of attesting similarity is already influenced during pre-training. Fig. 14 illustrates this: A search for Ludwig von Zumbusch's *Einsames Land* (1896) leads to different results depending on the embedding type selected by the user. While results for the models ImageNet and Arts show a clear preference for landscapes, the Poses model prefers images with persons and objects. In general, it can be observed that using the Poses model leads to worse results than utilizing the ImageNet and Arts models. We illustrate this by searching for Sperl's *Sommerlust* using all three embedding types (see Fig. 15). Models trained on ImageNet and Arts retrieve all three instances of Sperl's painting in the dataset, while the Poses model only finds one instance.

[61]

¹¹⁶ Cf. Kimmich 2017, pp. 18–19.

¹¹⁷ Cf. Kimmich 2017, p. 14. Kimmich discusses that in the context of social gender roles and biological sexes.

¹¹⁸ Cf. Kimmich 2017, p. 34.

¹¹⁹ Similitudo, for example, plays a central role for the genre of portraiture; a fundamental requirement for a portrait is that it resembles the subject / person (cf. Gaier et al. 2012). Mimesis has existed as a classical category since antiquity and is, according to Hartmut Winkler, connected to the question of how art and the media relate to the world (cf. Winkler 2021, p. 283). Iconicity then also refers to the fact that images (i. e. photos) are similar to what is shown (cf. Winkler 2021, p. 47). See the following sources for more information: Winkler 2021, pp. 11–12, 47, 283; Kimmich 2017, p. 15, 21; Gaier 2012.

¹²⁰ In his book, Winkler mentions that attesting similarity requires comparative processes; these processes, however, can happen unconsciously (cf. Winkler 2021, pp. 111–112).

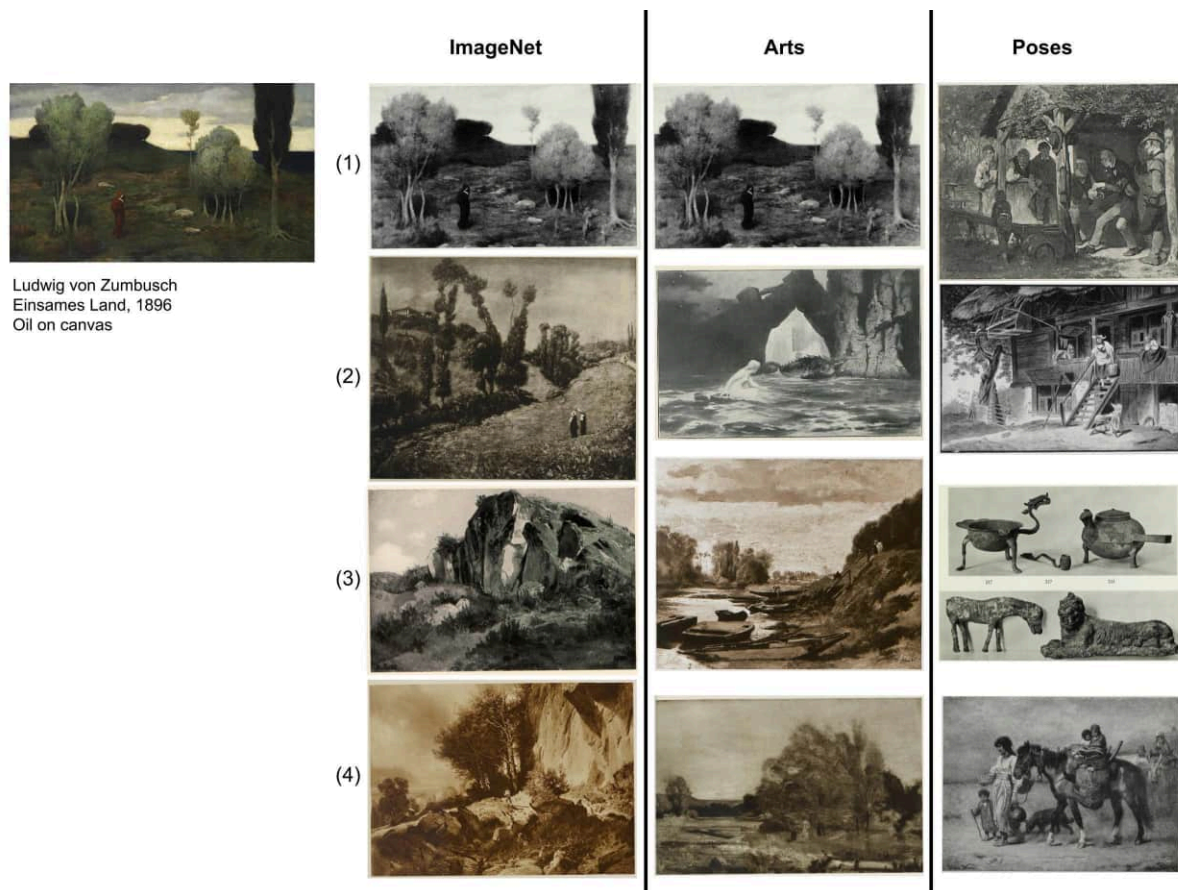


Figure 14: Search results using different embedding types; all models were trained on different data thus leading to different search results and allowing to focus on diverse similarity criteria (so our hypothesis). [Graphic: Mathias Zinnen / Sabine Lang 2024]

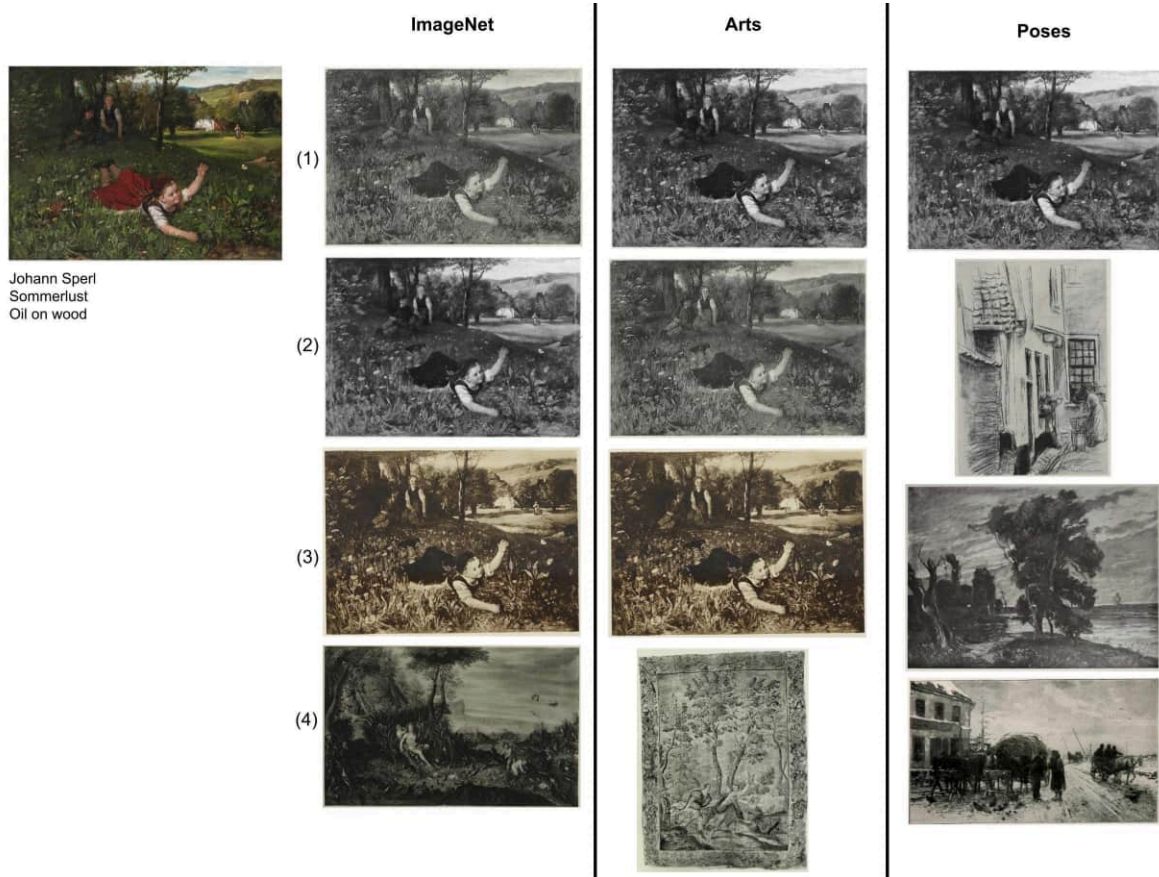


Figure 15: Search results using different embedding types; the models *ImageNet* and *Arts* perform best for the task of finding identical images to Johann Sperl's *Sommerlust*. [Graphic: Mathias Zinnen / Sabine Lang 2024]

In our method, we derive quantifiable, vector-shaped representations of image content and calculate the distance between two vectors using an arbitrary distance metric (i. e. Euclidean Distance). We assume that this distance approximates the degree of similarity between the images. However, the computation of the distance between two feature vectors does not provide any insights into interpretable similarity criteria; knowing the similarity distance function does not explain why two pictures might be perceived as similar, while others are seen as different.¹²¹ To understand what makes artworks similar from a data-driven perspective, we must consider the properties of the feature space. This requires discussing the function that translates pixel-based image representations into feature vectors. We refer to this process as quantification. [62]

7.2 Quantification

According to the *Duden* (dictionary of the German language), *quantification* means the transformation of qualities into quantities, for example the properties of something (here: an image) in numbers and measurable values.¹²² In this paper, quantification can relate to two different aspects, namely the process of translating an image into feature vectors and the fact that similarity is computed which requires a quantification process. [63]

¹²¹ While we can influence the distance computation by selecting a different distance metric (i. e. Euclidean Distance) to some degree, the general tendency will remain similar.

¹²² Cf. Dudenredaktion (ed.) 2024.

In order to search for identical and similar images, we utilize a neural network to extract feature vectors which are essentially numerical representations of a group of features which describe an image (such as colors, edges, or objects). Thus, these features quantify and represent the content of an image. During the search process, the feature vector of the given query image is compared to all feature vectors stored in the feature database and their distances are calculated using the Euclidean Distance. The closer the distance, the more similar the images are according to the algorithm. Therefore, similarity is not based on a subjective impression, but on measurable, comparable, and objective numerical values. Since feature vectors are extracted from input data – in our case digital images – the question arises whether the similarity measure depends on the image quality and can therefore vary even for images which appear identical or similar to the human eye (see case study on Johann Sperl, section 6.1). For example, different reproduction and digitization techniques might result in different color representations or image contrasts which might influence the feature vectors and calculations respectively. Fig. 10 shows that even for the same image the quality and color of the reproduction differs significantly. [64]

Cognitive psychology picks up on the idea that similarity is founded in and expressed by numerical values. Psychological processes are modeled using computers; this means that only processes that can be expressed in variables and algorithms are represented. Accordingly, features are variables to which numerical values are assigned allowing computation. Essentially, for Cognitive Psychology, similarity is characterized as a ›feature overlap¹²³: »People notice that a number of objects overlap substantially and proceed to form a category to include these items. [...] Categorization is justified by the observation that objects tend to cluster in terms of their attributes, be these physical features, linguistic labels [...].«¹²⁴ As Winkler concludes in his book on similarity, the computer provides the frame in which cognitivists think about similarity.¹²⁵ The same can be said about our project, since the method guides and essentially dictates how we think and write about similarity. [65]

7.3 Process of Abstraction

In his book, Winkler emphasizes that comparison and similarity separates things in aspects (features), which are similar and dissimilar.¹²⁶ Human observation, according to him, does not remain with the things themselves, but rather moves on to their properties and characteristics.¹²⁷ Accordingly, similarity abstracts something from things (the similar thing eventually results in a form; form then plays a crucial role for him). Thus, for Winkler, similarity is based on mechanisms of abstraction.¹²⁸ Accordingly, we understand the ›process of abstraction‹ as a process in which certain criteria of the input data are abstracted to determine if things are similar or dissimilar. As described previously in section 4.2, an image-based search requires the extraction of features from the data. Ideally, these features describe and represent the (content of the) data. This process aligns with Winkler's conception of similarity, as it is based on specific criteria which are context dependent. This process of abstraction in the machine mirrors the human approach to recognizing similarities, most evidently illustrated by the conversion of input data into a lower-dimensional vector representation during feature extraction. This representation retains the essential information needed for a given task – in this case, retrieval – while disregarding unimportant information. This process is thus associated with a loss of (certain) information and potentially with a loss of possible similarity criteria. [66]

¹²³ Cf. Winkler 2021, pp. 96–97.

¹²⁴ Anderson 1991, p. 411.

¹²⁵ Cf. Winkler 2021, p. 98.

¹²⁶ Cf. Winkler 2021, p. 257.

¹²⁷ Cf. Winkler 2021, p. 93.

¹²⁸ Cf. Winkler 2021, pp. 257–258.

Winkler also raises the question whether similarity implies that our attention is focused on certain criteria: [67]
 What exactly is guiding our attention and selection?¹²⁹ For humans, the cultural and societal context as well as personal experiences and preferences play a crucial role.¹³⁰ As mentioned before, the computer scientist can try to influence the type of similarity underlying the retrieval process by selecting a suitable pre-training method. We can assume, for example, that an extraction network trained to recognize body postures will put more weight on these similarity criteria (see results Fig. 14). If we interpret feature extraction as an abstraction in Winkler's sense, selecting a specific pre-training method determines the understanding of similarity underlying the retrieval system. Selecting a specific method can then be correlated to the type of information we are «omitting» during the translation of pixels into a feature vector. Whenever we select a specific method for this translation, we also decide which features are deemed irrelevant by the models. Similarity in this way is always influenced by the choice we make regarding the pre-training method. This means that the attention and selection process of the network is guided by human input, often motivated by the research question, task or personal interests. In comparison to humans, the attention and selection of similarity criteria in this context is much more conscious and goal-driven.

We conclude the following: Attesting similarity requires comparative processes separating things into aspects [68]
 (features), which are similar and dissimilar¹³¹, which in the context of this paper refer to image-internal components. These criteria are abstracted from the image and quantified, thus becoming measurable. We also noted that the image quality might influence the similarity measure. The conversion of input data into a lower-dimensional vector representation during feature extraction is associated with a loss of information and therefore might result in a loss of similarity dimensions. We also elaborated on the fact that the selection of the pre-training method can influence the type of similarity underlying the retrieval process. This section thus highlighted that the computer guides as well as limits our understanding of similarity. Discussing similarity in the context of this paper, with a particular focus on quantification and the process of abstraction, provided interesting insights and leaves room for further discussion and research.

8. Conclusion

This paper demonstrated the potential of applying machine learning methods for provenance research. We [69]
 showed how an image-based search in auction catalogs can circumvent the issue of missing information, varying titles or artist attribution and thus assist with the reconstruction of an object's provenance. The same method might be used to study similar material such as exhibition catalogs, a *catalogue raisonné*, magazines, newspapers or photographs. Future work also includes testing other models, such as *CLIP*, and creating a long-term accessible interface that will allow provenance researchers to use the method. Beyond technical aspects, the paper addressed the implications of machine learning-based retrieval for the concept of similarity, especially focusing on its quantification and the process of abstraction. In the context of this paper, similarity is based on visual criteria inherent in the image. Our proposed method enables us to explore different similarity dimensions by selecting a suitable pre-training method. We elaborated on the fact that quantification relates to two different aspects, namely the process of translating an image into feature vectors and that similarity is computed which requires a quantification process in the first place. Thus, similarity becomes a measurable, comparable and objective numerical value. We also highlighted a connection to Cognitive Psychology, where psychological processes are modeled using a computer and essentially are represented as numbers.¹³² The paper also emphasized that similarity requires abstraction processes. Here, we referred to Hartmut Winkler, who also wrote that similarity is based on mechanisms of abstraction, where comparison and similarity separate things into aspects (features) which are similar and

¹²⁹ Cf. Winkler 2021, p. 295.

¹³⁰ American philosopher Nelson Goodman famously stated that «Circumstances alter similarities» (quoted in Kimmich 2017, p. 24), thus emphasizing the context-dependency of similarity judgments (cf. Winkler 2021, p. 94).

¹³¹ Cf. Winkler 2021, pp. 111–112, 257.

¹³² Cf. Winkler 2021, p. 97.

dissimilar.¹³³ Accordingly, we understand the ›process of abstraction‹ as a process in which certain criteria of the input data are abstracted to determine if things are similar or not. In that sense the abstraction process equals the extraction of features from the data which describe and represent the (content of the) data and form the basis for an image-based search. Eventually, we also elaborated on the loss of information and potentially the loss of similarity criteria associated with feature extraction and asked what is guiding the criteria selection:¹³⁴ For humans, this process is influenced by the context;¹³⁵ for the machine, the selection process is determined by the computer scientist and the pre-training method selected for the retrieval process.

In the context of provenance research the question remains, if similar is enough. The *Provenance Research Manual* published by the German Lost Art Foundation states the following: »[...] [A]ll information about the identity of the piece (artist signatures, hallmarks, different attributions, variants, replicas or copies and recasts [...]) should be documented [...]»¹³⁶ and explanations must be given, if the identity is not clear.¹³⁷ Thus, the clear identity of the work is crucial when establishing its provenance. Being ›similar‹ thus might not seem enough for provenance researchers. However, visually similar objects might offer valuable clues on which research direction to pursue.

[70]

¹³³ Cf. Winkler 2021, pp. 257–258.

¹³⁴ Cf. Winkler 2021, p. 295.

¹³⁵ Cf. Winkler 2021, p. 94; Kimmich 2017, p. 24.

¹³⁶ German Lost Art Foundation et al. 2020, p. 40.

¹³⁷ Cf. German Lost Art Foundation et al. 2020, p. 81.

Bibliography

- Abubakar Abid / Ali Abdalla / Ali Abid / Dawood Khan / Abdulrahman Alfozan / James Zou: Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild. arXiv. 06.06.2019. DOI: [10.48550/arXiv.1906.02569](https://doi.org/10.48550/arXiv.1906.02569)
- John R. Anderson: The Adaptive Nature of Human Categorization. In: *Psychological Review* 98 (1991), No. 3, pp. 409–429. DOI: [10.1037/0033-295X.98.3.409](https://doi.org/10.1037/0033-295X.98.3.409)
- Astrid Bähr: German Sales 1930–1945. Bibliographie der Auktionskataloge aus Deutschland, Österreich und der Schweiz. Edited by Joachim Brand / Moritz Wullen. Berlin 2013. PDF. DOI: [10.11588/artdok.00002251](https://doi.org/10.11588/artdok.00002251)
- Peter Bell / Leonardo Impett: Ikonographie und Interaktion. Computergestützte Analyse von Posen in Bildern der Heilsgeschichte. In: *Das Mittelalter* 24 (2019), No. 1, pp. 31–53. DOI: [10.1515/mial-2019-0004](https://doi.org/10.1515/mial-2019-0004)
- Siwar Bengamra / Olfa Mzoughi / André Bigand / Ezzeddine Zagrouba: New Challenges of Face Detection in Paintings Based on Deep Learning. In: Giovanni Maria Farinella / Petia Radeva / Jose Braz / Kadi Bouatouch (eds.): Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP, Online, 08.–10.02.2021), Vol. 4. Vienna 2021, pp. 311–320. DOI: [10.5220/0010243703110320](https://doi.org/10.5220/0010243703110320)
- Siwar Bengamra / Olfa Mzoughi / André Bigand / Ezzeddine Zagrouba: A Comprehensive Survey on Object Detection in Visual Art: Taxonomy and Challenge. In: *Multimedia Tools and Applications* 83 (2024), No. 5, pp. 14637–14670. DOI: [10.1007/s11042-023-15968-9](https://doi.org/10.1007/s11042-023-15968-9)
- Valentine Bernasconi / Eva Cetinic / Leonardo Impett: A Computational Approach to Hand Pose Recognition in Early Modern Paintings. In: *Journal of Imaging* 6 (2023), No. 3. DOI: [10.3390/jimaging9060120](https://doi.org/10.3390/jimaging9060120)
- Britta Bommert: German Sales 1901–1929. Bibliographie der Auktionskataloge aus Deutschland, Österreich und der Schweiz. Edited by Joachim Brand. Berlin 2019. DOI: [10.11588/artdok.00006565](https://doi.org/10.11588/artdok.00006565)
- Nicolas Carion / Francisco Massi / Gabriel Synnaeve / Nicolas Usinier / Alexander Kirillov / Sergey Zagoruyko: End-to-End Object Detection with Transformers. In: Andrea Vedaldi / Horst Bischof / Thomas Brox / Jan-Michael Frahm (eds.): *Computer Vision – ECCV 2020. Conference Papers. Part I* (Online, 23.–28.08.2020). Cham, CH 2020, pp. 213–229. DOI: [10.1007/978-3-030-58452-8_13](https://doi.org/10.1007/978-3-030-58452-8_13)
- Giovanna Castellano / Gennaro Vessio: A Brief Overview of Deep Learning Approaches to Pattern Extraction and Recognition in Paintings and Drawings. In: Alberto Del Bimbo / Rita Cucchiara / Stan Sclaroff / Giovanni Maria Farinella / Tao Mei / Marco Bertini / Hugo Jair Escalante / Roberto Vezzani (eds.): *Pattern Recognition. ICPR International Workshops and Challenges. Proceedings (ICPR 2021, Online, 10–15.01.2021)*. Cham, CH 2021, pp. 487–501. DOI: [10.1007/978-3-030-68796-0_35](https://doi.org/10.1007/978-3-030-68796-0_35)
- Giovanna Castellano / Eufemia Lella / Gennaro Vessio: Visual Link Retrieval and Knowledge Discovery in Painting Datasets. In: *Multimedia Tools and Applications* 80 (2021), pp. 6599–6616. DOI: [10.1007/s11042-020-09995-z](https://doi.org/10.1007/s11042-020-09995-z)
- Eva Cetinic / James She: Understanding and Creating Art with AI: Review and Outlook. In: *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 18 (2022), No. 2, pp. 1–22. DOI: [10.1145/3475799](https://doi.org/10.1145/3475799)
- Wei Chen / Yu Liu / Weiping Wang / Erwin M. Bakker / Theodoros Georgiou / Paul Fieguth / Li Liu / Michael S. Lew: Deep Learning for Instance Retrieval. A Survey. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45 (2022), No. 6, pp. 7270–7292. DOI: [10.1109/TPAMI.2022.3218591](https://doi.org/10.1109/TPAMI.2022.3218591)
- Common Objects in Context. Last accessed: 30.07.2024. HTML. [\[online\]](#)
- Elliot J. Crowley / Andrew Zisserman: The State of the Art. Object Retrieval in Paintings using Discriminative Regions. In: Michel Valstar / Andrew French / Tony Pridmore (eds.): *Proceedings of the British Machine Vision Conference 2014 (BMVC 2014, Nottingham, UK, 01.–05.09.2014)*. Nottingham, UK 2014. PDF. DOI: [10.5244/C.28.38](https://doi.org/10.5244/C.28.38)
- Elliot J. Crowley / Andrew Zisserman: In Search of Art. In: Lourdes Agapito / Michael M. Bronstein / Carsten Rother (eds.): *Computer Vision – ECCV 2014 Workshops. Proceedings. Part I* (Zurich, 06.–12.09.2014). Cham, CH etc. 2015, pp. 54–70. DOI: [10.1007/978-3-319-16178-5_4](https://doi.org/10.1007/978-3-319-16178-5_4)
- Elliot J. Crowley / Andrew Zisserman: The Art of Detection. In: Gang Hua / Hervé Jégou (eds.): *Computer Vision – ECCV 2016 Workshops. Proceedings. Part I* (Amsterdam, 08.–10.10.2016 and 15.–16.10.2016). Cham, CH 2016, pp. 721–737. PDF. DOI: [10.1007/978-3-319-46604-0_50](https://doi.org/10.1007/978-3-319-46604-0_50)
- Gabriella Csurka / Christopher R. Dance / Lixin Fan / Jutta Willamowski / Cédric Bray: Visual Categorization with Bags of Keypoints. In: *Workshop on Statistical Learning in Computer Vision. Proceedings (ECCV 2004, Prague, 15.05.2004)*. PDF. [\[online\]](#)
- Jia Deng / Wei Dong / Richard Socher / Li-Jia Li / Kai Li / Li Fei-Fei: Imagenet. A Large-Scale Hierarchical Image Database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition. Proceedings (Miami, 20.–25.06.2009)*. Miami 2009, pp. 248–255. PDF. DOI: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848)
- Dudenredaktion (ed.): Quantifizierung. In: *Duden online*. Last accessed: 25.07.2024. HTML. [\[online\]](#)
- Mark Everingham / Luc Van Gool / Christopher K. I. Williams / John Winn / Andrew Zisserman: The PASCAL Visual Object Classes (VOC) Challenge. In: *International Journal of Computer Vision* (2010), No. 88, pp. 308–338. DOI: [10.1007/s11263-009-0275-4](https://doi.org/10.1007/s11263-009-0275-4)
- Pamela Fletcher / Anne Helmreich / David Israel / Seth Erickson: Local / Global: Mapping Nineteenth-Century London's Art Market. In: *Nineteenth-Century Art Worldwide* 11 (2012), No. 3. HTML. [\[online\]](#)
- Amalia Foka: Computer Vision Applications for Art History: Reflections and Paradigms for Future Research. In: *Proceedings of EVA London 2021. AI and the Arts: Artificial Imagination (EVA 2021, London, 05.–09.07.2021)*. London 2021, pp. 73–80. PDF. DOI: [10.14236/ewic/EVA2021.12](https://doi.org/10.14236/ewic/EVA2021.12)
- Martin Gaier / Jeanette Kohl / Alberto Saviello: Similitudo. Konzepte der Ähnlichkeit in Mittelalter und Früher Neuzeit. Paderborn 2012. [\[GVK record\]](#)
- Noah Garcia / George Vogiatzis: How to Read Paintings: Semantic Art Understanding with Multi-Modal Retrieval. arXiv. 23.10.2018. PDF. DOI: [10.48550/arXiv.1810.09617](https://doi.org/10.48550/arXiv.1810.09617)
- Raul Vicente-Garcia: Zum Ersten, zum Zweiten, zum Dritten – gefunden! In: *FUTUR* (2021), No. 2. Last accessed: 30.07.2024. HTML. [\[online\]](#)
- Leon A. Gatys / Alexander S. Ecker / Matthias Bethge: A Neural Algorithm of Artistic Style. arXiv. 26.08.2015. Version 2 from 02.09.2015. DOI: [10.48550/arXiv.1508.06576](https://doi.org/10.48550/arXiv.1508.06576)
- Leon A. Gatys / Alexander S. Ecker / Matthias Bethge: Image Style Transfer using Convolutional Neural Networks. In: *29th IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2016. Proceedings (Las Vegas, 27.–30.06.2016)*. Los Alamitos, US-CA etc. 2016, pp. 2414–2423. PDF. [\[online\]](#)
- Robert Geirhos / Patricia Rubisch / Claudio Michaelis / Matthias Bethge / Felix A. Wichmann / Wieland Brendel: ImageNet-Trained CNNs are Biased towards Texture. Increasing Shape Bias Improves Accuracy and Robustness. *OpenReview.net*. 21.12.2018. Last modified: 08.02.2026. PDF / HTML. [\[online\]](#)
- German Lost Art Foundation / Arbeitskreis Provenienzforschung e.V. / Arbeitskreis Provenienzforschung und Restitution – Bibliotheken / Deutscher Bibliotheksverband e.V. / Deutscher Museumsbund e.V. / ICOM Deutschland e.V. (eds.): *Provenance Resource Manual. To Identify Cultural Property Seized due to Persecution during the National Socialist Era*. 2020. PDF. [\[online\]](#)
- German Sales. Last accessed: 03.09.2025. HTML. DOI: [10.11588/portal.gs](https://doi.org/10.11588/portal.gs)

- German Sales: Project description. Last accessed: 24.07.2024. HTML. [\[online\]](#)
- Ross Girshick / Jeff Donahue / Trevor Darrell / Jitendra Malik: Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2014. Proceedings (Columbus, US-OH, 23.–28.06.2014). Los Alamitos, US-CA etc. 2014, pp. 580–587. PDF. [\[online\]](#)
- Ross Girshick: Fast R-CNN. In: 2015 IEEE International Conference on Computer Vision. ICCV 2015. Proceedings (Santiago, CL, 11.–18.12.2015). Los Alamitos, US-CA etc. 2015, pp. 1440–1448. PDF. [\[online\]](#)
- Nicolas Gonthier / Yann Gousseau / Saïd Ladjal / Olivier Bonfait: Weakly Supervised Object Detection in Artworks. arXiv. 05.10.2018. PDF. DOI: [10.48550/arXiv.1810.02569](#)
- Nicolas Gonthier / Saïd Ladjal / Yann Gousseau: Multiple Instance Learning on Deep Features for Weakly Supervised Object Detection with Extreme Domain Shifts. In: Computer Vision and Image Understanding 214 (2022), 103299. DOI: [10.1016/j.cviu.2021.103299](#)
- Dorothee Haffner: Provenienzforschung digital vernetzt. Ergebnisse sichtbar machen. In: Museumskunde 84 (2019), pp. 90–97. Last accessed: 29.07.2024. PDF. [\[online\]](#)
- Dorothee Haffner: Provenienzen in Sammlungsdatenbanken. Digitale und virtuelle Chancen für die Vermittlung. In: Deutsches Zentrum Kulturgutverluste (ed.): Digitale Provenienzforschung (=Provenienz & Forschung, 1). Dresden 2020, pp. 36–42. [\[GVK record\]](#)
- Junfeng He / Jinyuan Feng / Xianglong Liu / Tao Cheng / Tai-Hsu Lin / Hyunjin Chung / Shih-Fu Chang: Mobile Product Search with Bag of Hash Bits and Boundary Reranking. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2012. Proceedings (Providence, US-RI, 16.–21.06.2012). Los Alamitos, US-CA etc. 2012, pp. 3005–3012. DOI: [10.1109/CVPR.2012.6248030](#)
- Kaiming He / Xiangyu Zhang / Shaoqing Ren / Jian Sun: Deep Residual Learning for Image Recognition. In: 29th IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2016. Proceedings (Las Vegas, 27.–30.06.2016). Los Alamitos, US-CA etc. 2016, pp. 770–778. PDF. [\[online\]](#)
- Meike Hoffmann and Nicola Kuhn: Hitlers Kunsthändler: Hildebrand Gurlitt 1895–1956. Munich 2016. [\[GVK record\]](#)
- Meike Hopp: Kunsthandel im Nationalsozialismus: Adolf Weinmüller in München und Wien. Cologne etc. 2012. [\[GVK record\]](#)
- Meike Hopp: Provenienzforschung und digitale Forschungsinfrastrukturen in Deutschland: Bedürfnisse, Desiderate, Tendenzen. In: Eva Blimlinger / Heinz Schödl (eds.): ... (kein Ende in Sicht. 20 Jahre Kunstrückgabegesetz in Österreich (= Schriftenreihe der Kommission für Provenienzforschung, 8). Vienna 2018, pp. 35–59. DOI: [10.7767/9783205201274.37](#)
- Christian Huemer: The »German Sales 1930–1945« Database Project. In: Collections 10 (2014), No. 3, pp. 273–278. DOI: [10.1177/155019061401000306](#)
- Leonardo Impett / Sabine Süssstrunk: Pose and Pathosformel in Aby Warburg's Bilderatlas. In: Gang Hua / Hervé Jégou (eds.): Computer Vision – ECCV 2016 Workshops. Proceedings. Part I (Amsterdam, 08.–10.10.2016 and 15.–16.10.2016). Cham, CH 2016, pp. 888–902. HTML / PDF. DOI: [10.1007/978-3-319-46604-0_61](#)
- Leonardo Impett / Franco Moretti: Totentanz. Operationalizing Aby Warburg's Pathosformeln. In: New Left Review (2017), No. 107, pp. 68–97. [\[GVK record\]](#)
- Leonardo Impett: Analyzing Gesture in Digital Art History. In: Kathryn Brown (ed.): The Routledge Companion to Digital Humanities and Art History. London etc. 2020, pp. 386–407. [\[GVK record\]](#)
- Hyeong-Ju Jeon / Soonchul Jung / Yoon-Seok Choi / Jae Woo Kim / Jin Seo Kim: Object Detection in Artworks Using Data Augmentation. In: ICTC 2020. The 11th International Conference on Information and Communication Technology Convergence. Data, Network, and AI in the Age of »Untact« (Jeju, KR, 21.–23.10.2020). Jeju, KR 2020, pp. 1312–1314. DOI: [10.1109/ICTC49870.2020.9289321](#)
- Gesa Jeuthe: Kunstwerte im Wandel: Die Preisentwicklung der deutschen Moderne im nationalen und internationalen Kunstmarkt 1925 bis 1955. Vol. 7. Berlin 2014. [\[GVK record\]](#)
- Glenn Jocher / Ayusg Chaurasia / Jing Qiu: Ultralytics YOLO. GitHub. 10.01.2023. Version 8.4.53 from 22.05.2026. [\[online\]](#)
- Jeff Johnson / Matthijs Douze / Hervé Jégou: Billion-Scale Similarity Search with GPUs. In: IEEE Transactions on Big Data 7 (2019), No. 3, pp. 535–547. DOI: [10.1109/TBDATA.2019.2921572](#)
- Xuan Ju / Ailing Zeng / Jianan Wang / Qiang Xu / Lei Zhang: Human-Art. A Versatile Human-Centric Dataset Bridging Natural and Artificial Scenes. In: 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR 2023. Proceedings (Vancouver, 17.–24.06.2023). Los Alamitos, US-CA etc., pp. 618–629. PDF. [\[online\]](#)
- David Kadish / Sebastian Risi / Anders Sundnes Løvlie: Improving Object Detection in Art Images Using Only Style Transfer. In: 2021 International Joint Conference on Neural Networks. Proceedings (IJCNN, Online, 18.–22.07.2021). Shenzhen, CHN 2021. PDF. DOI: [10.1109/IJCNN52387.2021.9534264](#)
- Dorothee Kimmich: Ins Ungefähre: Ähnlichkeit und Moderne. Constance 2017. [\[GVK record\]](#)
- Ketterer Kunst (2018a). Auction on May 18, 2018, in Munich. Lot 15, Heinrich Bürkel »Der Kochensee mit den Häusern von Schlehdorf«. [\[online\]](#)
- Ketterer Kunst (2018b). Auction on November 23, 2018, in Munich. Lot 28, Adrian Ludwig Richter »Hirten am Feuer (Abendlandschaft)«. [\[online\]](#)
- Ketterer Kunst (2019). Auction on May 24, 2019, in Munich. Lot 57, Ludwig von Zumbusch »Einsames Land«. [\[online\]](#)
- Sabine Lang (2023a): Wie hat sich Provenienzforschung durch Digitalität verändert? In: Sebastian Finsterwalder (ed.): RETOUR. Freier Blog für Provenienzforschende. 07.08.2023. HTML. [\[online\]](#)
- Sabine Lang (2023b): »Mind the Gap«: Von Lücken in der Provenienzforschung und ihrer Präsenz im digitalen Raum. In: Peer Trilcke / Anna Busch / Patrick Helling (eds.): DHD 2023. Open Humanities, Open Culture. 9. Jahrestagung des Verbands Digital Humanities im deutschsprachigen Raum. Conference Abstracts. (Trier, Luxembourg, 13.–17.03.2023). Trier etc. 2023, pp. 212–217. PDF. DOI: [10.5281/zenodo.7715420](#)
- Tsung-Yi Lin / Michael Maire / Serge Belongie / Lubomir Bourdev / Ross Girshick / James Hays / Pietro Perona / Deva Ramanan / C. Lawrence Zitnick / Piotr Dollár: Microsoft COCO. Common Objects In Context. In: David Fleet / Tomas Pajdla / Bernt Schiele / Tinne Tuytelaars (eds.): Computer Vision – ECCV 2014. Proceedings. Part V (Zurich, 06.–12.09.2014). Cham, CH 2014, pp. 744–750. PDF. DOI: [10.1007/978-3-319-10602-1_48](#)
- Wei Liu / Dragomir Anguelov / Dumitru Erhan / Christian Szegedy / Scott Reed / Chen-Yang Fu / Alexander C. Berg: SSD. Single Shot MultiBox Detector. In: Computer Vision – ECCV 2016. Proceedings. Part I (Amsterdam, 11.–14.10.2016). Cham, CH 2016, pp. 21–37. DOI: [10.1007/978-3-319-46448-0_2](#)
- David G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints. In: International Journal of Computer Vision 60 (2004), pp. 91–110. DOI: [10.1023/B:VISI.0000029664.99615.94](#)
- Yue Lu / Chao Guo / Xingyuan Dai / Fei-Yue Wang: Image Captioning on Fine Art Paintings via Virtual Paintings. In: IEEE 1st International Conference on Digital Twins and Parallel Intelligence. DTPI 2021. Proceedings (Beijing, CHN, 15.07.–15.08.2021). Los Alamitos, US-CA etc., pp. 156–159. DOI: [10.1109/DTPI52967.2021.9540081](#)

- Prathmesh Madhu / Ronak Kosti / Lara Mührenberg / Peter Bell / Andreas Maier / Vincent Christlein: Recognizing Characters in Art History Using Deep Learning. In: Valerie Gouet-Brunet / Margarita Khokhlova / Liming Chen (eds.): SUMAC '19. Proceedings of the 1st Workshop on Structuring and Understanding of Multimedia Heritage Contents (Nice, FR, 21.10.2019). New York 2019, pp. 15–22. DOI: [10.1145/3347317.3357242](https://doi.org/10.1145/3347317.3357242)
- Prathmesh Madhu / Tilman Marquart / Ronak Kosti / Peter Bell / Vincent Christlein: Understanding Compositional Structures in Art Historical Images Using Pose and Gaze Priors. Towards Scene Understanding in Art History. In: Adrien Bartoli / Andrea Fusiello (eds.): Computer Vision – ECCV 2020 Workshops. Proceedings. Part II (Online, 23.–28.08.2020). Cham, CH 2020, pp. 109–125. DOI: [10.1007/978-3-030-66096-3_9](https://doi.org/10.1007/978-3-030-66096-3_9)
- Prathmesh Madhu / Tilman Marquart / Ronak Kosti / Dirk Suckow / Peter Bell / Andreas Maier / Vincent Christlein: ICC++. Explainable Feature Learning for Art History using Image Compositions. In: Pattern Recognition 136 (2023), p. 109153. DOI: [10.1016/j.patcog.2022.109153](https://doi.org/10.1016/j.patcog.2022.109153)
- Fabio Mariani: »Probably Sold to Paalen, Possibly by Exchange«: Vagueness, Incompleteness, Subjectivity and Uncertainty in Digital Art Provenance. In: Workshop on Computational Methods in the Humanities 2022 (COMHUM 2022, Lausanne, 09.–10.06.2022). Lausanne 2022. PDF. [\[online\]](#)
- Maria-Cristina Marinescu / Artem Reshetnikov / Joaquim Moré López: Improving Object Detection in Paintings Based on Time Contexts. In: Alfredo Cuzzocrea / Carlo Zaniolo (eds.): 2020 International Conference on Data Mining Workshops. Proceedings (ICDMW, Sorrento, IT, 17.–20.11.2020). Sorrento, IT 2020, pp. 926–932. PDF. DOI: [10.1109/ICDMW51313.2020.00133](https://doi.org/10.1109/ICDMW51313.2020.00133)
- Alexis Mermet / Asanobu Kitamoto / Chikahiko Suzuki / Akira Takagishi: Face Detection on Pre-modern Japanese Artworks using R-CNN and Image Patching for Semi-Automatic Annotation. In: Valerie Gouet-Brunet / Margarita Khokhlova / Ronak Kosti / Liming Chen / Xu-Cheng Yin (eds.): SUMAC'20. Proceedings of the 2nd Workshop on Structuring and Understanding of Multimedia Heritage Contents (Seattle, 12.10.2020). New York 2020, pp. 23–31. DOI: [10.1145/3423323.3423412](https://doi.org/10.1145/3423323.3423412)
- Kevin Musgrave / Serge Belongie / Ser-Nam Lim: A Metric Learning Reality Check. In: Andrea Vedaldi / Horst Bischof / Thomas Brox / Jan-Michael Frahm (eds.): Computer Vision – ECCV 2020. Proceedings. Part XXV (Online, 23.–28.08.2020). Cham, CH 2020, pp. 681–699. PDF. DOI: [10.1007/978-3-030-58595-2_41](https://doi.org/10.1007/978-3-030-58595-2_41)
- Neumeister. Last accessed: 24.07.2024. HTML. [\[online\]](#)
- David Nistér / Henrik Stewénius: Scalable Recognition with a Vocabulary Tree. In: Dan Huttenlocher / David Forsyth (eds.): 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2006. Proceedings. Volume 2 (New York, 17.–22.06.2006). Los Alamitos, US-CA etc. 2006, pp. 2161–2168. PDF. DOI: [10.1109/CVPR.2006.264](https://doi.org/10.1109/CVPR.2006.264)
- Fabian Offert / Peter Bell: IMG5.AI. A Multitodal Search Engine for Digital Art History. In: International Journal for Digital Art History 9 (2024), p. 5.28–5.39. DOI: [10.11588/dahj.2023.9.91295](https://doi.org/10.11588/dahj.2023.9.91295)
- OpenMMLab Pre-training Toolbox and Benchmark. Last accessed: 30.07.2024. HTML. [\[online\]](#)
- OpenMMLab Pose Estimation Toolbox and Benchmark. Last accessed: 30.07.2024. [\[online\]](#)
- Vishal Patoliya / Mathias Zinnen / Andreas Maier / Vincent Christlein: Smell and Emotion. Recognising Emotions in Smell-Related Artworks. arXiv. 05.07.2024. DOI: [10.48550/arXiv.2407.04592](https://doi.org/10.48550/arXiv.2407.04592)
- James Philbin / Ondrej Chum / Michael Isard / Josef Sivic / Andrew Zisserman: Lost in Quantization. Improving Particular Object Retrieval in Large Scale Image Databases. In: Linda Shapiro / Narendra Ahuja (eds.): 2008 IEEE Conference on Computer Vision and Pattern Recognition. Proceedings (CVPR, Anchorage, US-AK, 23.–28.06.2008). Los Alamitos, US-CA etc. 2008. PDF / HTML. DOI: [10.1109/CVPR.2008.4587635](https://doi.org/10.1109/CVPR.2008.4587635)
- PortApp. 2025. Last accessed: 05.02.2026. HTML. [\[online\]](#)
- PyMuPDF. Last accessed: 30.07.2024. HTML. [\[online\]](#)
- Alec Radford / Jong Wook Kim / Chris Hallacy / Aditya Ramesh / Gabriel Goh / Sandhini Agarwal / Girish Sastry / Amanda Askell / Pamela Mishkin / Jack Clark / Gretchen Krueger / Ilya Sutskever: Learning Transferable Visual Models from Natural Language Supervision. In: Marina Meila / Tong Zhang (eds.): Proceedings of the 38th International Conference on Machine Learning (PMLR 139, Online, 18.–24.07.2021). Cambridge, US-MA 2021, pp. 8748–8763. PDF. [\[online\]](#)
- Joseph Redmon / Santosh Divvala / Ross Girshick / Ali Farhadi: You Only Look Once. Unified, Real-Time Object Detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016, pp. 779–788. PDF. [\[online\]](#)
- Shaoqing Ren / Kaiming He / Ross Girshick / Jian Sun: Faster R-CNN. Towards Real-Time Object Detection with Region Proposal Networks. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (2016), No. 6, pp. 1137–1149. DOI: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)
- Artem Reshetnikov / Maria-Cristina Marinescu / Joaquim Moré López: DEArt. Dataset of European Art. In: Leonid Karlinsky / Tomer Michaeli / Ko Nishino (eds.): Computer Vision – ECCV 2022 Workshops. Proceedings. Part I (Tel Aviv, 23.–27.10.2022). Cham, CH 2022, pp. 218–233. PDF. DOI: [10.1007/978-3-031-25056-9_15](https://doi.org/10.1007/978-3-031-25056-9_15)
- Tal Ridnik / Emanuel Ben-Baruch / Asaf Noy / Lih Zelnik-Manor: ImageNet-21K. Pretraining for the Masses. arXiv. 22.04.2021. Version 4 from 05.08.2021. PDF. DOI: [10.48550/arXiv.2104.10972](https://doi.org/10.48550/arXiv.2104.10972)
- Lynn Rother / Fabio Mariani / Max Koss: Taking Care of History: Toward a Politics of Provenance Linked Open Data in Museums. In: Emily Lew Fry / Erin Canning (eds.): Perspectives on Data. Chicago 2022. HTML. DOI: [10.53269/9780865593152/06](https://doi.org/10.53269/9780865593152/06)
- Lynn Rother / Fabio Mariani / Max Koss: Hidden Value: Provenance as a Source for Economic and Social History. In: Jahrbuch für Wirtschaftsgeschichte / Economic History Yearbook 64 (2023), No. 1, pp. 111–142. DOI: [10.1515/jbwg-2023-0005](https://doi.org/10.1515/jbwg-2023-0005)
- Olga Russakovsky / Jia Deng / Hao Su / Jonathan Krause / Sanjeev Satheesh / Sean Ma / Zhiheng Huang / Andrej Karpathy / Aditya Kosla / Michael Bernstein / Alexander C. Berg / Li Fei-Fei: ImageNet Large Scale Visual Recognition Challenge. In: International Journal of Computer Vision 155 (2015), No. 3, pp. 211–252. PDF. DOI: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y)
- Artsiom Sanakoyeu / Dmytro Kotovenko / Sabine Lang / Björn Ommer: A Style-Aware Content Loss for Real-Time HD Style Transfer. In: Vittorio Ferrari / Martial Hebert / Christian Sminchisescu / Yair Weiss (eds.): Computer Vision – ECCV 2018. 15th European Conference. Proceedings. Part VIII (Munich, 08.–14.09.2018). Cham, CH 2018, pp. 698–714. HTML / PDF. DOI: [10.1007/978-3-030-01237-3_43](https://doi.org/10.1007/978-3-030-01237-3_43)
- Bruno Sartini / Aldo Gangemi: Towards the Unchaining of Symbolism from Knowledge Graphs. How Symbolic Relationships Can Link Cultures. In: Federico Boschetti / Angelo Mario Del Grosso / Enrica Salvatori (eds.): AIUCD 2021 – DHs for Society: E-Quality, Participation, Rights and Values in the Digital Age. Book of Extended Abstracts of the 10th National Conference (Pisa, 19.–22.01.2021). Pisa 2021, pp. 576–580. PDF. [\[online\]](#)
- Bruno Sartini / Sofia Baroncini / Marieke van Arp / Francesca Tomasi / Aldo Gangemi: ICON. An Ontology for Comprehensive Artistic Interpretations. In: ACM Journal on Computing and Cultural Heritage 16 (2023), No. 3, pp. 1–38. DOI: [10.1145/3594724](https://doi.org/10.1145/3594724)
- Hugo Scheithauer / Sarah Bénéière / Laurent Romary: Automatic Retro-Structuration of Auction Sales Catalogs Layout and Content. HAL Open Science. 15.04.2024. HTML / PDF. [\[online\]](#)
- Maximilian Schich / Christian Huemer / Piotr Adamczyk / Lev Manovich / Yang-Yu Liu: Network Dimensions in the Getty Provenance Index. arXiv. 09.06.2017. DOI: [10.48550/arXiv.1706.02804](https://doi.org/10.48550/arXiv.1706.02804)

- Benoit Seguin / Carlotta Striolo / Isabella diLenardo / Frederic Kaplan: Visual Link Retrieval in a Database of Paintings. In: Gang Hua / Hervé Jégou (eds.): *Computer Vision – ECCV 2016 Workshops. Proceedings. Part I* (Amsterdam, 08.–10.10.2016 and 15.–16.10.2016). Cham, CH 2016, pp. 753–767. PDF. DOI: [10.1007/978-3-319-46604-0_52](https://doi.org/10.1007/978-3-319-46604-0_52)
- Xi Shen / Alexei A. Efros / Mathieu Aubry: Discovering Visual Patterns in Art Collections with Spatially-Consistent Feature Learning. In: 2019 IEEE / CVF Conference on Computer Vision and Pattern Recognition. CVPR 2019. Proceedings (Long Beach, US-CA, 16.–20.06.2019). Los Alamitos, US-CA etc. 2019, pp. 9278–9287. PDF. [\[online\]](#)
- Stanislav Smirnov / Alma Eguizabal: Deep Learning for Object Detection in Fine-Art Paintings. In: 2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo, Cassino, IT, 22.–24.10.2018). Cassino, IT 2018, pp. 45–49. PDF. DOI: [10.1109/MetroArchaeo43810.2018.9089828](https://doi.org/10.1109/MetroArchaeo43810.2018.9089828)
- Matthias Springstein / Stefanie Schneider / Javad Rahnama / Eyke Hüllermeier / Hubertus Kohle / Ralph Ewerth: iART: A Search Engine for Art-Historical Images to Support Research in the Humanities. In: Heng Tao Shen / Yueting Zhuang / John R. Smith (eds.): *MM '21. Proceedings of the 29th ACM International Conference on Multimedia* (Online, 20.–24.10.2021). New York 2021, pp. 2801–2803. PDF. DOI: [10.1145/3474085.3478564](https://doi.org/10.1145/3474085.3478564)
- Mathias Springstein / Stefanie Schneider / Christian Althaus / Ralph Ewerth: Semi-supervised Human Pose Estimation in Art-historical Images. In: João Magalhães / Alberto del Bimbo / Shin'ichi Satoh / Nicu Sebe (eds.): *MM '22. Proceedings of the 30th ACM International Conference on Multimedia* (Lisbon, 10.–14.10.2022). New York 2022, pp. 1107–1116. PDF. DOI: [10.1145/3503161.3548371](https://doi.org/10.1145/3503161.3548371)
- Shaoyan Sun / Wengang Zhou / Qi Tian / Houqiang Li: Scalable Object Retrieval with Compact Image Representation from Generic Object Regions. In: *ACM Transactions on Multimedia Computing, Communications, and Applications* 12 (2015), No. 2. PDF. DOI: [10.1145/2818708](https://doi.org/10.1145/2818708)
- K. K. Thyagarajan / G. Kalaiarasi: A Review on Near-Duplicate Detection of Images Using Computer Vision Techniques. In: *Archives of Computational Methods in Engineering* 28 (2021), No. 3, pp. 897–916. PDF. DOI: [10.1007/s11831-020-09400-w](https://doi.org/10.1007/s11831-020-09400-w)
- Georgos Toliás / Ronan Sicre / Hervé Jégou: Particular Object Retrieval with Integral Max-Pooling of CNN Activations. arXiv. 18.11.2015. Version 2 from 24.02.2016. PDF. DOI: [10.48550/arXiv.1511.05879](https://doi.org/10.48550/arXiv.1511.05879)
- Nikolai Ufer / Max Simon / Sabine Lang / Björn Ommer: Large-Scale Interactive Retrieval in Art Collections Using Multi-Style Feature Aggregation. In: *PLoS One* 16 (2021), No. 11. HTML / PDF. DOI: [10.1371/journal.pone.0259718](https://doi.org/10.1371/journal.pone.0259718)
- Raul Vicente-García: Bildbasierte Suche in Auktionskatalogen. In: *Museumskunde* 89 (2024), no. 1 & 2, pp. 32–38. [\[GVK record\]](#)
- Sabrina Werner: Proveana. Zur Entwicklung der Forschungsdatenbank des Deutschen Zentrums Kulturgutverluste. In: *Deutsches Zentrum Kulturgutverluste* (ed.): *Digitale Provenienzforschung (=Provenienz & Forschung, 1)*. Dresden 2020, pp. 26–36. [\[GVK record\]](#)
- Nicholas Westlake / Hongping Cai / Peter Hall: Detecting People in Artwork with CNNs. In: Gang Hua / Hervé Jégou (eds.): *Computer Vision – ECCV 2016 Workshops. Proceedings. Part I* (Amsterdam, 08.–10.10.2016 and 15.–16.10.2016). Cham, CH 2016, pp. 825–841. PDF. DOI: [10.1007/978-3-319-46604-0_57](https://doi.org/10.1007/978-3-319-46604-0_57)
- Hartmut Winkler: Ähnlichkeit. Berlin 2021. [\[GVK record\]](#)
- Shuhei Yokoo / Kohei Ozaki / Edgar Simo-Serra / Satoshi Iizuka: Two-stage Discriminative Re-Ranking for Large-Scale Landmark Retrieval. In: *CVPRW 2020. 2020 IEEE / CVF Conference on Computer Vision and Pattern Recognition Workshops. Proceedings* (Online, 14.–19.06.2020). Los Alamitos, US-CA etc. 2020, pp. 4363–4370. DOI: [10.1109/CVPRW50498.2020.00514](https://doi.org/10.1109/CVPRW50498.2020.00514)
- Marchenko Yelizaveta / Chua Tat-Seng / Aristarkhova Irina: Analysis and Retrieval of Paintings Using Artistic Color Concepts. In: 2005 IEEE International Conference on Multimedia and Expo. Proceedings (Amsterdam, 06.07.2005). Los Alamitos, US-CA etc. 2005, pp. 1246–1249. PDF. DOI: [10.1109/ICME.2005.1521654](https://doi.org/10.1109/ICME.2005.1521654)
- Wengang Zhou / Houqiang Li / Qi Tian: Recent Advance in Content-Based Image Retrieval. A Literature Survey. arXiv. 19.06.2017. Version 2 from 02.09.2017. PDF. DOI: [10.48550/arXiv.1706.06064](https://doi.org/10.48550/arXiv.1706.06064)
- Mathias Zinnen / Prathmesh Madhu / Peter Bell / Andreas Maier / Vincent Christlein (2022a): Transfer Learning for Olfactory Object Detection. In: Ikki Ohmukai / Taizo Yamada (eds.): *DH 2022. Digital Humanities 2022. Responding to Asian Diversity. Conference Abstracts* (Online / Tokyo 25.–29.07.2022). Tokyo 2022, pp. 409–413. PDF. [\[online\]](#)
- Mathias Zinnen / Prathmesh Madhu / Ronak Kosti / Peter Bell / Andreas Maier / Vincent Christlein (2022b): ODOR. The ICPR 2022 Odeuropa Challenge on Olfactory Object Recognition. In: Michael Jenkin / Henrik I. Christensen / Cheng-Lin Liu (eds.): *ICPR 2022. 26th International Conference on Pattern Recognition. Proceedings* (Montreal, 21.–25.08.2022). Los Alamitos, US-CA etc. 2022, pp. 4989–4994. PDF. DOI: [10.1109/ICPR56361.2022.9956542](https://doi.org/10.1109/ICPR56361.2022.9956542)
- Mathias Zinnen / Azhar Hussian / Hang Tran / Prathmesh Madhu / Andreas Maier / Vincent Christlein: SniffyArt. The Dataset of Smelling Persons. In: Valerie Gouet-Brunet / Ronak Kosti / Li Weng (eds.): *SUMAC '23. Proceedings of the 5th Workshop on Analysis, Understanding and Promotion of Heritage Contents* (Ottawa, 02.11.2023). New York 2023, pp. 49–58. PDF. DOI: [10.1145/3607542.3617357](https://doi.org/10.1145/3607542.3617357)
- Mathias Zinnen / Azhar Hussian / Andreas Maier / Vincent Christlein: Recognizing Sensory Gestures in Historical Artworks. In: *Multimedia Tools and Applications. An International Journal* 84 (2025), pp. 39055–39083. PDF. DOI: [10.1007/s11042-024-20502-6](https://doi.org/10.1007/s11042-024-20502-6)
- Zhengia Zou / Keyan Chen / Zhenwei Shi / Yuhong Guo / Jieping Ye: Object Detection in 20 Years. A Survey. In: *Proceedings of the IEEE* 111 (2023), No. 3, pp. 257–276. PDF. DOI: [10.1109/JPROC.2023.3238524](https://doi.org/10.1109/JPROC.2023.3238524)
- Christoph Zuschlag: Einführung in die Provenienzforschung. Wie die Herkunft von Kulturgut entschlüsselt wird. Munich 2022. [\[GVK record\]](#)

Historical Auctions

- Auktionshaus Hugo Helbing (1927a). Auction on March 26, 1927, in Munich. German Sales. DOI: [10.11588/diglit.48878](https://doi.org/10.11588/diglit.48878)
- Auktionshaus Hugo Helbing (1927b). Auction on March 26, 1927, in Munich. Lot 106, Johann Sperl »Sommerlust«. German Sales. DOI: [10.11588/diglit.48878#0056](https://doi.org/10.11588/diglit.48878#0056)
- Auktionshaus Hugo Helbing (1927c). Auction on March 26, 1927, in Munich. Illustration Lot 106, Johann Sperl »Sommerlust«. German Sales. DOI: [10.11588/diglit.48878#0080](https://doi.org/10.11588/diglit.48878#0080)
- Auktionshaus Hugo Helbing (1931). Auction on April 14, 1931, in Munich. Lot 49, Ludwig von Zumbusch »Römische Ideallandschaft«. German Sales. DOI: [10.11588/diglit.49177#0023](https://doi.org/10.11588/diglit.49177#0023)
- Galerie Neupert. Auction on April 4, 1936, Zurich. Lot 40, Max Buri »Brienzersee«. German Sales. DOI: [10.11588/diglit.8670#0015](https://doi.org/10.11588/diglit.8670#0015)
- G. & L. Bollag, Zurich. Auction on October 23, 1931, in Zurich. Lot 57, Otto Frölicher »Barken«. German Sales. DOI: [10.11588/diglit.6826#0012](https://doi.org/10.11588/diglit.6826#0012)
- Internationales Kunst- und Auktionshaus (1933). Auction on August 1, 1933, in Berlin. Lot 433, Ludwig von Zumbusch »Landschaft mit Birken und Pappeln«. German Sales. DOI: [10.11588/diglit.6194#0025](https://doi.org/10.11588/diglit.6194#0025)

Kunsthau Lempertz (1928a). Auction on April 24, 1928, in Cologne. German Sales. DOI: [10.11588/diglit.17868](https://doi.org/10.11588/diglit.17868)

Kunsthau Lempertz (1928b). Auction on April 24, 1928, in Cologne. Lot 32, Johann Sperl »Kinder auf der Wiese«. German Sales. DOI: [10.11588/diglit.17868#0021](https://doi.org/10.11588/diglit.17868#0021)

Kunsthau Lempertz (1928c). Auction on April 24, 1928, in Cologne. Illustration Lot 32, Johann Sperl »Kinder auf der Wiese«. German Sales. DOI: [10.11588/diglit.17868#0059](https://doi.org/10.11588/diglit.17868#0059)

Kunstsalon Paul Cassirer (1925). Auction on March 3 and 4, 1925, in Berlin. German Sales. DOI: [10.11588/diglit.23253](https://doi.org/10.11588/diglit.23253)

Kunstsalon Paul Cassirer (1930a). Auction on November 14, 1930, in Berlin. German Sales. DOI: [10.11588/diglit.48920](https://doi.org/10.11588/diglit.48920)

Kunstsalon Paul Cassirer (1930b). Auction on November 14, 1930, in Berlin. Lot 88, Johann Sperl »Kind auf der Wiese«. German Sales. DOI: [10.11588/diglit.48920#0147](https://doi.org/10.11588/diglit.48920#0147)

Kunstsalon Dr. Störi (1928a). Auction on March 30 and 31, 1938, in Zurich. German Sales. DOI: [10.11588/diglit.24601](https://doi.org/10.11588/diglit.24601)

Kunstsalon Dr. Störi (1928b). Auction on March 30 and 31, 1938, in Zurich. Lot 26, Guillaumin Armand »Der Kran«. German Sales. DOI: [10.11588/diglit.24601#0010](https://doi.org/10.11588/diglit.24601#0010)

List of Figures and Tables

- Figure 1: A painting by Ludwig von Zumbusch was offered at three auctions, each time under a different title. [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: Ketterer Kunst 2019]
- Figure 2: *German Sales* offers different views of the data. Users can choose between the metadata, overview, facsimile, or OCR full text. [Screenshot: Mathias Zinnen / Sabine Lang 2024]
- Figure 3: Exemplary pages from different catalogs showing their heterogeneity regarding, for example, the number of images per page, the variety of offered lots, color of the pages and placement within the catalogs. [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]
- Figure 4: Word cloud visualizing the frequency of sales / object types found in the German Sales bibliographies. [Visualization: Mathias Zinnen / Sabine Lang 2024]
- Table 1: List of metadata extracted from the auction catalog bibliographies (cf. Bähr 2013; Bommert 2019).
- Figure 5: Complete pipeline of preprocessing steps applied to the auction catalogs prior to extracting features from the depicted objects. [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 6: Process employed to perform an image-based search for identical or similar images in auction catalogs. [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 7: Easy-to-use interface developed by using the open source software *Gradio*. [Screenshot: Mathias Zinnen / Sabine Lang 2024]
- Figure 8: Exemplary predictions from the image detection stage: most images are reliably detected. However, the algorithm has problems differentiating between different object types / elements as evidenced by detected carpets, sculptures, and ornamented text (bottom right). [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]
- Table 2: The table shows the metrics for our method where the extraction network was trained on the *ImageNet*, *Arts* or *Poses* datasets. While for the *top-1 accuracy* we achieve best results with the *Arts* model, the *ImageNet* model scores the best *top-5 accuracy* and *mAP*.
- Table 3: Evaluation of the models' performances when the query images are restricted to images of artworks with at least one person depicted. We see an increased performance in all models with the *Arts*-pre-trained models even achieving perfect retrieval metrics. Specifically the model trained for pose estimation performs considerably better and partly closes the gap compared to the two other pre-training schemes.
- Figure 9: These two artworks drastically lower the quantitative metrics: Adrian Ludwig Richter's *Hirten am Feuer*, (c. 1861) on the left and Heinrich Bürkel's *Der Kochensee mit den Häusern von Schlehdorf* (c. 1863 / 1867) on the right. [Graphic: Mathias Zinnen / Sabine Lang 2024, image sources: Ketterer Kunst 2018a; Ketterer Kunst 2018b]
- Figure 10: Crops of the target images of artworks as detected in the auction catalogs. Row (a) shows the images which could not be found by any of the extraction methods. The second row (b) shows detections of an artwork by Max Liebermann with a similarly bad image quality which were found (note also the difference in quality and color between the two identical images). [Graphic: Mathias Zinnen / Sabine Lang 2024, image source: German Sales 2025]
- Figure 11: Search results for Johann Sperl's *Sommerlust*: The work appeared in a catalog in (a) 1928, (b) 1930 and (c) 1927. [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 12: The first search results (excluding the first / identical) for Andreas Achenbach's *Fischerboot auf stürmischer See*: The painting was offered at auctions in (a) 1936, (b) 1931 and (c) 1928. [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 13: Search results for Liebermann's drawing *Einholung Bismarcks in Berlin* (1890). All results appear in an auction catalog published by the *Kunstsalon Paul Cassirer* in March 1925. [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 14: Search results using different embedding types; all models were trained on different data thus leading to different search results and allowing to focus on diverse similarity criteria (so our hypothesis). [Graphic: Mathias Zinnen / Sabine Lang 2024]
- Figure 15: Search results using different embedding types; the models *ImageNet* and *Arts* perform best for the task of finding identical images to Johann Sperl's *Sommerlust*. [Graphic: Mathias Zinnen / Sabine Lang 2024]