

Torsten Bendschus, Lara Mührenberg, Corinna Reinhardt,
Ronak Kost, Prathmesh Madhu, Ute Verstegen

Computer Vision und Deep Learning aus Perspektive der archäologischen Bildanalyse

→ Computer Vision, Deep Learning,
Christliche Archäologie,
Klassische Archäologie, Bildanalyse

Die Analyse und Interpretation historischer Bildwerke bildet einen Kernbestandteil des wissenschaftlichen Arbeitens archäologischer Disziplinen wie der Klassischen und der Christlichen Archäologie. Digitale Methoden der Computer Vision und des Deep Learning – beispielsweise im Bereich des automatisierten Erkennens von Objekten oder Körperhaltungen – können die bildanalytische Forschung in Zukunft maßgeblich unterstützen. Eine Herausforderung stellen jedoch die besonderen Phänomene und typischen Überlieferungsbedingungen von archäologischem Bildmaterial dar. Sie erschweren eine Übertragung von bereits entwickelten digitalen Modellen auf die Analyse dieser Artefakte, weil diese an andersartigen Bilddaten wie modernen Fotografien trainiert wurden.

Exemplarisch werden in diesem Beitrag vier zentrale Merkmale fachspezifischer Bilddaten und -datensets charakterisiert. Anschließend schlagen wir technische Problemlösungen vor, basierend auf Fallstudien aus einem laufenden Forschungsprojekt, das sich dezidiert mit der Entwicklung und Anwendung digitaler Methoden für die Analyse historischer Kunstwerke auseinandersetzt. Aus den testweise erzielten Resultaten werden Ansprüche an wie auch Potenziale für die weitergehende Forschung abgeleitet. Das methodische Vorgehen sowie mögliche Anwendungsgebiete aus archäologischer Perspektive werden in Zwischenfazits sowie abschließend kritisch reflektiert.

⇒ Zitierhinweis Early View
Torsten Bendschus, Lara Mührenberg,
Corinna Reinhardt, Ronak Kost, Prathmesh
Madhu, Ute Verstegen: Computer Vision
und Deep Learning aus Perspektive der
archäologischen Bildanalyse, in: Dieckmann
et al. (Hg.): 4D → Dimensionen | Disziplinen |
Digitalität | Daten, Heidelberg: arthistori-
cum.net, Advance online publication,
19.12.2022, [https://doi.org/10.11588/
arthistoricum.1100.c15427](https://doi.org/10.11588/arthistoricum.1100.c15427).

Einleitung

Das interdisziplinäre Forschungsprojekt **Iconographics. Computational Understanding of Iconography and Narration in Visual Cultural Heritage** widmet sich der Entwicklung und Anwendung neuartiger digitaler Methoden zur computergestützten Analyse historischer Bildwerke für die Kunst- und Geisteswissenschaften. ⁰¹ Den Schwerpunkt des seit April 2019 von der Emerging Fields Initiative der Friedrich-Alexander-Universität Erlangen-Nürnberg geförderten Projekts bildet die Erschließung komplexer Bildstrukturen und -zusammenhänge mit Hilfe von Computer Vision und Deep Learning. Dies geschieht auf Basis einer engen fächerübergreifenden Zusammenarbeit von (Digitaler) Kunstgeschichte, Klassischer Archäologie, Christlicher Archäologie und den Informatischen Wissenschaften. In gemeinsamer theoretischer und methodischer Reflexion wird auf Grundlage verschiedenartiger fachspezifischer Bildcorpora zu ikonografischen Fallbeispielen eine große Menge an Bildwerken analysiert und verglichen. Das Projekt versteht sich als exemplarische Grundlagenforschung, um die oft sehr komplexen kompositorischen, erzählerischen und semantischen Strukturen dieser Bilder für Computer Vision fassbar zu machen.

Die zugrundeliegenden Bildcorpora der beteiligten Fächer stellen durch ihre Verschiedenheit sehr unterschiedliche Anforderungen an die Entwicklung der Algorithmen und können jeweils als beispielhaft für eine Herangehensweise an historische Bildkulturen gelten. Hieraus ergeben sich auch für den Bereich der Computer Vision weitreichende Erkenntnisse für die Entwicklung neuartiger sowie für die Weiterentwicklung und Evaluierung bereits bestehender Methoden. Vor diesem Hintergrund behandelt der Beitrag einerseits dezidiert die speziellen Herausforderungen, die mit dem Bildmaterial der beteiligten archäologischen Disziplinen, der Klassischen Archäologie und der Christlichen Archäologie, für das Training von Modellen der digitalen Bilderkennung verknüpft sind. Aus diesem Grund sind die bild- und fachspezifischen Merkmale sowie die zugrunde liegende archäologische Artefaktbasis der Fallstudien eingangs zu charakterisieren. ⁰² Andererseits ist in der Analyse und Interpretation antiker und frühchristlicher bzw. byzantinischer Bildkulturen mit eigenen Darstellungskonventionen und Mitteln künstlerischen Ausdrucks zu rechnen. Zu diesem Zweck stehen mit Körperhaltungen und Objekten im Bild zwei Elemente im Fokus der Fallstudien, deren Bedeutung für die behandelten historischen Bildkulturen anschließend dargestellt wird. Bevor unterschiedliche Testsets relevanter Bildwerke im dritten Abschnitt des Beitrags dem exemplarischen Zugang zu den hier aufgeworfenen Fragestellungen dienen, wird die etablierte Methode der algorithmischen Ähnlichkeitsvisualisierung in UMAP-Plots exemplarisch erprobt, um die beschriebenen Herausforderungen eingangs an einem konkreten Anwendungsbeispiel deutlich zu machen. Mögliche technische Lösungen für die spezifischen Anforderungen, die mit den zuvor beschriebenen Merkmalen archäologischer Bildcorpora verbunden sind, werden in den folgenden Abschnitten sowohl vorgestellt als auch aus archäologisch-geisteswissenschaftlicher Perspektive reflektiert. Die Einzeluntersuchungen bilden die Basis, um abschließend den problembezogenen state-of-the-art zu resümieren und Potenziale sowie Grenzen zu beleuchten.

■ 01

Homepage: <https://www.izdigital.fau.de/forschung/efi-iconographics/>. An diesem transdisziplinären Forschungsprojekt sind Peter Bell und Dirk Suckow (Kunstgeschichte), Corinna Reinhardt und Torsten Bendschus (Klassische Archäologie), Ute Verstegen und Lara Mührenberg (Christliche Archäologie) sowie Andreas Meier, Vincent Christlein, Ronak Kostl und Prathmesh Madhu (Informatik) beteiligt.

■ 02

Um Missverständnisse hinsichtlich der Terminologie, die in den historischen Bildwissenschaften durchaus unterschiedlich genutzt wird, zu vermeiden, werden die Begriffe »Objekt« und »Material« in diesem Beitrag folgendermaßen genutzt: »Objekt« meint bildlich dargestellte Objekte (z. B. Schwert, Korb etc.), mit »Material« wird das Material bezeichnet, aus dem ein Artefakt besteht (z. B. Stein, Ton). Für archäologische Objekte im Allgemeinen wird der Begriff »Artefakt« verwendet, für archäologische Materialgattungen »Artefaktgattungen«.

Antike Bildträger und Bilder als Herausforderung für Computer Vision und Deep Learning

In den archäologischen Disziplinen wurde in der jüngeren Vergangenheit auf vielfältige Weisen erforscht, Methoden der Mustererkennung und generell der Computer Vision für das eigene Fachgebiet nutzbar zu machen. ⁰³ Dies geschieht beispielsweise in der Landschaftsarchäologie ⁰⁴, in der automatisierten Metadatenanreicherung großer Bildarchive ⁰⁵, für die numismatische Typenidentifikation und Stempelanalyse ⁰⁶, in der Kategorisierung und automatisierten Typenklassifikation von Keramikfragmenten ⁰⁷ oder in der computer-gestützten Analyse von Malerzuweisungen der griechischen Vasenmalerei ⁰⁸. Für Fragestellungen, die sich mit der Ikonografie und Narration historischer Bildwerke befassen, ist dies aber noch nicht der Fall. Aufgrund ihres repetitiven Erscheinens eignen sich aber auch die Einzelelemente antiker Bildwerke dafür, sie mittels Techniken der Computer Vision und des maschinellen Lernens miteinander zu vergleichen, um die bedeutenden ersten Schritte der archäologischen Bildanalyse auf vorikonografischer Ebene ⁰⁹ zu unterstützen. Dies umfasst zum einen die Komposition des Bildes, wie etwa die Anordnung von Figuren im Bildfeld und ihre Körperhaltungen, die helfen, Bezüge zwischen Figuren herzustellen und letztlich ihre Interaktion und Kommunikation miteinander deutlich machen. Zum anderen können die Techniken auch für das Erkennen von Objekten im Bild eingesetzt werden, die eine bestimmte Ikonografie kennzeichnen und somit beispielsweise bestimmte Akteur:innen identifizieren. ¹⁰ Beides spielt für die Bilderzählung eine entscheidende Rolle, weshalb im Rahmen des Projekts gezielt Bildcorpora zusammengestellt wurden, die aufgrund der Wiederholung signifikanter, auch kontext- und gattungsübergreifend auftretender Körperhaltungen von Protagonist:innen in ihrer Interaktion miteinander und der Wiederholung von Objekten Bildbezüge erwarten lassen, welche die antiken Betrachenden entschlüsseln und somit ein Bildverständnis ohne begleitende Erläuterungen entwickeln konnten.

■ 03

Rajdeep Chatterjee, Ankita Chatterjee, Rohit Halder, *Impact of Deep Learning on Arts and Archaeology: An Image Classification Point of View*, in: Manish Prateek et al. (Hg.), *Proceedings of International Conference on Machine Intelligence and Data Science Applications: Algorithms for Intelligent Systems, Singapore 2021*, https://doi.org/10.1007/978-981-33-4087-9_65. Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, Stuart James, *Machine Learning for Cultural Heritage: A Survey*, in: *Pattern Recognition Letters*, 133 2020, S. 102–108, <https://doi.org/10.1016/j.patrec.2020.02.017>. Simon H. Bickler, *Machine Learning Arrives in Archaeology*, in: *Advances in Archaeological Practice*, 9 (2) 2021, S. 186–191, <https://doi.org/10.1017/aap.2021.6>. Eine Übersicht für Anwendungsgebiete von Computer Vision in der Kunstgeschichte geben: Peter Bell, Björn Ommer, *Computer Vision und Kunstgeschichte: Dialog zweier Bildwissenschaften*, in: Piotr Kuroczynski et al. (Hg.), *Computing Art Reader: Einführung in die digitale Kunstgeschichte, Heidelberg 2018 (Computing in Art and Architecture, Band 1)*, S. 61–75, <https://doi.org/10.11588/arthistoricum.413>.

■ 04

Exemplarisch: Wouter Baernd Verschoof-van der Vaart, Karsten Lambers, *Learning to Look at LiDAR: The Use of R-CNN in the Automated Detection of Archaeological Objects in LiDAR Data from the Netherlands*, in: *Journal of Computer Applications in Archaeology*, 2 (1) 2019, S. 31–40, <http://doi.org/10.5334/jcaa.32>. Iban Berganzo-Besga, Hector A. Orengo, Felipe Lumbreras, Miguel Carrero-Pazos, João Fonte, Benito Vilas-Estévez, *Hybrid MSRM-Based Deep Learning and Multitemporal Sentinel 2-Based Machine Learning Algorithm Detects Near 10k Archaeological Tumuli in North-Western Iberia*, in: *Remote Sense*, 13 (20) 2021, 4181, <https://doi.org/10.3390/rs13204181>.

Verschiedene Fallstudien ermöglichen es, im Rahmen des Projekts zum einen fachspezifische Fragestellungen mit Hilfe der Computer Vision zu untersuchen, zum anderen aber auch, sich exemplarisch an die themenübergreifenden Herausforderungen archäologischen Bildmaterials für den Einsatz von Methoden der Computer Vision anzunähern. Hierfür wurden geeignete Bildcorpora zusammengestellt. Im Fall der Klassischen Archäologie handelt es sich um ein Corpus von circa 2.000 Bildern auf bemalten griechischen Keramikgefäßen, sogenannte Vasenbilder. Sie sind im sechsten und fünften Jahrhundert vor Christus gefertigt worden und im sogenannten schwarz- oder rotfigurigen Stil bemalt. Das Corpus beinhaltet Szenen der Frauenführung, der Verfolgung, des Raubs sowie des (mythologischen und agonalen) Ringkampfes und überliefert damit Bildschemata, für welche die körperliche Interaktion der Protagonist:innen zentral ist. Zum Training der Algorithmen wurden zusätzlich annähernd 12.000 weitere Bilder griechischer Vasenmalerei ausgewählt, die andere Themen zeigen. Ihre Zusammenstellung basiert auf Datenbankabfragen nach verschiedenen und in unterschiedlichen Ikonografien der antiken Kunst signifikanten Figuren und Objekten (z. B. Poseidon, Dreizack oder Hermes).

Das ikonografische Bildcorpus der Christlichen Archäologie umfasst ca. 1.000 Darstellungen der Verkündigung an Maria sowie der Magieranbetung, die anders als das Corpus der Klassischen Archäologie aus sehr unterschiedlichen Bildgattungen (z. B. Mosaik, Malerei, Relief, Textilien oder Toreutik) und von Bildträgern stark variierender Formate (vom Apsismosaik bis zum Ringstein) stammen. Sie decken ein chronologisches Spektrum vom dritten bis zum zehnten Jahrhundert nach Christus (im Westen) bzw. 15. Jahrhundert nach Christus (im Osten) ab. Da zudem das Bildcorpus im Teilprojekt der Kunstgeschichte dieselben Themen für die westliche Kunst berücksichtigt, ergänzen sich diese beiden Corpora. Weitere circa 7.000 Bilder wurden zu Trainingszwecken zusammengestellt. Zusätzlich zu einem großen Anteil von Bildwerken aus der römischen Kunst sind hier neben spätrömischen und byzantinischen auch beispielsweise armenische und koptische sowie islamische Darstellungen aufgenommen worden, so dass eine hohe kulturelle Heterogenität vorliegt.

Die fachspezifischen Bildcorpora stellen zwar nur jeweils kleine Ausschnitte des Gesamtbildbestands dar, ermöglichen es aber, einige Merkmale archäologischen (Bild-)Materials abzuleiten. Diese stellen die Computer Vision vor besondere Herausforderungen, für welche anhand von Fallbeispielen im Folgenden Lösungsvorschläge vorgestellt werden.

■ 05

Claudia Engel, Peter Mangiafico, Justine Issavi, Dominik Lukas, Computer Vision and Image Recognition in Archaeology, in: Proceedings of the Conference on Artificial Intelligence for Data Discovery and Reuse, New York 2019, <https://doi.org/10.1145/3359115.3359117>.

■ 06

Exemplarisch: Sebastian Zambanini, Martin Kappel, Mario Schlapke, On the Use of Computer Vision for Numismatic Research, in: Michael Ashley et al. (Hg.), VAST: The 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage, Braga 2008, <https://cvl.tuwien.ac.at/wp-content/uploads/2014/12/vast08.pdf>. Jessica Cooper, Ognjen Arandjelović, Learning to Describe: A New Approach to Computer Vision Based Ancient Coin Analysis, in: Sci 27 (2) 2020, <https://www.mdpi.com/2413-4155/2/2/27/htm>. Antike Münzen eignen sich nicht nur aufgrund ihrer massenhaften seriellen Produktion sehr gut für Anwendungen des maschinellen Lernens, hervorzuheben ist auch ihre vorbildliche Dokumentation nach hohen einheitlichen Qualitätsstandards, wie sie im Rahmen des Forschungs- und Digitalisierungsverbunds Netzwerk universitärer Münzsammlungen in Deutschland (NUMiD: <http://numid-verbund.de>) erfolgte. Seit Februar 2021 wird das Forschungsprojekt Data quality for Numismatics based on Natural language processing and Neural Networks von der Deutschen Forschungsgemeinschaft (DFG) gefördert, siehe <https://gepris.dfg.de/gepris/projekt/449148027>.

■ 07

Diesem Forschungsbereich widmet sich beispielsweise aktuell das an der Universität di Pisa beheimatete internationale Forschungsprojekt ArchAIDE, siehe <http://www.archaide.eu/>.

Small Data-Charakter der Datenbasis

Das erfolgreiche Training künstlicher Intelligenz erfordert sehr große Datenmengen und basiert daher in der Regel auf sehr umfangreichen Datensets moderner Fotografien. Über derartige Bilddatensets verfügen historische Kulturwissenschaften nicht. Für die antike Bilderwelt stellt etwa die Beazley Archive Pottery Database des Classical Art Research Centre der Universität Oxford mit annähernd 250.000 Bildern von griechischen Gefäßen und Vasenmalereien größtenteils des sechsten bis vierten Jahrhunderts vor Christus in ihrem Umfang bereits eine Ausnahme dar. ¹¹ Für andere Gattungen ist der Bildbestand der bestehenden Datenbanken in aller Regel deutlich geringer, was vor allem an den jeweils zum Teil geringen Objektmengen, aber auch an der modernen Dokumentation liegt.

■ 08

Dies wird aktuell im Projekt EGRAPHSEN. Möglichkeiten und Perspektiven der digitalen Malerzuweisung bei attischen Vasen vom Institut für Digital Humanities der Georg-August-Universität Göttingen in Zusammenarbeit mit dem Information Systems and Machine Learning Lab der Universität Hildesheim erforscht, siehe <https://www.uni-goettingen.de/de/598165.html>.

■ 09

Dieser Ausdruck referenziert das vom Kunsthistoriker Erwin Panofsky entwickelte Drei-Stufen-Interpretationsschema zur Deutung von Kunstwerken. Bei den drei Stufen handelt es sich um eine vorikonografische Beschreibung, eine ikonografische Analyse und eine ikonologische Interpretation, siehe Erwin Panofsky, Ikonographie und Ikonologie, in: Ekkehard Kaemmerling (Hg.), Bildende Kunst als Zeichensystem: Ikonographie und Ikonologie. Band 1: Theorien – Entwicklung – Probleme, Köln 1994, S. 207–225.

■ 10

Crowley und Zisserman stellten bereits 2013 einen Bild und Text verbindenden Ansatz vor, wie ausgehend von den in den Metadaten vorhandenen kurzen Beschreibungen wie »sitzender Zeus« visuell übereinstimmende Bereiche in den zugeordneten Bildern detektiert und daraus automatisierte Bildannotationen erzeugt werden können: Elliot J. Crowley, Andrew Zisserman, Of Gods and Goats: Weakly Supervised Learning of Figurative Art, in: British Machine Vision Conference, 2013, S. 1–11, <https://www.robots.ox.ac.uk/~vgg/publications/2013/Crowley13/crowley13.pdf>.

■ 11

<https://www.beazley.ox.ac.uk/carc/pottery/>. Für Verweise auf Datenbankeinträge in der Beazley Archive Pottery Database wird im Folgenden die Abkürzung BAPD mit Angabe der jeweiligen Vasennummer verwendet. Unter diesen Nummern sind auch alle aktuellen Objektdaten wie Aufbewahrungsorte und Inventarnummern zu finden.

Vielfalt der Bilder und Bildträger

Die Darstellungsweisen in historischen Zeugnissen visueller Kultur stellen Computer Vision-Methoden vor große Schwierigkeiten, auch wenn jene bereits erfolgreich aktuelle digitale Fotopools analysieren. Beispielsweise können historische Figurendarstellungen mitunter deutlich von anatomischen Gesetzmäßigkeiten abweichen. Gesichter sind bisweilen stark abstrahiert, Körper unnatürlich verdreht oder Gliedmaßen unproportional dargestellt. Der reiche Faltenwurf der Gewänder kann es erschweren, die eigentliche Körperhaltung der Figuren zu erkennen. Auch die künstlerische Wiedergabe von Objekten unterliegt gelegentlich bestimmten Darstellungskonventionen, die nicht den realen Vorbildern entsprechen müssen. Neben historischen Entwicklungen in den Darstellungsweisen sind auch in Hinblick auf die Gattungen große Unterschiede möglich, da diese jeweils eigene Bildfindungen überliefern, beispielsweise wenn dieselbe Figur oder Szene in einen kleinen Ringstein graviert, als Mosaik aus kleinen Steinen zusammengesetzt, oder als Sarkophagrelief ausgearbeitet ist.

Fragmentarität der Artefakte

Archäologisches Material ist oftmals nur fragmentiert überliefert. Für historische Bildwerke ist daher mit Beschädigungen, Fehlstellen oder gar nur mit Bruchstücken des Originals zu rechnen. Im Vergleich zu anderen kunstgeschichtlichen Disziplinen bilden derartige unvollständig erhaltene Bildwerke in den Archäologien einen nicht zu unterschätzenden Teil des Gesamtbestands. Aufgrund ihrer Kenntnis oft wiederholter Darstellungskonventionen ist es Expert:innen dennoch häufig möglich, nicht erhaltene Teile eines Bildes mit hoher Plausibilität zu rekonstruieren. Es ist zu prüfen, inwieweit auch Computer-Vision-Methoden imstande sind, auf Grundlage der verfügbaren Bildmengen vergleichbare Prognosen zu erstellen.

Dreidimensionalität der Artefakte

Digitale Bilddaten historischer Bildwerke liegen fast ausschließlich nur als zweidimensionale Abbilder vor, obwohl historische Bildträger häufig dreidimensionale Artefakte sind, die von ihren Betrachter:innen eine entsprechende Wahrnehmung fordern. Anschauliche Beispiele hierfür finden sich in der antiken Vasenmalerei, deren Bildträger sich durch die Krümmung des Gefäßkörpers auszeichnet. Ein Spiel mit der (Nicht-)Sichtbarkeit von Bildelementen durch das Drehen des Gefäßes oder durch eine Veränderung der eigenen Position oder Körperhaltung der Betrachtenden kann hierbei durchaus vom Vasenmaler beabsichtigt gewesen sein. Übliche fotografische Aufnahmen (oder auch Umzeichnungen/Abrollungen) können solche Bilder nicht unmittelbar als dreidimensionales Werk wiedergeben. Dies führt zu Problemen der digitalen Bilderkennung, wenn diese mit den Verzerrungen der Fotografien umzugehen hat, und wenn mit dieser Grundlage Aspekte des dreidimensionalen Bildes in der Bildanalyse Berücksichtigung finden sollen.

Körper und Objekte im Bild: Archäologische Fallbeispiele

Die methodische Herangehensweise an die beschriebenen Problemfelder wird anhand der digitalen Erkennung zweier Bildelemente exemplifiziert, deren systematische Analyse für die vorikonografischen Grundstrukturen der Bilder sehr bedeutsam ist: Körperhaltungen und Objekte. Die Annäherung an diese Elemente erfolgt mittels Fallstudien ausgewählter Ikonografien bzw. Bildschemata, da diese einerseits die Relevanz der Analyseaspekte konkret fassbar machen, andererseits dabei helfen, die spezifischen Anforderungen an die digitalen Methoden abzuleiten. Im Folgenden seien daher zwei Fallbeispiele aus der Projektarbeit kurz vorgestellt, um die fachspezifischen Fragestellungen und Herangehensweisen zu zeigen und in den folgenden Abschnitten die Probleme und Potenziale der digitalen Bildanalyse deutlich zu machen.

Körperhaltungen werden als Bildelemente begriffen, die auf genereller Ebene Interaktion anzeigen und somit alle Figuren sowie Objekte im Bild in einer Kausalbeziehung verbinden können. Körperhaltungen, inklusive Finger- und Handgesten, Blickrichtungen oder anderer Elemente der Körperdarstellung, geben den Betrachtenden bedeutende Anhaltspunkte, um unter anderem das Verhältnis mehrerer Figuren zueinander, ihre Handlungen und Reaktionen, aber auch die emotionale Verfasstheit einer Figur oder ihr Verhältnis zu nicht sichtbaren Adressat:innen (z. B. einer Gottheit) zu charakterisieren. Bei morphologischer Ähnlichkeit können Ausformungen von Körpersprache je nach Gattung, Epoche oder beispielsweise kultureller, religiöser oder sozialer Kontextualisierung zwar teils als feste Zeichen gelesen werden, doch ist eine solche semantische Konnotation nur eine Möglichkeit der Bildfunktion, weil sie auch in jedem Bild neu verhandelbar und damit dynamisch für die Bilderzählung nutzbar sind. **12**

■ 12

Zu »schemata« in der antiken griechischen Kunst siehe Maria L. Catoni, *La comunicazione non verbale nella Grecia antica: Gli schemata nella danza, nell'arte, nella vita*, Turin 2008.



□ 01

Frauenführungsszene, flankiert von einem Krieger (links) und einem Jüngling (rechts), auf einer schwarzfigurigen Halsamphora des Antimenes-Malers, um 520 v. Chr. (Foto: © The Trustees of the British Museum, Inv. 1836.2-24.10 [B 244])

Ein Beispiel derartiger Bildschemata bilden sogenannte Frauenführungs-szenen, die sich grundlegend durch das Führen oder Geleiten einer weiblichen durch eine männliche Figur auszeichnen. Merkmal einer Variante solcher Szenen ist der Griff des Mannes an den Rücken, an das Gewand ⁰¹ oder an die Hand oder das Handgelenk der Frau, während andere Bilder mit ähnlichen Körperhaltungen das Führen ohne Körperkontakt thematisieren. An dieser Stelle sei insbesondere die letztgenannte Variante, genannt *cheir' epi karpo* (zu Deutsch: Hand auf Handgelenk), herausgehoben, um die ikonografische Fixierung einer spezifischen Interaktionsgeste zwischen Mann und Frau und ihre Verwendung in unterschiedlichen Bildkontexten kurz vorzustellen. Für die jeweiligen Bilderzählungen der Frauenführungen sind jedoch zudem die bezeugten unterschiedlichen Interaktionen innerhalb dieses Bildschemas wichtig, da sie das Verhältnis zwischen den Protagonist:innen differenziert beschreiben.

Eines der vereinzelt sehr frühen Beispiele für den Griff ans Handgelenk findet sich bereits auf einem griechischen Luterion (Wasserbecken) aus dem achten Jahrhundert vor Christus ¹³, das womöglich die Entführung der Helena durch Paris darstellt. Aus dem sechsten und frühen fünften Jahrhundert vor Christus ist eine Vielzahl von sogenannten schwarzfigurigen Vasenbildern bekannt, die zentral ein Paar aus einer in ein dichtes Gewand (*peplos*) mit Mantel (*himation*) als Schleier gehüllten Frau in einem Entschleierungsgestus und einem sie am Handgelenk führenden Krieger in der Vollrüstung eines griechischen Hopliten zeigen. ¹⁴ Dieser in der modernen Forschung sogenannte *anakálypses*-Gestus der weiblichen Protagonistin beinhaltet einen in den Schleier gehüllten und parallel zur Gewandgürtung angewinkelten sowie einen auf Höhe der Schulter den Rand des Schleiers greifenden Arm. Der Name bezieht sich auf das Entschleiern der Braut im Hochzeitsritus (*anakalyptéria* ¹⁵), allerdings findet sich eine direkte Verbindung zwischen dem Begriff *anakálypses* und einer Hochzeitsdarstellung nur in der sehr viel späteren Beschreibung einer Malerei der Hochzeit von Pelops und Hippodameia beim griechischen Autor Flavius Philostratos aus dem frühen dritten Jahrhundert nach Christus. ¹⁶ Die Verbindung zum dargestellten Gestus stellte erst die moderne Forschung her. ¹⁷ Andere Interpretationen sehen darin u. a. ein Erkennungszeichen für weibliche sexuelle Reife ¹⁸ oder einen generellen Grußgestus ¹⁹. Dies eröffnet mehr Deutungsmöglichkeiten der Bilder, unter anderem auch diejenige, dass es sich hierbei nicht um konkrete mythologische Geschichten handelt, sondern um Bilder, in denen es um soziale Rollen und typische Episoden aus der antiken Lebenswelt geht. Auch in diesen wird anhand der unterschiedlichen Varianten mit unterschiedlichen Schwerpunkten erzählt – wenn etwa mit unterschiedlichen Handhaltungen oder ohne Körperkontakt mehr oder weniger Fokus auf die männliche, kriegsrische Dominanz gelegt wird.

■ 13

London, British Museum, Inv.-Nr. 1899.2-19.1. Zu dem Stück siehe u. a. Karl Scheffold, *Götter- und Heldensagen der Griechen in der früh- und hocharchaischen Kunst*, München 1993, S. 130.

■ 14

Beispiele: BAPD 310185, 320411 u. 20535.

■ 15

John H. Oakley, Rebecca H. Sinos, *The Wedding in Ancient Athens*, Madison, Wis. 1993, S. 30. Carola Reinsberg, *Ehe, Hetärentum und Knabenliebe im antiken Griechenland*, München 1993, S. 57–59. Amy C. Smith, *The Politics of Weddings at Athens: An Iconographic Assessment*, in: *Leeds International Classical Studies*, 4 (1) 2005, S. 1–32.

■ 16

Philostr. *imag.* 1, 17, 3.

■ 17

Eva C. Keuls, *The Reign of the Phallus: Sexual Politics in Ancient Athens*, New York 1995, S. 106–107. Ellen D. Reeder (Hg.), *Pandora: Women in Classical Greece*, Baltimore 1995, S. 339 m. Nr. 113. Viktoria Sabetai, *Women's Ritual Roles in the Cycle of Life*, in: Nikolaos Kaltsas et al. (Hg.), *Worshipping Women: Ritual and Reality in Classical Athens*, New York 2008, S. 289–297, insb. S. 296.

■ 18

Oakley, Sinos 1993, S. 30.

■ 19

Hedwig Kenner, *Weinen und Lachen in der griechischen Kunst*, Wien 1960, S. 17.

Insbesondere die Zusammensetzung der Beifiguren aus Personen der eigenen Familienstruktur (Jünglinge, Frauen, alte Männer) wie auch von Mitgliedern der Kampfgemeinschaft (Speerträger, Bogenschützen, andere Hopliten) sowie das Führen einer Frau in ihrer spezifischen in sich gekehrten Körperhaltung und auf das Gewand bezogenen Gestik legen nahe, dass diese Szenen Erwartungshaltungen an das altgriechische männliche Rollenbild und zentrale Aspekte seines Lebens (Eheschließung/Familie und kriegerische Sieghaftigkeit) verbildlichen sollen. Demgegenüber identifizieren signifikante Bildelemente wie Figurenanordnung oder Objekte eine Reihe gleichzeitiger Frauenführungsszenen als konkrete mythologische Narrative. Hierbei handelt es sich vor allem um Episoden aus dem trojanischen Sagenkreis, z. B. als Neoptolemos die trojanische Prinzessin Polyxene als Opfer zum Grab seines Vaters Achill führt und mit dem Griff ans Handgelenk die absolute Macht des Kriegers über die Frau gezeigt wird 02. 20

■ 20
BAPD 302032.



□ 02
Die Führung der trojanischen Prinzessin Polyxene durch Neoptolemos zum Grab seines Vaters Achill (im Bild rechts) auf einer schwarzfigurigen Hydria der Leagros-Gruppe, circa 525–500 v. Chr. (Foto: © Antikensammlung der Staatlichen Museen zu Berlin – Preußischer Kulturbesitz, Inv. F 1902, Fotograf: Johannes Laurentius)

In der späteren, nun rotfigurigen Malerei (zum Stilwechsel siehe unten, Abschnitt **Vielfalt der Bilder und Bildträger**) treten derartige Bilder mit unbenannten Krieger-Frau-Paaren kaum noch auf **21**, doch begegnet uns das Bildschema des Griffs ans Handgelenk in verschiedenen anderen Kontexten. Dies sind einerseits erneut Darstellungen mythologischer Paare, z. B. die Entführung der Helena durch Paris (siehe **03**, linkes Paar) **22**, ihre Heimführung aus Troja durch Menelaos **23**, Peleus mit Thetis **24** oder Agamemnon mit Briseis **25**.

■ 21

Beispiele: BAPD 16776, 204102, 203039 u. 202459.

■ 22

BAPD 204685.

■ 23

Auf Exemplaren in Boston (BAPD 201000) und Odessa (BAPD 200577) identifizieren Namensbeischriften das in diesem Schema dargestellte Paar als Menelaos und Helena.

■ 24

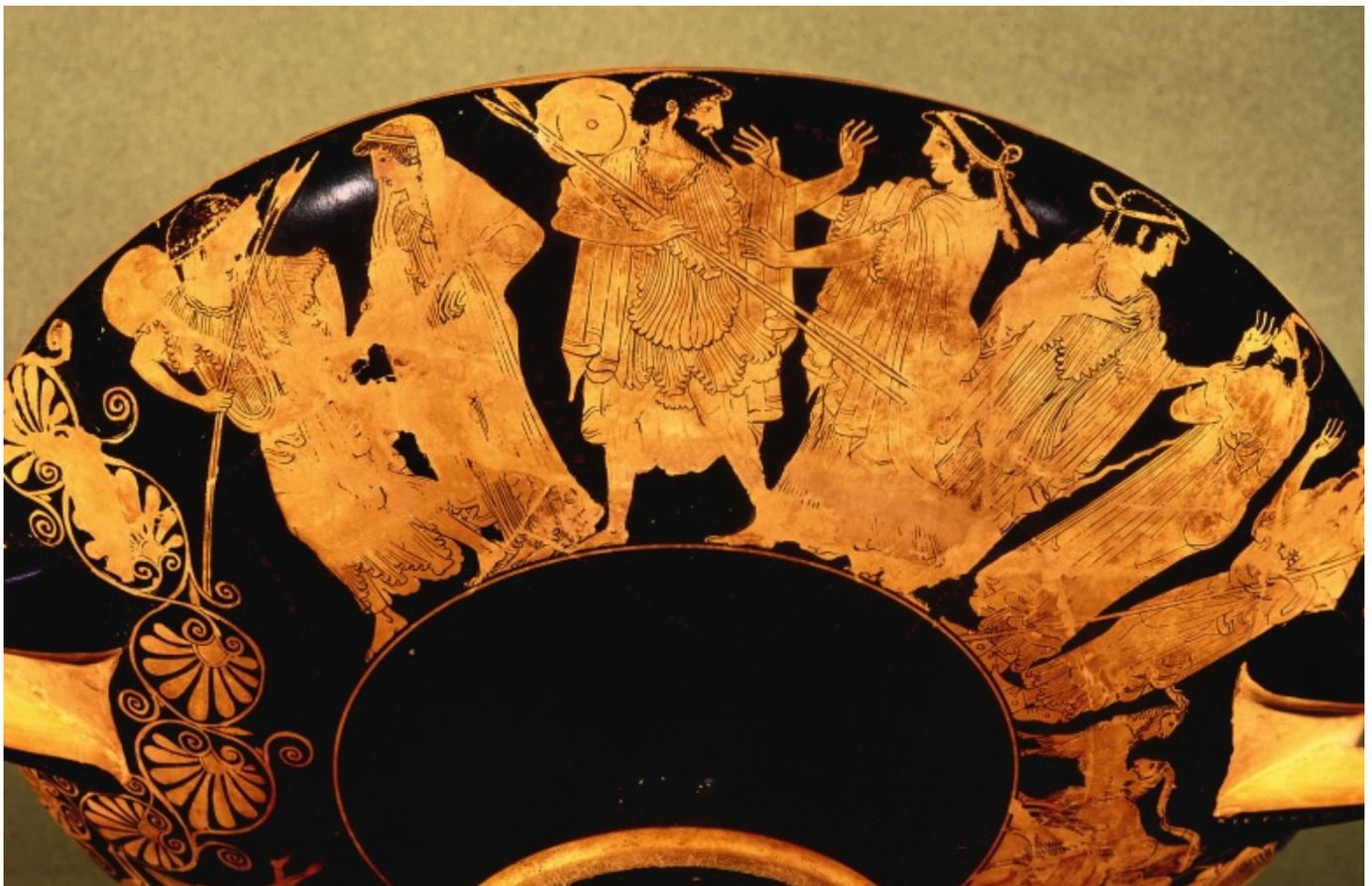
BAPD 200081 u. 202518.

■ 25

BAPD 204395. Eine Vase in Cerveteri (BAPD 13363) zeigt indes die Führung der Briseis durch Patroklos.

□ 03

Die Führung der Helena durch Paris (im Bild links) auf einer rotfigurigen Trinkschale des Makron, um 490/80 v. Chr. (Foto: © Antikensammlung der Staatlichen Museen zu Berlin – Preußischer Kulturbesitz, Inv. F 2291, Fotograf: Johannes Laurentius)



■ 26

Reinsberg 1993, S. 59–61. Oakley, Sinos 1993, S. 26–28. Elke Hartmann, Heirat, Hetärentum und Konkubinat im klassischen Athen, Frankfurt – New York 2002, S. 85–86.

■ 27

Beispiele: BAPD 211247, 214470 u. 213434.

Andererseits wird das Führen am Handgelenk zu einem festen Motiv in nicht-mythischen Hochzeitsprozessionen, deren regulärer Bestandteil die Brautführung zum Haus des Bräutigams war 04. 26 Die Bilder zeigen den Bräutigam, der seiner Braut voranschreitet, seinen Blick auf sie zurück wendet und sie am Handgelenk fortführt. Das Brautpaar wird mitunter von einer großen Anzahl an Fackelträgerinnen, Familienangehörigen und Musikern flankiert, von Gottheiten und/oder einem fliegenden, die Eheleute bekränzenden Eros begleitet und von Hochzeitsbehältnissen, Kränzen und Architektur ergänzt. 27



□ 04

Brautführung auf einer rotfigurigen Loutrophore aus Athen, um 425 v. Chr. (Foto: © The Trustees of the British Museum, Inv. 1910,0711.1)

Obgleich die Präsenz des Eros mitunter als Verbildlichung der romantischen Beziehung des Paares gelesen wird ²⁸, sehen die meisten Interpretationen im »cheir' epi karpō«-Gestus eine »herrschende Geste« ²⁹, den »rituelle[n] Ausdruck der Besitzergreifung« ³⁰, »etwas Herrisches und Gewalttätiges« ³¹ oder einen »act of violence« im Rahmen einer rituellen Brautentführung ³². Vor dem Hintergrund der Bildrelationen zu den genannten früheren Vasenbildern mit den Krieger-Frau-Paaren und ihrer variantenreicheren Interaktionsdynamik erscheint diese Konnotation mit Gewalttätigkeit und männlicher Dominanz plausibel. Dieses Beispiel zeigt zum einen, wie wichtig es ist, unterschiedliche Interaktionen anhand der variablen Körperhaltungen innerhalb eines generellen Schemas zu detektieren, um die jeweilige Bilderzählung zu verstehen. Es zeigt zum anderen aber auch, dass bildkontext- und bildstilübergreifend ein Aspekt wie der Griff ans Handgelenk wiedererkannt werden muss, um die Semantik von Bildbezügen verstehen zu können, die dem Horizont der zeitgenössischen Sehgewohnheiten und Darstellungskonventionen entspricht.

Auch **Objekte** können stark in Bildhandlungen involviert sein, z. B. als Ausgangspunkt oder Ziel von Aktionen der Figuren, als Verlängerung ihrer Körpersprache (z. B. ein Zeigestock) oder als Ursprung sensorierlicher Stimuli, die Reaktionen der Akteur:innen hervorrufen (z. B. Räuchergefäße oder Lebensmittel). Sie zählen somit ebenso zu den Bildelementen, deren komplexe Analyse grundlegend für das Bildverständnis ist. Während Elemente wie Körperhaltungen die Figuren einer Bildhandlung in eine Beziehung zueinander setzen und spezifische Situationen charakterisieren, erlauben es häufig erst Objekte (z. B. Götterattribute), die dargestellte Handlung als beispielsweise ein bestimmtes mythologisches Narrativ zu identifizieren. Mitunter werden mythologische bzw. biblische oder heilige Personen erst durch die ihnen beigegebenen Objekte erkennbar. Derartige Objekte sind daher Grundvoraussetzung für das sichere Identifizieren einer Ikonografie, für die auch das quantitative Vorkommen der Objekte, ihre Kombinationen und ihre bildräumliche Positionierung konstituierend sein können. Daraus ergeben sich Thesen zur bildimmanenten Funktion von Objekten für die Narration, aber auch zu Wandlungsprozessen und Konjunkturen kulturräumlichen Ausprägungen.

■ 28

Robert F. Sutton, *Pornography and Persuasion on Attic Pottery*, in: Amy Richlin (Hg.), *Pornography and Representation in Greece and Rome*, New York – Oxford 1992, S. 1–33, insb. S. 26–27.

■ 29

Susanne Moraw, *Bilder, die lügen: Hochzeit, Tieropfer und Sklaverei in der klassischen Kunst*, in: Günter Fischer et al. (Hg.), *Die andere Seite der Klassik: Gewalt im 5. und 4. Jahrhundert v. Chr.*, Stuttgart 2005, S. 73–88 (Zitat S. 83).

■ 30

François Lissarrague, *Frauenbilder*, in: Pauline Schmitt Pantel (Hg.), *Geschichte der Frauen. Bd. 1: Antike*, Frankfurt – New York 2003, S. 177–254 (Zitat S. 185).

■ 31

Gerhard Neumann, *Gesten und Gebärden in der griechischen Kunst*, Berlin 1965, S. 59.

■ 32

Anthi Dipla, Dimitris Paleothodoros, *Selected for the Dead: Erotic Themes on Grave Vases from Attic Cemeteries*, in: Ing-Marie Back Danielsson et al. (Hg.), *Encountering Imagery: Materialities, Perceptions, Relations*, Stockholm 2012, S. 209–233 (Zitat S. 216).

Die tragende Rolle, die Objekte bei der bildlichen Umsetzung von Geschichten spielen, die auch durch schriftliche Quellen überliefert sind, zeigen die beiden von der Christlichen Archäologie zuvorderst in den Blick genommenen Ikonografien – die Verkündigung an Maria und die Anbetung der Magier. ³³ Diese sind Szenen aus dem Marienleben bzw. der Kindheit Jesu und haben beide ihren Ursprung in Texten des Neuen Testaments und seines Umfelds. Darüber hinaus setzen die Bilder eigene Schwerpunkte in der Narration und gehen inhaltlich über die Textquellen hinaus.

Das Lukasevangelium berichtet davon, dass der Engel Gabriel von Gott zu der Jungfrau Maria geschickt wird, um ihr zu verkündigen, dass der Heilige Geist über sie kommen und sie ein Kind empfangen werde. ³⁴ Maria befindet sich im Moment des Besuchs im Inneren des Hauses in Nazareth. Nachdem Gabriel Maria den Auftrag Gottes geschildert und Maria ihre Bereitwilligkeit signalisiert hat, verlässt der Bote den Ort bereits wieder. Während die anderen kanonischen Evangelien nichts über diese Begebenheit berichten, sind weitere außerbiblische (sogenannte apokryphe) Texttraditionen bekannt, die die Verkündigung an Maria thematisieren, insbesondere der Bericht des Protevangeliums nach Jakobus. ³⁵ Da Maria hier die Protagonistin ist, erfährt der:die Leser:in mehr über den Kontext: Maria ist zum Wasserschöpfen zum Brunnen gegangen und hört dort zum ersten Mal die Stimme des Engels, der sich ihr aber noch nicht zeigt. Daraufhin geht sie zurück ins Haus, setzt sich und beginnt mit dem Spinnen des purpurnen Garns für den Vorhang des Jerusalemer Tempels, zu dem sie durch das Los bestimmt worden ist. In diesem Moment erscheint ihr der Verkündigungengel. Beide Verkündigungstraditionen – im Haus und an der Quelle – bilden die Grundlagen für die sich entwickelnden Ikonografien dieser Szene in den bildlichen Darstellungen, in denen die Details der Berichte zusammengeführt werden.

■ 33

Die Veröffentlichungen zu den bildlichen Darstellungen dieser Szenen und Maria, der Protagonistin, im Besonderen, sind zahlreich, vgl. zuletzt: Robin M. Jensen, *The Apocryphal Mary in Early Christian Art*, in: Andrew Gregory et. al. (Hg.), *The Oxford Handbook of Early Christian Apocrypha*, Oxford 2015, S. 289–305. Leslie Brubaker, Mary B. Cunningham (Hg.), *The Cult of the Mother of God in Byzantium: Texts and Images*, London u. a. 2016. Catherine Gines Taylor, *Late Antique Images of the Virgin Annunciate Spinning: Allotting the Scarlet and the Purple*, Leiden 2018. Thomas Arentzen, Mary B. Cunningham (Hg.), *The Reception of the Virgin in Byzantium: Marian Narratives in Texts and Images*, Cambridge 2019. François Boespflug, *L'Annunciazione a Maria nell'arte d'Oriente e d'Occidente*, Mailand 2020.

■ 34

Lk 1, 26–38.

■ 35

Protev. Jac. 11. Ev. PsMt. folgt diesem Bericht. Vgl. auch im Koran, Sure 19, 16–21. Zur auf apokryphen Quellen basierenden Marienikonographie: David R. Cartlidge, J. Keith Elliott, *Art and the Christian Apocrypha*, London – New York 2001, S. 21–26. Jensen 2015.

Als früheste Darstellung der Verkündigung wird in der Forschung traditionell eine Szene aus der Priscillakatakomben in Rom angesprochen. ³⁶ Ob es sich hierbei tatsächlich um eine Darstellung der Verkündigung handelt, ist jedoch nicht mit Sicherheit zu sagen. Die Verkündigung im Haus ist zwar die weitaus häufigere Darstellung, wird aber je nach Gattung, Region und Epoche mit anderen Akzenten versehen. Spätestens seit Maria auf dem Konzil von Ephesos 431 nach Christus der Titel der Theotokos (Gottesgebärende) verliehen worden ist, spielen z. B. imperialisierende Motive in die Darstellungen hinein: Die Gottesmutter wird nun thronend gezeigt, und die hierarchische Stellung zwischen ihr und dem Boten verhandelt ³⁷. ³⁷ Die Hauptbestandteile der Szene, die immer wieder – aber nicht notwendigerweise – vorkommen, sind: Maria und Gabriel/ein Engel, ein Sitzmöbel, ein Korb, die Handlung des Spinnens oder Spinnengeräte und Wolle. Gabriel ist an den Flügeln als Engel zu erkennen; ³⁸ Maria hingegen ist nicht besonders ausgezeichnet. Es ist beispielsweise nicht durchgängig ein Nimbus oder eine bestimmte Farbe ihres Gewands gegeben. Bestimmend ist aber einerseits die Nähe zu einem Sitzmöbel bzw. ihr Sitzen, wobei die Art des Sitzmöbels stark divergiert: Maria kann auf einem Korbsessel platziert sein, einem Stuhl oder auch auf einem Thron. Zum anderen wird Maria als weibliche Person gekennzeichnet, was weniger über physische Merkmale als über ihre Kleidung und Frisur, und insbesondere durch die weiblich-tugendhafte Tätigkeit des Spinnens angezeigt wird. ³⁹ Dass es sich bei der Frau bei oder auf dem Sitzmöbel, der ein Korb mit Wolle beigeordnet sein kann, um Maria und bei der Person mit Flügeln um den Engel mit dem Namen Gabriel handelt, wird erst durch den Kontext und das Zusammenspiel der unterschiedlichen Objekte deutlich.

Die Schilderung, dass Magier im Osten aufbrechen und einem Stern folgen, um dem neugeborenen König ihre Aufwartung zu machen, findet sich im Matthäusevangelium. ⁴⁰ Die Magier treten in das Haus ein, sehen dort Maria mit dem Neugeborenen und fallen vor ihm nieder, um ihm zu huldigen und die mitgebrachten Gaben – Gold, Weihrauch und Myrrhe – zu überreichen. Das apokryphe Protevangelium des Jakobus folgt diesem Bericht, ersetzt aber das Haus durch eine Höhle. ⁴¹ In beiden Textzeugnissen wird die Personenanzahl der Magier nicht genannt, wohl aber die Dreizahl der Geschenke. Aus ikonografischer Sicht besonders interessant ist, dass das Pseudo-Matthäus-Evangelium berichtet, der Besuch der Magier habe sich erst zwei Jahre nach der Geburt zugetragen – Jesus ist hier also bereits ein Kleinkind. ⁴² Die Magier finden den Jesusknaben auf dem Schoß der Mutter sitzend vor und überreichen ihm selbst Goldstücke. Die ältesten erhaltenen Darstellungen dieses Bildthemas finden sich in den stadtrömischen Katakomben und auf Sarkophagen.

■ 36

Sog. Cubiculum der Verkündigung, Priscilla 15; Datierung: 2. Hälfte 3. Jh. n. Chr. Vgl. Barbara Mazzei, *Il cubicolo dell'Annunciazione nelle catacombe di Priscilla in Roma*, in: *Rivista di Archeologia Cristiana*, 75 1999, S. 233–280.

■ 37

Vgl. Johannes Emminghaus, s. v. Verkündigung an Maria, in: *LCI IV (1972)*, Sp. 422–437. Jensen 2015, 299–303. Maria Lidova, *Embodied Word: Telling the Story of Mary in Early Christian Art*, in: Thomas Arentzen, Mary B. Cunningham (Hg.), *The Reception of the Virgin in Byzantium: Marian Narratives in Texts and Images*, Cambridge 2019, S. 17–43, insb. S. 43.

■ 38

Vgl. Johann Michl, s. v. Engel IV (christlich), in: *RAC V (1962)*, Sp. 109–200, insb. Sp. 125.

■ 39

Zum Spinnen Mariens vgl. Maria Evangelatou, *The Purple Thread of the Flesh: The Theological Connotations of a Narrative Iconographic Element in Byzantine Images of the Annunciation*, in: Antony Eastmond (Hg.), *Icon and Word: The Power of Images in Byzantium: Studies Presented to Robin Cormack*, Burlington 2003, S. 261–280 sowie Gines Taylor 2018.

■ 40

Mt 2,1–12.

■ 41

Protev. Jac. 21. Vgl. zur Höhlen- vs. Stallgeburt Pia Wilhelm, s. v. Geburt Christi, in: *LCI II (1968)*, Sp. 86–120, insb. Sp. 95–96.

■ 42

Ev. PsMt. 16, 1.



□ 05

Die mit kaiserlichen Insignien ausgestattete Maria in der Verkündigungsszene des Mosaiks am ehemaligen Apsisstirnbogen in S. Maria Maggiore. Rom, Mitte 5. Jh. n. Chr. (Foto: Wilpert 1916, Taf. 53–55)

■ 43

Vgl. dazu Jutta Dresken-Weiland, *Bild, Grab und Wort: Untersuchungen zu Jenseitsvorstellungen von Christen des 3. und 4. Jh., Regensburg 2010* (Hdb. zur Geschichte des Todes im frühen Christentum und seiner Umwelt, Band 2). Ute Versteegen, *Die vier Weisen aus dem Morgenland: Eine frühchristliche Darstellung der Magierhuldigung*, Vortrag am 21.12.2006 an der Theologischen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg, 2006, <http://archiv.ub.uni-heidelberg.de/artdok/volltexte/2008/474/>.

Im Gegensatz zur Verkündigung steht die Ikonografie der Anbetung vornehmlich auf den Sarkophagen in der Spätantike relativ schnell fest: ⁴³ Maria sitzt mit dem Jesuskind auf dem Arm oder Schoß auf einem Sitzmöbel, zumeist einem Lehnstuhl. Von der Seite nähern sich ihnen in gebückter Haltung die Magier, die analog zu den drei Gaben zu dritt gezeigt werden und ihre Geschenke präsentieren ⁴⁴. Die Hauptbestandteile, die die Szene identifizierbar machen, sind neben Maria, dem Kleinkind bzw. Baby und dem Sitzmöbel vor allem die Magier: Sie sind als Bewohner der Gebiete östlich des Römischen Reiches gezeigt, tragen Hosen, Mäntel und phrygische Mützen. Diese Darstellung greift auf Motive zurück, die aus dem Kontext der Herrscherverehrung durch besiegte Völker stammen. ⁴⁴ In byzantinischer Zeit bleibt diese Ikonografie zwar in den Grundzügen bestehen, die Szene wird aber in die Geburtshöhle verlegt; der Kontext wird reicher ausgestaltet und Maria auf einem Bett liegend gezeigt.



■ 44

Vgl. Theodor Klauser, s. v. *Aurum coronarium*, in: *RAC I* (1950), Sp. 1010–1020. Johannes G. Deckers, *Göttlicher Kaiser und kaiserlicher Gott: Die Imperialisierung des Christentums im Spiegel der Kunst*, in: Franz Alto Bauer et. al. (Hg.), *Epochenwandel? Kunst und Kultur zwischen Antike und Mittelalter*, Mainz am Rhein 2001, S. 3–16. Robin M. Jensen, *Allusions to Imperial Rituals in Fourth-Century Christian Art*, in: Lee M. Jefferson et. al. (Hg.), *The Art of Empire: Christian Art in Its Imperial Context*, Minneapolis 2015, S. 13–47.

Die im Bild dargestellten Objekte haben also auch hier unterschiedliche Funktionen, die sich überlagern können: Sie ermöglichen durch attributiven Einsatz die Identifikation der Protagonist:innen (z. B. die Flügel des Engels) sowie der Szene an sich (z. B. die Verbindung von Flügeln und Korb), sind Teil der Narration, indem sie in die dargestellten Handlungen eingebunden sind (z. B. der Vorgang des Spinnens) und fügen zudem übergeordnete Bedeutungsebenen hinzu (z. B. die Kleidung der Magier oder der Thron Mariens).



□ 06

Nicht nur die für sog. Barbaren aus dem Osten typische Bekleidung (Hosen, Mäntel und phrygische Mützen), auch Motive, die aus dem Kontext der Herrscherverehrung durch besiegte Völker stammen, werden in die frühchristliche Bildschöpfung der Magieranbetung übernommen. Links: Darstellung der Anbetung auf einem Sarkophag, Rom (Cimitero anonimo presso S. Lorenzo o di Novaziano), 1. Viertel 4. Jh. n. Chr., Rep. I,662 (Foto: Wilpert 1932, Taf. CLXXIX,2). Rechts: Abführung eines gefangenen Parthers. Rom, Bogen des Septimius Severus, 203 n. Chr. (Foto: Anthony Majanlahti, Wikimedia Commons, lizenziert unter Creative-Commons-Lizenz by-sa-2.0-de, https://commons.wikimedia.org/wiki/File:Forum_arch_of_septimius_severus.jpg)

Zum Einstieg: Automatische Analyse archäologischen (Bild-)Materials am Beispiel digitaler Plots

Die eingangs beschriebenen Herausforderungen archäologischen (Bild-)Materials betreffen nicht einzelne Artefaktgattungen oder konkrete Fragestellungen, sondern sind Merkmale des Gesamtbestands des archäologischen Corpus bzw. der wissenschaftlichen Auseinandersetzung in den jeweiligen Disziplinen. Dementsprechend sind es auch oft mehrere Faktoren, die eng miteinander verbunden sind und folglich die Anwendbarkeit digitaler Methoden beeinflussen. Dieser Umstand soll an einem Anwendungsbeispiel, der digitalen Ähnlichkeitsanalyse und Visualisierung in Plots, kurz veranschaulicht werden:

Modelle der Computer Vision bieten die gewinnbringende Möglichkeit, Bilder anhand vieler Dimensionen von Ähnlichkeiten zu vergleichen und sie somit differenziert zueinander in Bezug zu setzen. ⁴⁵ Die Visualisierung der Bilder und ihrer Ähnlichkeitsabstufungen in – wegen des zugrundeliegenden Algorithmus – sogenannten UMAP-Plots (Uniform Manifold Approximation and Projection) ermöglicht es Forschenden, sowohl Makrostrukturen zu erkennen und zu analysieren, als auch Mikrostrukturen innerhalb ihrer Position darin, wie Cluster oder andere Muster im Plot, zu detektieren und zu vergleichen. ⁴⁶ Die Verbindung mit den jeweiligen Metadaten der Bilder erlaubt es zudem, inhaltliche Aussagen beispielsweise über chronologische Entwicklungen zu treffen. Der Gewinn liegt nicht nur in einer Bewältigung großer Bildmengen im Sinne eines **distant viewing** ⁴⁷, die für Forschende kaum oder nur sehr mühsam realisierbar wäre, sondern auch in der andersartigen Betrachtung der Bilder durch das maschinelle Sehen abseits individueller oder fachspezifischer Erwartungshaltungen, wodurch bislang unerkannte Bezüge aufgedeckt und neue Forschungen angeregt werden können.

Die automatisierte Sortierung der Bildcorpora zeigt bei den archäologischen Artefakten der Klassischen wie der Christlichen Archäologie hingegen bislang wenig Erfolg. Im Fall der christlich-archäologischen Artefakte ist zu beobachten, dass die KI (Künstliche Intelligenz) sehr stark auf die Farbigkeit der Abbildungen, die Binnenstruktur der Fotos bzw. der abgebildeten Artefakte sowie deren Umrissform fokussiert ^[07]. Dadurch entsteht in der Regel eine Sortierung nach archäologischen Gattungen bzw. Materialien (z. B. Sarkophage oder Buchmalerei), die jedoch nicht durch eine erkannte Ähnlichkeit in der ikonografischen Ausformung oder gar narrativen Übereinstimmungen begründet ist. Auch Abbildungen antiker Gefäße umfassen teils den vollständigen Gefäßkörper, geben teils aber auch nur ein konkretes Vasenbild als Ausschnitt wieder. Die daraus resultierende Irritation des Computer-Vision-Modells kann umgangen werden, indem anstelle von Plots der ursprünglichen Abbildungen lediglich Annotationen (z. B. bestimmter Objekte) auf diese Weise visualisiert werden. Fotografische Verzerrungen, die durch die Position des Bildes auf dem Gefäßkörper entstehen und durch das Modell als Ähnlichkeit rezipiert werden können sowie die unterschiedliche Farbigkeit der Abbildungen bleiben als Problem allerdings weiterhin bestehen, wie ein Test anhand von 9.713 Annotationen von Objekten der übergeordneten Objektkategorie Waffen zeigt

■ 45

Peter Bell, Leonardo Impett, **Ikonographie und Interaktion: Computergestützte Analyse von Szenen der Evangelien**, in: *Das Mittelalter*, 24 (1) 2019, S. 31–53, <https://doi.org/10.1515/mial-2019-0004>. Grundlegend hierzu: Lev Manovich, **Cultural Analytics**. Cambridge – London 2020.

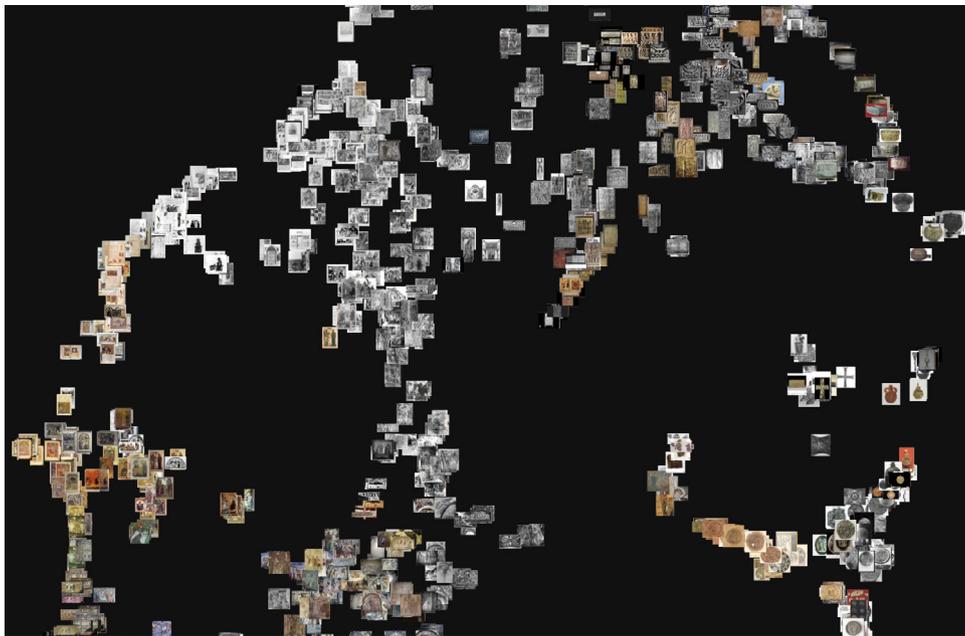
■ 46

Zum Clustering: Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, Stuart James 2020, S. 106. Zum UMAP-Algorithmus speziell: Leland McInnes, John Healey, **UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction**, <https://api.semanticscholar.org/CorpusID:3641284>.

■ 47

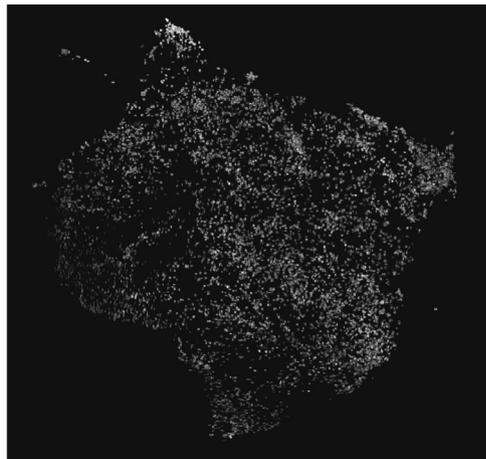
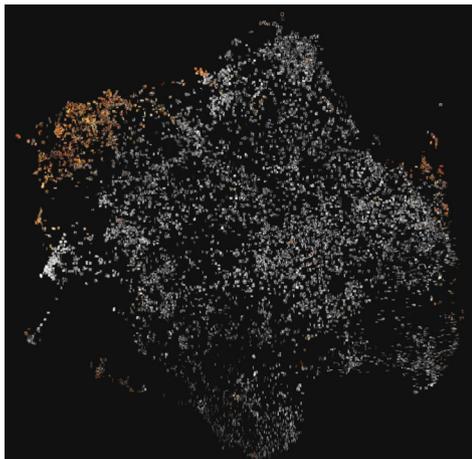
Vgl. insbesondere die Forschung von Manovich 2020. Der Begriff ist angelehnt an den des sog. »distant reading«. Hierzu: Franco Moretti, **Distant Reading**, Konstanz 2016.

(siehe [08], links). Eine vorherige Bearbeitung des Sets zu Graustufen-Abbildungen resultiert in Hinblick auf die Frage nach ähnlichen Objektdarstellungen in präziseren Clustern (siehe [08], rechts). Werden nur schwarz-weiße Abbildungen genutzt, hätte dies auch für das christlich-archäologische Bildmaterial den Vorteil, dass so die ungewollte Sortierung nach beispielsweise farbigen Hintergründen, auf denen die Artefakte fotografiert worden sind, entfallen. Hierdurch entsteht aber zugleich ein Verlust, denn die Farbgebung spätantiker und byzantinischer Bilder kann für ihre Rezeption und Interpretation eine entscheidende Rolle spielen (z. B. durch den Einsatz von Purpur als kaiserlicher Farbe für die Gewandung Mariens).



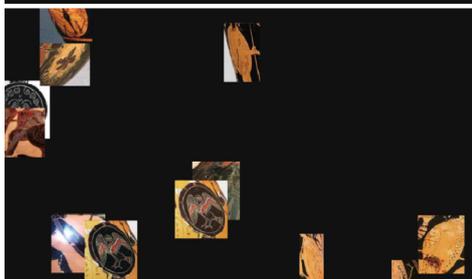
□ 07

Die Bilder der Verkündigung und Anbetung werden nach Farbe (unten links: Buchmalerei, Ikonenmalerei, Mosaik), Binnenstruktur (obere Hälfte: Sarkophage, Elfenbein) und Form (unten rechts: runde Kleinfunde und Textilornamente) geclustert, was eine Sortierung nach archäologischen Gattungen zur Folge hat. (Plot: Ronak Kostı und Prathmesh Madhu)



□ 08

Die Visualisierung auf Basis von über 9.700 Annotationen der Oberkategorie Waffen ohne (links) und mit (rechts) vorheriger Bildbearbeitung zu Graustufen-Abbildungen ergibt ganz unterschiedliche Plots. Die untere Reihe zeigt jeweils Ausschnitte einzelner Cluster mit Darstellungen von Schilden. (Plots: Ronak Kostı und Prathmesh Madhu)



Die Anwendbarkeit von digitalen Methoden zur Visualisierung von Ähnlichkeiten in Plots und Clustern ist somit für archäologisches Bildmaterial noch stark eingeschränkt. Dies liegt an einer Vielzahl von Gründen: Zwar ist die geringe Bildmenge nicht für den Erfolg des Verfahrens ausschlaggebend, jedoch bieten die Praktiken der Zusammenstellung archäologischer Bildcorpora, die sich zumeist aus vielen unterschiedlichen Datenbanken und/oder durch Scans aus der Fachliteratur speisen, keine optimalen Voraussetzungen. Die verfügbaren Abbildungen unterliegen keinen einheitlichen Standards und weisen daher hinsichtlich der Farbigkeit, Bildqualität, Aufnahmewinkel etc. große Unterschiede auf. Zudem nimmt das Modell, wie der Test mit christlich-archäologischem Material zeigt, Sortierungen anhand von Bildcharakteristika vor, die mit den abgebildeten Artefaktgattungen verknüpft sind, was dazu führen kann, dass gattungsübergreifend auftretende darstellerische Ähnlichkeiten nicht adäquat repräsentiert werden. Ebenso können ähnliche fotografische Verzerrungen, die in der Materialität des Bildträgers begründet sind, als vom Modell bildimmanente Ähnlichkeiten interpretiert werden, was z. B. bei Bildelementen, die aufgrund ihrer Position auf dem Gefäßkörper fotografisch ähnlich erfasst wurden, zu einer entsprechenden Positionierung in einem Cluster führen kann. Ähnliche Irritationen sind durch fragmentarische Erhaltungszustände zu erwarten, da das Modell die (gedankliche) Rekonstruktion des ursprünglichen Bildes nicht leisten kann.

Als Zukunftsperspektive ist den beschriebenen Schwierigkeiten mit Kombinationen verschiedener im Folgenden in diesem Beitrag vorgestellter Verfahren zu begegnen. Die digitale Analyse von Körperhaltungen und deren Visualisierung in Plots ist beispielsweise anhand von digital erzeugten Posen-skeletten für Bildmaterial der Kunstgeschichte bereits erprobt worden. ⁴⁸ Auf diesem Weg können sämtliche Irritationen, d. h. sowohl Bildfarbigkeit und -qualität, aber auch Kleidung, Ornamente, Stil oder Textur, in der digitalen Analyse unberücksichtigt bleiben und somit Plots erzeugt werden, die ausschließlich die Figurenhaltung und -bewegung erfassen. Methoden wie das style transfer learning könnten diese Analyse auch stil- und gattungsübergreifend ermöglichen, während Methoden wie die Entzerrung von Fotografien dreidimensionaler Objekte bereits im Vorfeld durch die Materialität des Bildträgers erzeugte Unschärfen in der Bestimmung von Gelenkpunkten und Posen ausgleichen können.

■ 48

Peter Bell, Leonardo Impett 2019,
S. 42–45 m. Abb. 1 u. 2.

Small Data-Charakter der Datenbasis

In Hinblick auf die Menge der zur Verfügung stehenden Bildwerke ist mit mehreren herausfordernden Faktoren umzugehen. Während die einen im archäologischen Material selbst begründet liegen, betreffen die anderen wissenschaftspolitische Hürden, welche die organisatorische und rechtliche Durchführung von Forschungsprojekten an großen Bilddatenpools erschweren, wie sie Computer-Vision-Anwendungen erfordern.

Aktuell stellen nur wenige große Museen einen Teil ihres Bildbestands als lizenzfreie Daten über eigene Museumsbilddatenbanken zur Verfügung. Die größte Onlinedatenbank für antike Vasenmalerei, die Beazley Archive Pottery Database, bietet zwar eine automatisierte Abrufbarkeit des gesamten Archivs an ⁴⁹, besitzt aber nicht die Nutzungsrechte der Abbildungen. ⁵⁰ Für andere archäologische Artefaktgattungen sind derartige Datenbanken mitunter gar nicht vorhanden. Es gibt beispielsweise keine Datenbank, die das christlich-archäologische Material umfassend bereitstellt. In aller Regel geschieht die Zusammenstellung ikonografischer Bildcorpora daher noch immer durch Einzelabfragen in einschlägigen Bildarchiven, Museums- und Forschungsdatenbanken sowie durch manuelles Scannen aus der Fachliteratur. Dieses Bildmaterial ist urheberrechtlich geschützt und eine Publikation daher meist nur nach Lizenzerwerb bei den jeweiligen Museen oder Verlagen gestattet. Erschwerend kommt hinzu, dass die umfangreiche Recherche und das anschließende Scannen des Bildmaterials einen hohen personellen Aufwand bedeuten. Aus diesem Grund ist eine schnelle gebündelte Bereitstellung umfassender Bilddaten, beispielsweise in Form annotierter Bilddatensets, anhand derer digitale Methoden über den Rahmen einzelner Forschender, Projekte und Institutionen hinaus erprobt und weiterentwickelt werden können, gerade im archäologischen Bereich aktuell nicht realisierbar. ⁵¹ Der weitere Ausbau einer umfangreichen Bereitstellung von Abbildungen und Metadaten nach der in ihrem Anliegen explizit auf wissenschaftliche Datenbanken zielenden CC0-Lizenzklärung ⁵² wäre wünschenswert und zeitgemäß. ⁵³

Aber auch das archäologische Bildcorpus selbst stellt eine Herausforderung für das Trainieren der Computer-Vision-Modelle dar, der im Rahmen des Forschungsprojekts auf unterschiedliche Weise begegnet wurde. So beträgt die Anzahl der für das klassisch-archäologische Fallbeispiel zusammengetragenen Frauenführungsszenen nur knapp über 200 Vasenbilder. Jedoch handelt es sich bei der Frauenführung um ein durch Figureninteraktion konstituiertes Bildschema, das kontextübergreifend und somit in ganz unterschiedlichen Ikonografien auftreten kann, für die wiederum ganz unterschiedliche Objekte bedeutsam sein können. Im Kontext profaner Hochzeitsprozessionen können dies beispielsweise Kränze, Fackeln oder Architekturelemente (z. B. Türen und Säulen) sein, während zum Beispiel Waffen in mythologischen Kontexten zu erwarten sind, in denen die eine Frau führende männliche Figur als ein bestimmter Krieger (z. B. Menelaos) charakterisiert werden soll. Da zudem griechische Vasenbilder bereits in einem umfangreichen Bildcorpus, der Beazley Archive Pottery Database, vorliegen, war es von Beginn an möglich, ein breites Spektrum unterschiedlicher Objekte anhand einer hohen Zahl von Bildern zu trainieren.

■ 49

Siehe <https://www.carc.ox.ac.uk/xdb/ASP/CARCOpenData.pdf>.

■ 50

Die Bildrechtsproblematik betrifft auch die umfangreiche archäologische Objektdatenbank iDAI.objects arachne, die durch das Deutsche Archäologische Institut und das Archäologische Institut der Universität zu Köln administriert werden, siehe <https://idai.world/what/images>. Zu den Herausforderungen der rechtlichen Rahmenbedingungen vgl. jüngst Friederike Fless, Peter Baumeister, Benedikt Boyxen, Camilla Colombi, Benjamin Ducke, Martina Düntzer, Reinhard Förtsch, Jana Littmann, Fabian Riebschläger, Henriette Senst, Sabine Thänert, Juliane Watson, Kathleen Wennrich, Agnes Wilczek, Die iDAI.world vor dem Hintergrund der neuen Digitalgesetze, in: *Forum for Digital Archaeology and Infrastructure*, 2021/1, S. 1–51, <https://doi.org/10.34780/s2ne-t268>.

■ 51

Für spätere kunsthistorische Perioden ist die Akkumulation von Datensets bereits weiter fortgeschritten und vorrangig auf Malerei konzentriert. Siehe z. B. zu OmniArt: Gjorgji Strezoski, Marcel Worrang, Omniart: Multi-Task Deep Learning for Artistic Data Analysis, 2017, arXiv preprint arXiv:1708.00684. Zu IconArt: Nicolas Gonthier, Yann Gousseau, Said Ladjal, Olivier Bonfait, Weakly Supervised Object Detection in Artworks, in: Laura Leal-Taixé et. al. (Hg.), *Computer Vision: ECCV 2018 Workshops*, Cham 2019, S. 692–709, https://doi.org/10.1007/978-3-030-11012-3_53.

■ 52

<https://de.creativecommons.net/2009/03/17/neu-im-programm-cc0/>.

■ 53

Dieses Desiderat thematisieren auch: Marco Fiorucci, Marina Khoroshiltseva, Massimiliano Pontil, Arianna Traviglia, Alessio Del Bue, Stuart James 2020, S. 103–104.

In Hinblick auf die digitale Objekterkennung konnte auf diesem Weg ein state-of-the-art RetinaNet-Modell (Resnet-50) als Objekterkennungsmethode angewandt und dieses anhand von circa 42.500 Annotationen in annähernd 12.000 Bildern antiker Vasenmalerei mittels eines supervised training-Ansatzes speziell auf dieses Material trainiert werden. Beim supervised training-Ansatz werden in einer Zusammenstellung von Bildern (dataset) relevante Objektklassen durch aus der Fachwissenschaft stammende Expert:innen annotiert sowie in einem festgelegten Arbeitsablauf digital erzeugte Ergebnisse durch diese Expert:innen korrigiert. ⁵⁴ Das Training der digitalen Objekterkennung im klassisch-archäologischen Teilprojekt umfasste insgesamt 81 Objektklassen, darunter Waffen wie Schwert und Speer, Götterattribute (z. B. Kerykeion und Dreizack), kultische Objekte (z. B. Fackel, Kranz und Thymiaterion), Musikinstrumente (z. B. Aulos und Lyra) sowie Architektur (z. B. Säulen und Altäre). Dieses Objekterkennungsmodell war imstande, für die digitale Erkennung von Objekten in antiken Vasenbildern Ergebnisse mit einer durchschnittlichen Genauigkeit (average precision) von 32% zu liefern. Dieser Wert wird ermittelt, indem die Zusammenstellung annotierter Bilder in ein Trainings- und ein Testset unterteilt wird. Während die Objekterkennung anhand der Annotationen im Trainingsset erlernt wird, können die digital erzeugten Resultate für die Bilder des Testsets (die dem Modell im Vorfeld noch unbekannt sind) mit den Expert:innen-Annotationen des Testsets verglichen und somit Angaben zur Korrektheit der automatisch erzeugten Prognosen gewonnen werden. Mit einer average precision von 32% liegt die Performanz des genutzten Modells für antike Vasenbilder somit nur geringfügig unter der Leistung des gleichen Modells für die Objekterkennung in modernen Fotografien (35,9%). ⁵⁵ Durch die zusätzliche Anwendung von abstract classes, unter denen in Form von Oberkategorien (z. B. Waffen) mehrere visuell ähnliche Objektklassen (z. B. Speer und Schwert) zusammengefasst sind, konnte die durchschnittliche Präzision auf 42.3% verbessert werden. ⁵⁶

■ 54

Zum Vorgang des Annotierens
 grundsätzlich: Fotis Jannidis, Hubertus Kohle, Malte Rehbein (Hg.), *Digital Humanities: Eine Einführung*, Stuttgart 2017, S. 253–263. Christoph Sager, Christian Janiesch, Patrick Zschech, *A Survey of Image Labelling for Computer Vision Applications*, in: *Journal of Business Analytics*, 2021, S. 1–20. Exemplarisch zur Nutzung sogenannter bounding boxes für die Annotation kulturhistorischer Bildcorpore: Mitchell Van Zuijlen, Hubert Lin, Kavita Bala, Sylvia Pont, Maarten W.A. Wijntjes, *Materials In Paintings (MIP): An Interdisciplinary Dataset for Perception, Art History, and Computer Vision*, in: *PLoS ONE*, 26 (8) 2021, e0255109, <https://doi.org/10.1371/journal.pone.0255109>. Zur Problematik der Kohärenz archäologischer Klassifikationssysteme im Machine Learning: Dylan S. Davis, *Defining What We Study: The Contribution of Machine Automation in Archaeological Research*, in: *Digital Applications in Archaeology and Cultural Heritage*, 18 2020, e00152, <https://doi.org/10.1016/j.daach.2020.e00152>. Zu einem Beispiel automatisierter Annotationen in der Vasenmalerei: Elliot J. Crowley, Andrew Zisserman 2013.

■ 55

https://paperswithcode.com/sota/object-detection-on-coco?tag_filter=3, Modell #40 (RetinaNet).

■ 56

Prathmesh Madhu, Ronak Kosti, Corinna Reinhardt, Torsten Bendschus, Vincent Christlein, Andreas Maier, Peter Bell, *Deep Learning Based Attribute Representation in Ancient Vase Paintings*, in: *Digital Humanities 2020: Intersections/Carrefours 2020, Ottawa 2020, Abstract 287*, <https://doi.org/10.17613/va2g-h523>.

Für die Bearbeitung der christlich-archäologischen Artefakte wurde eine andere Herangehensweise gewählt: Die Objekterkennung erfolgte hier ausgehend von den dargestellten Objekten, die für die im Rahmen des Projekts in den Blick genommenen Ikonografien – die Verkündigung und die Anbetung – am ausschlaggebendsten sind. Das im Rahmen des Projekts behandelte christlich-archäologische Bildcorpus umfasst für diese beiden Ikonografien nur um die 1.000 Abbildungen von ca. 500 Einzelobjekten. Es handelt sich dabei um annähernd den gesamten bekannten Bestand. Ein zweiter schwerwiegender Faktor ist, dass die christlichen Bildfindungen teilweise einen langandauernden Prozess durchlaufen haben, in dessen Entwicklung sich feste Ikonografien erst herausgebildet haben [09]. Welche Bestandteile der in den Details durchaus divergierenden Textgrundlagen ins Bild gesetzt werden, wie der Bildaufbau konstruiert wird und welche Personen oder Objekte besonders in den Fokus gerückt werden, stand folglich nicht von Beginn an fest. Dies führt zu Ikonografien, die nur in einem bestimmten regionalen und zeitlichen Rahmen aus denselben Standardelementen konstruiert und in diesem Kontext als kanonisch ausgeformt werden. Durch diesen Umstand verringert sich der Bildbestand aus Sicht der Computer Vision nochmals.

□ 09

Diese Anbetung der Magier aus der Domitilla-Katakombe zeigt eine abweichende Bildkomposition mit vier Magiern. Rom, Domitilla-Katakombe, 4. Jh. n. Chr. (Foto: Wilpert 1903, Taf. 116,1)



Die geringe Anzahl der überlieferten frühchristlichen und byzantinischen Bildzeugnisse macht es unmöglich, die Modelle nur anhand dieser Bilder zu trainieren. Glücklicherweise kommen hier die spätantiken Bildfindungsprozesse selbst zur Hilfe, denn die christlichen Bildentwürfe sind keine creatio ex nihilo, sondern rekurrieren auf den paganen Motivschatz [10]. Aus der römisch-kaiserzeitlichen, nicht-christlichen Ikonografie bekannte Elemente werden neu kombiniert, in einen anderen Kontext gesetzt oder schlicht in ihrer Aussage beibehalten. Viele Motive bestehen auch in späteren Epochen fort. So ist es möglich gewesen, die Datenbasis für das Trainieren der Modelle zu erweitern. Neben nicht-christlichen römischen, frühchristlichen und byzantinischen Darstellungen, die andere Szenen zeigen, sind auch westliche mittelalterliche

Bildwerke genutzt worden, die eine Darstellung der Objekte beinhalten, welche für die Ikonografie der Verkündigung und Anbetung typisch sind. So konnte ein Corpus mit annähernd 9.000 Abbildungen und fast 18.000 Annotationen von sitzenden Frauen (mit und ohne Kind auf dem Schoß), Sitzmöbeln, Flügeln, Wolle, Spinngeräten, dem Spinnvorgang, der Wasserquelle bzw. Brunnen, phrygischen Mützen, Hosen und Mänteln zusammengestellt werden – Objekte, die bei der Identifizierung der geschilderten Ikonografien große Bedeutung besitzen. Der Mehrwert, der hierbei entsteht, ist, dass auf diese Weise einzelne Motive und Motivkombinationen über einen langen Zeitraum und in unterschiedlichen Regionen verfolgt werden können.

□ 10

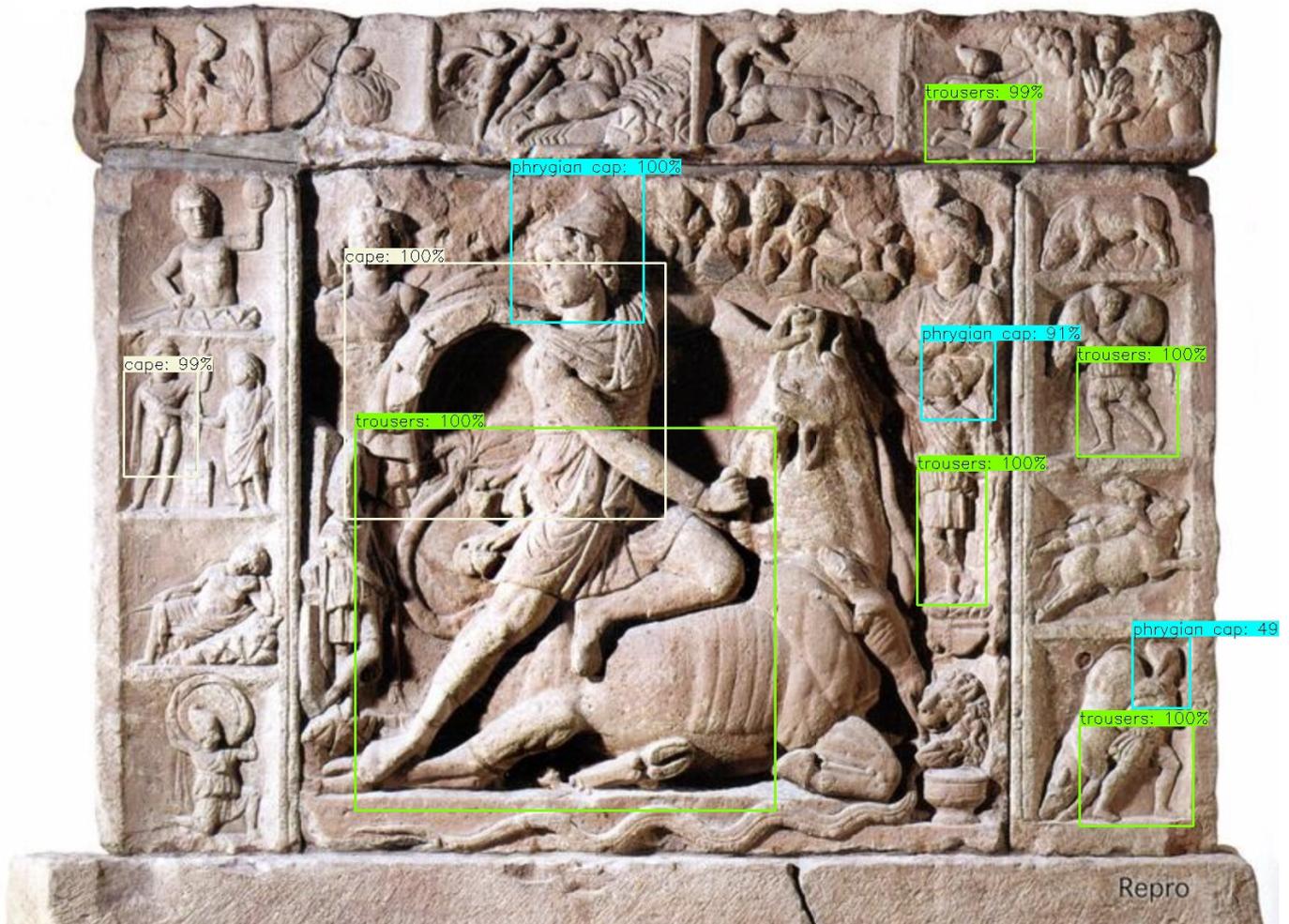
Darstellung des Motivtransfers der für die Ikonografie der Anbetung wichtigsten Objekte (Grafik: Lara Mührenberg) (Fotos: Sarkophag [Mitte]: Wilpert 1932, Taf. CLXXIX,2; Relief [unten rechts]: Anthony Majanlahti, Wikimedia Commons, lizenziert unter CreativeCommons-Lizenz by-sa-2.0-de, https://commons.wikimedia.org/wiki/File:Forum_arch_of_septimius_severus.jpg; alle weiteren Abbildungen: © New York, Metropolitan Museum of Art, lizenziert unter CC0, <https://www.metmuseum.org>)



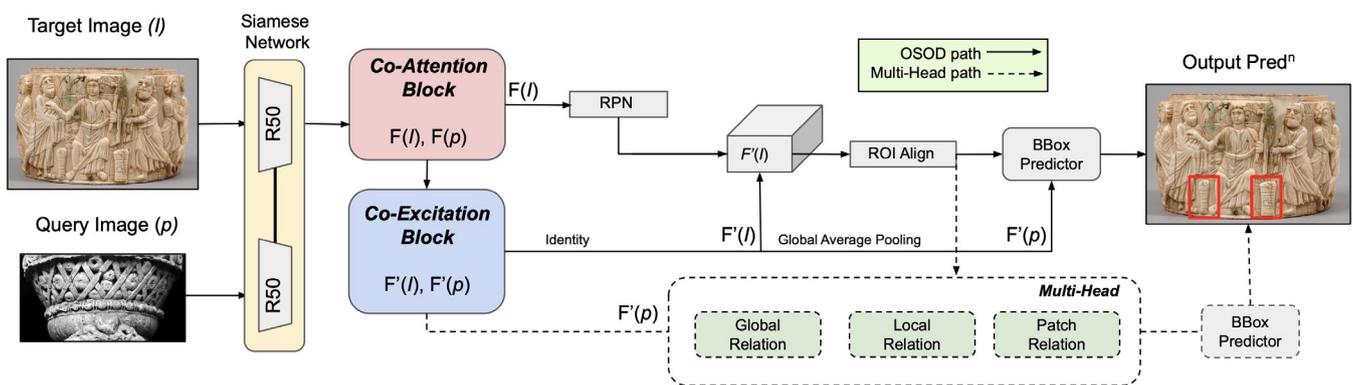
Dennoch stellt dies noch immer eine für Computer-Vision-Verhältnisse sehr geringe Bildmenge für die Entwicklung und Anwendung entsprechender Modelle dar. Hinzu kommt die starke stilistische Heterogenität der christlich-archäologischen Artefakte (vgl. hierzu unten, Abschnitt **Vielfalt der Bilder und Bildträger**). Die Implementierung einer state-of-the-art-Objekterkennungsmethode mit supervised-learning-Ansatz – ähnlich der oben für das klassisch-archäologische Bildcorpus geschilderten und basierend auf Annotationen auf Objektebene – lieferte für das Material der Christlichen Archäologie daher nur eine Performanz von 23% (siehe [11](#)). Zudem würden bei dieser Methode für jegliche Anwendung auf ein abweichendes Set von Objekten stets neue Annotationen benötigt werden. Ohne solch ein erneutes Training wäre die Anwendbarkeit dieses Modells wiederum erheblich eingeschränkt, weshalb es keinen adäquaten Problemlösungsansatz darstellt. Stattdessen ist eine Methode zu bevorzugen, die anhand einer nur geringen Anzahl von Annotationen trainiert werden kann und gleichzeitig fähig ist, eigenständig neue Objektklassen zu erkennen.

□ 11

Objekterkennung auf einem Relief aus Heidelberg-Neuenheim mit Szenen aus dem Mithrasmythos, Badisches Landesmuseum Karlsruhe, 2. Jh. n. Chr. (Objekterkennung durchgeführt von: Ronak Kosti und Prathmesh Madhu) (Foto: © Badisches Landesmuseum, Karlsruhe, Inv. C 16, Fotograf: Thomas Goldschmidt. Publikation mit frdl. Genehmigung)



Im Rahmen des Projekts wurde aus diesen Gründen für das christlich-archäologische Bildcorpus testweise ein few-shot-object-detection-Ansatz (FSOD) verfolgt, bei dem das Objekterkennungsmodell für ein Set von Objektklassen trainiert und anschließend anhand eines sich von diesem unterscheidenden Set von Objekten getestet wird [12]. 57 Hierfür werden dem (zuvor trainierten) Modell zwei Bilder gezeigt: Zum einen ein Abfragebild (query image), das das Objekt selbst darstellt, zum anderen ein Zielbild (target image), in dem dieses Objekt digital erkannt werden soll. Abhängig von der Zahl an Abfragebildern wird diese Methode als N (one/two/three/few)-shot object detection bezeichnet, wobei im Falle mehrerer Abfragebilder diese immer das gleiche Objekt zeigen müssen. Theoretisch erlaubt es dieses Verfahren, ein trainiertes Modell zu nutzen, um beliebige Objekte in einem Zielbild zu erfassen, auf welche das Modell zuvor nicht explizit trainiert wurde.



□ 12
Architektur des FSOD-Frameworks.
(Grafik: Ronak Kosti und Prathmesh Madhu)

In der Praxis ist allerdings auch für dieses Modell zu beobachten, dass die Leistung deutlich in Abhängigkeit zur Anzahl der Abfragebilder steht und sich steigert, je mehr Abfragebilder zur Verfügung stehen. Verglichen mit der Leistung bei modernen Fotografien (22%) ist die Performanz dieses Modells in Teststudien mit dem Bildmaterial der Christlichen Archäologie noch als gering einzuschätzen (3,58%). Die Resultate der erstmaligen Anwendung des few-shot object detection-Ansatzes auf Bilddaten der Christlichen Archäologie bestätigen somit momentan die Schwierigkeiten im Umgang mit dem **small data**-Charakter archäologischer Bilddatensets.

Aus archäologischer Sicht stellt der few-shot object detection-Ansatz gegenüber denjenigen Verfahren, die eine große Menge an Bildern und Annotationen benötigen, dennoch aus mehreren Gründen einen enormen Vorteil dar. Insbesondere nimmt diese Methode auf den Umstand Rücksicht, dass nur eine geringe Anzahl archäologischer Artefakte die Jahrtausende (mehr oder weniger) unbeschadet überstanden hat und der wissenschaftlichen Erforschung zur Verfügung steht. Selbst für den Fall, dass überdurchschnittlich viele Neufunde in den kommenden Jahren oder Jahrzehnten hinzukommen sollten, ist nicht damit zu rechnen, dass sich die Summe der materiellen Hinterlassenschaften in der Art erweitern ließe, wie es für die herkömmlichen Methoden der Computer Vision notwendig wäre. Darüber hinaus werden für diesen Ansatz weit weniger

■ 57
Prathmesh Madhu, Anna Meyer, Mathias Zinnen, Lara Mührenberg, Dirk Suckow, Torsten Bendschus, Corinna Reinhardt, Peter Bell, Ute Versteegen, Ronak Kosti, Andreas Maier, Vincent Christlein, One-Shot Object Detection in Heterogeneous Artwork Datasets, in: 2022 Eleventh International Conference on Image Processing Theory, Tools and Applications (IPTA), <https://ieeexplore.ieee.org/document/9784141>.

personelle Ressourcen auf fachwissenschaftlich-archäologischer Seite benötigt, denn das Annotieren tausender einzelner Objekte erfordert ein archäologisch geschultes Auge und eine hohe fachliche Kompetenz. Mit den für archäologische Forschung zur Verfügung stehenden Mitteln ist ein solcher Personalaufwand nicht über einen längeren Zeitraum leistbar. Die few-shot object detection-Methode macht es hingegen möglich, auch ohne hohen personellen und zeitlichen Aufwand in kurzer Zeit Ergebnisse zu erzielen und Bilder zu Testzwecken durchsuchen zu können. Diese Euphorie wird natürlich von dem Umstand gebremst, dass auch bei dem few-shot object detection-Ansatz bessere Resultate erzielt werden, wenn mehr Abfragebilder zu Verfügung stehen. Die Tatsache, dass die Resultate zum jetzigen Zeitpunkt noch weit hinter denen zurückbleiben, die durch die Anwendung auf moderne Fotografien erzielt werden konnten, ist hingegen aufgrund der Besonderheit des Materials zu erwarten gewesen und stellt keinen Grund dar, das Verfahren an sich infrage zu stellen. Darüber hinaus ist ein interessanter Aspekt für zukünftige Unternehmungen, dass das Verfahren theoretisch erlaubt, beliebige Objekte in einem Bild zu erfassen, auf die die Modelle im Vorfeld weder hingewiesen noch trainiert worden sind. Dies ermöglicht nicht nur eine weitere Entlastung hinsichtlich personeller Ressourcen im Vorfeld, sondern es können bei einer Objektsuche auch Objekte in den Fokus der archäologischen Betrachtung gerückt werden, die zuvor bei der Annotation noch keine Beachtung gefunden haben.

Vielfalt der Bilder und Bildträger

Die bereits angesprochene Vielfalt des christlich-archäologischen Artefaktcorpus' im untersuchten Fallbeispiel zeigt sich auf ganz unterschiedlichen Ebenen. Nicht nur die ikonografischen Ausformungen innerhalb einer gezeigten Szene können sehr heterogen sein, das Corpus beinhaltet zudem Artefakte, die zeitlich und regional sehr unterschiedlich einzuordnen sind ^[13]. Die Darstellungen datieren in einen Zeitraum zwischen dem dritten und 15. Jahrhundert nach Christus und umfassen westliche Bilder ebenso wie byzantinische, armenische, koptische und islamische. Diese Bilder folgten unterschiedlichen Konventionen und sind für Betrachtende mit ganz unterschiedlichen Sehgewohnheiten gemacht worden. Eine weitere Herausforderung stellen die unterschiedlichen Artefaktgattungen dar, die jeweils eigene Anforderungen an die Ausformungen der Ikonografie stellen. Dies ist zum einen durch die Beschaffenheit des Materials bedingt, zum anderen durch die jeweilige Ästhetik der Artefaktgattung und -größe sowie die technische Ausführung und nicht zuletzt den Nutzungskontext, für den unterschiedliche Aspekte des Themas in den Vordergrund gerückt werden konnten.

□ 13

Das christlich-archäologische Bildcorpus umfasst Artefakte unterschiedlichster Gattungen, Zeitstellungen und geographischer Räume. Links: Verkündigungsszene in einer äthiopischen Evangelienhandschrift. Äthiopien, spätes 14. / frühes 15. Jh. n. Chr. (Foto: © New York, Metropolitan Museum of Art, Inv. 1998.66, lizenziert unter CC0, <https://www.metmuseum.org/art/collection/search/317618>). Rechts: Verkündigung an Maria auf einem Armreif. Ägypten, 5.-7. Jh. n. Chr. (Foto: © New York, Metropolitan Museum of Art, Inv. 12.182.100, lizenziert unter CC0, <https://www.metmuseum.org/art/collection/search/446227>)



■ 58

Einführend: John Boardman, Schwarzfigurige Vasen aus Athen: Ein Handbuch, Mainz 1994 (Kulturgeschichte der Antiken Welt, Band 1). John Boardman, Rotfigurige Vasen aus Athen: Ein Handbuch. Die archaische Zeit, Mainz 1994 (Kulturgeschichte der Antiken Welt, Band 4). John Boardman, Rotfigurige Vasen aus Athen: Ein Handbuch. Die klassische Zeit, Mainz 1991 (Kulturgeschichte der Antiken Welt, Band 48).

■ 59

Zu bilinguen Gefäßen: Beth Cohen, Attic Bilingual Vases and Their Painters, New York 1978.

■ 60

Zur Methode: Prathmesh Madhu, Ronak Kosti, Lara Mührenberg, Peter Bell, Andreas Maier, Vincent Christlein, Recognizing Characters in Art History Using Deep Learning, in: SUMAC 2019: The 1st Workshop on Structuring and Understanding of Multimedia heritAge Contents. Nizza, 21.-25. Oktober 2019, Nizza 2019, S. 15–22, arXiv:2003.14171v2. Prathmesh Madhu, Atharva Pawar, Ronak Kosti, Dirk Suckow, Peter Bell, Andreas Maier, Vincent Christlein, Impact of Styles on Iconography Recognition in Art History (in Vorbereitung).

■ 61

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 et al. (Hg.), Computer Vision: ECCV 2014, Cham 2014, S. 740–755 (Lecture Notes in Computer Science), arXiv:1405.0312.

■ 62

Xun Huang, Serge Belongie, Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization, in: 2017 IEEE International Conference on Computer Vision (ICCV), Venedig 2017, S. 1510–1519, arXiv:1703.06868.

Die griechische Vasenmalerei des sechsten und fünften Jahrhunderts vor Christus, wie sie dem klassisch archäologischen Fallbeispiel zugrunde liegt, beinhaltet als kohärente Gattung deutlich weniger Vielfalt in Hinblick auf die künstlerischen Stile, die medialen Gesetzmäßigkeiten der Gattungen und die Darstellungskonventionen als das archäologische Bildmaterial in seiner gesamten Breite. Sie bietet sich aber als Ausgangspunkt technischer Lösungen an, weil sie in erster Linie nur einen der genannten Aspekte – nämlich verschiedene künstlerische Stile (neben den individuellen Stilen der Vasenmaler) – präsentiert: Ab circa 525 vor Christus setzte sich langsam eine neue Maltechnik für Vasenbilder durch. Während zuvor die Bildelemente mit verdünntem Tonschlicker auf den Tongrund aufgemalt wurden und der Tonschlicker durch seine spezifischen Inhaltsstoffe und die Brenntechnik (dem sogenannten Dreiphasenbrand) eine schwarze Färbung annahm, malte man nun den Hintergrund mit eben jenem Tonschlicker und sparte die Bildelemente davon aus. ⁵⁸ Ergebnis hiervon waren zunächst schwarze Figuren auf dem rötlichen Tongrund des Gefäßes (sogenannte schwarzfigurige Vasenmalerei), später rotgrundige Figuren mit schwarzem Hintergrund (sogenannte rotfigurige Vasenmalerei). Zwar waren beide Maltechniken eine Zeit lang noch simultan beliebt und es existieren aus der Übergangsphase gar sogenannte bilingue Vasen, d. h. Gefäße, deren Dekoration auf zwei unterschiedlichen Seiten in unterschiedlichen Stilen ausgeführt ist ⁵⁹, doch gehen mit der historischen Entwicklung der griechischen Vasenmalerei vom sechsten bis fünften Jahrhundert vor Christus auch mitunter große beispielsweise kompositorische Unterschiede zwischen Bildern schwarz- und rotfigurigen Stils einher. Dennoch können einzelne Bildelemente wie z. B. die signifikante körpersprachliche Interaktion zweier Protagonist:innen über einen langen Zeitraum und unterschiedliche Stile hinweg mit erheblicher Konstanz auftreten (siehe oben, Fallbeispiel Frauenführungsszenen).

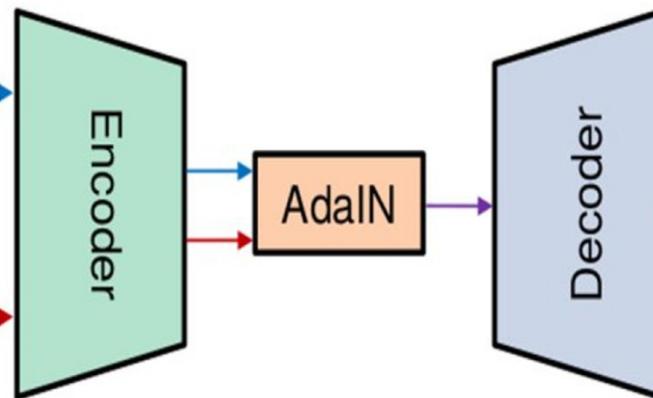
Aufgrund der im Verhältnis zum christlich-archäologischen Bildcorpus deutlich geringeren Heterogenität in den bildlichen Darstellungen der griechischen Vasenmalerei ist im Hinblick auf ihre Erschließung für Computer Vision-Verfahren die Anwendung der sogenannten style transfer learning-Methode sinnvoll ¹⁴. ⁶⁰ Dabei werden die zugrunde liegenden Bilddaten zunächst in zwei Bildgruppen aufgeteilt, die wir Stilbilder (style image) nennen. Das erste Set (RB-set) besteht ausschließlich aus 50 rot- und 50 schwarzfigurigen Vasenbildern. Im zweiten Set (CA-set) wird das gesamte Set der klassisch-archäologischen Vasenbilder als Stilbilder betrachtet. Der nächste Schritt besteht darin, anhand dieser Bildersets ihren Stil auf einen realen Datensatz wie COCO (Common Objects in COntext, content image) zu übertragen. ⁶¹ Es werden also zwei verschiedene COCO-Bilddatensätze mit den jeweiligen spezifischen Stilen des RB-Sets und des CA-Sets erzeugt. Dies ist mit einer echtzeitfähigen style transfer-Methode wie Adaln ⁶² leicht möglich. Wird das ursprüngliche Modell nun erneut auf die neuen COCO-Bilder trainiert, deren Stil transferiert worden ist, lernt das Modell den Inhalt im Kontext der RB/CA-Stile.

Um die Effekte dieser Stilübertragung zu vergleichen, wird das trainierte Modell erneut an unabhängigen (annotierten) Bilddaten der Klassischen Archäologie getestet. Mit dem RB-Stil wird für characters, d. h. menschliche oder anthropomorphe Figuren, die vom Modell methodisch wie ein Objekt behandelt werden, eine average precision von 31% erreicht, mit dem CA-Stil eine von 32%. Dieser geringe Unterschied in der Performanz ist motivierend, denn er bedeutet, dass der style transfer nicht für alle Vasenbilder angewandt werden muss, sondern wenige repräsentative Bilder ausreichen.

Content Image



Style Image



Styled Image

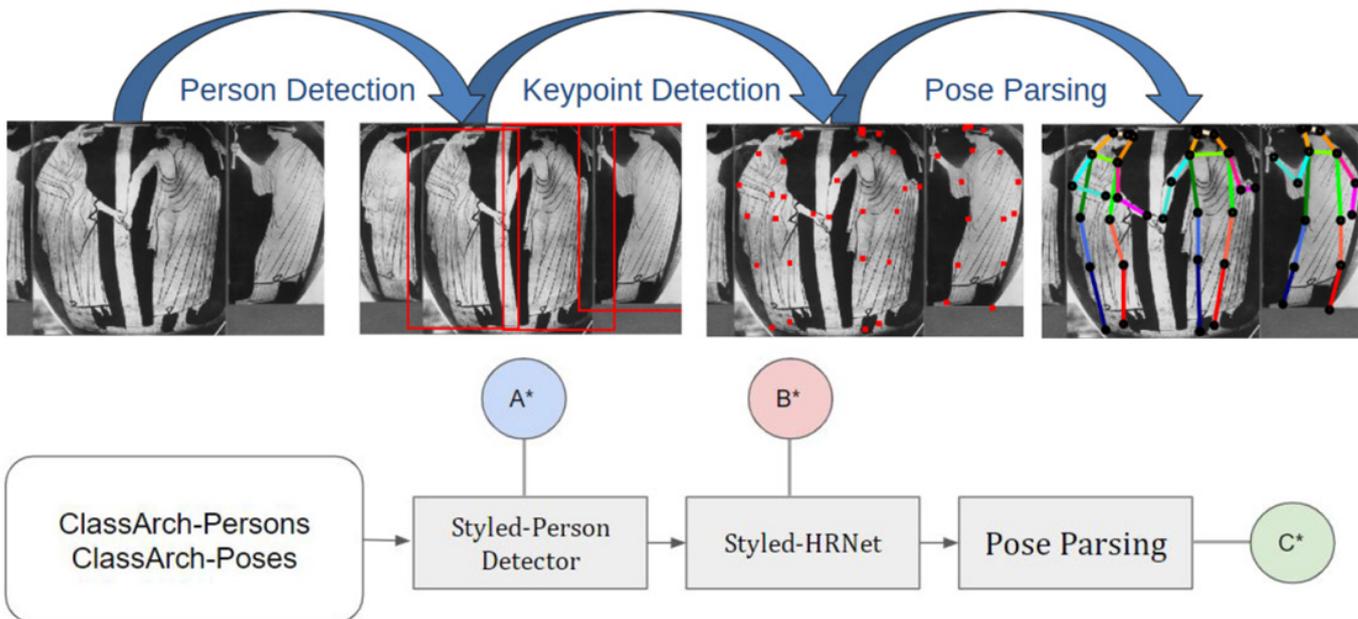


□ 14

Visualisierung der style transfer learning-Methode. (Grafik: Ronak Kostl und Prathmesh Madhu)

In Hinblick auf das Bildmaterial des christlich-archäologischen Fallbeispiels gestaltet sich ein Stiltransfer jedoch weniger einfach, da die behandelten Bildwerke eine deutlich höhere Vielfalt an Artefaktgattungen, Bildfindungen, technischen Ausführungen etc. aufweisen. Für dieses Bildmaterial scheitert daher das style transfer learning. Des Weiteren ist in Hinblick auf historische Bildwerke im Allgemeinen zu beachten, dass künstlerische Konventionen oder konkrete Darstellungsabsichten dazu führen können, dass beispielsweise bestimmte Objekte im Bildwerk optisch nur bedingt ihren realen Vorbildern entsprechen. Auch können Gegenstände in Bildwerken gezeigt werden, die nicht (mehr) in der Lebenswelt vorkommen, daher nicht im zugrundeliegenden Bilderpool dargestellt sind und folglich auch nicht an diesem trainiert werden. Antike Räuchergefäße wie das Thymiaterion, der kultisch bedeutsame Dreifuß (ein dreifüßiger Metallkessel) oder spezielle, häufig in Symposionsszenen

auftretende Gefäßformen wie Kylikes (spezielle Trinkschalen) sind nur wenige Beispiele für in griechischer Vasenmalerei gängige Objekte, die in modernen Fotografien nicht zu erwarten sind. Aus diesem Grund wurde das style transfer learning zwar für die digitale Objekterkennung von characters getestet (siehe oben), eine Anwendung auf eine breite Auswahl von Objekten allerdings von Beginn an ausgeschlossen.



□ 15

Visualisierung der digitalen Erkennung von Körperhaltungen. (Grafik: Ronak Kostı und Prathmesh Madhu)

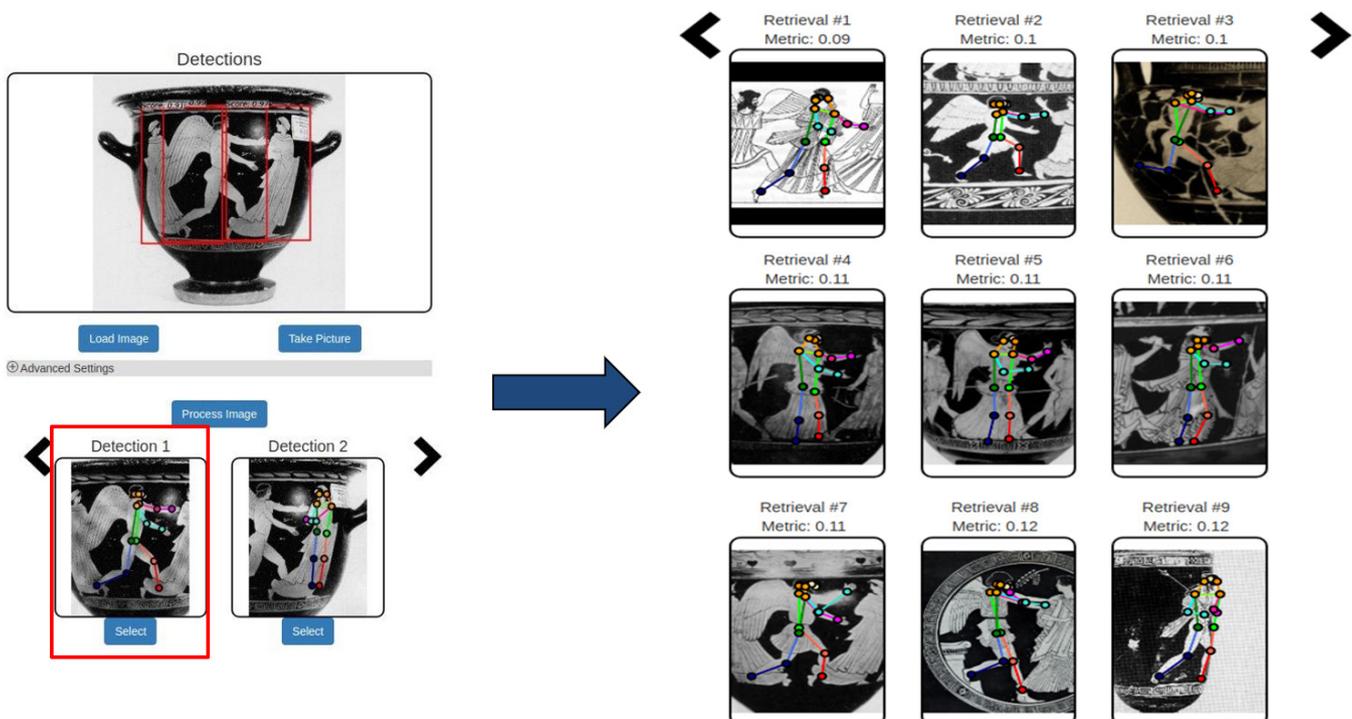
■ 63

Prathmesh Madhu, Angel Villar-Corrales, Ronak Kostı, Torsten Bendschus, Corinna Reinhardt, Peter Bell, Andreas Maier, Vincent Christlein, Enhancing Human Pose Estimation in Ancient Vase Paintings via Perceptually-grounded Style Transfer Learning, *Journal on Computing and Cultural Heritage*, Just Accepted (November 2022). <https://doi.org/10.1145/3569089>. Aus diesem Artikel stammen die Abb. 14 und 15. Siehe dort ausführlicher zum Thema sowie zu den Resultaten des style transfer learning für die digitale Objekterkennung.

Ein sinnvollerer Anwendungsgebiet findet sich hingegen im Bereich der digitalen Erkennung von Körperhaltungen (pose estimation) [15]. Hier kann ein verfügbares Bildersetz wie COCO-Persons mittels style transfer-Verfahren bearbeitet und einem state-of-the-art-Modell der pose estimation, wie zum Beispiel OpenPose, zur Verfügung gestellt werden, um es auf eine bestimmte künstlerische Gattung wie die griechische Vasenmalerei zu trainieren. [63] Da die stilistische Umwandlung moderner Fotografien zu stilisierten Vasenbildern aber größtenteils Aspekte wie Farbschema und Konturendicke betrifft, das auf diese Weise erzeugte Bild jedoch noch immer in der Figurendarstellung stark von originalen Vasenbildern abweicht (vgl. style image und styled image in [14]), wurden zur weiteren Spezialisierung Annotationen von bis zu 18 Gelenkpunkten in Figuren aus über 2.000 Bildwerken angefertigt. Auf diesem Weg konnte ein Modell entwickelt bzw. adaptiert werden, das zunächst automatisch die Figuren in einem Vasenbild erkennt (person detection), daraufhin digital ihre Gelenkpunkte ermittelt (keypoint detection) und diese mittels pose parsing zu Posenskeletten verbindet.

Auf Grundlage einer derartigen digitalen Erkennung von Körperhaltungen für historische Bildwerke sind künftig Anwendungsprogramme ähnlich der sogenannten pose-based image retrieval application vorstellbar, die im Projekt bereits testweise entwickelt wurde [16]. Hierbei ist es den Nutzer:innen

möglich, eine beliebige Abbildung hochzuladen, in welcher das Programm automatisch zunächst die Figuren erkennt (detections) und anschließend ihre Posen bestimmt. Die erzeugten Posenskelette werden daraufhin mit denen der Figuren von Vasenbildern eines großen Bilderpools verglichen und Figurendarstellungen mit ähnlichen Posenskeletten vorgeschlagen (retrievals). Auf diese Weise könnten Forschende erheblich dabei unterstützt werden, auch stil- und kontextübergreifend Bildrelationen zu erkennen, die sich in der Körpersprache der Figuren äußern und daher beispielsweise mit Einzelabfragen in den gängigen textbasierten Bilddatenbanken aktuell nicht ohne Weiteres auffindbar sind.



□ 16

Ausschnitt aus der pose-based image retrieval application. (Tool: Angel Villar-Corrales, Screenshot: Ronak Kosti und Prathmesh Madhu)

Die Ergebnisse zeigen, dass das style transfer learning ein sinnvoller Schritt im Training der Computer-Vision-Modelle sein kann, um deren Leistungsfähigkeit zu verbessern. Voraussetzung hierfür ist allerdings einerseits, dass für einen bestimmten zu trainierenden Stil ausreichend Bildmaterial vorhanden ist. Andererseits stößt die Anwendbarkeit dort auf ihre Grenzen, wo die spezifischen Charakteristika künstlerischer Darstellungen zu gravierenden Unterschieden zu den modernen, zu stilisierenden Fotografien führen, denn (Spät-) Antike und byzantinische Bilder funktionieren nicht wie moderne Alltagsfotografien, weil sie in der Regel nicht die bloße Wiedergabe der Realität zum Ziel haben. Diese Schwierigkeiten können indes teilweise, z. B. für die Darstellungen menschlicher oder anthropomorpher Figuren, durch ein zusätzliches spezielles Training mittels Annotationen überwunden werden.

Fragmentarität der Artefakte

Da das Training der Computer-Vision-Modelle zumeist auf Grundlage vollständig erhaltener Bilder oder Bildelemente erfolgt, stellt sich die Frage, wie sich die Performanz der auf das archäologische Bildmaterial trainierten Modelle hinsichtlich fragmentierter Erhaltungszustände verhält. Diesem Problem wurde am Beispiel der digitalen Objekterkennung nachgegangen. Spezielle Testsets wurden hierfür aus den Bildcorpora beider beteiligter Archäologien erstellt, da die Überlieferungssituation für sie ein gleichermaßen relevanter Faktor der Beschäftigung mit materieller Kultur ist. ⁶⁴

Anders als in benachbarten archäologischen Disziplinen stammt ein Großteil der christlich-archäologischen Fundstücke nicht aus modernen Grabungen, sondern ist bereits seit Beginn der Neuzeit zusammengetragen und archiviert worden. ⁶⁵ Mitunter sind die Stücke auch Teil von Kirchenschätzen und Sammlungen oder sind etwa als liturgische Geräte bis in die Neuzeit in Gebrauch gewesen – als Ausstattungstücke von Bauten sind sie dies zu einem großen Teil bis heute. Diese Umstände haben dazu geführt, dass naturgemäß zunächst besonders schöne, aufwendige und vollständige Artefakte und Ausstattungselemente die Zeit überdauert haben. Hinzu kommt, dass das Fach Christliche Archäologie sich zunächst auf genuin in einen christlichen Entstehungs- oder Nutzungskontext einzuordnende Stücke fokussiert hat und sich die Forschung erst in den letzten Jahrzehnten stärker auch nicht (eindeutig) als christlich anzusprechenden Funden und Befunden der spätantiken und byzantinischen Zeit zugewandt hat. ⁶⁶ Dies hat eine eindeutige Identifizierung der Stücke als christlich – durch auf den Stücken gezeigte Motive oder Inschriften oder einen eindeutig christlichen Kontext – notwendig gemacht.

■ 64

Zu methodischen Implikationen des Aspekts der überlieferungsbedingten Unvollständigkeit von Artefakten: William Tronzo (Hg.), *The Fragment: An Incomplete History*, Los Angeles 2009. Kay Malcher, Stephan Müller, Katharina Philipowski, Antje Sablotny (Hg.), *Fragmentarität als Problem der Kultur- und Textwissenschaften*, Paderborn 2013. Im digitalen Kontext zuletzt: Ursula Schädler-Saub (Hg.), *Das Fragment im digitalen Zeitalter: Möglichkeiten und Grenzen neuer Techniken in der Restaurierung*, Berlin 2021.

■ 65

Vgl. zur Geschichte des Faches: Michael Schmauder (Hg.), *Spiegel einer Wissenschaft: Zur Geschichte der christlichen Archäologie vom 16. bis 19. Jahrhundert dargestellt an Autoren und Büchern: Eine Ausstellung des Christlich-Archäologischen Seminars in der Universitätsbibliothek Bonn, September – Dezember 1991, Bonn 1991*. William H. C. Frend, *The Archaeology of Early Christianity: A History*, London 1996. Sowie: Sabine Schrenk, Ute Verstegen (Hg.), *Forschungsgeschichte als Aufbruch: Beiträge zur Geschichte der Christlichen Archäologie und Byzantinischen Kunstgeschichte, XXIV. Tagung der Arbeitsgemeinschaft Christliche Archäologie, Bonn, 10.–12. Mai 2018, Heidelberg 2022*, <https://doi.org/10.11588/propylaeum.952>.

■ 66

Vgl. dazu die Beiträge in Schrenk, Verstegen (Hg.) 2022.

Unter den hier behandelten Stücken, die mit Darstellungen der Verkündigung oder Anbetung versehen sind, sind folglich weniger fragmentierte Stücke als man zunächst annehmen möchte. Besonders viele dieser Fragmente sind Bruchstücke von Sarkophagen. Dies mag mehrere Gründe haben: Zum einen sind sie aufgrund ihres robusten Materials auch in fragmentiertem Zustand besonders langlebig, zum anderen wurden römische Sarkophage in der Antike gerne außerhalb der Städte entlang der Gräberstraßen, in Mausoleen, in den Katakomben oder den Umgangsbasiliken platziert, die im Mittelalter weit außerhalb der sehr viel kleiner gewordenen Stadtkerne gelegen waren. Dort konnten auch Fragmente überdauern, bis sie in der Neuzeit von interessierten Sammler:innen von Altertümern entdeckt wurden. Da die Darstellung der Anbetung auf Sarkophagen häufig vorkommt und einer festen Ikonografie folgt, kann die Szene auch auf Bruchstücken leicht identifiziert werden ¹⁷. **67**

■ 67

Dresken-Weiland nennt 114 bekannte Objekte, siehe Dresken-Weiland 2010, S. 20.



□ 17

Aufgrund der hohen Standardisierung ist eine Rekonstruktion fragmentierter Sarkophagbruchstücke mit der Darstellung der Magieranbetung mit hoher Wahrscheinlichkeit möglich. (Foto: Wilpert 1932, Taf. CCXXII)

■ 68

Vgl. Sabine Schrenk, *Textilien des Mittelmeerraumes aus spätantiker bis frühislamischer Zeit. Gewebeanalysen: Regina Knaller, Riggisberg 2004 (Die Textilsammlung der Abegg-Stiftung 4)*, S. 319.

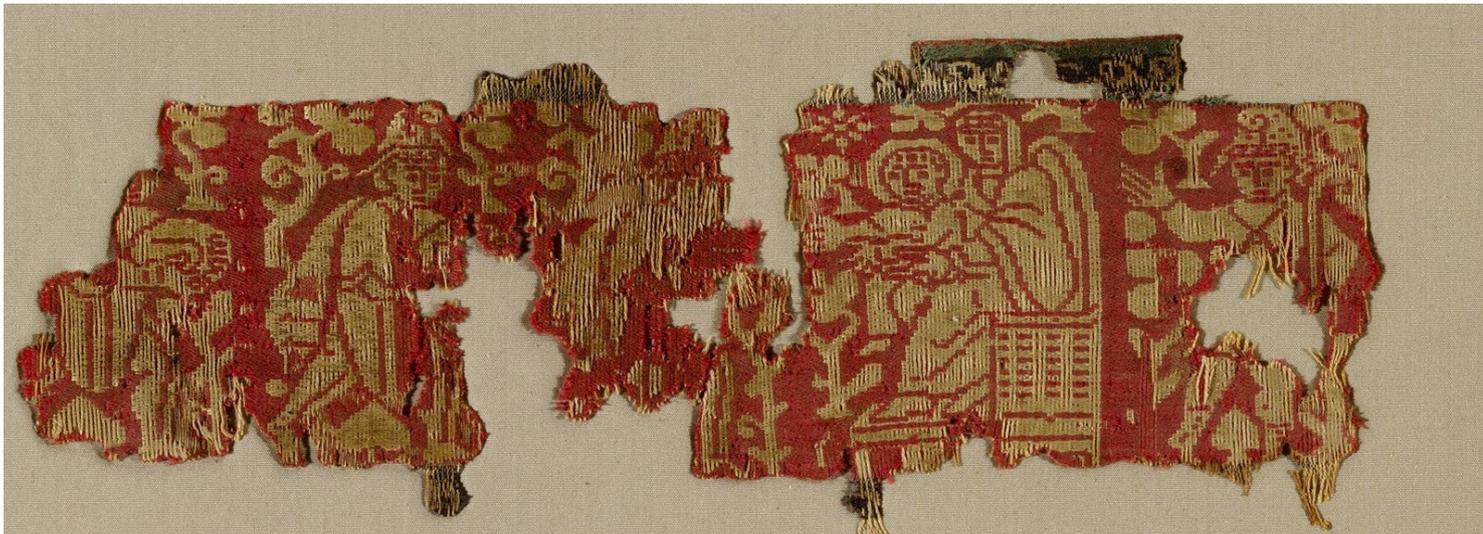
■ 69

Dies gilt bspw. für die Verkündigungsszene auf der sog. Marienseide, die in das späte 4. / frühe 5. Jh. n. Chr. datiert und sich heute in der Abegg-Stiftung, Riggisberg (Inv. 3100b) befindet. Zum magischen Aspekt des Musterrapports vgl. Henry Maguire, *Garments Pleasing to God: The Significance of Domestic Textile Designs in the Early Byzantine Period*, in: *Dumbarton Oaks Papers*, 44 1990, S. 215–224, insb. S. 220. Anders: Arne Effenberger, *Maria als Vermittlerin und Fürbitterin: Zum Marienbild in der spätantiken und frühbyzantinischen Kunst Ägyptens*, in: Leena Mari Peltomaa et. al. (Hg.), *Presbeia Theotokou: The Intercessory Role of Mary Across Times and Places in Byzantium*, Wien 2015, S. 49–108, insb. S. 62.

■ 70

Vgl. dazu bspw. die Obergadenmosaiken in S. Apollinare Nuovo in Ravenna aus dem 1. Viertel des 6. Jhs. n. Chr., die im 19. Jh. fehlerhaft rekonstruiert worden sind. Vgl. dazu Emanuela Penni Iacco, *La basilica di S. Apollinare Nuovo di Ravenna attraverso i secoli*, Bologna 2004.

Wesentlich fragiler und damit seltener sind spätantike und byzantinische Textilien, von denen einige wenige erhaltene Stücke die Verkündigung und die Anbetung zeigen ^[18]. Hierbei stellt sich die, auch webtechnisch bedingte ⁶⁸, spätantike Vorliebe für Rapportmuster als Glücksfall heraus: Auch bei stärker fragmentierten Stücken lässt sich die Darstellung so leicht identifizieren. ⁶⁹ Ebenfalls nur unvollständig überliefert sind einige der Wandmalereien, die die betreffenden Szenen zeigen. Anders als bei Artefaktgattungen wie der Skulptur oder Kleinkunst ist dabei nicht der Bildträger beschädigt, sondern die darauf aufgetragenen Pigment- und mitunter Putzschichten. Nur selten sind die Malereien aus den teilweise bis heute genutzten Kirchenräumen im originalen Zustand erhalten, weil sie meist über die Jahrhunderte immer wieder übermalt oder sogar gänzlich ersetzt wurden. Dabei ist es im Laufe der Auffrischungs- und Restaurierungsarbeiten immer wieder zu Abänderungen und fehlerhaften Rekonstruktionen gekommen. ⁷⁰ Computergestützte Verfahren könnten in Zukunft durch einen Abgleich mit vergleichbarem Material die Chance bieten, Indizien für fehlerhafte Restaurierungen oder Fälschungen zu liefern.



□ 18

In Syrien gefertigtes Textil mit Darstellung der Magieranbetung (links) und der Verkündigung (rechts außen). Fundort: Ägypten, 5./6. Jh. n. Chr. (Foto: © New York, Metropolitan Museum of Art, Inv. 90.5.11a-e, lizenziert unter CC0, <https://www.metmuseum.org/art/collection/search/443592>)



Für die griechische Keramik liegen aufgrund der Langlebigkeit des Tonmaterials und der Beständigkeit der eingebrannten Tonschlickermalerei zwar gute Erhaltungsbedingungen vor, doch sind meist nur zerbrochene Gefäße überliefert, die oft teils beträchtliche Fehlstellen aufweisen oder gar nur als einzelne Fragmente erhalten sind ^[19]. Grund hierfür ist bei bemalter Feinkeramik – je nach Größe und Gefäßform – die mitunter sehr geringe Dicke des Gefäßkörpers und die damit verbundene Fragilität des Bildträgers. ⁷¹ Ist nur ein Fragment erhalten, erlauben Form, Dicke und Krümmung des Bruchstückes oft eine sehr präzise Rekonstruktion der ursprünglichen Gefäßform. Demgegenüber steht aber eine große Anzahl vollständig erhaltener Gefäße, die aus ungestörten Grabkontexten stammen. Diese sind als Grundlage für die Ergänzung von Fragmenten von entscheidender Bedeutung, da man die Darstellung anhand signifikanter Bildelemente durch Darstellungskonventionen und den repetitiven Gebrauch von Bildschemata häufig mit hoher Plausibilität rekonstruieren kann.

■ 71

Einführend zu den gängigsten Gefäßformen in der griechischen Keramik bspw. Ingeborg Scheibler, Gefäße/Gefäßformen-Typen, in: Der Neue Pauly (DNP) Band 4, Stuttgart 1998, Sp. 853–858.



□ 19

Zwei Beispiele nur fragmentarisch erhaltener Frauenführungsszenen: das Ergreifen einer Frau durch einen Krieger auf schwarzfigurigen Amphorabruchstücken (links), 540–520 v. Chr., und das Ergreifen einer Braut durch den Bräutigam auf Fragmenten einer rotfigurigen Loutrophore des Frauenbad-Malers (rechts), 430–400 v. Chr. (beide Fotos: © The Trustees of the British Museum, Inv. 1888,0208.114 [B 125.13] bzw. 1896.12-17.11)

Die erstellten Testsets mussten die geschilderten Bedingungen adäquat repräsentieren, um belastbare Ergebnisse zu produzieren. Zudem musste berücksichtigt werden, welche Objektklassen zuvor bereits trainiert wurden und somit dem Modell bekannt waren. Das Bildcorpus der Christlichen Archäologie wurde daher hinsichtlich nur eines Objekts, der phrygischen Mütze, untersucht. Diese ist ein für die Ikonografie der Anbetung äußerst wichtiges Objekt, mit dem die Magier aus dem Osten versehen sind. Die Kopfbedeckung eignet sich nicht nur wegen ihrer charakteristischen Form, die auf allen Medien leicht erkennbar ist, ausgesprochen gut, sie ist zudem in nahezu allen Gattungen vertreten. Darüber hinaus sind rund 1.500 Abbildungen mit Darstellungen phrygischer Mützen für das Vergleichscorpus (bestehend aus nicht-christlichen römischen Artefakten, frühchristlichen und byzantinischen Stücken, die andere Szenen zeigen, sowie westlichen mittelalterlichen Bildwerken) zusammengetragen worden. Es ist ein Trainingsset von 130 Abbildungen zusammengestellt worden, von denen 30 genuine Fragmente zeigen und weitere 100 durch Bearbeitungen mit einem Bildbearbeitungsprogramm erzeugt worden sind. Hierbei ist darauf geachtet worden, den Verlauf der künstlichen Bruchkanten

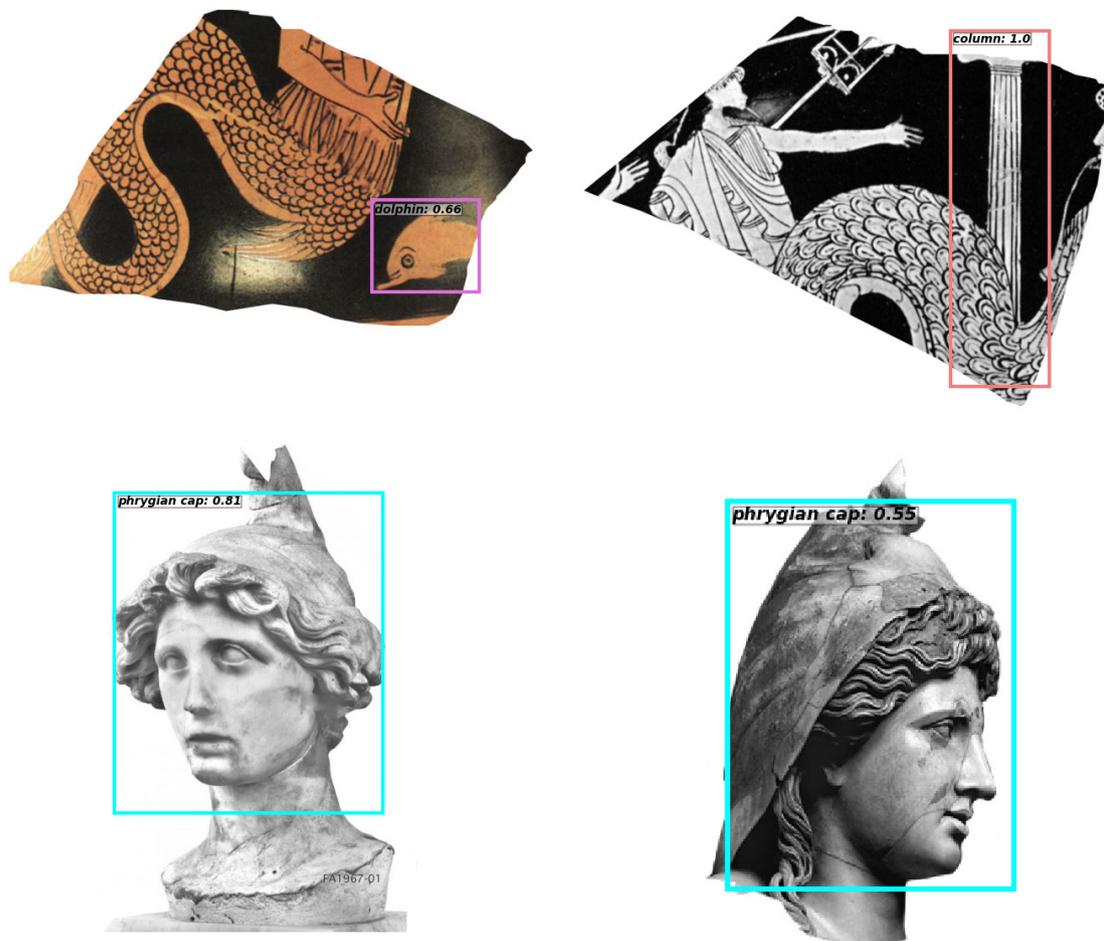


■ 72

Die Prozentangaben in Abb. 20 geben demgegenüber lediglich die Zuversicht (confidence) an, die der Algorithmus seinen eigenen Prognosen zuschreibt. Für das Bildmaterial der Christlichen Archäologie liegt diese für die Originalbilder genuiner Fragmente bei 68%, für die künstlich erzeugten Fragmentbilder bei 73%. Im Fall der Klassischen Archäologie liegt die durchschnittliche confidence beim gesamten Testset fragmentierter Bilder bei 76%. Einen Aussagewert für die Korrektheit der Prognosen besitzen diese Angaben nicht und können daher im Folgenden ignoriert werden.

möglichst naturalistisch aussehen zu lassen. Diese Methode ermöglicht außerdem die Simulation unterschiedlicher Erhaltungszustände (auch auf Grundlage des gleichen Bildes), um die Performanz anschließend vergleichen zu können. Ähnlich wurde mit dem Bildmaterial der Klassischen Archäologie verfahren, bei dem 400 künstlich simulierte Fragmente durch 50 Abbildungen von (realen) Gefäßbruchstücken ergänzt wurden. Entsprechend dem Training auf ein breites Spektrum an Objekten (siehe oben) wurden für diese Bilder Darstellungen ausgewählt, die eines oder mehrere der 81 trainierten Objektklassen zeigen.

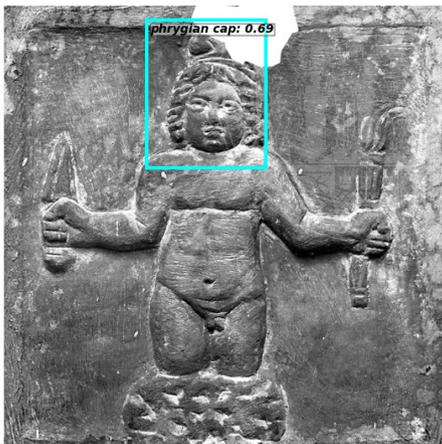
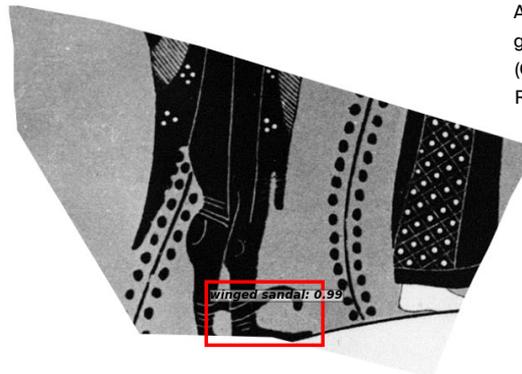
Um die Leistungsfähigkeit der Objekterkennungsmodelle zu ermitteln, ist das entwickelte und auf die Objektklassen der Christlichen und Klassischen Archäologie trainierte object detection-Verfahren unverändert auf die realen und simulierten Fragmente angewandt worden [20]. Für numerische Aussagewerte zur durchschnittlichen Genauigkeit (average precision) der Prognosen wäre indes ein aufwändigerer Versuch inkl. Annotationen der Fragmentbilder notwendig gewesen, wie er zur digitalen Objekterkennung beschrieben wurde (siehe oben). 72



Vor dem Hintergrund des relativ kleinen Bildersets von Fragmenten wurde die manuelle Durchsicht der Resultate durch aus dem Fach stammende Expert:innen als hinreichend erachtet. Diese Sichtung der Ergebnisse legt nahe, dass die Performanz der digitalen Objekterkennung für unvollständige Objekte deutlich schlechter ist als für vollständige. Diese Beobachtung entspricht den Erwartungen, da auch die Leistung menschlicher Expert:innen bei der Erkennung von Objekten durch einen schlechteren Erhaltungszustand in der Regel deutlich eingeschränkt ist. Zu erwarten wäre ebenso, dass ein umfangreicheres Training eine bessere Befähigung des Modells bei unvollständigen Darstellungen gewährleisten sollte. Überraschend ist in dieser Hinsicht aber die Beobachtung, dass beispielsweise das Objekt Säule im Testset der Klassischen Archäologie gegenüber anderen Objektklassen am häufigsten falsch prognostiziert wurde, obgleich es im Training unter den am häufigsten annotierten Objekten an fünfter Stelle steht.

□ 20

Unterschiedliche Beispiele für die erfolgreiche digitale Objekterkennung trotz fragmentarischer Darstellungen. In der oberen Reihe wurden verschiedene Objekte (Delfin, Säule, Kerykeion u. Flügelschuh) in Bildern griechischer Vasenmalerei, in der unteren Reihe die Phrygische Mütze in verschiedenen Abbildungen aus dem christlich-archäologischen Bildcorpus korrekt erkannt. (Objekterkennung durchgeführt von: Ronak Kosti und Prathmesh Madhu)



■ 73

Nada A. Rasheeda, Jan Nordin,
**Classification and Reconstruction
 Algorithms for the Archaeological
 Fragments**, in: *Journal of King Saud
 University - Computer and Information
 Sciences*, 32 (8) 2020, S. 883–894,
[https://doi.org/10.1016/j.
 jksuci.2018.09.019](https://doi.org/10.1016/j.jksuci.2018.09.019).

Der Test der digitalen Objekterkennung wirft demnach viele Fragen für die weitere methodische Forschung und Entwicklung auf. Sie bestätigen aber auch die Notwendigkeit weiterer Forschungen auf diesem Gebiet, denn schlechte oder fragmentierte Erhaltungszustände stellen eine grundlegende Herausforderung im Umgang mit archäologischem Bildmaterial dar. Dementsprechend vielfältig und gewinnbringend können die Anwendungsmöglichkeiten für Computer Vision-Modelle sein, die imstande sind, automatisiert zuverlässige Rekonstruktionen zu erzeugen. Neuronale Netzwerke zur Klassifizierung und Rekonstruktion antiker Gefäße anhand von Keramikfragmenten befinden sich beispielsweise in Entwicklung und haben das Potenzial, eine sinnvolle Unterstützung für die Bestimmung fragmentiert überlieferter Keramikobjekte zu bilden. ⁷³ Ausgangspunkt hierfür sind die geometrischen Merkmale der Gefäßkörper, wofür durch den massenhaften Erhalt unverzierter Gebrauchskeramik und die starke Standardisierung ihrer Formen und Größen eine deutlich bessere Grundlage als für die Bilder bemalter Feinkeramik vorliegt. Eine robuste Objekterkennung anhand unvollständiger Darstellungen von Objekten ist dennoch ein erster Schritt, auf ähnliche Weise weitere Bildelemente bis hin zu ganzen Ikonografien automatisiert zu rekonstruieren. Ergänzungen von Bildwerken, sowohl durch Expert:innen als auch durch Algorithmen, bleiben zwar immer hypothetisch, können aber hohe Plausibilität beanspruchen, wenn sie auf der Kenntnis repetitiv gebrauchter Schemata und Darstellungskonventionen basieren. Denkbar ist eine Applikation, die Bildelemente digital erfasst und nach Abgleich mit einem umfassenden Bildcorpus auf Basis des Vorkommens bestimmter Körperhaltungen und/oder Objekte Vorschläge für ähnliche Darstellungen unterbreitet.

Dreidimensionalität der Artefakte

Im Gegensatz zu den digitalen Bilddaten, die keine Materialität besitzen, sind historische Bilder mit ihrem Träger und seiner dreidimensionalen Gestalt eng verbunden. Seine Form beeinflusste gleichermaßen die Bildkonzeption und Bildrezeption ⁷⁴, wie einige Beispiele aus der antiken Kunst zeigen. Vor allem mobile Bildträger mit gekrümmter Oberfläche wie Gefäße aus verschiedensten Materialien waren hier eine Herausforderung für die Handwerker/Künstler, mussten sie doch die antike Betrachtungsweise der Artefakte mit einbeziehen, wenn sie einen bestimmten Eindruck erzeugen wollten. Bei griechischer Keramik handelt es sich oftmals um Gefäße, die beim gemeinsamen, fest im Sozialleben verankerten Gelage, dem sogenannten Symposion, benutzt wurden – zum Weinmischen, Schöpfen oder zum Trinken. Nahm man beispielsweise eine Trinkschale in die Hand und führte sie zum Mund, sah man zunächst die bemalte Außenseite, blickte entlang des eingezogenen Randes nach oben auf den Wein und sah spätestens nach dem Austrinken schließlich das Innenbild. Durch die Benutzung des Gefäßes, sein Drehen und Wenden, wurde das Bild darauf demnach ursprünglich dreidimensional, also in mehreren Ansichten, gesehen und die verschiedenen Ansichten komplementär verstanden. ⁷⁵

■ 74

Welche Nutzungsmöglichkeiten ein Gegenstand durch seine Gestaltung (und seinen Bilddekor) anbietet, wird unter dem Begriff der Affordanz untersucht, siehe Richard Fox, Diamantis Panagiotopoulos, Christina Tsouparopoulou, Affordanz, in: Thomas Meier et. al. (Hg.), *Materiale Textkulturen: Konzepte – Materialien – Praktiken*, Berlin – Boston 2015 (*Materiale Textkulturen*, Band 1), S. 63–70, <https://doi.org/10.1515/9783110371291.63>. Zum teils fehlenden Zusammenhang zwischen Typologie-Zugehörigkeit, Affordanz und Bilddekor griechischer Gefäße: Nikolaus Dietrich, *Affordanzen, Typen und Bilddekor im Widerstreit: Zu einem Phänomen antiker materieller Kultur an Beispielen aus der archaischen Luxuskeramik und der kaiserzeitlichen Wohnarchitektur*, in: Elisabeth Günther et. al. (Hg.), *Mehrdeutigkeiten, Rahmentheorien und Affordanzkonzepte in der archäologischen Bildwissenschaft*, Wiesbaden 2021 (*Philippika 147*), S. 105–139. Eine exemplarische Studie zur sog. Exekias-Schale: Martina Sauer, *Von Bildimpulsen zu Vitality Semiotics: Affordanz und Rahmen (frames) aus kunstgeschichtlicher Sichtweise am Beispiel der Exekias-Schale* in München, in: Elisabeth Günther et. al. (Hg.), *Mehrdeutigkeiten, Rahmentheorien und Affordanzkonzepte in der archäologischen Bildwissenschaft*, Wiesbaden 2021 (*Philippika 147*), S. 79–104.

■ 75

Zum partiellen Sehen von Bildern auf antiker Symposions-Keramik siehe Nikolaus Dietrich, *Levels of Visibility and Modes of Viewing in Attic Vase-Painting*, in: Achim Lichtenberger (Hg.), *The Diversity of Classical Archaeology*, Turnhout 2017 (*Studies in Classical Archaeology*, Band 1), S. 303–322.

■ 76

Zu den Standards der Gefäßfotografie durch das *Corpus Vasorum Antiquorum* siehe Hansjörg Bloesch, *Conseils pour photographier les vases grecs*, in: Charles Dugas et. al. (Hg.), *Annexe II. Colloque international sur le Corpus vasorum antiquorum*, Lyon, 3–5 juillet 1956, Lyon 1957, S. 31–33.

Expert:innen ergänzen oft gedanklich den historischen Rezeptionsprozess und ziehen beispielsweise weitere ausgewählte Fotografien hinzu, wenn das Original nicht direkt verfügbar ist, um das Bild in seiner Vollständigkeit zu verstehen und zu zeigen; bisweilen werden in den Publikationen auch Abrollungen besonders langer Bildfriese angeboten, die den Bilddekor komplett und entzerrt zeigen.

Eine vorherige manuelle Auswahl von Fotografien ist bei einer automatisierten Verarbeitung einer großen Anzahl von Bilddaten nicht sinnvoll leistbar. Aber auch wenn vorher jeweils eine Ansicht gewählt würde, auf der alle Bildelemente bestmöglich zu sehen sind, entsteht für die Computer Vision das Problem, dass die notgedrungen fotografisch verzerrten Wiedergaben der Bildfelder und ihrer Bildelemente in ihren Grundformen rezipiert, analysiert und verglichen werden. Ähnlichkeiten und Unterschiede werden somit im Bildvergleich auf dieser problematischen Grundlage trainiert und ermittelt. Die darauf aufbauende Analyse – etwa der Körperhaltungen – birgt ebenso Probleme: So kann durch einen nicht-planen Bildträger in einer Fotografie ein falscher Eindruck einer Gestalt in Hinblick auf ihre Körperproportionen oder generell von Figurenzusammenhängen durch ihre Körperhaltungen entstehen. Beispielsweise erscheinen auf den Aufnahmen der Vasenbilder die Oberkörper von Figuren oftmals verkürzt, wenn sie – zum Beispiel bei Amphoren oder Hydrien – über die Krümmung von Schulter zum Gefäßbauch gezeichnet sind oder wenn die Trinkschalen vollständig auf den stark gekrümmten Außenseiten bemalt sind. Durch schräge Blickwinkel beim Fotografieren können zwar häufig Partien mit sichtbar gemacht werden, die sonst durch eine (nur) frontale Aufnahme unsichtbar wären. ⁷⁶ Auch ein solcher Blickwinkel verzerrt jedoch beispielsweise den Eindruck von Figurenzusammenhängen, da die Figuren durch die Krümmung des Gefäßkörpers stärker zueinander geneigt erscheinen als sie es tatsächlich sind. Dies geschieht besonders häufig außerhalb der Bildmitte bei den rahmenden (Bei-)Figuren am rechten und linken Bildrand von quereckigen Bildfeldern, wie sie etwa auf Amphoren oder Hydrien, aber auch auf Trinkschalen üblich sind (vgl. ⁰¹, rechte Figur im Bildfeld). Durch die Rundung des Gefäßes und seine zusätzliche Krümmung nach oben scheinen sich die Figuren auf den Fotografien zur Bildmitte hin zu neigen. Ohne eine Modifikation des Bildes oder die Benutzung mehrerer Aufnahmen lassen sich daher beispielsweise Fragen, die sich auf bestimmte Körperhaltungen der Protagonist:innen innerhalb eines Bildes und deren Wiederholung durch Beifiguren beziehen, kaum mittels einer computergestützten Erkennung von Körperhaltungen beantworten. Gerade dieser Aspekt ist jedoch bei den oben genannten Frauenführungsszenen für das Verständnis des Bildes wichtig, weil bisweilen ein zweiter Krieger das Krieger-Frau-Paar rahmt (vgl. ⁰¹) oder ein zweites Paar hinzugefügt wird.

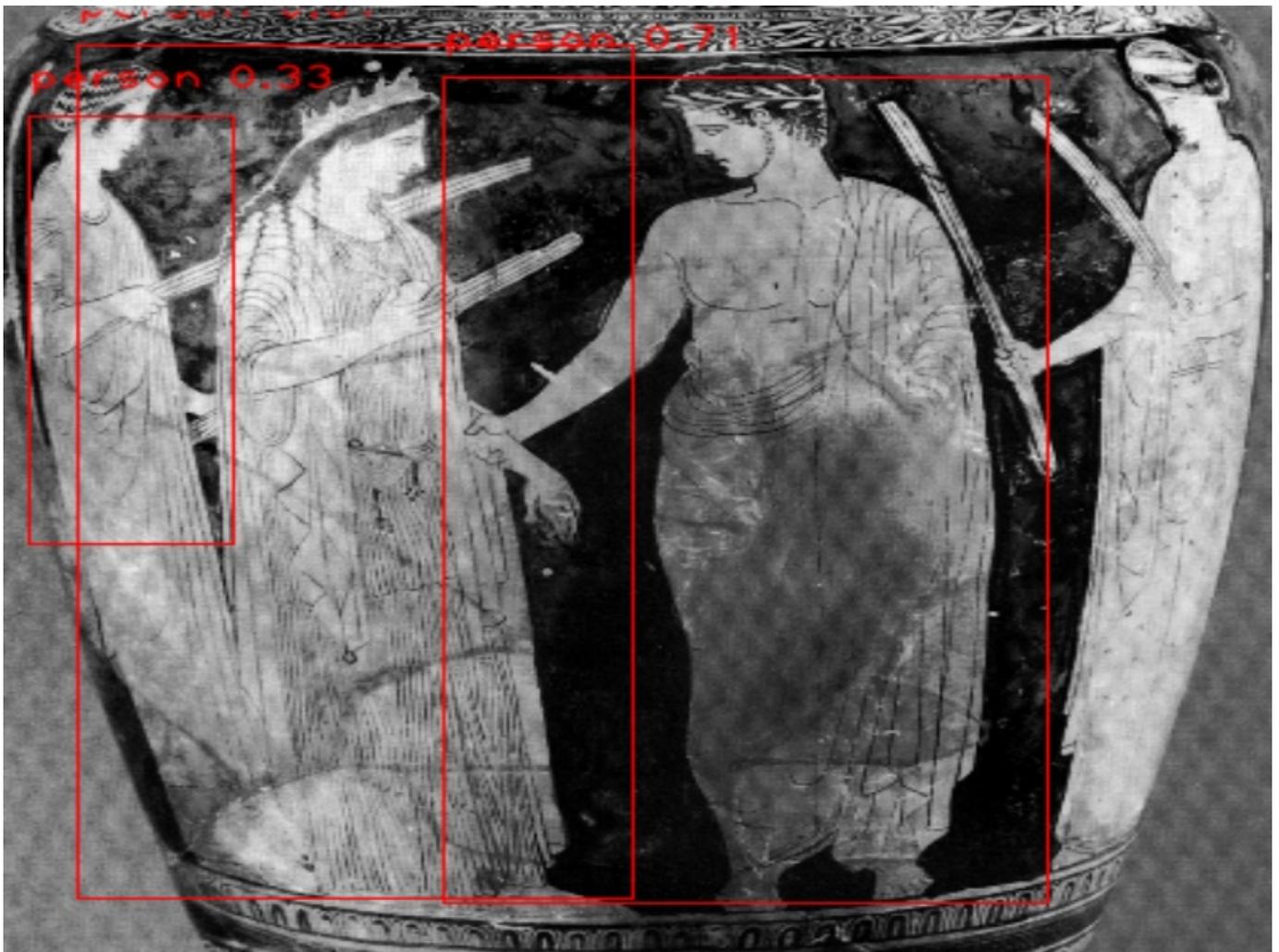
■ 77

Eine Variante der oben beschriebenen Frauenführungsszenen in der schwarzfigurigen Vasenmalerei ist das Führen am Gewand (vgl. [01]). Dabei greift der Krieger den Schleier der Frau ungefähr auf Hüfthöhe. Auf einer Halsamphora im Museo Archeologico Nazionale von Neapel (BAPD 9027204) fasst der Krieger den Schleier hingegen auf Schulterhöhe, d. h. dort, wo die Hand der Frau im anakálypses-Gestus zu erwarten wäre. In einer anderen populären Variante ist der Arm des Kriegers hinter die Frau an deren Rücken geführt, im Innenbild einer Schale in New York (BAPD 13330) hingegen zwischen Schleier und Peplos der Frau gestreckt.

Es ist zwar davon auszugehen, dass auch der antike Betrachter:innenblick eine gewisse Unschärfe in Hinblick auf Ähnlichkeitsabstufungen umfasst, weil in den Bildern durchaus deutliche Varianzen innerhalb sicher zu vergleichender Körperhaltungen belegt sind. [77] Diese decken sich jedoch wohl kaum mit den Bildverzerrungen, mit denen die automatische Bildanalyse umzugehen hat. Bisher stand bei den Bilderkennungen vor allem die kompositorisch hervorgehobene und weniger verzerrte Bildmitte im Vordergrund, während zu den Seiten hin sowohl Objekte als auch überhaupt die Figuren und ihre Körperhaltungen nur zum Teil oder gar nicht erkannt wurden [21].

□ 21

Figurenerkennung beim Vasenbild einer rotfigurigen Loutrophore des Polygnotos in Toronto (Royal Ontario Museum of Archaeology, Inv. 929.22.3), 450–430 v. Chr. In der Frauenführungsszene einer Hochzeitsprozession werden Braut und Bräutigam (Bildmitte) erkannt, während die flankierenden Fackelträgerinnen nicht (rechts) bzw. nur teilweise (links) detektiert werden. (Figurenerkennung durchgeführt von: Ronak Kostı und Prathmesh Madhu)



Ähnlich verhält es sich auch bei christlich-archäologischen Artefakten unterschiedlicher Gattungen. Eine Vielzahl spätantiker und byzantinischer Bilder sind auf Bildträgern aufgebracht, die entweder selbst bewegt werden sollten oder um die herum eine Bewegung der Betrachter:innen stattgefunden hat. Handliche Gegenstände wie Pyxiden, Weihrauchgefäße oder Silberkästchen sind oft rundum mit bildlichen Darstellungen ausgestaltet. Diese sind nicht alle gleichzeitig sichtbar, sondern es bedarf wie bei den genannten griechischen Gefäßen der Drehung des jeweiligen Artefakts. Ein Beispiel ist eine figürlich dekorierte Silberflasche in New York ⁷⁸, welche eine Magieranbetung mit Sonderikonografie zeigt: Die Magier werden durch einen Engel geführt, was bei ausschnitthafter Darstellung von Maria und Engel zur Verwechslung mit einer Verkündigungsszene führen kann. Wandbilder, Bodenmosaiken oder Sarkophage hingegen sind Gegenstände, um die herum bzw. über die sich die Betrachter:innen bewegen müssen, um die einzelnen Aspekte der Bilder sehen, entschlüsseln und im Rezeptionsprozess miteinander verknüpfen zu können. Die Ausgestaltung von Kirchenräumen mit Malerei und Mosaik ist mitunter bewusst so angelegt, dass es einer Bewegung des Menschen wie der Drehung des Kopfes bedarf, um zusammengehörige Szenen sehen zu können, etwa wenn Maria und Gabriel in der Szene der Verkündigung auf einander gegenüberliegende Wandpartien gesetzt sind ²². ⁷⁹ Blickbeziehungen oder Zeigegesten von Figuren, die sich auf andere Bildorte im selben Raum beziehen, können ohne körperliche Wendung der Betrachtenden nicht erfasst werden. Fotos, die beispielsweise den Bilderfries einer Pyxis in abgerollter Form zeigen oder zwei in der Realität nicht direkt nebeneinander angebrachte Malereien derselben Szene kompilieren, vermögen es nicht, die Zusammenhänge der Figuren und Objekte so abzubilden, wie sie sich in der realen Betrachter:innenperspektive eigentlich darstellen. Hinzu kommt, dass die Bildorte der Darstellungen in architektonischen Räumen teilweise sehr weit oben liegen (etwa in einem Kirchenraum) oder Bilder für die Draufsicht (etwa bei einem Fußbodenmosaik) konzipiert sind. Bei entsprechenden Bildorten ist zu berücksichtigen, dass wahrnehmungsbedingte Verzerrungen in der künstlerischen Konzeption bewusst ausgeglichen worden sein können. Eine Fotografie, die einen entsprechenden Bildort frontal aufnimmt (beispielsweise eine Drohnenaufnahme eines Mosaikbodens), vermag es nicht, diese Effekte zu berücksichtigen. ⁸⁰ Umgekehrt entstehen bei in Schrägansicht aufgenommenen Fotografien möglicherweise Verzerrungen, die vor Anwendung automatisierter Verfahren erst ausgeglichen werden müssten.

■ 78

New York, Metropolitan Museum of Art, Inv. 1984.196. Zu dem Stück siehe Helen C. Evans, Robert Hallman, Melanie Holcomb (Hg.), *The Arts of Byzantium*, in: *The Metropolitan Museum of Art Bulletin*, n.s. 58 (4) 2001, S. 23.

■ 79

Zum Bildort der Verkündigung in byzantinischen Kirchausstattungen: Maréva U, *Images et passages dans l'espace ecclésial à l'époque médiobyzantine*, in: Sulamith Brodbeck et al. (Hg.), *Visibilité et présence de l'image dans l'espace ecclésial: Byzance et Moyen Âge occidental*, Paris 2019 (Publications de la Sorbonne, Série Byzantina Sorbonensia, Band 30), S. 303–327, <http://books.openedition.org/psorbonne/39842>.

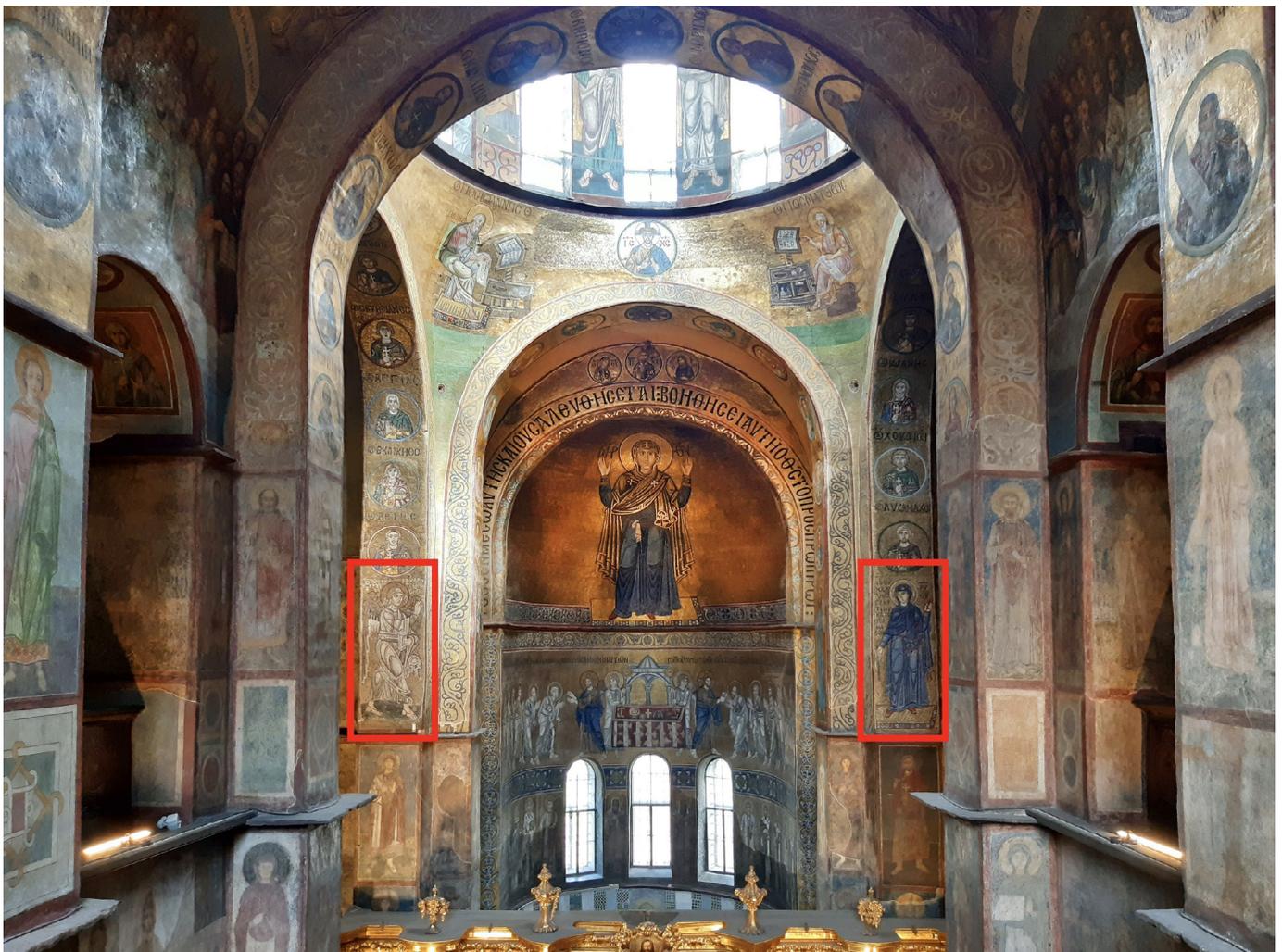
■ 80

Vgl. dazu Lukas Oberem, *Zur Kopfhaltung bei der Betrachtung von Deckenmalerei*, in: Stephan Hoppe et al. (Hg.), *Digitale Raumdarstellung: Barocke Deckenmalerei und Virtual Reality*, Heidelberg 2020 (Computing in Art and Architecture, Band 4), S. 288–307, <https://doi.org/10.11588/arthistoricum.774.c10133>.

Um eine größtmögliche automatisierte Vergleichbarkeit von Bildern auf dreidimensionalen Artefakten oder in architektonischen Räumen zu gewährleisten, wäre daher ein Rückgriff auf 3D-Scandaten der Monumente und Artefakte wünschenswert.

□ 22

Die Verkündigungsszene der die Apsis rahmenden Bogenlaibungen der Zentralkuppel (links Gabriel, rechts Maria) spannt sich über den Raum auf und erfordert die Drehung des Kopfes der Rezipient:innen. Kiew, Sophienkathedrale, 11. Jh. n. Chr. (Foto: Rasal Hague, Wikimedia Commons, lizenziert unter Creative-Commons-Lizenz BY-SY-4.0, [https://commons.wikimedia.org/wiki/File:Интер%27ер_Софія_\(20210903\).jpg?uselang=de](https://commons.wikimedia.org/wiki/File:Интер%27ер_Софія_(20210903).jpg?uselang=de))



■ 81

Zur Anwendung von 3D-Mustererkennung archäologischer Artefakte: Martin Langner et al., Schemata – 3D Classification Methods and Archaeological Identification Criteria: An Interdisciplinary Collaboration Using the Example of Ancient Terracotta Statuettes, in: *International Journal for Digital Art History*, 8 2021, S. 22–37, <https://doi.org/10.11588/dah.2021.E1.83925>.

■ 82

Grundlegend zu Potentialen und Grenzen von 3D-Scans archäologischer Monumente und Artefakte die Publikationen und Richtlinien der Projekte 3D-ICONS und 3D-COFORM: <http://3dicons-project.eu/guidelines-and-case-studies>, <https://cordis.europa.eu/project/id/231809/results/de>. Zur Problematik des Materials des zu scannenden Artefakts: Pedro Santos, et al., 3D Mass Digitization: A Milestone for Archeological Documentation, in: *Virtual Archaeology Review*, 8 (16) 2017, <https://doi.org/10.4995/var.2017.6321>. Jüngste Ansätze zum erfolgreichen Scannen von hochglänzenden Metalloberflächen: Reimar Tausch et al., Towards 3D Digitization in the GLAM (Galleries, Libraries, Archives, and Museums) Sector: Lessons Learned and Future Outlook, in: *IPSI Transactions on Internet Research*, 16 (1) 2020, S. 45–53. Chiara E. Catalano et al., Challenges in the Digitisation of a High-Reflective Artwork, in: Alan Chalmers et al. (Hg.), *GCH 2021: Eurographics Workshop on Graphics and Cultural Heritage*, 2021, S. 141–149. Siehe auch thematisiert in: Peter Houska et al., Direct Elastic Unrollings of Painted Pottery Surfaces from Sparse Image Sets, in: Alan Chalmers et al. (Hg.), *GCH 2021. Eurographics Workshop on Graphics and Cultural Heritage*, 2021, S. 131–140, insb. S. 132, <https://diglib.eg.org/bitstream/handle/10.2312/gch20211417/131-140.pdf>. Siehe generell den Sammelband: Kate Kelley, Rachel K. L. Wood (Hg.), *Digital Imaging of Artefacts: Developments in Methods and Aims*, Oxford 2018.

■ 83

Peter Houska et al., 2021.

■ 84

Peter Houska et al., 2021, S. 137–138 m. Abb. 12–13.

Ausgangspunkt der hier angesprochenen Computer-Vision-Modelle können allerdings momentan nur die digitalen Fotografien sein, die in großer Anzahl durch langjährige systematische Fotokampagnen vorliegen. Dreidimensionale Scans sind demgegenüber bislang nicht nur wegen des hohen Aufwands, der zum Teil großen Stückzahlen und wegen des Schutzaspekts wenig üblich, auch wenn archäologische Artefakte zunehmend in 3D digitalisiert werden. ⁸¹ Gerade für einige Gattungen, die im Projekt relevant sind, gibt es aktuell noch schwerwiegende Problematiken, die einen 3D-Scan erschweren, zum Beispiel wenn die genannten griechischen Gefäße eine hochglänzende Oberfläche zeigen (Glanztomalerei), oder andere Artefakte aus durchsichtigem Material wie Glas oder Edelsteinen oder aus hochglänzenden Edelmetallen gefertigt sind. ⁸²

An einer technischen Problemlösung ist im Rahmen des Projekts noch nicht geforscht worden, so dass an dieser Stelle nur Überlegungen theoretischer Natur folgen können. Das Problem ist in der Forschung jedoch bekannt und ein Lösungsansatz für die Verzerrungen in den Fotografien griechischer Vasenbilder ist jüngst vorgestellt worden. ⁸³ Die Methode, die zum Ziel hat, Darstellungen der Bilder digital entzerrt zu erzeugen, basiert auf einem Set aus mehreren Ansichten des gleichen Vasenbildes, wobei die Aufnahmen gleichzeitig die Rotationssymmetrie des Gefäßkörpers beibehalten müssen. Während auf diesem Weg komplette Abrollungen der Vasenbilder entstehen, die auch Bildabschnitte beinhalten, welche durch die Betrachtenden nicht aus einer einzigen Perspektive heraus erfassbar wären, wurde die Methode auch erfolgreich für einzelne Abbildungen und Ansichten aus unterschiedlichen Aufnahmewinkeln angewandt. ⁸⁴

Inwiefern diese Methode für unsere Fragestellungen gewinnbringend eingesetzt werden kann, ist noch exemplarisch zu prüfen. Für bestimmte Fragen an antike Vasenbilder wird sie einen wichtigen Fortschritt bedeuten. So entstehen durch das Entrollen vollständiger Vasenbilder auf Basis mehrerer Fotografien Rekonstruktionen, die es ermöglichen, das Vasenbild als Ganzes zu analysieren und mögliche (fotografische) Verzerrungen durch die Materialität des Gefäßkörpers auszublenden. Mit der gleichen Intention werden bereits seit langer Zeit ergänzend zu Fotografien gelegentlich auch entrollte Zeichnungen der Vasenbilder in Fachpublikationen abgedruckt. In dieser Hinsicht können automatisiert erstellte Rekonstruktionen eine mit deutlich geringerem Aufwand verbundene Alternative zu ohnehin schwierig zu erstellenden 3D-Scans darstellen.

Problematisch für eine umfassende Anwendung im Bereich der Bilderkennung ist dagegen, dass für viele Vasenbilder keine derartige Bilderreihe vorliegt und auch nicht entsprechende Metadaten in den digitalen Fotografien hinterlegt sind. Daher könnte die Methode auf Basis von Bilderreihen nur auf eine geringe Anzahl von Vasenbildern angewandt werden. In Hinblick auf das Training von Objekt- und Figurenerkennungsmodellen ist diese Anzahl nicht ausreichend. Jedoch ist hierfür auch keine vollständige und zusammenhängende Abrollung des gesamten Bildfeldes notwendig. Würde nur die einzelne Fotografie der jeweiligen Vasenmalerei entzerrt und zum Training genutzt, wäre bereits dem Umstand Rechnung getragen, dass Darstellungen von Figuren oder Objekten in den Fotografien trotz sehr ähnlicher künstlerischer Ausführung deutliche Unterschiede aufweisen können, wenn sie z. B. verkürzt, gebogen oder in die Länge gezogen erscheinen. Dies erschwert das Training des Algorithmus, auch wenn gewisse Verzerrungen oder andere Blickwinkel sich wiederum positiv

auf das Training der Modelle auswirken können, um auch diese Unschärfen und Variationen bereits im Training berücksichtigen zu können. Die Umwandlung von Bildern auf Trägern mit starker Materialität in 2D-Wiedergaben ließe, beispielsweise als Zwischenschritt im Rahmen der Objekterkennung, signifikante Verbesserungen der Performanz erwarten und könnte als notwendiger Zwischenschritt etabliert werden, wenn Fotografien als Ausgangslage benutzt werden.

Ein Nachteil hinsichtlich einer Implementierung des diskutierten Modells zur Entzerrung von Vasenbildern für die digitale Figuren- und Objekterkennung liegt darin, dass die Methode nicht auf maschinellem Lernen, sondern auf individuellen Werten für jedes zu entzerrende Vasenbild und dessen Bildträger basiert (was die Ergebnisse durch die Individualität sehr präzise macht). Um das Entzerren auf große Bildmengen anzuwenden, ist jedoch ein maschineller Lernprozess zu bevorzugen. Da in den Fotografien die Gefäße für gewöhnlich auf einem monochromen Hintergrund erscheinen, ist es möglich, die eigentlichen Gefäßkörper automatisiert im Bildfeld zu lokalisieren und zu identifizieren.⁸⁵ Hinzu kommt die generelle Annahme, dass die stärksten geometrischen Verzerrungen der Vasenbilder an den Seiten des Gefäßkörpers auftreten. Zwar sind in Bilddatenbanken in der Regel Profilzeichnungen und genaue Maßangaben einzelner Gefäße nicht hinterlegt, doch können anhand dieser für die kanonischen Gefäßtypen (z. B. Halsamphora, Kelchkrater etc.) Näherungswerte bestimmt werden, die einer weiteren Präzisierung der Entzerrung dienen können, auch wenn durch verschiedene Aufnahmewinkel der Fotografien weiterhin mit Unschärfen zu rechnen ist.

Unabhängig dieser technischen Aspekte ist zur Diskussion zu stellen, dass eine Entzerrung aller Fotografien aus den Bildcorpora des Projekts je nach bildwissenschaftlicher Fragestellung mitunter gar nicht zielführend ist. Es muss berücksichtigt werden, dass zum Beispiel die Malfläche auf einem Gefäß natürlich als dreidimensional gewölbt begriffen wurde und die Konzeption des Bildes daraufhin ausgerichtet war⁸⁶, oder dass Bildorte im Raum bewusst auf wahrnehmungsbedingte Verzerrungen hin konzipiert wurden (siehe oben). Eine Entzerrung vorliegender Fotografien kann daher wiederum auch solche intendierten Effekte nivellieren (wenn Fotografien sie überhaupt korrekt visuell vermitteln können), was in der Gesamtheit für das Training wohl kaum Unterschiede bedeutet, aber für das einzelne Artefakt und seine Analyse von Bedeutung sein kann. In einem denkbaren Anwendungstool sollte sie deshalb optional bleiben, denn auch der Verzicht auf Entzerrung kann eine valide Voraussetzung zur Beantwortung von Forschungsfragen sein. Da die Resultate von Objekt- wie auch Figurenerkennung in prozentualen Angaben wiedergegeben werden, wären auf diesem Weg etwa belastbare quantitative Ergebnisse denkbar, die beispielsweise die Sichtbarkeit von Beifiguren bei zentraler Betrachtung der Bildmitte eines Vasenbildes (entsprechend der gängigen Perspektive von Fotografien) betreffen.

■ 85

Zu einem shape based vase retrieval-Tool der Visual Geometry Group (University of Oxford), das Vasenformen automatisiert anhand von Abbildungen ermittelt, siehe <https://www.robots.ox.ac.uk/~vgg/research/vases/index.html>.

■ 86

Beispielsweise durch Staffelung der Figuren entlang der Wölbung. Siehe zu diesen Strategien Didier Martens, *Une esthétique de la transgression: Le vase grec de la fin de l'époque géométrique au début de l'époque classique*, Bruxelles 1992, S. 82–87.

Fazit

Computer-Vision-Modelle bieten in den historischen Bildwissenschaften vielfältige Anwendungsmöglichkeiten, die große Potenziale haben, die fachspezifischen Forschungen gewinnbringend zu unterstützen. Dabei sind jedoch eine Vielzahl spezifischer Merkmale der Bildcorpora zu beachten, für welche im vorliegenden Beitrag anhand von Fallbeispielen Beschreibungen aus archäologischer Perspektive vorgelegt und vorläufige technische Lösungsansätze diskutiert wurden. Die Ausführungen machen deutlich, dass eine sinnvolle disziplinäre Einbettung der Modelle auf Grundlage eines intensiven transdisziplinären Austauschs und eingehender methodisch-theoretischer Reflexion geschehen muss. Dies bedeutet einerseits, dass eine Etablierung digitaler Modelle in der wissenschaftlichen Praxis der Fächer eine eigene aktive Beteiligung der Disziplinen an den neuartigen technischen Entwicklungen voraussetzt, da eine Adaption der Techniken ohne gewissenhafte Berücksichtigung der spezifischen Merkmale der archäologischen Artefakte nicht möglich ist. Andererseits fordert dies auch auf Seiten der Archäologien eine Auseinandersetzung mit der Funktionsweise dieser Modelle und dem Zustandekommen ihrer Resultate, um letztere als Ausgangspunkt für bildwissenschaftliche Forschungen sinnvoll nutzen zu können. Es ist auch deswegen von archäologischer Seite wichtig, sich an dieser Grundlagenforschung aktiv zu beteiligen, da zu befürchten ist, dass diese ansonsten von Personen oder Einrichtungen betrieben werden wird, die keine geisteswissenschaftliche oder archäologische Expertise besitzen. Zudem darf nicht vergessen werden, dass in den vergangenen Jahrhunderten in der archäologischen Wissenschaft stets die neuesten technischen Errungenschaften erprobt worden sind (z. B. die Fotografie), die für das Fach zunächst nutzbar gemacht werden mussten. Da – wie unsere Tests gezeigt haben – die Performanz der Modelle meist im direkten Zusammenhang mit der Menge der zum Training zur Verfügung gestellten Bilddaten steht, ist innerhalb der Fächerkultur, Museums- und Sammlungslandschaft zudem eine Transparenz im Hinblick auf die Zusammenstellung und Bereitstellung großer Bildersets sowie für nächste Entwicklungsschritte auch von 3D-Scan-Respositorien wünschenswert, um so eine schnelle und unkomplizierte forschungsrelevante Nutzung und Veröffentlichung umfangreicher (z. B. annotierter) Bildcorpora zu ermöglichen.