

Projektvorstellung aus:

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

Titel:

Ornamentale Schwünge und Neuronale Netze. Ähnlichkeitsanalysen Augsburgs Rocaille-Drucke des 18. Jahrhunderts

Autor*in:

Ines Röckl

Kontakt: ines.roeckl@ur.de

Institution: Universität Regensburg

GND: [124504592X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0009-0004-7807-6782](https://orcid.org/0009-0004-7807-6782)

Contribution (CRediT): [Conceptualization](#) | [Data curation](#) | [Investigation](#) | [Methodology](#) | [Project administration](#) | [Validation](#) | [Visualization](#) | [Writing – original draft](#) | [Writing – review & editing](#)

Autor*in:

Thomas Hudcovic

Kontakt: hudo@uni-bremen.de

Institution: Universität Bremen

GND: [139402584X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0009-0005-9970-3228](https://orcid.org/0009-0005-9970-3228)

Contribution (CRediT): [Formal analysis](#) | [Investigation](#) | [Software](#)

Autor*in:

Gabriel Zachmann

Kontakt: zach@cs.uni-bremen.de

Institution: Universität Bremen

GND: [12349981X](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9) ORCID: [0000-0001-8155-1127](https://orcid.org/0000-0001-8155-1127)

Contribution (CRediT): [Funding acquisition](#) | [Project administration](#) | [Supervision](#) | [Writing – review & editing](#)

DOI des Beitrags:

[10.17175/sb008_004](https://doi.org/10.17175/sb008_004)


Nachweis im OPAC der Herzog August Bibliothek:

[1950651703](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0011-9)

Erstveröffentlichung:

21.05.2026

Lizenz:

Sofern nicht anders angegeben 

Letzte Überprüfung aller Verweise:

25.03.2026

Format:

PDF ohne Paginierung, Lesefassung

GND-Verschlagwortung:

[Datenbank](#) | [Frühe Neuzeit](#) | [Kunstgeschichte](#) | [Maschinelles Lernen](#) | [Multimodales System](#) | [Ornament](#) | [Segmentierung](#)

Empfohlene Zitierweise:

Ines Röckl / Thomas Hudcovic / Gabriel Zachmann: Ornamentale Schwünge und Neuronale Netze. Ähnlichkeitsanalysen Augsburgs Rocaille-Drucke des 18. Jahrhunderts. In: Hartmut Beyer / Thomas Mandl (Hg.): Bildähnlichkeit und Bildsuche: Geistes- und informationswissenschaftliche Zugänge zu historischem Material (= Zeitschrift für digitale Geisteswissenschaften / Sonderbände, 8). Wolfenbüttel 2026. 21.05.2026. HTML / XML / PDF. DOI: [10.17175/sb008_004](https://doi.org/10.17175/sb008_004)

Ines Röckl / Thomas Hudcovic / Gabriel Zachmann

Ornamentale Schwünge und Neuronale Netze. Ähnlichkeitsanalysen Augsburger Rocaille-Drucke des 18. Jahrhunderts

Abstract

Versteht man das Bemühen um die Kategorie der Ähnlichkeit als kunsthistorische Aufgabe, stößt man bei der ornamentalen Form des Rokokos auf ein Paradox: Einerseits scheint die Rocaille das Prinzip der Ähnlichkeit zu verkörpern, da sie etwa die Stofflichkeit von Muscheln oder Rinde aufgreifen kann. Andererseits fällt es extrem schwer, die enorm wandelbaren Rocaille-Kompositionen untereinander zu vergleichen. Unser Projekt widmet sich diesem Problem mittels Analysen von Augsburger Druckgrafiken des 18. Jahrhunderts. Über Verfahren des Deep Learning werden annotierte Klassen des Ornaments in neuronalen Netzen trainiert. Ziel des Projekts ist eine Anwendung, die via Bild- oder Texteingabe durch ein Vision-Language-Model Ähnlichkeitsvorschläge generiert.

If one understands the category of similarity as an art historical task, a paradox arises in ornamental forms of Rococo: On one hand, the Rocaille seems to embody the principle of similarity, by mimicing for example the texture of shells or bark. On the other hand, it is very difficult to compare the highly variable Rocaille compositions with each other. Our project focuses on this problem through the analysis of Augsburg prints from the 18th century. Using deep learning techniques, annotated classes of ornaments are trained in neural networks. The goal of the project is an application that generates similarity suggestions via image or text input through a vision-language model.

1. Befragungen der Rocailles in der Augsburger Druckgrafik des 18. Jahrhunderts

»Ist es denn also nöthig, eine Sache mit lauter Hirngespinsten zu verzierern, mit Dingen, die in der Welt nicht zu finden sind, oder die sich nicht schicken? Und da ein jeder Verständiger, der da fraget, was es denn eigentlich vorstellen soll, sich mit der Antwort begnügen muß: es sey Rocaille, Grottesque, Arabesque, a la Chinoise, oder en gout baroque: kurz um, es sey so Mode!«¹

[1]

Die harsche Kritik, die Friedrich August Krubsacius 1759 verlauten ließ, adressiert die Ornamentform der Rocaille², die im 18. Jahrhundert im Medium der Druckgrafik von Frankreich ihren Weg nach Augsburg fand und sich formal aus der französischen Regencé-Groteske entwickelt hat.³ In der Reichsstadt Augsburg wurde das Ornament ebenfalls im Druck zunächst kopiert und anschließend rezipiert, aber auch in zahlreichen Materialien wie Stuck, Porzellan, Holz oder Beschlagwerk umgesetzt. Die Augsburger Blätter wurden im 18. Jahrhundert so zum Schauplatz der Rocaille und es entstanden zwischen 1737⁴ und 1770 tausende an zumeist undatierten Drucken, die in Serien von meist vier bis zwölf organisiert waren.⁵

[2]

¹ Krubsacius 1759, S. 35. Die Ergebnisse dieses Beitrages sind aus einem laufenden Forschungsprojekt an der Universität Bremen und der Universität Regensburg erwachsen, Projektleitung Julian Jachmann (Universität Regensburg, Kunstgeschichte) und Gabriel Zachmann (Universität Bremen, Computer Vision), Wissenschaftliche Mitarbeiter*innen: Ines Röckl (Universität Regensburg, Kunstgeschichte), Thomas Hudcovic (Universität Bremen, Computer Vision); Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projektnummer 461631274, Geschäftszeichen JA 1976/2-1. Projekt und Datenbank: <https://rocaille-ornament.de/>

² Zum Begriff Rocaille vgl. Irmscher 2009, S. 340.

³ Vgl. Bauer 1962, S. 3–11; Augustyn 1997, S. 789–794; Irmscher 2009, S. 343–346.

⁴ Exemplarisch ist die Rocaille in der Augsburger Druckgrafik um 1736 bei Gottfried Bernhard Göz nachweisbar (vgl. Wildmoser 1985, S. 316–319).

⁵ Zum Druckzentrum Augsburger vgl. Gier / Janota (Hg.) 1997; Paas (Hg.) 2001; Paas et al. (Hg.) 2013.

Diffamierende Schriften wie von Krubsacius spiegeln demzufolge nicht nur den Zeitgeist des anklingenden Klassizismus,⁶ sondern liefern neben der Bewertung auch eine zeitgenössische Beschreibung,⁷ die Rückschlüsse auf die Charakterisierung der Ornamentform gibt: Wenn die Rocaille aus »Dingen [bestünde], die in der Welt nicht zu finden sind, oder die sich nicht schicken«⁸, ist von freien und fantastischen Formationen die Rede, die zwischen Natur und Kunst⁹ zu verorten sind und dem vernunftgetriebenen und naturverständigen Klassizisten nur negativ aufstoßen können. Auf Quellen wie die von Krubsacius bezugnehmend, lassen sich in der Rocaille-Forschung des 20. und 21. Jahrhunderts¹⁰ neben dem Changieren zwischen Natur und Kunst eine Vielzahl von Eigenschaften ablesen, deren Ausprägungen sich jedoch je nach Künstler*in und Technik sehr verschieden gestalten. [3]

Als wesentlich für die Zielsetzung der Ähnlichkeit im Kontext von Morphologien, gilt das Moment des Proteischen.¹¹ So stellt sich die Rocaille als überaus wandelbar dar und kann verschiedene Materialien und Formen adaptieren.¹² Wie ein »alchemistischer Tiegel«¹³ werden so Ähnlichkeitsbeziehung zu Naturformen virulent, wie etwa Assoziationen zu Schaum, Wasser, Muschel, Fels, Erde, Feder, Fledermausflügel oder Rinde.¹⁴ Gleichzeitig erschwert diese Wandelbarkeit die Vergleichbarkeit der verschiedenen Form-Ausprägungen untereinander, was sich nicht zuletzt als ekphrastisches Problem entpuppt. Während auf mikromorphologischer Ebene die Formen und Beschaffenheiten der Rocaille personalstilistische Ausprägungen abbilden können, ist für das Meta-Gerüst der Rocaille eine Struktur in »c-, seltener s-förmige Volutenspange[n]«¹⁵ ausschlaggebend. Dieses Volutenspangengerüst kann zwar – im proteischen Sinne – rocaillisiert¹⁶ auftreten, ermöglicht jedoch die Vergleichbarkeit verschiedenster Rocaille-Formen. Ebenso können die »konvexe[n] Seiten«¹⁷ im weitesten Sinne als Kämmen beschrieben werden, die nicht zwingend muschelartig ausfallen müssen.¹⁸ [4]

Die Klassifizierung in Volutenspange und Kamm erlaubt verschiedene kunsthistorische Befragungen: Ab wann kann von einer Ähnlichkeit hinsichtlich der Gesamtkomposition gesprochen werden? Wie verhält es sich mit Übereinstimmungen, die lediglich in Teilen erfolgt? Wie lassen sich stilistische, materielle oder zeitliche Entwicklungen durch den Vergleich ablesen? Mittels Verfahren der *Computer Vision* wird sich diesen Fragestellungen angenähert. Gleichzeitig ergeben sich auf Seiten der *Computer Vision* ebenso relevante Problemstellungen, etwa das Training von Encoder-Decoder-Netzen auf einem winzigen Datensatz mit Grauwertbildern, was Farbinformationen als zusätzlich diskriminierende Faktoren ausschließt. Hinzu kommt das Problem der nicht eindeutig identifizierbaren Segmentierungselemente, da die Umgebung, in der ein Element eingebettet ist, bei der Klassifizierung eine Rolle spielt. Die Segmentierungsobjekte sind also nicht »semantisch invariant«. Ein weiteres Problem auf Pixelebene ist der nicht-balancierte Datensatz, der den Netzen ein Bias (»es ist wahrscheinlicher, dass es keine Rocaille ist, da es nicht viele [5]

⁶ Vgl. Bauer 1962, S. 41–47, 63–69; Irmscher 2009, S. 410–411.

⁷ Es ist bereits von Herrmann Bauer angemerkt worden, dass es »keine gleichzeitige Theorie des Rocaille-Ornaments« (Bauer 1962, S. 41) gibt. Weiter vermerkt er: »Interessant ist, daß [sic] in diesem Bereich all die Phänomene, wie sie hier [= Bauers Publikation *Rocaille. Zur Herkunft und zum Wesen eines Ornament-Motivs*] analysiert wurden, schon gesehen und mit großem Ekel, auch in bornierter edler Einfalt registriert wurden« (Bauer 1962, S. 41).

⁸ Krubsacius 1759, S. 35.

⁹ Vgl. Irmscher 2009, S. 339–340.

¹⁰ Neben Bauer 1962 und Irmscher 2009 sind monografische Publikationen zu einzelnen Protagonist*innen zu nennen. Zu Franz Xaver Habermann: Krull 1977; zu Gottfried Bernhard Göz: Isphording 1982; Isphording 1984; Wildmoser 1984; Wildmoser 1985; zu Johann Esaias Nilson: Helke 2005.

¹¹ Aus der proteischen Qualität lassen sich die Bildwertigkeit und die Eigenplastizität ableiten, die als zentrale Eigenschaften in der Sekundärliteratur zur Rocaille Eingang gefunden haben (vgl. Irmscher 2009, S. 339–340).

¹² Zudem wird von Irmscher die »mögliche Asymmetrie (contraste)« (Irmscher 2009, S. 339–340) angefügt. Zusätzlich ist die mikromegalische Struktur für die Rocaille charakteristisch, mit der irrational proportionale Größenverhältnisse beschrieben werden können (vgl. Bauer 1962, S. 20–21).

¹³ Irmscher 2009, S. 339.

¹⁴ Die Vielseitigkeit der Materialassoziationen schlagen sich auch in der Kritik Krubsacius' nieder, so beschreibt er die Rocaille als »Mischmasch« verschiedener Dinge, wie »[z]erbrochenen Muscheln, Lappen« oder »Federn« (Krubsacius 1759, S. 36).

¹⁵ Irmscher 2005, S. 141.

¹⁶ Vgl. Irmscher 2009, S. 340.

¹⁷ Irmscher 2005, S. 141.

¹⁸ Weiter nennt er neben dem Muschelrand »phantastische Palmetten, Fledermausflügel, gereifte Godronen und ähnliche kammartig-bizarre Bildungen« (Irmscher 2005, S. 141).

gibt) antrainieren würde. Das zentrale Anliegen ist jedoch die Untersuchung zwischen der visuellen Morphologie im Bildraum einzelner Rocaille oder Rocaille-Primitive und der von den Netzen gelernten internen Wissensrepräsentation im sogenannten *Latent Space*. Genauer stellt sich die Frage, ob sich Mannigfaltigkeiten oder Submannigfaltigkeiten im Latent Space identifizieren lassen, die Rocailles bzw. Rocaille-Primitive im Bild beschreiben können, und ob es gegebenenfalls Isomorphismen zwischen verschiedenen (Sub-)Mannigfaltigkeiten gibt.

Diese Fragestellungen schränken die Vorgehensweise ein: Zwecks der Untersuchung der morphologischen Ähnlichkeiten zwischen Rocaille bzw. Rocaille-Primitiven können komplexere Netzarchitekturen nicht verwendet werden, um die Interpretierbarkeit der intern aufgebauten Wissensrepräsentation im *Latent Space* möglichst einfach zu halten. [6]

2. Projektbeschreibung, Digitalisate und Metadaten

Diesen Fragen widmet sich seit April 2022 das DFG-Projekt ›Digitale Morphologie der Ornamentik – Entwicklung von Verfahren an der Schnittstelle zwischen Kunstgeschichte und Computer Vision zur Analyse, Modellierung und Recherche von Ornamentformen am Beispiel der Augsburger Rocaille-Drucke des 18. Jhs., das an der Universität Bremen (Gabriel Zachmann, Thomas Hudcovic) und der Universität Regensburg (Julian Jachmann, Ines Röckl) angesiedelt ist. Im Projekt werden mit Hilfe von Verfahren der Computer Vision, Künstlichen Intelligenz und neuronaler Netze Rocaille-Formen in Kompositionen identifizierbar gemacht, die zuvor manuell annotiert wurden. Zielsetzung des Projekts ist eine digital zugängliche Datenbank, die neben den üblichen Kategorisierungen nach kunsthistorischen Eckdaten eine Befragung hinsichtlich Bildähnlichkeit ermöglicht. Bezüglich der erforschten Morphologien können interessierte Nutzer*innen und Forscher*innen eigene Bilder von Rocaille – in Grafik aber auch anderen materiellen und medialen Umsetzungen, wie etwa Holz, Porzellan, Beschlagwerk oder Stuck – in der Suchmaschine uploaden, woraufhin eine Auswahl an ähnlichen Augsburger Rocaille-Drucken vorgeschlagen wird. Dies adressiert das bereits angesprochene Desiderat bezüglich Kategorisierung und Beschreibung der Rocaille. [7]

Als Datenkorpus dienen Digitalisate ornamentaler Druckgrafiken – vornehmlich Radierungen – die den Augsburger Rocaille-Drucken zugeordnet werden.¹⁹ Zu jedem Digitalisat wurde ein Meta-Datensatz erstellt, der über die folgenden Kategorisierungen inklusive recherchierter Inhalte verfügt: ID-Nummer, Titel, Nummerierung in Folge, Zeichner*in / Entwerfer*in, Stecher*in, Verleger*in, Anmerkungen, Bestand, Literatur, Digitales Bild. Durch diese sowohl visuelle wie auch textuelle Herangehensweise, soll auf Basis eines Vision-Language-Models die Vergleichbarkeit und somit die Befragung nach Ähnlichkeiten untereinander ermöglicht werden. [8]

3. Druckgrafik und Datenkorpus: Von der Genese der Rocaille zur Segmentierung

Der Datenkorpus besteht aus ca. 1.600 Digitalisaten, die aus verschiedenen Beständen in Deutschland zusammengetragen wurden. Alle Drucke sind in Augsburg im 18. Jahrhundert herausgegeben worden, zumeist undatiert sowie nicht coloriert und zeigen Rocaille-Ornament. So wie Technik und Farbe einheitlich im Datensatz auftreten, sind es die Kompositionen und Morphologien der Rocaille auf den ersten Blick keineswegs: Um eine Vorstellung über die historische Genese sowie die Struktur und Morphologie der Rocaille in der Druckgrafik zu bekommen, wird die Bandbreite und der Facettenreichtum des Datenkorpus [9]

¹⁹ Die Digitalisierung der Drucke erfolgte bereits im Kontext des DFG-Projekts ›Augsburger Architektur- und Ornamentstichwerke. Der Anteil der Kunsthandlungen an der Formierung und Verbreitung ästhetischer Modelle während des 18. Jahrhunderts‹ (Prof. Dr. Katharina Krause, Förderzeitraum: 2005–2010, Projektnummer 19185709).

vorgestellt. Dies ist nicht zuletzt deshalb zentral, da in der Annotation eine möglichst breite Streuung des Datensatzes hinsichtlich Autorschaft und Gattung forciert wurde, um die formale Breite des Datensatzes zu berücksichtigen.

Für die zunächst in Frankreich entwickelte Form der Rocaille und deren Verbreitung im 18. Jahrhundert ist das Medium der Druckgrafik seit den Anfängen von zentraler Bedeutung. Auch wenn die Rocaille als Namensgeberin des Rokokos gilt, sind einige Entwicklungsstränge zu ihrer formalen Herausbildung nicht abschließend geklärt. Einigkeit in der Forschung herrscht jedoch bezüglich der zentralen Rolle des Goldschmieds und königlichen Hofkünstler Juste-Aurèle Meissonniers, der die Form in jenem Medium erstmals populär machte. Als Ornamentbegriff lässt sich der Terminus zum ersten Mal bei einer Anzeige zu Meissonniers *Livre d'ornemens* in der März-Ausgabe des *Mercure de France* von 1734 nachweisen.²⁰ Doch nicht nur begrifflich, auch formal weisen Meissonniers Entwürfe, die kurz vor 1730 entstanden sind, Tendenzen auf, die sich von der vorherigen Tradition der Regencé-Groteske mit vorwiegend Laub- und Bandelwerk absetzt. Während Meissonniers Werk gemeinsam mit den Entwicklungen bei Jacques de Lajoue und Nicolas Pineau als wegbereitend für die Rocaille gilt,²¹ sind es vor allem Künstler wie Antoine Watteau, Claude Gillot, Claude III Audran, Nicolas Pineaus, Jean Mondon, Jacques de Lajoue, Pierre Edme Babel und François Boucher, deren Drucke in Augsburg kopiert wurden.²² Eine wesentliche Rolle kam in Augsburg den Verleger*innen zu, die teils auch selbst entwarfen oder stachen und als Kunstverlage auftraten. Persönliche Kontakte und Handelsbeziehungen waren dabei sowohl beim Import französischer Blätter wie auch beim Export der Augsburger Drucke zentral.²³ Die in Augsburg nachgestochenen Drucke wurden zumeist spiegelverkehrt umgesetzt sowie in der Beschriftung und im Privileg verändert.²⁴ Ein Beispiel für einen in Augsburg kopierten französischen Entwurf zeigt die figurativ angereicherte allegorische Szenerie (vgl. Abbildung 1).

[10]

²⁰ Vgl. Irmscher 2009, S. 340, 346–347. Vgl. weiter: Bauer 1962, S. 18–19.

²¹ Vgl. Irmscher 2009, S. 346. Vgl. auch zur Entstehung des Rokokos: Kimball 1980 [1943], S. 152–174.

²² Vgl. Irmscher 2009, S. 357–358.

²³ Vgl. Irmscher 2009, S. 357.

²⁴ Zu den Verbreitungswegen und möglichen Übernahmen gibt es bereits einzelne Forschungen. Jedoch hofft das Projekt genau hier als ein Recherchetooll anzusetzen, um Übernahmen von ornamentalen Entwürfen nachverfolgen zu können. Studien zur Verbreitung und Übernahmen von Rocaille, v. a. im Kontext von Augsburg u. a.: Mandroux-França 1973, S. 342–445; Seeger 2014, S. 285–294; Krause 2015, S. 185–199.



Abb. 1: »L'Amour Couronne. / Der gekrönte Cupido«. Sechstes Blatt einer Folge von sieben Blättern, entworfen von Jean Mondon, herausgegeben bei Johann Georg Merz in Augsburg. [Fotografie: Staats- und Stadtbibliothek Augsburg, Graph Merz, J.G. 41]

Das Blatt erschien unter der Verlagsnummer 10 bei I. G. Merz (= Johann Georg Merz) in Augsburg und ist ein Entwurf nach Jean Mondon, der zuvor in Frankreich unter der Verlagsnummer C erschienen ist.²⁵ Obwohl hier keine spiegelverkehrte Umsetzung vorliegt, wurden neben den Nummerierungen die Beischriften angepasst; so ist der Titel zweisprachig angegeben. Die bildliche Übereinstimmung ist eindeutig, jedoch besticht im direkten Vergleich die französische Ausführung in ihrer zeichnerischen Qualität. Inhaltlich wird der Sieg Cupidos verhandelt, der durch zwei Putti dargestellt wird.²⁶ Die Figuren sind mit einem Köcher mit Pfeil und Bogen ausgestattet sowie mit einem (Lorbeer-)Kranz. Ob es sich bei der Szenerie um eine Krönung oder ein sanftes Ringen um die Krone handelt, wird durch die deutsche Bildunterschrift ausgedeutet: »Der gekrönte Cupido.« Dies würde bedeuten, dass der Moment hier die Krönung des vorderen Puttos zeigt, jedoch lässt das Wortspiel im französischen auch die reflexive Interpretation zu, was eine Selbst-Krönung nicht ausschließt und so die beiden Putti im zweiten Moment zu Kontrahent*innen macht. Drapiert sind die Putti auf einer Wolke, die auf einem Volutenaufbau liegt: Dieser schält sich unangekündigt aus der Erde und erschließt sich durch proportional große C- und S-Schwünge von rechts unten nach links oben und nimmt so einen Großteil des Bildraums ein. Begleitet werden die teils aus Volutenbündeln bestehenden Schwünge von muschelartigen Kämmen, die dem Volutenschwung folgen und konzentrisch zu diesen angeordnet sind. In ähnlicher Logik entstehen so kleine Rocaille-Kartuschen, die additiv unterhalb der Wolke (mit einem Herz)

[11]

²⁵ Ein gründlicher Abgleich des Französischen und Augsburger Drucks kann hier nicht erfolgen. Abweichungen sind jedoch vor allem in Nummerierung, Beischrift und Anordnung der Blätter zu konstatieren. Vgl. Digitalisat gallica.bnf.fr. Letzter Zugriff: 17.06.2024. Vgl. Schmitz 1939, Nr. 403.

²⁶ Ggf. handelt es sich bei dem zweiten Putto um Psyche. Diese Vermutung stützt sich auf die Geschichte von Amor und Psyche in den *Metamorphosen* des Apuleius (vgl. *Apul. met.* 4,28–6,24).

und links unterhalb eines weiteren Bildfelds angebracht sind. Rechts unterhalb des hinteren Puttos, lassen sich auch noch architektonische Elemente ausmachen. Hinzu kommen neben der Rocaille aus Voluten und Kämmen, Akanthus und Blumenranken. Akanthushäubchen bekrönen die Volutenköpfe, was vor allem am oberen linken Bildrand deutlich wird. Hier formieren sich die Volutenbündel zu einem angeschnittenen Bildfeld – einer halben Kartusche – die ihren unteren Abschluss in einem palmettenähnlichen gedrehten Muschel-Blatt findet. Innerhalb der Kartusche zeigt sich lediglich skizzenhaft eine Szenerie. Hier könnte es sich analog zu Apuleius Geschichte aus den *Metamorphosen* um die Schlüsselszene aus der Verfolgung von Apoll und Daphne bei Ovid handeln.²⁷ Diese Narration wird in dem Blatt als skizzenhafte Nebenszene am Bildrand angeordnet. Zwar lenken die Volutenstränge zum Bildfeld oben links, jedoch liegt der Fokus auf den eigentlichen Nebenfiguren: Putti wie Ornament rücken in den Fokus der Handlung. Dies zeigt sich nicht zuletzt in dem Rocaille-Aufbau, der den Großteil des Blattes einnimmt und beide amourösen mythologischen Geschichten verbindet. Die Rocaille erscheint bildwertig und bildet sich eigenplastisch aus – beides Eigenschaften, die in dem proteischen Charakter begründet sind²⁸ – und durch die metamorphen Szenerien selbstreferentiell ihren wandelbaren Charakter aufzeigen. Gegliedert wird die Narration durch die dominierenden Volutenspangen. So wird die angedeutete Szenerie in der Kartusche oben links zum Beiwerk anstatt als Hauptszene aufzutreten. Die Putti, die sonst häufig als szenisches Echo der Hauptnarration wirken, werden in ihrer frivolen Art mit dem im Bild vorherrschenden Hell-Dunkel-Kontrast hervorgehoben. Dies wird zusätzlich durch den fast leeren Himmel rechts oben mit leichten Wolkenschwaden betont. Das Beispiel Mondons deckt mit der allegorisch-mythologischen Szenerie in einem Rocaille-Aufbau bereits eine Sparte der im Ornamentdruck verhandelten Thematik ab und macht die Verdichtung von Ornament und Narration deutlich.²⁹

Neben den capricciosen Aufbauten zeigen eine Vielzahl der Blätter isolierte Kartuschen und Schilde. Ergänzend muss neben der kunstgewerblichen Sparte, die in Augsburg im 18. Jahrhundert zusammen mit der Druckgrafik florierte, noch auf die geistlichen Themen in den Drucken hingewiesen werden.³⁰ In der bikonfessionellen Handelsmetropole³¹ wurden christliche Themen etwa vom Verlag der katholischen Gebrüder Klauber herausgegeben.³² Besondere Popularität erfuhr auch das Werk des protestantischen Entwerfers, Stechers und Verlegers Johann Esaias Nilson. In den zumeist gesellschaftlichen Darstellungen in Kombination mit Versen treten morphologisch eher erd- oder rindenähnliche³³ Rocaille-Formen auf.

[12]

Bei den verschiedenen (personal-)stilistischen Ausprägungen³⁴ und Themen finden sich als kontinuierlichste Kategorisierung die C- und S-Volutenspangen mit den Kämmen. Auch wenn die Ausprägung der Voluten und Kämmen je nach Blatt unterschiedlich ausfällt und im Projekt bereits über weitere Klassifizierungen nachgedacht wird, ermöglicht dieser erste Zugriff eine breite Vergleichbarkeit der Rocaille, die über das Medium der Druckgrafik hinausgeht.

[13]

Um diese Zielsetzung an der Schnittstelle zur Computer Vision zu erreichen, werden die Kategorisierungen der Rocaille als pixelweise Klassifizierungen pro Bild (Labels) annotiert. Diese bilden die Ground Truth, mit denen sich die neuronalen Netze während des Trainingsprozesses evaluieren lassen. Das Training von neuronalen Netzen ist mathematisch gefasst ein Optimierungsprozess, bei dem ein Netz das Optimierungsobjekt darstellt, das letztendlich einer Funktion angenähert werden soll, die ein bestimmtes Problem löst, wie z. B. das Klassifizieren von Pixeln in Bildern. Die Ground Truth-Daten dienen zusammen

[14]

²⁷ Vgl. Ov. met. 1,452–567.

²⁸ Vgl. Irmscher 2009, S. 339–340.

²⁹ Zum Verhältnis zwischen Ornament und Zeitlichkeit vgl. Leonhard 2011, S. 63–85.

³⁰ Vgl. Augustyn 1997, S. 791.

³¹ Vgl. Irmscher 2009, S. 357.

³² Vgl. Augustyn 1997, S. 820–830.

³³ Vgl. Bauer 1962, S. 52–57; Helke 2005, S. 155; Irmscher 2009, u. a. S. 339, 359–360.

³⁴ Auffällig sind bei Franz Xaver Habermann die Durchgitterung der Kämmen bzw. die geperlten Kämmen (vgl. Abbildung 2; vgl. Krull 1977, S. 21–22, 46, 49). Bei Gottfried Bernhard Göz finden sich etwa lose Kämmen, die ohne Volute existieren, wie etwa in der Folge *Die Schrecken des Krieges* (vgl. Isphording 1982, S. 47–48, 249–251; Wildmoser 1985, S. 186–189. Siehe exemplarisch collections.vam.ac.uk. Letzter Zugriff: 17.06.2024.)

mit dem Output des Netzes als Argumente für die Kostenfunktion (Loss-Funktion), die abhängig vom Trainingsziel ist, die die derzeitige Qualität des Netzoutputs evaluiert, d. h. eine ›Distanz‹ formuliert zwischen dem Sollwert (Ground Truth) und dem Istwert (Netzoutput). Die Rocaille-Formen inklusive der jeweiligen Annotationen liegen als Trainingsdatensatz einem neuronalen Netz zugrunde, dessen Aufgabe die semantische Segmentierung der Rocaille-Kategorien innerhalb eines Bildes ist. Dieser Ansatz wird multimodal erweitert mit zusätzlichen Textbeschreibungen und räumlichen Informationen wie Bounding Boxes, die die Trainingsbasis für ein *Vision Language Model* bilden. So können zusätzlich zu der visuellen Erkennung auch Textassoziationen und andere Metainformationen sowohl erhalten als auch als Nutzereingabe beziehungsweise Frage an das Modell gestellt werden. Das zentrale Ziel ist jedoch die Erfassung von Ähnlichkeiten zwischen Rocailles bzw. Rocaille-Elementen und eine sich daraus ableitende Morphologie dieser. Dabei soll der Ähnlichkeitsbegriff durch Erfassung von möglichen Mannigfaltigkeiten oder ähnlich greifbaren Features in der von den neuronalen Netzen erlernten Wissensrepräsentation abgeleitet werden.

4. Morphologie der Rocaille: Volutenspangen und Käämme als Klassen im Labeling

Die Identifizierung der Volutenspangen und Käämme soll anhand eines prototypischen Blatts aus einer Folge von vier Blättern von Franz Xaver Habermann herausgearbeitet werden (vgl. Abbildung 2).³⁵

[15]



Abb. 2: Rocaillekompositionen. Zweites Blatt einer Folge von vier Blättern, entworfen von Franz Xaver Habermann, herausgegeben von Johann Georg Hertel in Augsburg. [Fotografie: © Staatsgalerie Stuttgart, Graphische Sammlung, alter Bestand]

³⁵ Vgl. u. a. Schmitz 1939, Nr. 135.

Die Folge zeigt isolierte Rocailleentwürfe mit je acht bis neun Rocaillekompositionen pro Blatt. Während das Betrachten der ganzen Folge Fragen nach Variation und Erfindungsreichtum in den Vordergrund rückt, dient die morphologische Analyse eines einzelnen Blatts der Identifizierung der Klassen. Als Grundgerüst der Rocaille bilden sich die C-Spangen im Blatt deutlich ab, vereinzelt finden die als C-Schwünge geformten Spangen ihr Ende in Palmwedelornamentik oder Muscheln. Teils bilden sich Volutenspangen als bandartige, gerollte oder tordierte Schwünge in ihrer Kontur mit Volutenköpfen aus, teils verschleifen diese mit den Kämmen oder sind überlagert. Auch an Stellen, wo die Volutenköpfe nicht direkt sichtbar sind und das Volutenband verdeckt ist, modelliert sich durch die Schwünge eine konvex-konkave Torsion. [16]

S-Spangen und -Schwünge entstehen im Blatt durch tangentielle Addition von abwechselnd gegeneinander verlaufenden C-Schwüngen. Dadurch wird eine longitudinale Grundstruktur ermöglicht, wie sie sich links unten parallel zum Druckplattenrand ergibt. Durch zueinander gerichtete C-Schwünge und -Spangen werden wiederum runde, kartuschenähnliche Formen erzeugt, wie etwa rechts oben. Teils geben geschwungene C-Spangen die gesamte Ausrichtung vor, was sich in der Komposition rechts unten nachvollziehen lässt. So lassen die Strukturen verschiedene Anwendungen und Funktionen denkbar werden: Das longitudinale Ornament ließe sich als eine Rahmenleiste umsetzen, während die Kartusche oben rechts z. B. als Eckstück einen Richtungswechsel einleiten könnte. Allgemein verlaufen die Kämmen häufig im näheren Kontext zu den Schwüngen in konzentrischer Biegung. [17]

Neben der immer wieder auftretenden Palmwedelornamentik und Muscheln, die teils in die Schwünge übergehen oder sie überlagern, finden sich in dem Blatt auch Blumengirlanden, filigrane vegetabile Gehänge und Akanthus. Diese Ornamente »kontaminieren« im Sinne der Rocaillesierung die Kämmen: So treten die Kämmen im Blatt teils mit Perlen auf, erscheinen muschelähnlich, erinnern an Akanthus, oder es bilden sich in spangenlosen Kämmen baumrindenähnliche Formen aus. Weiter lassen sie sich die Kämmen in ihrer Textur als Kamm-Inneres sowie in ihrer Beschaffenheit an den Kamm-Rändern zu ihrer Kontur befragen. Dabei ergibt sich Textur durch die Schraffur, die sich durch die Technik der Radierung in das digitalisierte Büttenpapier übertragen hat. Kreuz- und Parallelschraffur können so Rhythmisierungen oder Pointierung in den Kämmen erzeugen. Die so generierten Kammrillen können weiter mit ungleichmäßigen Durchbrechungen, gleichmäßigen Durchgitterungen oder wie hier mit Perlen besetzt sein. Dies lässt sich sowohl beim Akanthus als auch in den muschelartigen Kämmen nachweisen. [18]

Diese – hier exemplarisch ausgeführte – Klassifizierung der Morphologie der Rocaille in dem Blatt dient als Grundvoraussetzung für das Labeling. Um einen idealtypischen Zugriff der oben ausführlich beschriebenen Rocailledefinition zu übersetzen, wurden von den ca. 1.600 digitalisierten schwarz-weiß Drucken 753 ausgewählt die C- und S-Spangen und Kämmen abbilden und dennoch die morphologische Varianz der Spangen und Kämmen im gesamten Datensatz widerspiegeln. Ebenso verdeutlicht diese Auswahl die Bandbreite der Autorschaft, Themen und Gattungen. Für das Training wurde in einem ersten Schritt pixelgenau annotiert (119 images) und als Spangen die Volutenköpfe mit dem bandartig sich abzeichnenden S- bzw. C-Schwung markiert. Als problematisch gestalteten sich dabei vor allem die sehr kleine Datenmenge von gelabelten Blättern, die topologisch nicht eindeutigen Segmentierungsziele, da die Volutenspangen allein wenig Aussagekraft besitzen, sowie der unbalancierte Datensatz, zumal in jedem Blatt eine unterschiedliche Anzahl von Labels gesetzt wurde. Diese Probleme erschwerten das erfolgreiche Training des Netzes, weshalb das Netz einem Vortraining unterzogen wurde, für das die gesamten nicht-gelabelten Daten verwendet wurden, zuzüglich frühneuzeitlicher Drucke mit ornamentalen Schwüngen wie es im Roll- und Beschlagwerk oder dem Knorpelwerk vorkommt. Da das Netz eine U-Net-artige Encoder-Decoder-Architektur aufweist, bot es sich an, es als die U-Net-Komponente im *Denosing Diffusion Probabilistic Model (DDPM)*³⁶ einzusetzen. Ein DDPM ist ein generatives Modell, welches in der Lage ist, den Trainingsdaten ähnliche Samples zu erzeugen. In diesem Fall Bilder, die den Bildern im Datensatz visuell möglichst ähnlich sind. Durch das Vortraining baut sich das Netz bereits Vorwissen über visuelle Features auf; von der Schraffur bis hin zu Strukturen, wie z. B. [19]

³⁶ Vgl. Ho et al. 2020, S. 6840–6851.

Komponenten der Rocaille. Intuitiv gesehen dient das bereits aufgebaute Vorwissen dem Netz beim Training als Ausgangsbasis für das Lernen der eigentlichen Aufgabe, dabei werden effektiv weniger Trainingsdaten (Labels) benötigt. Mathematisch lässt sich das als ein Modell betrachten, bei dem die im Training zu lernenden Parameter bereits so initialisiert sind, dass sie schneller und stabiler konvergieren. Dennoch sind *False Negatives* und *False Positives* noch möglich. Um die problematischen Punkte noch weiter zu reduzieren, wurden daher in einem zweiten Schritt in einem Zug mit den Voluten zusätzlich auch die Kämme gelabelt (184 images). Die C- und S-Schwünge wurden erweitert als das aufgefasst, was als klare, lineare Verbindung in den Blättern auftaucht. Für die Kämme ist wesentlich, dass sie in ihrer Textur konzentrisch dem Schwung folgen und eine geschlossene Kontur aufweisen. Auch wurden keine isolierten Kämme gelabelt, sondern lediglich die Kämme, die einer Volute zuzuordnen sind. In der *Semantic Segmentation* wurde mit diesen zwei Segmentierungsklassen gearbeitet. ›Volute‹ und ›Kamm‹ bilden die Rocaille als ›Metaklasse‹ ab und stehen bidirektional in Relation zueinander. Die Digitalisate wurden im Labelprozess jeweils in der Folge gelabelt, was in komplexen und verdichteten Blättern eine dreistellige Anzahl von Labels pro Blatt hervorbringt.

5. Die Rocaille als Wissensrepräsentation im neuronalen Netz

Wie in Abschnitt 1 dargelegt, weist die Segmentierung von Rocaille-Elementen in den Augsburgern Drucken mehrere signifikante Herausforderungen auf. Dadurch sind herkömmliche Machine-Learning-Methoden, wie z. B. *Random Forests*, sowie neuronale Netze wie *DeepLabV3+*³⁷ oder auf Transformern basierende Segmentierungsnetze wie *SETR*³⁸ oder *Mask2Former*³⁹ ungeeignet. Diese fielen bereits wegen der hohen Komplexität der zu segmentierenden Elemente und der gleichzeitig geringen Datenmenge heraus. [20]

Gleichzeitig ist eine einfache Aneinanderkettung der grundlegendsten Basisblöcke (*Multilayer Perceptrons*) aufgrund des Speicherbedarfs, bedingt durch die Parameteranzahl, ebenfalls nicht rentabel. Bereits die Entscheidung, das Problem als reine Segmentierungsaufgabe zu betrachten, stellt einen Kompromiss dar. Optimal wäre es, das Problem als *Instance Segmentation* zu behandeln, bei der ein neuronales Netz nicht nur eine pixelweise Klassifizierung vornimmt, sondern jedem klassifizierten Pixel auch eine spezifische Instanz zuordnet. Hierfür wären Netzarchitekturen wie *Mask R-CNN*⁴⁰ einsetzbar. Da jedoch Instance Segmentation eine schwierigere Aufgabenstellung als reine Segmentierung ist, stellt auch hier die geringe Datenmenge ein Problem dar. Der übliche Ansatz mit *Transfer Learning*, also ein vortrainiertes Netz auf einem üblichen Datensatz wie *ImageNet* zu verwenden, erweist sich ebenfalls als schwierig, da zum einen die Augsburgern Drucke gegenüber Datensätzen wie *ImageNet* doch sehr unterschiedlich sind und zum anderen die *Latent Spaces* für die Untersuchung nicht weiter durch fremde Informationen an zusätzlicher Komplexität gewinnen sollen. [21]

Die gravierendste Herausforderung ist die geringe Datenmenge: Üblich sind mehrere Zehntausend bis Millionen Bilder bzw. Datenpunkte im Trainingsdatensatz, je nach Aufgabe und Anzahl der zu segmentierenden Klassen. Ein weiterer kritischer Faktor ist die fehlende semantische Invarianz der Segmentierungsobjekte, da der semantische Kontext der Umgebung, in die die Rocaille-Elemente im Bild eingebettet sind, die Klassifizierung beeinflusst. Zudem ist der Datensatz stark unausgewogen: Das Verhältnis von Nicht-Rocaille-Pixeln zu Rocaille-Pixeln beträgt etwa 31:1, was unbehandelt zu einem Bias führen und die Netze zu mehr *False Negatives* bewegen würde. Darüber hinaus sollte die zu verwendende Architektur möglichst einfach und nachvollziehbar sein, um die Untersuchung von möglichen topologischen Mannigfaltigkeiten in der von dem Netz intern aufgebauten Wissensrepräsentation, dem *Latent Space* (vgl. Abbildung 4), und möglicher Isomorphismen zwischen diesen Mannigfaltigkeiten und den Rocailles im Bildraum zu erleichtern. [22]

³⁷ Vgl. Chen et al. 2018, S. 801–818.

³⁸ Vgl. Zheng et al. 2021, S. 6881–6890.

³⁹ Vgl. Cheng et al. 2022, S. 1290–1299.

⁴⁰ Vgl. He et al. 2017, S. 2961–2969.

*U-Nets*⁴¹ sind eine neuronale Netzwerkarchitektur, die insbesondere für medizinische Bildverarbeitung und Segmentierungsaufgaben und deren oftmals ebenfalls relativ kleinen Datensätze (einige Tausend) entwickelt wurden. Diese bilden in der Kategorie der Encoder-Decoder-Netze eine eigene Unterart. Allgemein lässt sich die Encoder-Decoder-Architektur mit zwei Hauptkomponenten beschreiben: dem *Encoder* und dem *Decoder*. Der Encoder nimmt die Eingabedaten, wie etwa ein Bild, auf und transformiert diese durch eine Serie von Schichten – typischerweise Faltungs- und Pooling-Schichten in der Bildverarbeitung – in eine zunehmend komprimierte, aber semantisch hochdimensionale Repräsentation, auch Wissensrepräsentation oder *Latent Space* genannt. Diese stellt eine abstrakte und semantisch reiche Darstellung der ursprünglichen Daten dar, die wesentliche Merkmale und Informationen extrahiert hat. Der Decoder hat die Aufgabe, diese komprimierte Repräsentation wieder in eine Ausgabe zu transformieren, die eine ähnliche Struktur wie die ursprüngliche Eingabe hat, jedoch oft mit einer spezifischen Modifikation oder Anreicherung, wie etwa der Markierung bestimmter Bildbereiche. Bei der Bildsegmentierung hat der Decoder somit die Aufgabe, bestimmte Regionen oder Objekte innerhalb des Bildes klar und präzise zu identifizieren. [23]

U-Nets erweitern die grundlegende Encoder-Decoder-Architektur durch den symmetrischen Aufbau von *Encoder* und *Decoder* bezüglich der eingesetzten Schichten und deren Reihenfolge. Diese Schichten sind durch zusätzliche Verbindungen zwischen dem Encoder und dem Decoder, den sogenannten *Skip-Connections*, verbunden, die es dem Netzwerk ermöglichen, detaillierte Informationen aus den frühen Schichten der Enkodierung direkt in die entsprechenden Schichten der Dekodierung zu übertragen und dadurch mit weniger Trainingsdaten zurechtzukommen. Im Rahmen dieses Projekts wurden mehrere Variationen von *U-Nets* entwickelt, die sich durch den Aufbau und die Größe einzelner Schichten unterscheiden, um die Leistungsfähigkeit zu erhöhen. [24]

U-Nets zählen auch zur Kategorie der *Convolutional Neural Networks (CNN)*, da einzelne Schichten auch Faltungsoperationen beinhalten. Abbildung 3 stellt die grundlegende Funktionsweise von *U-Nets* und deren Aufbau des *Latent Space* dar. [25]

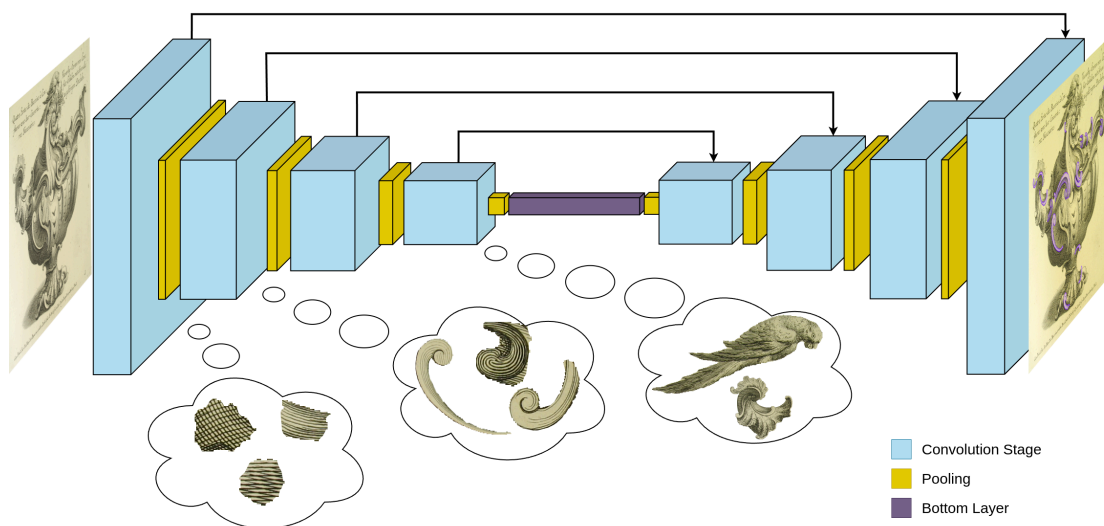


Abb. 3: Modellarchitektur des U-Net: Inspiriert vom biologischen visuellen Kortex ermöglicht der hierarchische Aufbau von Faltungsschichten (Convolutional Layers), visuelle Merkmale in unterschiedlicher semantischer Komplexität zu erfassen und zu erlernen. [Grafik: Thomas Hudcovic 2026]

⁴¹ Vgl. Ronneberger et al. 2015, S. 234–241.

Die vordersten Schichten lernen kleinste Bildelemente wie Linien und Kurven, während spätere Schichten Kompositionen von Elementen bzw. Primitiven aus vorherigen Schichten erlernen. Das Prinzip der hierarchischen Komposition von komplexeren Objekten / Primitiven oder Motiven aus Objekten / Primitiven aus vorigen Schichten zieht sich durch das ganze Netz und ist ein grundlegendes Merkmal von CNNs. Je tiefer man in das Netzwerk vordringt, desto komplexer und umfassender werden die erkannten Merkmale. Diese hierarchische Vorgehensweise entspricht dem *Perceptual Field* in einem CNN: Die frühen Schichten nehmen kleine, lokale Details wahr, während die späteren Schichten eine immer globalere und abstraktere Sichtweise entwickeln. Die räumliche Dimension der gesamten Schicht wird dabei immer komprimierter, während die Dimensionalität der Wissensrepräsentation immer weiter zunimmt. [26]

Im tiefsten Punkt im *U-Net* bzw. die innersten Schichten (*Bottom Layer*) kodieren dabei semantisch das gesamte Gesehene und sind der Kern des *Latent Space* bzw. der Wissensrepräsentation (vgl. Abbildung 4) und sind für die Morphologieuntersuchung der Rocailles am interessantesten. [27]

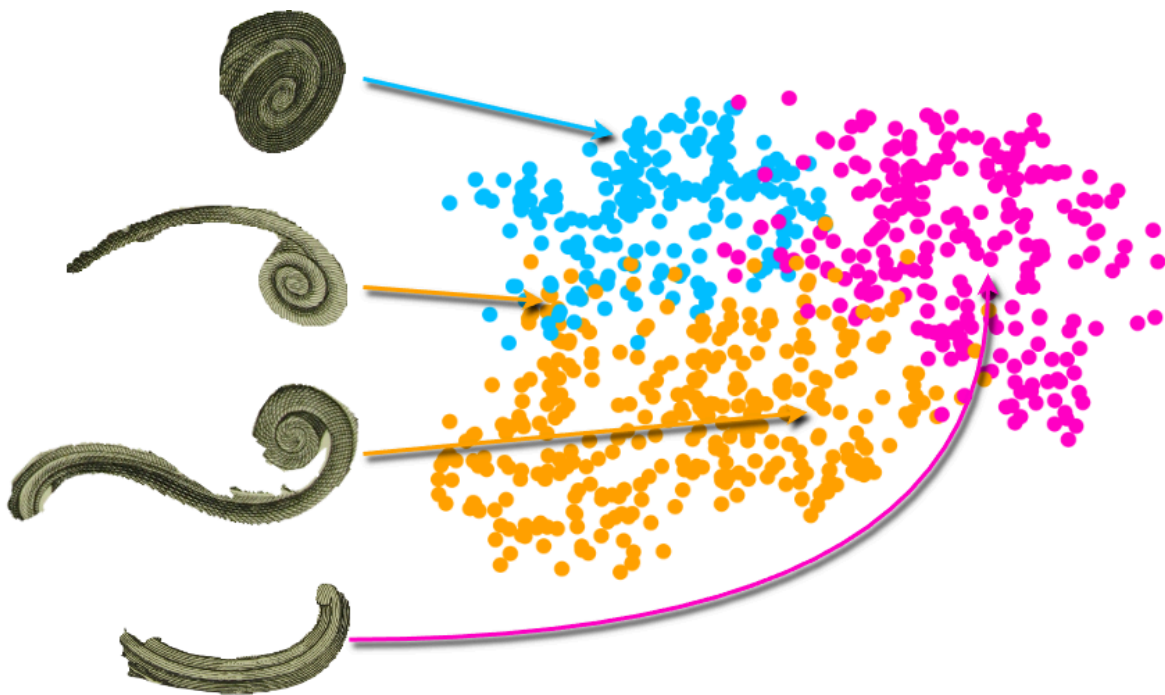


Abb. 4: 2D-Repräsentation eines Latent Space: Encoder-Decoder-Modelle erlernen die (pixelweise) Klassifikation visueller Elemente. Mathematisch entspricht dies der Platzierung nicht-linearer Trenngrenzen (Decision Boundaries) innerhalb eines hochdimensionalen Datensatzes. [Grafik: Thomas Hudcovic 2026]

Dafür müssen sich zunächst Mannigfaltigkeiten identifizieren lassen, also topologisch greifbare Objekte bestehend aus Punktmengen, die in ihrer Konstellation und Ausprägung das Wissen repräsentieren. Falls dies möglich ist, so wäre der nächste Schritt, Korrespondenzen zwischen diesen Mannigfaltigkeiten und den Objekten im Bild zu finden, sodass beispielsweise die lokale Manipulation einer Mannigfaltigkeit, die eine Volute repräsentieren könnte, auch die Volute im Bild entsprechend manipuliert (zum Beispiel verformt). Dadurch ließen sich – so die Hoffnung – Morphologien zwischen Bildern und personalstilistischen Ausprägungen herstellen. Hier begründet sich auch die Wahl der *U-Nets* als Netzarchitektur: Durch den einfachen Aufbau, aber die gleichzeitig relativ hohe Segmentierungsleistung, lassen sich Analysen im *Latent Space* einfacher nachvollziehen. Gleichzeitig dienen die eigens entwickelten Varianten der *U-Nets* als guter Vergleich für die Analyse der Differenz zwischen den jeweiligen Latent Spaces der *U-Nets*, bedingt durch den Einsatz anderer Schichtkomponenten. [28]

Da selbst für *U-Nets* der Trainingsdatensatz vergleichsweise klein ist, wurden dafür geeignete Methoden entwickelt. Diese geringe Menge hat oft zur Folge, dass wichtige Merkmale unzureichend identifiziert und erlernt werden und dass das Netz nur unzureichende Diskriminierungsleistung von Merkmalen oder Kompositionswissen aufbauen kann. Dies wird auf zwei Arten behandelt: Zum Ersten wird die Menge an Trainingsdaten durch *Augmentierung* vervielfacht. Dabei werden Teile von Bildern mit anderen Teilen algorithmisch und zufällig (aber semantisch sinnvoll) gemixt, z. B. mittels *CutMix*,⁴² um so ein neues Bild zu erhalten. *CutMix* wurde für dieses Projekt speziell für Segmentierungsaufgaben und Bildkontexterhaltung modifiziert. Zum Zweiten wird für das Training eine sogenannte *Loss-Funktion* angewandt, die sich als Komposition aus einem pixel- bzw. verteilungsbasierten *Loss* (dem *Focal Loss*)⁴³ und einem regionsbasiertem *Loss* (*Dice Loss*) zusammensetzt. Dadurch werden *False Negatives*, aber auch *False Positives* stärker im Trainingsprozess bestraft und gleichzeitig regionale Kohärenz gefördert. Die nicht vorhandene semantische Invarianz der Rocaille kann realistischerweise nur durch mehr Trainingsdaten effektiv behandelt werden, daher ist der größte direkte Schritt die Augmentierung des Datensatzes. Jedoch hilft die Loss-Funktion sowie das implizite Lernen der Umgebung der markierten Rocaille-Pixel in den Trainingsdaten, kontextbehaftete Diskriminatoren im Wissen aufzubauen. Ein weiterer, zukünftiger Untersuchungsgegenstand, ist demzufolge die Frage danach, ob man dieses mangelhafte oder gar fehlende Wissen im *Latent Space* identifizieren kann.

[29]

Eine weitere Maßnahme gegen die geringe Datenmenge ist das Vortraining der Netze ohne genaues Lernziel. Tatsächlich ist dies ein eigener Forschungszweig im Bereich des *Deep Learnings*, bekannt als sogenanntes *Self-Supervised Learning*. Dieses basiert auf Theorien aus der Neurokognition (die Motivation lässt sich dabei vom realen Lernen eines Kleinkinds ableiten),⁴⁴ aus denen im Rahmen dieses Projektes eine neuartige Vortrainingsmethode entwickelt wurde. Dafür werden die *U-Nets* als Teil von *Denoising Diffusion Probabilistic Models (DDPM)* verwendet (also einer Architektur, in der ein *U-Net* nur ein Baustein ist),⁴⁵ mit dem Ziel, Bilder zu generieren, wie es z. B. bei *Stable Diffusion*⁴⁶ genutzt wird. Das Generieren von Bildern muss einem DDPM ebenfalls antrainiert werden, benötigt aber in seiner Grundauführung keine Labels. Das Netz bekommt ein Bild, welches schrittweise mehr und mehr verrauscht wird. Als Aufgabe soll das Netz das Rauschen in jedem Schritt wieder entfernen und das Originalbild wiederherstellen. Das Rauschen wird dabei schrittweise bis zur vollständigen Unkenntlichkeit des Originalbildes gesteigert; die Aufgabe wird also sukzessive schwerer gestaltet. Dabei bildet das Netz visuelles Wissen der Bilder im *Latent Space* aus. Im Zuge dieses neuartigen Vortrainings werden auch verschiedene Varianten von *U-Nets* und anderen Encoder-Decoder-Netzen auf deren Qualität und Nutzbarkeit für DDPM evaluiert.

[30]

Auf den gesamten, ungelabelten Rocaille-Bilddatensatz angewandt, kann somit ein Netz bereits Vorwissen aufbauen, bevor es gezielt für die Segmentierungsaufgabe auf den Bildern trainiert wird. Die Segmentierungsergebnisse zeigen eine signifikante Anzahl von *False Negatives*, wobei eher wenige *False Positives* auftreten. Ohne Vortraining liegt der *Intersection over Union-Wert* (IoU) bei etwa 0.5, während er mit Vortraining auf etwa 0.7 steigt. Der IoU bezeichnet dabei die Überlappung der Segmentierungsausgabe des Netzes mit den tatsächlichen Segmentierungsmarkierungen in den Labels (vgl. *Abbildung 5*) und ist eine herkömmliche Qualitätsmetrik bei Segmentierungsaufgaben.

[31]

⁴² Vgl. Yun et al. 2019, S. 6022–6031.

⁴³ Vgl. Lin et al. 2017, S. 2980–2988.

⁴⁴ Vgl. Meta AI (Hg.) 2023.

⁴⁵ Vgl. Ho et al. 2020, S. 6840–6851.

⁴⁶ Vgl. Rombach et al. 2022, S. 10674–10685.

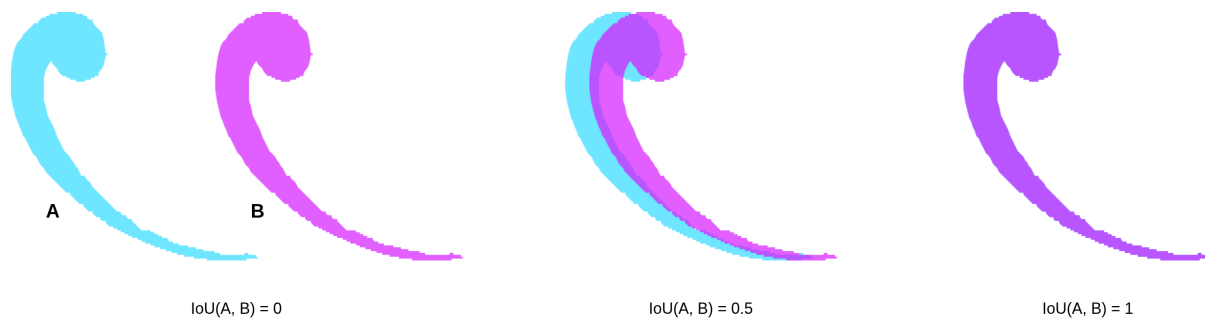


Abb. 5: Die Metrik ›Intersection over Union‹ quantifiziert die Qualität der Segmentierung. Sie bildet das Verhältnis der Abdeckung zwischen Vorhersagemaske (Prediction Mask) und Referenzmaske (Ground Truth) auf einem geschlossenen Intervall von 0 bis 1 ab. [Grafik: Thomas Hudcovic 2026]

Diese Beobachtungen lassen sich durch die Eigenheiten kleiner und unausgewogener Datensätze sowie die Vorteile des Vortrainings erklären. Der IoU-Wert von 0.5 ohne Vortraining deutet darauf hin, dass das Modell Schwierigkeiten hat, die verschiedenen Klassen korrekt zu segmentieren. Dies ist insbesondere bei einem kleinen Datensatz verständlich, da das Modell nicht genügend Beispiele hat, um die Merkmale jeder Klasse zuverlässig zu lernen. Die hohe Anzahl von *False Negatives* weist darauf hin, dass das Modell viele relevante Objekte nicht erkennt, was typisch für unausgewogene Datensätze ist, in denen einige Klassen unterrepräsentiert sind. Die geringe Anzahl von *False Positives* in beiden Szenarien lässt darauf schließen, dass das Modell grundsätzlich gut darin ist, Bereiche zu identifizieren, die nicht zu den Zielklassen gehören. [32]

6. Integration mit einer Datenbank und Nutzbarmachung

Auch wenn durch die Klassen ›Volute‹ und ›Kamm‹ bereits eine Vergleichsbasis generiert werden kann, wird bei der Mikroanalyse von Details hinsichtlich der Morphologie (vor allem personalstilistische Aspekte) nicht nur auf bildliche, sondern auch auf die textuelle Ebene zurückgegriffen. Daher werden im Meta-Datensatz besonders die Beschreibungen aus der Sekundärliteratur berücksichtigt und als Textbasis genutzt. Ergänzt wurden diese durch eigene Beschreibungstexte. Diese textuelle Ebene lässt sich mit der Online-Datenbank an das Projektziel koppeln: So ist es Nutzer*innen möglich, Textsuchen zu tätigen und Informationen zu den Bildern im Datensatz zu bekommen. In einem zusätzlichen Ansatz wird daher auf *Large Language Models (LLMs)* wie OpenAIs ChatGPT⁴⁷ oder Metas Llama⁴⁸ zurückgegriffen, die sich in der sehr jungen Vergangenheit als Netzarchitekturen mit viel Potenzial herausgestellt haben. Basierend auf diesen LLMs hat sich ein weiterer, multimodaler Ansatz entwickelt, die sogenannten *Vision Language Models (VLMs)*,⁴⁹ um die Fähigkeiten von LLMs nicht nur auf Text, sondern auch auf andere Modi, wie zum Beispiel Bilder, zu erweitern. Dabei werden LLM-Architekturen, oft bereits vortrainiert, herangezogen und anschließend mit passenden Text-Bild-Paaren weiter trainiert. Die primäre Herausforderung ist hier das Überführen von Text und Bild in eine gemeinsame Datenrepräsentation, die dann als Eingabe für ein LLM dienen kann. Dies ist auch der wesentlich diskriminierende Faktor von verschiedenen VLM-Architekturen: Das sogenannte *Embedding* von zusätzlichen Inputmodi (Text, Bild, etc.) zum bereits vorhandenen Modus. [33]

VLMs ermöglichen es, Text- und / oder Bildfragen zu stellen und Text- und / oder Bildantworten zu bekommen. Zum Beispiel könnte ein*e Nutzer*in bestimmte Merkmale in einer Rocaille als Textbeschreibung liefern und das VLM versucht dann, passende Bilder, die die Rocaille mit den beschriebenen Merkmalen enthalten, im Datensatz zu identifizieren. [34]

⁴⁷ Vgl. Achiam et al. 2023.

⁴⁸ Vgl. Touvron et al. 2023.

⁴⁹ Vgl. Zhang et al. 2024.

So erlaubt die digital zugängliche Datenbank Anfragen via Bildersuche: Beim Upload eines Fotos mit Rocaille [35] kann eine Erkennung der Volutenspangen und Kämme so erfolgen, dass das Foto möglichst nah in den Stil der Augsburger Drucke mittels *Style Transfer*⁵⁰ überführt wird. So werden etwa die Volutenspangen der Schnitzarbeit an der Türe der Neuen Residenz in Passau erkannt (vgl. Abbildung 6).



Abb. 6: Fotografie einer Schnitzarbeit an der Neuen Residenz in Passau (links), Volutenerkennung und Styletransfer (rechts). [Grafik: Thomas Hudcovic 2026]

Folglich bietet die Suchmaschine eine Volltextsuche mit *Fuzzy-Search* sowie eine semantische Suche auf Grundlage der hochgeladenen Bilder. Als Ergebnis werden ähnliche Augsburger Druckgrafiken präsentiert. Die inhaltlichen Bezüge können sich dabei auf Typ, Stil, Rocaille-Komposition, Figuren oder einzelne Motive beziehen. [36]

⁵⁰ Vgl. Gatys et al. 2016, S. 2414–2423.

Bibliografie

Josh Achiam / Steven Adler / Sandhini Agarwal / Lama Ahmad / Ilge Akkaya / Florencia Leoni Aleman / Diogo Almeida / Janko Altmerschmidt / Sam Altman / Shyamal Anadkat / Red Avila / Igor Babuschkin / Suchir Balaji / Valerie Balcom / Paul Baltescu / Haiming Bao / Mohammad Bavarian / Jeff Belgum / Irwan Bello / Jake Berdine / Gabriel Bernadett-Shapiro / Christopher Berner / Lenny Bogdonoff / Oleg Boiko / Madelaine Boyd / Anna-Luisa Brakman / Greg Brockman / Tim Brooks / Miles Brundage / Kevin Button / Trevor Cai / Rosie Campbell / Andrew Cann / Brittany Carey / Chelsea Carlson / Rory Carmichael / Brooke Chan / Che Chang / Fotis Chantzis / Derek Chen / Sully Chen / Ruby Chen / Jason Chen / Mark Chen / Ben Chess / Chester Cho / Casey Chu / Hyung Won Chung / Dave Cummings / Jeremiah Currier / Yunxing Dai / Cory Decareaux / Thomas Degry / Noah Deutsch / Damien Deville / Arka Dhar / David Dohan / Steve Dowling / Sheila Dunning / Adrien Ecoffet / Atty Eleti / Tyna Eloundou / David Farhi / Liam Fedus / Niko Felix / Simón Posada Fishman / Juston Forte / Isabella Fulford / Leo Gao / Elie Georges / Christian Gibson / Vik Goel / Tarun Gogineni / Gabriel Goh / Rapha Gontijo-Lopes / Jonathan Gordon / Morgan Grafstein / Scott Gray / Ryan Greene / Joshua Gross / Shixiang Shane Gu / Yufei Guo / Chris Hallacy / Jesse Han / Jeff Harris / Yuchen He / Mike Heaton / Johannes Heidecke / Chris Hesse / Alan Hickey / Wade Hickey / Peter Hoeschele / Brandon Houghton / Kenny Hsu / Shengli Hu / Xin Hu / Joost Huizinga / Shantanu Jain / Shawn Jain / Joanne Jang / Angela Jiang / Roger Jiang / Haozhun Jin / Denny Jin / Shino Jomoto / Billie Jonn / Heewoo Jun / Tomer Kaftan / Łukasz Kaiser / Ali Kamali / Ingmar Kanitscheider / Nitish Shirish Keskar / Tabarak Khan / Logan Kilpatrick / Jong Wook Kim / Christina Kim / Yongjik Kim / Jan Hendrik Kirchner / Jamie Kiros / Matt Knight / Daniel Kokotajlo / Łukasz Kondraciuk / Andrew Kondrich / Aris Konstantinidis / Kyle Kopic / Gretchen Krueger / Vishal Kuo / Michael Lampe / Ikai Lan / Teddy Lee / Jan Leike / Jade Leung / Daniel Levy / Chak Ming Li / Rachel Lim / Molly Lin / Stephanie Lin / Mateusz Litwin / Theresa Lopez / Ryan Lowe / Patricia Lue / Anna Makanju / Kim Malfacini / Sam Manning / Todor Markov / Yaniv Markovski / Bianca Martin / Katie Mayer / Andrew Mayne / Bob McGrew / Scott Mayer McKinney / Christine McLeavey / Paul McMillan / Jake McNeil / David Medina / Aalok Mehta / Jacob Menick / Luke Metz / Andrey Mishchenko / Pamela Mishkin / Vinnie Monaco / Evan Morikawa / Daniel Mossing / Tong Mu, Mira Murati / Oleg Murk / David Mély / Ashvin Nair / Reiichiro Nakano / Rajeev Nayak / Arvind Neelakantan / Richard Ngo / Hyeonwoo Noh / Long Ouyang / Cullen O’Keefe / Jakub Pachocki / Alex Paino / Joe Palermo / Ashley Pantuliano / Giambattista Parascandolo / Joel Parish / Emy Parparita / Alex Passos / Mikhail Pavlov / Andrew Peng / Adam Perelman / Filipe de Avila Belbute Peres / Michael Petrov / Henrique Ponde de Oliveira Pinto / Michael (Rai) Pokorny / Michelle Pokrass / Vitvichyr H. Pong / Tolly Powell / Alethea Power / Boris Power / Elizabeth Proehl / Raul Puri / Alec Radford / Jack Rae / Aditya Ramesh / Cameron Raymond / Francis Real / Kendra Rimbach / Carl Ross / Bob Rotsted / Henri Roussez / Nick Ryder / Mario Saltarelli / Ted Sanders / Shibani Santurkar / Girish Sastry / Heather Schmidt / David Schnurr / John Schulman / Daniel Selsam / Kyla Sheppard / Toki Sherbakov / Jessica Shieh / Sarah Shoker / Pranav Shyam / Szymon Sidor / Eric Sigler / Maddie Simens / Jordan Sitkin / Katarina Slama / Ian Sohl / Benjamin Sokolowsky / Yang Song / Natalie Staudacher / Felipe Petroski Such / Natalie Summers / Ilya Sutskever / Jie Tang / Nikolas Tezak / Madeleine B. Thompson / Phil Tillet / Amin Tootoonchian / Elizabeth Tseng / Preston Tuggle / Nick Turley / Jerry Tworek / Juan Felipe Cerón Uribe / Andrea Vallone / Arun Vijayarvija / Chelsea Voss / Carroll Wainwright / Justin Jay Wang / Alvin Wang / Ben Wang / Jonathan Ward / Jason Wei / CJ Weinmann / Akila Welihinda / Peter Welinder / Jiayi Weng / Lilian Weng / Matt Wiethoff / Dave Willner / Clemens Winter / Samuel Wolrich / Hannah Wong / Lauren Workman / Sherwin Wu / Jeff Wu / Michael Wu / Kai Xiao / Tao Xu / Sarah Yoo / Kevin Yu / Qiming Yuan / Wojciech Zaremba / Rowan Zellers / Chong Zhang / Marvin Zhang / Shengjia Zhao / Tianhao Zheng / Juntang Zhuang / William Zhuk / Barret Zoph: Gpt-4 Technical Report. arXiv. 15.03.2023. Version 6 vom 04.03.2023. DOI: [10.48550/arXiv.2303.08774](https://arxiv.org/abs/10.48550/arXiv.2303.08774)

Michael von Albrecht (Hg.): Publius Ovidius Naso. Metamorphoseon libri quindecim. Lateinisch / Deutsch. Stuttgart 2013. [\[Nachweis im GVK\]](#)

Wolfgang Augustyn: Augsburgs Buchillustration im 18. Jahrhundert. In: Helmut Gier / Johannes Janota (Hg.): Augsburgs Buchdruck und Verlagswesen von den Anfängen bis zur Gegenwart. Wiesbaden 1997, S. 791–861. [\[Nachweis im GVK\]](#)

Herrmann Bauer: Rocaille. Zur Herkunft und zum Wesen eines Ornament-Motivs. Berlin 1962. [\[Nachweis im GVK\]](#)

Edward Brandt / Wilhelm Ehlers (Hg.): Apuleius. Der goldene Esel. Metamorphosen Libri XI. Lateinisch-deutsch. 6., überarbeitete Auflage. Berlin 2012. [\[Nachweis im GVK\]](#)

Liang-Chieh Chen / Yukun Zhu / George Papandreou / Florian Schroff / Hartwig Adam: Encoder-decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: ECCV 2018. 15th European Conference on Computer Vision. Proceedings (München, 08.09.–14.09.2018). München 2018, S. 801–818. PDF. [\[online\]](#)

Bowen Cheng / Ishan Misra / Alexander G. Schwing / Alexander Kirillov / Rohit Girdhar: Masked-Attention Mask Transformer for Universal Image Segmentation. In: IEEE / CVPR 2022. Conference on Computer Vision and Pattern Recognition. Proceedings (New Orleans, 18.06.–24.06.2022). New Orleans 2022, S. 1290–1299. PDF. [\[online\]](#)

Leon A. Gatys / Alexander S. Ecker / Matthias Bethge: Image Style Transfer Using Convolutional Neural Networks. In: IEEE / CVPR 2016. Conference on Computer Vision and Pattern Recognition. Proceedings (Las Vegas, 27.06.–30.06.2016). Las Vegas 2016, S. 2414–2423. PDF. [\[online\]](#)

Helmut Gier / Johannes Janota (Hg.): Augsburgs Buchdruck und Verlagswesen. Von den Anfängen bis zur Gegenwart. Wiesbaden 1997. [\[Nachweis im GVK\]](#)

Kaiming He / Georgia Gkioxari / Piotr Dollár / Ross Girshick: Mask R-CNN. In: IEEE / ICCV 2017. International Conference on Computer Vision. Konferenzabstracts (Venedig, 22.10.–29.10.2017). Venedig, IT 2017, S. 2961–2969. PDF. DOI: [10.1109/ICCV.2017.322](https://doi.org/10.1109/ICCV.2017.322)

Gun-Dagmar Helke: Johann Esaias Nilson (1721–1788). Augsburgs Miniaturmaler, Kupferstecher, Verleger und Kunstakademiedirektor. München 2005. [\[Nachweis im GVK\]](#)

Jonathan Ho / Ajay Jain / Pieter Abbeel: Denoising Diffusion Probabilistic Models. In: Hugo Larochelle / Marc’Aurelio Ranzato / Raia Hadsell / Maria Floriana Balcan / Hsuan-Tien Lin (Hg.): NeurIPS 2020. 34th Annual Conference on Neural Information Processing Systems. Proceedings (Online, 06.12.–12.12.2020). Vancouver 2020, S. 6840–6851. PDF. [\[online\]](#)

Günter Irmscher: Style rocaille. In: Barockberichte 51 / 52 (2009), S. 339–414. [\[Nachweis im GVK\]](#)

Günter Irmscher: Ornament in Europa. 1450–2000. Eine Einführung. Köln 2005. [\[Nachweis im GVK\]](#)

Eduard Isphording: Gottfried Bernhard Göz 1708–1774. 2 Bände. Weissenhorn 1982. Band 1: Ölgemälde und Zeichnungen (Textband). [\[Nachweis im GVK\]](#)

Eduard Isphording: Gottfried Bernhard Göz 1708–1774. 2 Bände. Weissenhorn 1984. Band 2: Ölgemälde und Zeichnungen (Tafelband). [\[Nachweis im GVK\]](#)

Fiske Kimball: The Creation of the Rococo Decorative Style. New York 1980 [1943]. [\[Nachweis im GVK\]](#)

Katharina Krause: Sans théorie, sans raisonnement, sans goût, sans invention. Ornamentstich als Medium von Erfindung und Verbreitung von Ideen im Kunsthandwerk des 18. Jahrhunderts. In: Christof Jeggle (Hg.): Luxusgegenstände und Kunstwerke vom Mittelalter bis zur Gegenwart. Produktion – Handel – Formen der Aneignung. Konstanz 2015, S. 185–199. [\[Nachweis im GVK\]](#)

Friedrich August Krubsacius: Gedanken von dem Ursprunge, Wachstume und Verfall der Verzierungen in den schönen Künsten, d. i. Der Bau- Schnitz- Maler- und Kupferstecherkunst. Leipzig 1759. [\[Nachweis im GVK\]](#)

Ebba Krull: Franz Xaver Habermann (1721–1796). Ein Augsburgs Ornamentist des Rokoko. Augsburg 1977. [\[Nachweis im GVK\]](#)

Karin Leonhard: Ornament und Zeitlichkeit. Kartusche, Rocaille, Arabeske. In: Thomas Kisser (Hg.): Bild und Zeit. Temporalität in der Kunst und Kunsttheorie seit 1800. München 2011, S. 63–85. [\[Nachweis im GVK\]](#)

Patrick Lewis / Ethan Perez / Aleksandra Piktus / Fabio Petroni / Vladimir Karpukhin / Naman Goyal / Heinrich Küttler / Mike Lewis / Wen-tau Yih / Tim Rocktäschel / Sebastian Riedel / Douwe Kiela: Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In: Hugo Larochelle / Marc’Aurelio

- Ranzato / Raia Hadsell / Maria Floriana Balcan / Hsuan-Tien Lin (Hg.): NeurIPS 2020. 34th Annual Conference on Neural Information Processing Systems. Proceedings (Online, 06.12.–12.12.2020). Vancouver, CA 2020, S. 9459–9474. PDF. [\[online\]](#)
- Tsung-Yi Lin / Priya Goyal / Ross Girshick / Kaiming He / Piotr Dollár: Focal Loss for Dense Object Detection. In: IEEE / ICCV 2017. International Conference on Computer Vision. Proceedings (Venedig, 22.10.–29.10.2017). Venedig 2017, S. 2980–2988. PDF. [\[online\]](#)
- Marie-Thérèse Mandroux-França: Information artistique et ›masse-media‹ au 18 siècle. La diffusion de l'ornement gravé rococo au Portugal. In: Bracara Augusta 27 (1973), S. 412–445. [\[Nachweis im GVK\]](#)
- Meta AI (Hg.): Yann LeCun on a Vision to Make AI Systems Learn and Reason like Animals and Humans. Letzte Aktualisierung 23.02.2023. HTML. [\[online\]](#)
- John Roger Paas (Hg.): Augsburg, die Bilderfabrik Europas. Essays zur Augsburger Druckgraphik der frühen Neuzeit. Augsburg 2001. [\[Nachweis im GVK\]](#)
- John Roger Paas / Josef H. Biller / Marie-Luise Hopp-Gantner (Hg.): Gestochen in Augsburg. Forschungen und Beiträge zur Geschichte der Augsburger Druckgrafik. Hommage à Wolfgang Seitz zum 90. Geburtstag 2011 und Festgabe zum 40. Jahrestag des von ihm gegründeten Augsburger Forscherkreises 1973–2013 (= Schwäbische Geschichtsquellen und Forschungen, 29). Augsburg 2013. [\[Nachweis im GVK\]](#)
- Robin Rombach / Andreas Blattmann / Dominik Lorenz / Patrick Esser / Björn Ommer: High-resolution Image Synthesis with Latent Diffusion Models. In: IEEE / CVPR 2022. Conference on Computer Vision and Pattern Recognition. Proceedings (New Orleans, 18.06.–24.06.2022). New Orleans, US-LA 2022, S. 10674–10685. DOI: 10.1109/CVPR52688.2022.01042
- Olaf Ronneberger / Philipp Fischer / Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Nassir Navab / Joachim Hornegger / William Wells / Alejandro F. Frangi (Hg.): MICCAI 2015. Medical Image Computing and Computer-Assisted Intervention. Lecture Notes in Computer Science 9351 (2015), S. 234–241. PDF. DOI: 10.1007/978-3-319-24574-4_28
- Hermann Schmitz: Katalog der Ornamentstichsammlung der Staatlichen Kunstbibliothek Berlin. Berlin u. a. 1939. [\[Nachweis im GVK\]](#)
- Ulrike Seeger: Abgekupfert, aufgefächert, angewandt: Maria Philippina Küssels »gemalter Himmel«. Zur Rezeption französischer Ornamentstiche via Augsburg. In: Sabine Frommel / Eckhard Leuschner (Hg.): Architektur- und Ornamentgraphik der Frühen Neuzeit. Migrationsprozesse in Europa. Rom 2014, S. 285–294. [\[Nachweis im GVK\]](#)
- Hugo Touvron / Thibaut Lavril / Gautier Izacard / Xavier Martinet / Marie-Anne Lacaux / Timothee Lacroix / Baptiste Rozière / Naman Goyal / Eric Hambro / Faisal Azhar / Aurelien Rodriguez / Armand Joulin / Edouard Grave / Guillaume Lample: LLaMA. Open and Efficient Foundation Language Models. arXiv. 27.02.2023. DOI: [10.48550/arXiv.2302.13971](#)
- Rudolf Wildmoser: Gottfried Bernhard Göz (1708–1774) als ausführender Kupferstecher. Untersuchung und Katalog der Werke. In: Jahrbuch des Vereins für Augsburger Bistumsgeschichte e. V. 18 (1984), S. 257–340. [\[Nachweis im GVK\]](#)
- Rudolf Wildmoser: Gottfried Bernhard Göz (1708–1774) als ausführender Kupferstecher. Untersuchung und Katalog der Werke. Teil II: Katalog. In: Jahrbuch des Vereins für Augsburger Bistumsgeschichte e.V. 19 (1985), S. 140–296. [\[Nachweis im GVK\]](#)
- Sangdoon Yun / Dongyoon Han / Seong Joon Oh / Sanghyuk Chun / Junsuk Choe / Youngjoon Yoo: CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features. In: IEEE / ICCV 2019. International Conference on Computer Vision. Proceedings (Seoul, 27.10.–02.11.2019). Seoul 2019, S. 6022–6031. PDF. DOI: 10.1109/ICCV.2019.00612
- Jingyi Zhang / Jiaxing Huang / Sheng Jin / Shijian Lu: Vision-Language Models for Vision Tasks: A Survey. In: IEEE Transactions on Pattern Analysis and Machine Intelligence 46 (2024), H. 8, S. 5625–5644. PDF. DOI: 10.1109/TPAMI.2024.3369699
- Sixiao Zheng / Jiachen Lu / Hengshuang Zhao / Xiatian Zhu / Zekun Luo / Yabiao Wang / Yanwei Fu / Jianfeng Feng / Tao Xiang / Philip H. S. Torr / Li Zhang: Rethinking Semantic Segmentation from a Sequence-to-sequence Perspective with Transformers. In: IEEE / CVF 2021. Conference on Computer Vision and Pattern Recognition. Proceedings (Nashville, 20.06.–25.06.2021). Nashville 2021, S. 6881–6890. PDF. [\[online\]](#)

Abbildungsverzeichnis

- Abb. 1: ›L'Amour Couronne. / Der gekrönte Cupido‹. Sechstes Blatt einer Folge von sieben Blättern, entworfen von Jean Mondon, herausgegeben bei Johann Georg Merz in Augsburg. [Fotografie: Staats- und Stadtbibliothek Augsburg, Graph Merz, J.G. 41]
- Abb. 2: Rocaillekompositionen. Zweites Blatt einer Folge von vier Blättern, entworfen von Franz Xaver Habermann, herausgegeben von Johann Georg Hertel in Augsburg. [Fotografie: © Staatsgalerie Stuttgart, Graphische Sammlung, alter Bestand]
- Abb. 3: Modellarchitektur des U-Net: Inspiriert vom biologischen visuellen Kortex ermöglicht der hierarchische Aufbau von Faltungsschichten (Convolutional Layers), visuelle Merkmale in unterschiedlicher semantischer Komplexität zu erfassen und zu erlernen. [Grafik: Thomas Hudcovic 2026]
- Abb. 4: 2D-Repräsentation eines Latent Space: Encoder-Decoder-Modelle erlernen die (pixelweise) Klassifikation visueller Elemente. Mathematisch entspricht dies der Platzierung nicht-linearer Trenngrenzen (Decision Boundaries) innerhalb eines hochdimensionalen Datensatzes. [Grafik: Thomas Hudcovic 2026]
- Abb. 5: Die Metrik ›Intersection over Union‹ quantifiziert die Qualität der Segmentierung. Sie bildet das Verhältnis der Abdeckung zwischen Vorhersagemaske (Prediction Mask) und Referenzmaske (Ground Truth) auf einem geschlossenen Intervall von 0 bis 1 ab. [Grafik: Thomas Hudcovic 2026]
- Abb. 6: Fotografie einer Schnitzarbeit an der Neuen Residenz in Passau (links), Volutenerkennung und Styletransfer (rechts). [Grafik: Thomas Hudcovic 2026]