

Universität Heidelberg  
Zentrum für Europäische Kultur- und Geschichtswissenschaften  
Historisches Seminar  
Sommersemester 2025

## **Vom Bildnis zum Bit**

### Deep Learning zur Analyse der Ikonographie antiker Münzen

Arbeit zur Erlangung des Grades eines Master of Arts

vorgelegt von  
Mattia Celisi

vorgelegt bei Prof. Dr. Cord Arendes  
und Prof. Dr. Fred Hamprecht

Mattia Celisi

Global History, M.A.

6. Fachsemester

Matrikelnummer: 4123310

Apothekergasse 1, 69117 Heidelberg

[mattia.celisi@stud.uni-heidelberg.de](mailto:mattia.celisi@stud.uni-heidelberg.de)

1.	Einleitung .....	1
1.1.	Problemstellung und Potential .....	1
1.2.	Forschungsstand .....	7
1.2.1	Natural Language Processing .....	7
1.2.2	Bildererkennung durch neuronale Netze .....	12
2.	Von Münzen zu Forschungsdaten .....	20
2.1.	Bestandteile antiker Münzen .....	21
2.2.	Entwicklung ihrer Klassifizierung .....	26
2.3.	Technische Umsetzung – FAIRe Forschungsdaten .....	28
3.	Von Forschungsdaten zur automatisierten Münzklassifikation .....	32
3.1.	Münzbeschreibungen und Named Entity Recognition .....	32
3.1.1	Umsetzung durch den Knowledge Graph .....	36
3.1.2	Umgang mit Mehrdeutigkeit .....	41
3.1.3	Fehlerquellen und Analyseprobleme .....	48
3.2.	Münzbildvorbereitung .....	54
3.3.	Das Convolutional Neural Network .....	59
3.3.1	Training und Gütekriterien .....	59
3.3.2	Qualitative Überprüfung .....	67
4.	Schluss .....	70
5.	Abkürzungsverzeichnis: .....	74
6.	Literaturverzeichnis .....	75
6.1.	Quellenverzeichnis .....	75
6.2.	Literaturverzeichnis .....	77
7.	Abbildungsverzeichnis .....	81

# 1. Einleitung<sup>1</sup>

## 1.1. Problemstellung und Potential

Die Numismatik befasst sich als historische Grundwissenschaft mit Münzen, anderen Geldformen und verwandten Objekten.<sup>2</sup> Eine ihrer grundlegenden Aufgaben ist die Bestimmung und Klassifizierung ihrer Quellen, ein aufwendiger Abgleich mit Beschreibungen in Referenzkatalogen.<sup>3</sup> Deshalb wundert es nicht, dass das kleine Universitätsfach seit rund zwanzig Jahren eine kontinuierlich wachsende Zahl an Publikationen verzeichnet, die eine Automatisierung dieses Prozesses durch den Einsatz von Bilderkennungsprogrammen, die von der Informatik zur Verfügung gestellt werden, in den Blick nehmen. Nach wie vor sind die Ergebnisse diesbezüglich jedoch noch nicht vollständig überzeugend und das Gelingen dieser Automatisierung wäre die ersehnte Revolution, die zu einer Umstrukturierung des Arbeitsfeldes führen würde. Kapazitäten würden frei für Forschungsfragen, die seither aufgrund der verzögerten Datengewinnung nur langsam oder überhaupt nicht in Angriff genommen werden konnten. Zudem tritt der Münzmarkt nach wie vor mit einem Anwendungsinteresse an die Wissenschaft heran, der als Motor für akademische Bemühungen zwar nicht im Vordergrund steht, aber dennoch als Ergänzung für das Gesamtbild nicht ungenannt bleiben darf.

Antike Münzen sind im Vergleich zu anderen archäologischen Quellen Massenprodukte und stehen der Forschung in besonders großem Umfang zur Verfügung. Das macht die Quelle nicht nur für die Beschreibung im klassisch ikonografischen Sinn interessant, sondern sie bildet auch einen veritablen Korpus für quantitative Analysen. Diese Tatsache musste das interdisziplinäre Interesse am Gegenstand, namentlich der Austausch zwischen Geschichts- und informationstechnischer Wissenschaft, befördern. So verfügen die Autorinnen und Autoren der in dieser Arbeit zitierten Studien alle über das informatische Fachwissen, das es ihnen erlaubt, erprobte Bilderkennungsalgorithmen zu nutzen, neuerdings auch auf Basis von künstlicher Intelligenz (KI).

Der Quellenbestand der Numismatik und insbesondere an (antiken) Münzen ist zahlenmäßig unüberschaubar, unübersichtlich und damit auch schwierig zu ordnen. Seit dem 19. Jahrhundert

---

<sup>1</sup> Der Autor dankt der Unterstützung durch das Land Baden-Württemberg für die Nutzung des Hochleistungsrechners BwUniCluster2.0, durch den das Training der KI möglich war.

<sup>2</sup> WEISSER, Bernhard: Numismatiker in Deutschland. Zur Situation im Jahre 2019, in: Geldgeschichtliche Nachrichten, hrsg. v. Gesellschaft für Internationale Geldgeschichte, S. 264–277, hier S. 264.

<sup>3</sup> GUO, Zhongliang u.a.: A Siamese Transformer Network for Zero-Shot Ancient Coin Classification, in: Journal of Imaging 9, 6 (2023), hier S. 1.

hat sich die Numismatik ein standardisiertes Vokabular erarbeitet, um ihrer Quellen beschreibend Herr zu werden. Dieses Vokabular beinhaltet objektive Kategorien wie Maß oder Material sowie interpretative Kategorien zur Beschreibung und Deutung der Abbildungen auf den Münzen. Mittels dieser Beschreibungskategorien werden Münzen auch über lange Zeiträume hinweg vergleichbar und bei ausreichender Datenmenge auch für wirtschafts- und sozialhistorische Fragestellungen interessant. Zu diesen gehört zum Beispiel der Grad der Monetarisierung und die Zirkulation von Münzen in den jeweiligen münzprägenden Gesellschaften. Diese Fragestellungen wurden und werden anhand der nach objektiven Kriterien gewonnenen Daten beantwortet. Die Abbildungen auf den Münzen spielen in der quantitativen Forschung im Grunde keine Rolle,<sup>4</sup> da die Sprache, die diese Abbildungen zu erfassen versucht, trotz allgemeiner und anerkannter Beschreibungsregeln zu variationsreich ist.

Das Kardinalproblem ist also: Quantitativen Untersuchungen fehlt der direkte Zugang zu den numismatischen Objekten, es kann nur untersucht werden, was über die Münzen geschrieben wird. Die Beschreibungen sind jedoch zu variationsreich und die Menge der Objekte ist zu groß, als dass eine sinnvolle Klassifizierung vorgenommen werden kann. Es muss eine Methode gefunden werden, bei der, von numismatischen Beschreibungsregeln ausgehend, die Münzen möglichst einheitlich, so präzise wie möglich und hinreichend genau beschrieben und mithin klassifiziert werden können. Dies ist insbesondere dann wichtig, wenn Daten über Münzen verlangt werden, die von unterschiedlichen Institutionen stammen. Es wird noch zu zeigen sein, dass die zugrunde liegenden numismatisch-ikonographischen Konventionen für die Beschreibung erheblich voneinander abweichen, je nachdem welcher Katalog bzw. welche Datenbank herangezogen wird. Dem Forschungsfeld der automatisierten Bilderkennung gelingt noch keine Vereinheitlichung dieser Beschreibungstexte. Häufig werden hier die Algorithmen darauf trainiert, standardisiert erfasste und maschinenlesbare Kriterien wie den Herausgeber oder die Münzstätten zu erkennen. Eine KI, die die herkömmliche händische Klassifikation von Münzen komplett ersetzen möchte, muss allerdings auch viele interpretative, nicht strukturiert verfasste Texte und Textteile miteinschließen. Gerade in ihnen befinden sich die meisten Informationen, die für die umfassende Klassifikation notwendig sind. Nur so kann die KI den semantischen Zusammenhang von Münzelementen – ihre Korrelationen und Kausalitäten – verstehen.<sup>5</sup>

---

<sup>4</sup> HIJMANS, Steven: Art and Economy: Images on Roman Coins, in: Sessions 4–5, Single Contributions (2023), hier S. 521.

<sup>5</sup> COOPER, Jessica/ARANDJELOVIĆ, Ognjen: Understanding Ancient Coin Images, 2019, S. 1, online: <http://arxiv.org/pdf/1903.02665v2> [25.08.2025].

Die vorliegende Arbeit nimmt diese interpretativen Münzbeschreibungen in den Blick, um daraus eine systematische und maschinenverwertbare Ikonographie abzuleiten. Diese systematisierende Arbeit ist notwendig, da die Beschreibungen nur Deutungen der Darstellungen auf Münzen sind und sich damit zwangsläufig unterscheiden.<sup>6</sup> Für diesen Vorgang ist es hilfreich, wenn die Daten der Ausgangslage bereits relativ standardisiert sind und damit einen geringeren Vereinheitlichungsaufwand erfordern. Die Numismatik, die an den Fürstenhöfen der Renaissance entstand, beschäftigt sich bis heute am ausführlichsten mit antiken Münzen, die deshalb besser beschrieben sind als die Münzen anderer Epochen.<sup>7</sup> Um den Untersuchungsgegenstand hinreichend räumlich einzugrenzen, wurden zusätzlich nur Münzen der römischen Antike ausgewählt, die wiederum innerhalb der Antike die bestbeschriebenen Münzen vorweisen kann. Münzen der römischen Antike umfassen konventionell den Zeitraum zwischen dem Beginn der römischen Münzprägung circa 320 v. Chr. und dem Amtsantritt des oströmischen Kaisers Anastasios I. im Jahr 491 n. Chr., mit dessen Währungsreform fachintern das Ende der Spätantiken Numismatik markiert wird.

Für Bilderkennungsalgorithmen gilt grundsätzlich, dass Daten in hoher Qualität und Anzahl verfügbar sein müssen. Moderne Münzen können bereits seit Jahrzehnten mit optischen Algorithmen automatisiert erkannt werden, die jedoch bei in Handarbeit hergestellten antiken Münzen nicht funktionieren.<sup>8</sup> Allerdings schreitet die Digitalisierung von Münzsammlungen stetig voran. Und auch hier gilt der Zusammenhang: je mehr Daten vorliegen, desto mehr sind Variationen innerhalb einzelner Phänomene in ihnen abgebildet. Es ist zu erwarten, dass auch eine automatisierte Klassifikation historischer Münzen bei einer ausreichenden Datengrundlage möglich ist.

Die vorliegende Arbeit zielt aus dem genannten Grund darauf ab, möglichst alle verfügbaren und ausreichend beschriebenen Münzdaten zu nutzen, wobei gezeigt sein wird, dass nicht alle dieser Datenquellen auch die notwendigen Qualitätsanforderungen erfüllen. Die infrage kommenden Datenherkünfte lassen sich in drei Hauptkategorien unterteilen: erstens museale und universitäre Einrichtungen, zweitens Archive und Denkmalämter und drittens der Münzhandel. Die Digitalisierung von Münzbeständen wird von diesen Institutionen unterschiedlich priorisiert. Museale und universitäre Sammlungen (GLAM: Galleries, Libraries, Archives,

---

<sup>6</sup> PAVLEK, Barbara/WINTERS, James/MORIN, Olivier: Standards and Quantification of Coin Iconography. Possibilities and Challenges, in: Digital Scholarship in the Humanities 37, 1 (2022), S. 202–217, hier S. 208.

<sup>7</sup> HOWGEGO, Christopher: Ancient history from coins (Approaching the ancient world), London, New York, 1995.

<sup>8</sup> NÖLLE, Michael u.a.: Dagobert - A New Coin Recognition and Sorting System, 2003.

Museums) haben die Digitalisierung ihrer Bestände zu einer ihrer Kernaufgaben gemacht.<sup>9</sup> Die in GLAM-Institutionen erfassten Münzen werden mit spezialisierten Erfassungssystemen katalogisiert und unterliegen einer mehrstufigen Qualitätskontrolle. Diese beinhaltet, dass viele Datenfelder standardisiert und normdatenbasiert<sup>10</sup> sind. Durch diese Praxis ist die Numismatik eine Vorreiterin der Digitalisierung innerhalb der Geisteswissenschaften.<sup>11</sup>

Im Gegensatz zu GLAM-Institutionen arbeiten Denkmalämter ihre Bestände aus Personalmangel oftmals nicht auf und es fehlen darüber hinaus zentrale Internetportale zum Abruf der Daten.<sup>12</sup> Der Münzhandel besitzt wiederum einen immensen, wenn auch schwer bezifferbaren Anteil an Münzdaten – allein das Vergleichsportal Coin Archives listet über 2,5 Millionen Münzen.<sup>13</sup> Sie unterliegen allerdings nicht dem gleichen Grad an Vernetzung und Qualitätskontrolle wie institutionelle Daten.<sup>14</sup> Seit die institutionellen Sammlungen große Datenmengen zur Verfügung stellen, werden die Datenbanken des Münzhandels nicht mehr für automatisierte Münzerkennungen verwendet.<sup>15</sup> Mittelfristig ist es dennoch unerlässlich, die Daten des Münzhandels heranzuziehen, da manche besonders seltene Münzen nur dort zu finden sind. Der Aufwand, diese Daten zu bereinigen, ist allerdings kaum abzuschätzen. Die vorliegende Arbeit verwendet deshalb ausschließlich digitalisierte Münzdaten von GLAM-Institutionen.

Die Entwicklung einer präzisen und zuverlässigen Münzerkennungs-KI bietet ein hohes praktisches Potential für die Münzerfassung. Die durch die Automatisierung um ein Vielfaches gesenkten Klassifikationskosten erlauben die Katalogisierung derjenigen Bestände, die durch den

---

<sup>9</sup> WEISSER, Numismatiker in Deutschland, S. 268. In den deutschen Webportalen für Sammlungsmünzen Ikmk.net und Kenom.de befinden sich beispielsweise 200 000 Münzen unterschiedlicher Epochen. Auch international wurden großen Sammlungen wie die der American Numismatic Society (800 000 Münzen), des British Museums (480 000 Münzen) und der Bibliothèque Nationale de France digitalisiert.

<sup>10</sup> Zu Normdaten und deren Nutzung in dieser Arbeit vgl. Kapitel 1.2.1 „Natural Language Processing“.

<sup>11</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 212.

<sup>12</sup> Eine Ausnahme bildet zum Beispiel die systematische Erfassung der Münzen aus Sachsen-Anhalt im SESAM-Projekt (<https://www.sesam-projekt.de>). Diese Daten werden in das Portal KENOM eingespielt.

<sup>13</sup> Coin Archives sammelt Daten zu vergangenen und künftigen Auktionen der weltweit größten Auktionshäuser. Vgl. <https://www.coinarchives.com>.

<sup>14</sup> Die Beschreibungen der Münzen des Münzhandels folgen nicht immer dem wissenschaftlichen Standard. Teilweise werden Auktionsbilder nicht einzeln, sondern in einem Raster hochgeladen und sind damit nicht nutzbar. Außerdem liegen die Textdaten nicht nur auf Deutsch und Englisch, sondern auch auf Französisch, Spanisch und Italienisch vor, was die Auswertung um ein Vielfaches verkompliziert. Auch der Download der Daten ist erschwert, da die Webseiten des Münzhandels im Vergleich zu denen musealer Institutionen ein stärkeres Bedürfnis, ihre Seiten vor Crawlern zu schützen, und Sicherheitsfeatures dagegen eingebaut. Mittelfristig werden aber wohl auch die Daten des Münzhandels nutzbar. Erste Ansätze zur Einbindung von Auktionsmünzen wurden bereits gemacht. Der RPC-online verknüpft teilweise Daten aus dem Münzhandel mit ihren Münztypen.

<sup>15</sup> Eine Ausnahme bilden COOPER, Jessica/ARANDJELOVIĆ, Ognjen: Understanding Ancient Coin Images, 2019, online: <http://arxiv.org/pdf/1903.02665v2>. Sie verwenden einen Datensatz von etwa 100.000 republikanischen Münzen aus dem Münzhandel. Dieser Ansatz ist für die vorliegende Arbeit nicht übertragbar.

Mangel an finanziellen Mitteln bisher nicht systematisch erfasst werden konnten. So profitieren von einer Münzerkennungs-KI zunächst einmal wissenschaftliche Institute und Denkmalämter, die vor einem erheblichen Münzbestand stehen, der noch nicht vollständig klassifiziert ist. Mittelfristig könnten so auch die unzähligen Münzen im Privatbesitz nach wissenschaftlichen Standards erfasst werden. Insgesamt führt dies zu einer verbesserten Bewertung der Quellen, wovon neben der Forschungsgemeinschaft auch der Münzhandel profitiert.

Sowohl Forschung als auch Münzhandel benötigen grundsätzlich eine exakte Klassifikation einer Münze, die zuverlässig nur von Spezialisten vorgenommen werden kann.<sup>16</sup> Bisherige KI-Systeme sind noch zu ungenau und wurden mit zu wenig aussagekräftigen Daten trainiert, als dass sie für den praktischen Einsatz geeignet wären. Es hat sich gezeigt, dass sie die Effizienz des Arbeitsablaufs bisher noch nicht maßgeblich unterstützen können.<sup>17</sup> Wahrscheinlich ist das erst dann der Fall, wenn die KI zwei zentrale Anforderungen erfüllt: Die erste betrifft die statistische Validität. Je höher die statistische Zuverlässigkeit der KI ist, desto gerechtfertigter ist es, ihre Ergebnisse zu nutzen. Die zweite – möglicherweise noch wichtigere – betrifft die Erklärbarkeit. Die KI muss klar definierte Arbeitsschritte übernehmen und ihre Ergebnisse müssen für den Menschen verständlich und überprüfbar sein.

Für die Numismatik bietet eine automatisierte Münzerfassung vielfältige Möglichkeiten, ihre Prioritäten anzupassen und sich wissenschaftsorganisatorisch neu aufzustellen. Die Numismatik ist in zweierlei Hinsicht spezialisiert. Zum einen, weil sie einen eigenen, eng abgesteckten Quellenbestand und eine eigene Methodik hat, zum anderen, weil ihre Vertreter häufiger an Museen als an Universitäten angestellt sind.<sup>18</sup> Bestanden in den 1990ern noch Lehrstühle in Düsseldorf und Saarbrücken, gibt es heute in Deutschland nur noch einen Lehrstuhl für antike Numismatik in Frankfurt.<sup>19</sup> Andere Epochen sind in der Universitätslandschaft nicht mehr repräsentiert. Numismatik findet hauptsächlich bei den diesbezüglichen musealen und universitären Sammlungen statt. Durch Automatisierung freiwerdende Zeit und Ressourcen können genutzt werden, um die Numismatik wieder stärker in die universitäre Lehre einzubinden. Außerdem kann die in der Numismatik im Vergleich zu anderen Teilbereichen der Archäologie und

---

<sup>16</sup> ANWAR, Hafeez/ZAMBANINI, Sebastian/KAMPEL, Martin: A Bag of Visual Words Approach for Symbols-Based Coarse-Grained Ancient Coin Classification, 2013, S. 1, online: <http://arxiv.org/pdf/1304.6192v1> [25.08.2025].

<sup>17</sup> MELNYKOV, Oleksandr/KAPELESHCHUK, Andrei: A CNN-Assisted Decision Support System for Ancient World Coin Classification, in: Computer systems and information technologies, 1 (2024), S. 99–107, hier S. 100.

<sup>18</sup> HOWGEGO, Christopher: Ancient history from coins (Approaching the ancient world), London, New York, 1995, S. XII.

<sup>19</sup> WEISSER, Numismatiker in Deutschland, S. 265f.

Geschichte heute schon erheblich größere Vernetzung mit der Informatik weiter ausgebaut und vermittelt werden.

Ein weiteres Potential liegt in der verbesserten Qualität der durch KI gewonnen numismatischen Daten, insbesondere ihre Interoperabilität.<sup>20</sup> Die Vorhersagen der vorliegenden KI sind im Gegensatz zu menschlicher Katalogisierung nicht einzelsprachlich gebunden und sind damit unmittelbar mit internationalen Datenbanksystemen vernetzbar. Wenn, wie in der vorliegenden Arbeit vorgenommen, Konzepte auf Münzen erkannt werden, dann werden dadurch Informationen aus Münzbeschreibungen zugänglich gemacht, die bisher aufgrund der hohen sprachlichen Varianz in den Beschreibungen nicht auswertbar waren.<sup>21</sup> Auf diese Weise werden gänzlich neue Forschungsdaten generiert, die ein tieferes Verständnis von Münzen beinhalten, welches für künftige Forschung genutzt werden kann.

Alle Möglichkeiten, die sich durch die Nutzung einer umfassenden KI-Münzklassifikation bieten, sind heute noch nicht in Gänze absehbar. Die in den letzten Jahrzehnten im Münzhandel gehandelten Münzen sind fotografiert und digitalisiert. Wenn diese Münzen gespeichert und von der KI wiedererkannt werden, kann der illegale Handel mit diesen Stücken erkannt und unterbunden werden. Auch Diebesgut kann auf diese Weise wieder auf den ursprünglichen Besitzer rückgeführt werden.<sup>22</sup> Diese Ideen, welche die Ergebnisse der Münzforschung einem gesellschaftlichen Nutzen zuführen, wurden bisher aufgrund mangelnder Daten nicht weiterverfolgt.

Die vorliegende Arbeit ist in drei Hauptkapitel gegliedert, die von der theoretischen Grundlage bis hin zur praktischen Umsetzung und Auswertung der KI führen. Zunächst wird ausführlich der Forschungsstand in den relevanten Bereichen Natural Language Processing (NLP) und KI-Bilderkennung dargestellt, um die Methodenauswahl zu rechtfertigen. Das zweite Kapitel widmet sich den Münzdaten als Quellen und beschreibt, wie die Numismatik die Bestandteile antiker Münzen klassifiziert. Außerdem wird diskutiert, welche dieser Münzdaten für die vorliegende Arbeit genutzt werden können und wie der Download und die Vereinheitlichung dieser

---

<sup>20</sup> Interoperabilität ist eine der Elemente der FAIR-Kriterien, die als Anforderung an wissenschaftliche Daten formuliert wurden. Vgl. (<https://www.nfdi.de/>).

<sup>21</sup> So ergibt die Suche nach dem Ort Lyon 2970 Treffer im Portal IKMK, wird nach der lateinischen Schreibweise Lugdunum gesucht, werden nur 2412 Treffer erzielt, obwohl der Normdateneintrag zum Ort beide Schreibweisen enthält und damit über beide Begriffe auffindbar sein müsste.

<sup>22</sup> ARANDJELOVIĆ, Ognjen: Automatic Attribution of Ancient Roman Imperial Coins, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2010), S. 1728–1734, hier S. 1731. Die von Arandjelović verwendeten algorithmischen Methoden können Münzen voneinander unterscheiden, auch wenn sie aus einem anderen Winkel oder mit einer anderen Belichtung fotografiert sind.



Daten erfolgen kann. Darauf aufbauend behandelt das dritte Kapitel die Überführung dieser Forschungsdaten in eine zu entwickelnde automatisierte Münzklassifikation. Hier werden die Münzbeschreibungen mittels einer eigens entwickelten Named Entity Recognition auf Basis eines Knowledge Graphs maschinenlesbar gemacht und auf ihre Qualität überprüft. Außerdem werden die Münzbilder für das Training vorbereitet. Im Folgenden wird die Funktionsweise der KI und ihr Trainingsprozess samt seiner quantitativen Gütekriterien beschrieben. Den Abschluss bildet das vierte Kapitel, das die erzielten Ergebnisse kritisch reflektiert, die Grenzen der verwendeten Daten sowohl im Hinblick auf ihre numismatische Klassifikation als auch auf die technische Datenaufbereitung diskutiert und schließlich Perspektiven für zukünftige Forschung aufzeigt.

## 1.2. Forschungsstand

### 1.2.1 Natural Language Processing

Das Natural Language Processing (NLP) umfasst eine Vielzahl an Vorgängen und Methoden, die auf die automatisierte Verarbeitung natürlicher Sprache abzielen. Im Zusammenhang mit der Erfassung bildlicher Darstellungen auf Münzen sind hierbei Methoden des Teilbereichs der Named Entity Recognition (NER) relevant. Dabei geht es darum, in den Münzbeschreibungen die richtigen, das heißt die tatsächlich auf der Münze abgebildeten Darstellungen (Named Entities) zu erkennen, auch wenn sie in unterschiedlichen sprachlichen Varianten auftreten oder sogar nur implizit erwähnt werden.

Bisherige Vorarbeiten, die die Möglichkeiten der NER auf Münzen testen, können diese Anforderung nur teilweise erfüllen. Die einfachste Methode ist die Keyword Search (Stichwortsuche). Hier werden Münzbeschreibungen mit einem vorab erstellten Katalog bekannter Darstellungen und alternativer Schreibweisen abgeglichen. Anwar et al. extrahieren auf diese Weise drei Darstellungen aus ihrem Datensatz von 550 Münztypen (Münzserien).<sup>23</sup> Cooper und Arandjelović generieren fünf Darstellungen aus einem mehrsprachigen Datensatz.<sup>24</sup> Mit drei respektive fünf extrahierten Konzepten erreichen beide Vorarbeiten nur einen Bruchteil aller tatsächlich vorhandenen Darstellungen auf Münzen. Außerdem werden flektierte Wortformen und Rechtschreibfehler nicht als solche erkannt, wodurch nicht korrekt extrahiert wird und die

---

<sup>23</sup> ANWAR, Hafeez/ZAMBANINI, Sebastian/KAMPEL, Martin: A Bag of Visual Words Approach for Symbols-Based Coarse-Grained Ancient Coin Classification, 2013, S. 2, online: <http://arxiv.org/pdf/1304.6192v1> [25.08.2025].

<sup>24</sup> COOPER, Jessica/ARANDJELOVIĆ, Ognjen: Understanding Ancient Coin Images, 2019, S. 4f., online: <http://arxiv.org/pdf/1903.02665v2> [25.08.2025].

gewonnenen Daten unvollständig sind. Die für die Stichwortsuche genutzten Thesauri müssen erweitert werden, um eine vollständige NER von Münzdaten zu ermöglichen.

Die vorliegende Arbeit nutzt einen sogenannten Knowledge Graph, um den angesprochenen Schwächen der Thesauri zu begegnen. Dieser ist ein etabliertes Mittel, um komplexe NER-Fragestellungen zu beantworten.<sup>25</sup> Beim Knowledge Graph handelt es sich um ein Netz von Konzepten, die über sinnhaltige Verbindungen miteinander verbunden sind. NER auf Basis von Knowledge Graphs ist deterministisch und eignet sich bei hinreichender Qualität des Graphs besser als alternative Methoden. Eine solche ist die von Deligio und Gencer entwickelte NER-KI, die mit deutschen und englischen Beschreibungen provincialrömischer Münzen aus dem Portal Corpus Nummorum (CN) als numismatische Datenbasis trainiert wurde.<sup>26</sup> Das Modell stößt allerdings bei Daten aus mehreren Datenbanken aus dem gesamten Spektrum der römischen Numismatik an seine Grenzen, da es deren Darstellungen teilweise nicht kennt.<sup>27</sup>

In der vorliegenden Arbeit entsprechen die im Knowledge Graph vernetzten Konzepte den Darstellungen auf Münzen, etwa *Schwert* und *Speer*. Mithilfe der semantisch sinnvollen Relationen zwischen den Konzepten – *Schwert* und *Speer* teilen beispielsweise das Oberkonzept *Waffe* – können nur implizit genannte oder auf ungewöhnlichere Weise beschriebene Konzepte in Texten erkannt werden. Für die vorliegende Arbeit werden diese Verbindungen mithilfe der Fachliteratur aus den Münzbeschreibungen selbst erstellt. Pavlek und Winters extrahieren auf diese Weise eine nicht genau definierte Anzahl an Elementen aus einem Datensatz von etwa 6700 griechischen Münzen. Sie verweisen jedoch bereits in diesem vergleichsweise überschaubaren Datensatz auf die Schwierigkeit, für alle Münzen korrekte Relationen zwischen den Konzepten aus den Beschreibungen zu generieren.<sup>28</sup> Für die hier untersuchte, deutlich größere Datenbasis

---

<sup>25</sup> DIETZ, Laura/KOTOV, Alexander/MEIJ, Edgar: Utilizing Knowledge Graphs for Text-Centric Information Retrieval, in: The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, hrsg. v. Kevyn COLLINS-THOMPSON u.a. 2018, S. 1387–1390, hier S. 1387.

<sup>26</sup> Genau genommen handelt es sich um zwei verschiedene Modelle, einem deutschen und einem englischen, die nicht zu einer einheitlichen sprachungebundenen Systematik zusammengeführt werden. DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

<sup>27</sup> DIES.: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, S. 114f., online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025]. Bei der Anwendung des englischsprachigen Modells auf einen Datensatz aus dem OCRE-Portal, werden nur Genauigkeiten von knapp unter 70% erzielt.

<sup>28</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 205.

kann nicht jeder Eintrag manuell überprüft werden, sodass ihre Vorgehensweise angepasst werden muss.

Die Schwierigkeit, einen Knowledge Graph zu erstellen, ergibt sich daraus, dass Münzbeschreibungen aus numismatischer Fachsprache bestehen, die nur durch Fachwissen verstanden werden kann.<sup>29</sup> Sie unterliegt fachmethodisch historisch bedingt nicht den strengen Kriterien der Univozität, sondern verwendet Begriffe aus dem Griechischen, Lateinischen und der jeweiligen Landessprache wie Synonyme. So bezeichnen zum Beispiel die Begriffe *Signum* und *Standarte* den gleichen bildlichen Gegenstand und müssen im Knowledge Graph als Synonyme verbunden werden. Hinzu kommt, dass die Numismatik auch innerhalb der Archäologie eine besondere Bildsprache hat. Damit diese zuverlässig von der KI erkannt werden können, müssen im Knowledge Graph alle visuell identischen Darstellungen in einem einzigen Konzept gespeichert werden, auch wenn es sich auf der Deutungsebene um unterschiedliche Darstellungen handelt. Ein Beispiel ist die Nymphe Amalthea, die teilweise als Frau, teilweise als Ziege dargestellt wurde. Auf römischen Münzen tritt sie nur in Gestalt einer Ziege auf.<sup>30</sup> In den Beschreibungen römischer Münzen findet sich aber oft nur *Amalthea*.<sup>31</sup> Eine NER für römische Münzen muss demnach bei dem Wort *Amalthea* die Darstellung einer Ziege erkennen und dieser Zusammenhang muss im Knowledge Graph gespeichert werden. Das zeigt, dass es für das Anliegen der vorliegenden Arbeit essenziell ist, zunächst mittels numismatischen Fachwissens und auf Grundlage von Münzbeschreibungen eine geeignete umfassende Wissensbasis zu schaffen, die alle Darstellungen auf Münzen zusammen mit ihren unterschiedlichen Beschreibungsvarianten umfasst. Für diesen Schritt gibt es noch keine überzeugenden Automatisierungsmöglichkeiten.<sup>32</sup> Eine hermeneutische, händische Vorarbeit ist notwendig.

In der Vorarbeit können allerdings bereits bestehende Wissenssammlungen genutzt werden. Dazu gehören in erster Linie strukturierte Datensammlungen eindeutiger und einheitlicher Konzepte, sogenannte Normdatensysteme. Sie können unmittelbar genutzt werden, da sie wie ein Knowledge Graph aufgebaut sind.<sup>33</sup> Besonders nützlich sind hierbei die bestehenden

---

<sup>29</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 205.

<sup>30</sup> STOLL, Heinrich Wilhelm: Amalthea, in: Ausführliches Lexikon der Griechischen und Römischen Mythologie, hrsg. v. Wilhelm Heinrich ROSCHER, Leipzig 1884, S. 262–266, hier S. 262f.

<sup>31</sup> Vgl. RIC II, 32, Hadrian 2829; RIC V Gallienus 393/394.

<sup>32</sup> Denkbar ist eine automatisierte Extraktion der numismatischen Zusammenhänge aus der Fachliteratur mittels Large Language Models. Hierzu gibt es allerdings noch keine Vorarbeiten.

<sup>33</sup> DIETZ u.a., Utilizing Knowledge Graphs for Text-Centric Information Retrieval, S. 1388.

Normdatensysteme der Münzdatenbanken.<sup>34</sup> In den Münzdaten werden diese Konzepte, häufig Personen und Orte, in bestimmten Datenbankfeldern anstelle ihrer umfangreicheren und wenig standardisierten textuellen Entsprechung gespeichert. So wird der Kaiser Augustus beispielsweise mit dem Permalink „<http://nomisma.org/id/augustus>“ angegeben.<sup>35</sup> Mit einem solchen Link sind die Daten zwischen mehrsprachigen Datenbanken vergleichbar und erlauben eine schnelle Datenabfrage.<sup>36</sup> Numismatische Normdatensysteme sind noch nicht vollständig interoperabel und teilweise fehlerhaft. Dennoch können sie als Ausgangspunkt für die Erstellung einer Wissensbasis zu Personen und Orten in Beschreibungen dienen. Normdaten enthalten häufig auch alternative Schreibweisen. Im Portal Wikidata beinhaltet der Normdateneintrag zu *Augustus* zum Beispiel als Alternativname *Gaius Octavius*.<sup>37</sup> Es ist zu erwarten, dass diese Alternativen auch im vorliegenden Datensatz auftauchen, sodass sie für den Knowledge Graph übernommen werden können. Die Nutzung der Normdaten im Knowledge Graph hat den weiteren Vorteil, dass die Vorhersagen der in der vorliegenden Arbeit entwickelten KI direkt von den bestehenden Datenbanken interpretiert werden können. Dadurch ergeben sich Möglichkeiten einer Qualitätsüberprüfung. Die aus den Beschreibungstexten mittels NER extrahierten Konzepte können mit den bereits mit dem Münzdatensatz verknüpften Normdaten verglichen werden.

Zu Personen und Orten liefern die Normdatensysteme der Numismatik bereits einige Informationen für den Knowledge Graph, allerdings enthalten sie keine Informationen zu Tieren und Objekten auf Münzen.<sup>38</sup> Allgemeine Informationen kann der Thesaurus des Deutschen Archäologischen Instituts liefern.<sup>39</sup> Diese müssen allerdings im numismatischen Kontext neu geprüft werden. Insgesamt liefern die Normdatensysteme nur Informationen zu Darstellungen auf Münzen und ihren alternativen Schreibweisen. Die Konzepte werden nicht in eine Hierarchie

---

<sup>34</sup> Jede der Datenbanken hat eigene Normdatenportale, die miteinander vernetzt sind. Durch den höchsten Grad der Vernetzung zeichnet sich das internationale Nomisma-Projekt aus. Es sammelt alle Personen und Orte der antiken Numismatik und verknüpft sie mit den jeweiligen Normdatensystemen der großen Münzdatenbanken. Wenn vorhanden, werden auch allgemeine Normdatensysteme wie das der Deutschen Nationalbibliothek (GND) oder das internationale Virtual International Authority File (VIAF) mit Nomisma verknüpft. Diese verknüpften Daten haben die gleiche Form wie Knowledge Graphs und können deshalb al

<sup>35</sup> Vgl. zum Beispiel RIC I<sup>2</sup> Augustus 1A, online: [https://numismatics.org/ocre/id/ric.1\(2\).aug.1A](https://numismatics.org/ocre/id/ric.1(2).aug.1A).

<sup>36</sup> Theoretisch könnte auch hier eine Begriffssuche verwendet werden, dieser Ansatz ist bei Linked Open Data allerdings schwierig, da eine Volltextsuche zu lange dauern würde. TOLLE, Karsten u.a.: Semantic Search Based on Natural Language Processing. A Numismatic Example, in: *Journal Of Ancient History And Archaeology* 5, 3 (2018), S. 68–79, hier S. 68.

<sup>37</sup> Vgl. <https://www.wikidata.org/wiki/Q1405>.

<sup>38</sup> Nur das Portal Corpus Nummorum stellt zumindest ein Glossar aller Objekte und Tiere zur Verfügung, allerdings nicht in einem maschinenlesbaren Format. Vgl. <https://www.corpus-nummorum.eu/pdf/Glossary.pdf>

<sup>39</sup> <https://vocabs.acdh.oeaw.ac.at/DAI/de>

eingearbeitet und sie werden nicht über semantisch sinnvolle Relationen miteinander verknüpft. Diese beiden Elemente sind, wie zu zeigen sein wird, notwendig, um eine NER mit ausreichender Genauigkeit zu ermöglichen.<sup>40</sup> Um diese Verbindungen in den Knowledge Graph einzupflegen, bedarf es genauerer numismatischer Systematiken.

Einen Ausgangspunkt dafür liefert das Portal DIANA, das einen Überblick über gleiche Darstellungen auf antiken Münzen über größere zeitliche und räumliche Distanzen hinweg gibt.<sup>41</sup> Das Portal Diana ist wohl aufgrund fehlender Mittel nicht mehr aufrufbar und die Daten sind größtenteils verloren.<sup>42</sup> Für die vorliegende Arbeit relevant ist allerdings seine Grundidee, die Normdaten in eine Hierarchie einzuordnen, nämlich in die Gruppen Personen, Objekte, Tiere und Pflanzen.<sup>43</sup> *Mars hält Speer* wird dadurch auf der Abstraktionsebene zu *Mann hält Waffe*. Für den Knowledge Graph ergibt diese Hierarchisierung die Möglichkeit, *Speer* über das Prädikat *hat Oberkonzept* mit dem Begriff *Waffe* zu verbinden. Diese Form der Hierarchisierung ist flexibel und es gilt für die vorliegende Arbeit, so viele ähnliche visuelle Darstellungen wie möglich zu einer Oberkategorie zusammenzufassen. Pavlek und Winters nutzen die Hierarchisierungsmöglichkeit auf drei verschiedene Weisen. Erstens werden seltene Motive durch häufigere Äquivalente ersetzt. Vorgeschlagen wird beispielsweise, den auf Münzen nur selten abgebildeten Apfel durch den häufigeren Begriff *Granatapfel* zu ersetzen. Zweitens werden ähnliche Darstellungen auf einem höheren Abstraktionsgrad zusammengefasst. So zum Beispiel im Fall von Stachelschnecken und Kammermuscheln, die durch den allgemeinen Begriff *Muscheln* ersetzt werden. Drittens werden Personen groben Kategorien zugeordnet (etwa König und Nymphe).<sup>44</sup> Auf dieser Basis wird ein Knowledge Graph erstellt, der mit möglichst wenigen Kategorien die Darstellungen auf Münzen möglichst genau beschreibt.

Die vorliegende Arbeit verwendet neben der Hierarchisierung eine zweite Funktionalität des Knowledge Graphs, die bisher noch nicht für Münzbeschreibungen genutzt wurde. Diese ist,

---

<sup>40</sup> Vgl. Kapitel 3.1 Münzbeschreibungen und Named Entity Recognition.

<sup>41</sup> CALTABIANO, Maria u.a.: DIANA. An Approach to Coin Iconography According to Time and Space Through Digital Maps, in: 2013 Digital Heritage International Congress (DigitalHeritage) (2013), S. 413–416, hier S. 413.

<sup>42</sup> Ein großer Nachteil bei der Erstellung von DIANA war, dass jeder Eintrag händisch angelegt werden musste. Vgl. Ebd., S. 415. Wahrscheinlich scheiterte das Projekt auch an dieser aufwendigen Katalogisierung.

<sup>43</sup> DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, S. 3, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025]. Deligio und Gencer übernehmen diese Einteilung. Andere Systematiken, die Personen in Gottheiten und Menschen aufteilen, dabei aber mythische Kreaturen wie Zentauren den Gottheiten zuordnen, sind in diesem Fall schwieriger nutzbar. Visuell ähneln die mythischen Kreaturen eher Tieren, von denen sie oft inspiriert sind und werden wohl auch von der KI leichter dieser Gruppe zugeordnet werden.

<sup>44</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 209.

wie sich zeigen wird, zentral, um die nur implizit genannten Darstellungen zu erkennen. Aus dem oben genannten Beispiel lässt sich nämlich auch *Mars* mit dem Prädikat *kommt zusammen vor* mit dem Konzept *Speer* verbinden. Das daraus ableitbare Wissen wird in Kapitel 3.1.2 „Umgang mit Mehrdeutigkeit“ diskutiert.

### 1.2.2 Bilderkennung durch neuronale Netze

Die automatisierte Erkennung von Münzen entwickelte sich als Forschungsfeld in den 1990er Jahren. Die Fragestellungen und eingesetzten Methoden wandelten sich dabei im Zuge des technischen Fortschritts in der Informatik. Kamen zunächst analytische Methoden zum Einsatz, dominieren seit Mitte der 2010er Jahre KI-Modelle, insbesondere Convolutional Neural Networks (CNNs). Analytische Methoden beruhen auf festgelegten, nachvollziehbaren Regeln zur Auswertung von Daten, während KI datengetriebene, lernbasierte Ansätze nutzt, um Muster zu erkennen. Mit zunehmendem Abstraktionsgrad der Modelle veränderten sich auch die Ziele der Forschung: Ging es zunächst um die Erkennung spezifischer oder stempelgleicher Münzen, richtete sich der Einsatz von CNNs auf die Erkennung von gemeinsamen Elementen auf unterschiedlichen Münzen und von Münztypen.

Analytische Methoden erreichen hohe Genauigkeiten bei der Unterscheidung einzelner Münzen. Moderne Münzen können damit sogar eindeutig klassifiziert werden.<sup>45</sup> Die zitierte und andere Arbeiten nutzten den SIFT-Algorithmus, der gleiche Elemente auf Bildern unabhängig von der Rotation, dem Zoom oder dem Kamerawinkel erkennt.<sup>46</sup> Allerdings geht SIFT davon aus, dass die Münze rund ist und damit ein definierbares Zentrum hat. Das ist bei antiken Münzen nicht der Fall. Diese Ansätze scheiterten auch bei antiken Münzen bereits grundsätzlich, da hier die Variation innerhalb eines Typs durch die manuelle Prägung mitunter sehr groß ist.<sup>47</sup> Je weniger visuelle Gemeinsamkeiten es gibt, desto schwieriger ist es für SIFT, diese Konzepte

---

<sup>45</sup> NÖLLE, Michael u.a.: *Dagobert - A New Coin Recognition and Sorting System*, 2003, S. 329. Die frühe Forschung entsteht dabei aus wirtschaftlichen Überlegungen im Kontext der Euro-Einführung. Als durch die Umstellung auf den Euro die alten Münzen der Landeswährungen eingetauscht wurden, brachten die Menschen oft auch Kleingeld aus Urlauben zum Umtauschen. Bestehende Systeme, die Münzen zählen konnten, konnten nicht alle umtauschbare Währungen erkennen. Es mussten also 12 Währungen mit etwa 100 verschiedenen Münzen aus etwa 1000 Serien klassifiziert werden. Nölle et al. entwickeln eine Sortiermaschine, die mit einer Genauigkeit von über 95% moderne Münzen auseinanderhält. Sie enthält allerdings auf eine technische Vorsortierung auf Basis von Gewicht und Größe, sodass das Modell nicht alle Münzbilder gleichzeitig vergleichen musste. Vgl. DIES.: *Dagobert - A New Coin Recognition and Sorting System*, 2003, S. 336

<sup>46</sup> ARANDJELOVIĆ, Automatic Attribution of Ancient Roman Imperial Coins, S. 1732.

<sup>47</sup> ANWAR, Hafeez u.a.: *Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention*, 2019, S. 5, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

zu erkennen. Das ursprüngliche Ziel, auch antike Münztypen zu erkennen, geriet durch diese Schwierigkeit zunächst in den Hintergrund.

Bis Mitte der 2010er Jahre fokussieren Studien deshalb auf die Unterscheidung einzelner Münzen voneinander. Daran anschließend wurden auch Studien zur Stempelgleichheit durchgeführt.<sup>48</sup> Bei diesen Arbeiten wurden häufig die Reverse der Münzen untersucht, da sie eine höhere Varianz abbilden.<sup>49</sup> Am umfassendsten wurde die Frage der Unterscheidung einzelner Münzen von Kampel und Huber-Mörk untersucht, die anhand von 240 Münzen desselben Münztyps mit jeweils 10 Bildern pro Münze aus unterschiedlichen Belichtungen und Winkeln die jeweiligen Münzen zu 95% unterscheiden konnten.<sup>50</sup> Es bleibt allerdings unklar, ob für andere ähnlichere Münztypen die Genauigkeit genauso hoch wäre, und es zeigt durch die hohen Datenanforderungen, wie unpraktikabel dieser Ansatz bei einem größeren Münzbestand ist. Um im Sinne einer Erkennung von allen Darstellungen auf Münzen praktisch nutzbar zu sein, müssten von allen römischen Münzen mehrere Bilder vorliegen.

Ab Mitte der 2010er Jahre rückten mit der Entwicklung von Bildklassifikations-KIs Fragen von Münztypen und auf Münzen dargestellten Personen wieder in den Fokus. Diese Studien verwenden meist die bis heute leistungsstarken CNNs. Diese zeichnen sich dadurch aus, dass sie dem eigentlichen neuronalen Netzwerk sogenannte Convolutional-Layer voranstellen. Diese Schichten extrahieren spezifische Muster aus dem Input-Bild und verarbeiten sie dabei effizient. Mit diesen Modellen war es möglich, bei hinreichend guten Daten eine Klassifikation zu ermöglichen. Andere Ansätze mit anderen Arten von Machine Learning Algorithmen wie bei Guo et al. haben sich nicht durchgesetzt.<sup>51</sup> Für Fragen, die wie bei der Lokalisierung einer Münze mit der Klassifikation verwandt sind, haben auch andere Modelle gute Ergebnisse gezeigt.<sup>52</sup> Die Zahl der Publikationen ab Mitte der 2010er-Jahre bis heute zeigt, dass das Interesse an einer Münzklassifikation anhaltend hoch ist und auch internationale Forschergruppen beschäftigt.

---

<sup>48</sup> KAVELAR, Albert/ZAMBANINI, Sebastian/KAMPEL, Martin: Reading the Legends of Roman Republican Coins, in: *Journal on Computing and Cultural Heritage* 7, 1 (2014), S. 1–20, hier S. 3.

<sup>49</sup> ANWAR, Hafeez/ZAMBANINI, Sebastian/KAMPEL, Martin: A Bag of Visual Words Approach for Symbols-Based Coarse-Grained Ancient Coin Classification, 2013, S. 1, online: <http://arxiv.org/pdf/1304.6192v1> [25.08.2025].

<sup>50</sup> KAMPEL, Martin/HUBER-MÖRK, Reinhold/ZAHARIEVA, Maia: Image-Based Retrieval and Identification of Ancient Coins, in: *IEEE Intelligent Systems* 24, 2 (2009), S. 26–34, hier S. 26f.

<sup>51</sup> GUO u.a., A Siamese Transformer Network for Zero-Shot Ancient Coin Classification.

<sup>52</sup> DELIGIO, Chrisowalandis/TOLLE, Karsten/WIGG-WOLF, David (Hrsg.): Supporting the Analysis of a Large Coin Hoard with AI-Based Methods, 2024.

**Abbildung 1: Forschungsstand zu CNNs in der antiken Numismatik**

Jahr	Autoren	Modell	Bilder (Avers/Revers)	Gegenstand	Klassen	Genauigkeit
2015	Kim, Pavlovic <sup>53</sup>	SuperVision	~4500 (Av+Rv)	Portrait/Münztyp	96/314	76%
2017	Schlag, Arandjelović <sup>54</sup>	Eigenes	78 778 (Av)	Portrait	83	82%
2018	Loyal <sup>55</sup>	VGG-16	26 096 (Av)	Portrait/Münztyp	69/195	91/66%
2019	Anwar et al. <sup>56</sup>	Eigenes	18 285 (Rv)	Reversmotiv	228	98%
2019	Cooper, Arandjelović <sup>57</sup>	Alexnet	unklar	Reversmotiv	5	72-84%
2021	Gampe <sup>58</sup>	VGG-16	~126 000 (Av)	Portrait	66	82%
2024	Gampe, Tolle <sup>59</sup>	VGG-16	~40 000 (Av+Rv)	Münzstätte	98	79%
2024	Melnykov, Kapeleshchuk <sup>60</sup>	Eigenes	369 (Av+Rv)	Münztyp	4	90-99%

Bei der Modellauswahl zeigen sich zwei verschiedene Ansätze: einerseits die Entwicklung eines eigenen kleinen CNNs und andererseits das sogenannte Finetuning von bestehenden Modellen wie VGG-16, VGG-19 und Alexnet. Dabei hat sich letzterer Ansatz als erfolgreicher erwiesen, da diese komplexer sind, das heißt mehr rechnerische Ebenen (engl. Layers) haben. Die bestehenden Modelle wurden auf großen Bilddatensätzen wie ImageNet<sup>61</sup> vortrainiert. Da CNNs in anfänglichen Layern abstrakte Konzepte wie Kanten und Farben lernen, können die mit dem allgemeinen ImageNet-Datensatz erlernten Features in der Münzerkennung weiterverwendet werden. Die Resultate der Studien zeigen, dass das Finetuning erfolgsversprechend und dabei ressourcenschonend ist.<sup>62</sup> Bei Münzdatensätzen zeigt Gampe, dass sich das komplexere

<sup>53</sup> KIM, Jongpil/PAVLOVIC, Vladimir: Discovering Characteristic Landmarks on Ancient Coins using Convolutional Networks, in: 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico (2016), S. 1595–1600.

<sup>54</sup> SCHLAG, Imanol/ARANDJELOVIĆ, Ognjen: Ancient Roman Coin Recognition in the Wild Using Deep Learning Based Recognition of Artistically Depicted Face Profiles, in: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW) (2017), S. 2898–2906.

<sup>55</sup> LOYAL, Antje: Maschinelles Lernen angewendet auf Bilder antiker Münzen, Frankfurt, 2018, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit\\_Antje\\_Loyal.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit_Antje_Loyal.pdf) [25.08.2025].

<sup>56</sup> ANWAR, Hafeez u.a.: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention, 2019, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

<sup>57</sup> COOPER, Jessica/ARANDJELOVIĆ, Ognjen: Understanding Ancient Coin Images, 2019, online: <http://arxiv.org/pdf/1903.02665v2> [25.08.2025].

<sup>58</sup> Theoretisch könnten die komplexen Modelle wie Alexnet und VGG-16/19 auch ohne vorheriges Training genutzt werden. Der Trainingsaufwand dafür ist allerdings prohibitiv. Vgl. GAMPE, Sebastian: Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen, Frankfurt, 2021, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf) [25.08.2025].

<sup>59</sup> GAMPE, Sebastian/TOLLE, Karsten (Hrsg.): Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, Frankfurt, 2024.

<sup>60</sup> MELNYKOV/KAPELESHCHUK, A CNN-Assisted Decision Support System for Ancient World Coin Classification.

<sup>61</sup> Vgl. <https://www.image-net.org>.

<sup>62</sup> GAMPE/TOLLE, Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, S. 34f.



VGG-16-Modell besser eignet als Alexnet.<sup>63</sup> Die vorliegende Arbeit nutzt deshalb das komplexeste der hier zitierten Modelle, nämlich VGG-19, das auf der gleichen Architektur wie VGG-16 aufbaut, allerdings mehr Layer hat.

Neben den unterschiedlichen Modellen, die verwendet werden, fällt in der Forschungsübersicht auf, wie groß die Varianz der angegebenen Genauigkeiten ist.<sup>64</sup> Diese erklärt sich durch die unterschiedlichen Datensätze, mit denen sie jeweils trainiert wurden. Sie sagt nichts über die Robustheit des Modells aus. Bisherige Modelle betrachten immer nur einen Teilbereich der römischen Numismatik, entweder provinzialrömische, republikanische, oder kaiserliche Prägungen. Jeder Teilbereich bringt eigene Herausforderungen mit sich und die Gütekriterien lassen sich deshalb kaum vergleichen oder verallgemeinern. Wenn nämlich die auf dem Ursprungsdatensatz trainierte KI auf anderen Datensätzen wesentlich schlechtere Ergebnisse erzielt, ist sie für eine allgemeine automatisierte Münzerkennung untauglich. Anwar et al. unterscheiden beispielsweise die Reversmotive republikanischer Münzen mit einer beeindruckenden Erfolgsquote von 98%. Dieser Datensatz ist jedoch zu einfach, da sich die gewählten Motive wie *Löwe* und *Greif* stark voneinander unterscheiden, und zudem jeweils in ausreichender Menge vorhanden sind. Auf einem anderen Datensatz, wo die Reversbilder nicht das zentrale Element sind, erzielt die KI von Anwar et al. schlechtere Genauigkeiten.<sup>65</sup> Außerdem liegt die hohe Genauigkeit nicht an dem gewählten Modell, da auch ein VGG-16 mit Finetuning eine ähnlich hohe Genauigkeit erzielt.<sup>66</sup>

Loyal, die mit einem ähnlich großen Datensatz auf dem Avers dargestellte Personen auf kaiserzeitlichen Münzen untersucht, und dabei sogar weniger Klassen als Anwar et al. bildet, erreicht eine deutlich reduzierte Genauigkeit. Ihr Datensatz umfasst dabei alle römischen Stücke im Bestand der American Numismatic Society (ANS) und ist somit realitätsnaher als der eigens erstellte Datensatz von Anwar et al.<sup>67</sup> Ein realitätsnaher Datensatz ist wichtig, da er die

---

<sup>63</sup> GAMPE, Sebastian: Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen, Frankfurt, 2021, S. 34, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf) [25.08.2025].

<sup>64</sup> Die Genauigkeit (engl. Accuracy) ist der Begriff für die richtigen positiven Antworten des Modells. Obwohl diese Metrik nicht ausreichend ein Modell beschreibt, wird es in der Literatur am häufigsten angegeben und wird deswegen hier für Vergleichszwecke verwendet. Für die Qualitätsmaße dieser Arbeit siehe Kapitel 3.3.1 „Training und Gütekriterien“.

<sup>65</sup> ANWAR, Hafeez u.a.: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention, 2019, S. 27, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

<sup>66</sup> DIES.: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention, 2019, S. 21, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

<sup>67</sup> LOYAL, Antje: Maschinelles Lernen angewendet auf Bilder antiker Münzen, Frankfurt, 2018, S. 38, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit\\_Antje\\_Loyal.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit_Antje_Loyal.pdf) [25.08.2025].

tatsächlichen Darstellungen auf Münzen widerspiegelt. Durch die Einbeziehung aller Münzen, auch stark beschädigter Exemplare, kann ein umfassenderes Bild der kaiserzeitlichen Prägungen erstellt werden. Zudem ermöglicht dies eine realistischere Bewertung der Genauigkeit von Klassifikationsverfahren. Um möglichst realitätsnah zu sein, wird der Datensatz der vorliegenden Arbeit nicht im Vorfeld überprüft. Das bedeutet, dass die Genauigkeit, die KI auf dem vorliegenden Datensatz erzielen kann, zwangsläufig niedriger ist als bei ‚gepflegten‘ Datensätzen.

Datensätze römischer Münzen, die keine manuelle Vorauswahl treffen, sehen sich allerdings mit einer neuen Herausforderung konfrontiert: der mangelnden Ausgeglichenheit der Daten. Manche Elemente auf Münzen sind in realitätsnahen Daten stark überrepräsentiert, während andere nur sehr selten auftauchen. Während des Trainings müssen diejenigen Klassen aussortiert werden, die zu selten im Datensatz vorkommen und deswegen nicht erkannt werden können. Gampe und Tolle wählen zum Beispiel bei ihrem CNN von den 122 Münzstätten, die in ihrem Datensatz auftauchen, nur diejenigen 98 aus, zu denen mehr als 20 Bilder vorliegen.<sup>68</sup> Das liegt an der Quellenlage der Münzen, die sich gerade dadurch auszeichnen, sowohl mit seltenen, als auch mit gewöhnlicheren Motiven versehen zu sein. Besonders betroffen von seltenen Darstellungen sind provinzialrömische Münzen, die in kleinerer Serie geprägt wurden und deshalb sogar noch regelmäßig neue Typen auftauchen.<sup>69</sup> Was für den Münzhandel förderlich ist, nämlich die vielen Münztypen mit sehr raren Exemplaren, erschwert die automatisierte Klassifikation. Auch bei republikanischen oder Reichsprägungen gibt es noch lange nicht für alle Klassen ein digitalisiertes Bild, geschweige denn mehrere pro Münztyp, so wie sie für das Training einer KI benötigt werden.<sup>70</sup> Beim Machine Learning werden häufigere Klassen besser und häufiger erkannt als seltenere Klassen.<sup>71</sup> Dabei kann das sogenannte Overfitting auftreten, was die fehlende Generalisierbarkeit des Modells beschreibt. Bei Anwendung auf einem größeren Datensatz erzielt das Modell dann schlechtere Ergebnisse, weil es zu spezifisch auf den nicht ausgeglichenen Datensatz trainiert wurde. Gampe und Tolle adressieren das Problem der unausgeglichenen Daten, indem sie mit unterschiedlichen Ansätzen die größeren Klassen in

---

<sup>68</sup> GAMPE/TOLLE, Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, S. 1.

<sup>69</sup> HEKSTER, Olivier/MANDERS, Erika: Coinage in the Roman Provinces: the RPC and CHRE Projects, in: *Journal of Roman Studies* 112 (2022), S. 199–213, hier S. 201.

<sup>70</sup> GUO u.a., A Siamese Transformer Network for Zero-Shot Ancient Coin Classification, S. 2.

<sup>71</sup> Das liegt daran, dass die Fehlklassifikation seltener Klassen die im Modelltraining zentralen, globalen Qualitätsmaße nur wenig beeinflusst. Vgl. HERRERA, Francisco u.a.: *Multilabel Classification*, Cham, 2016, S. 133.

kleinere aufteilen. Allerdings können sie Overfitting nicht verhindern.<sup>72</sup> Dieser Ansatz der Aufteilung ist bei größeren Datensätzen nicht praktikabel, allein, weil dadurch zu viele Klassen generiert würden, um sinnvoll ausgewertet werden zu können. Das gilt insbesondere für den vorliegenden Datensatz, der mit über 500.000 Münzen eine wesentlich größere ikonographische Varianz als der Datensatz von Gampe und Tolle abbildet.

Um dennoch Overfitting zu reduzieren, wird in der vorliegenden Arbeit eine bisher in der Münzklassifikation nicht verwendete Art von CNNs verwendet. Grundlage ist das Ergebnis Gampes und Tolles, dass die automatisierte Erkennung von Münzstätten besser funktioniert, wenn die Unterklassen anhand der Darstellung auf einzelnen Münzen gebildet werden.<sup>73</sup> Die KI lernt so nicht nur, auf Basis diffuser Kriterien Münzstätten von Münzen zu erkennen, sondern aufgrund eindeutiger Motive. Diese Grundfunktion ist sogenannten Multilabel-Modellen inhärent. Im Gegensatz zu der bisherigen Forschung, die sogenannte Multiclass-Modelle nutzt, zeichnet sich eine Münze nicht durch ein einziges Label aus, sondern durch deren Kombination (Abb. 2).

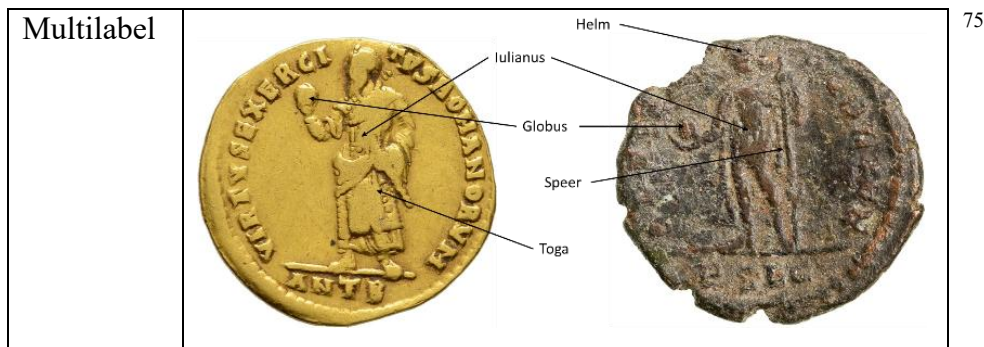
Abbildung 2: Multiclass und Multilabel Modelle



<sup>72</sup> GAMPE/TOLLE, Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, S. 1.

<sup>73</sup> Ebd., S. 3f. In ihrer Münzstättenenerkennung ist die Münzstätte Pergamon am häufigsten vertreten, sie unterteilen die in Pergamon geprägten Münzen in 16 Untertypen, die sich in den aus den Beschreibungen extrahierten Konzepten unterscheiden.

<sup>74</sup> Eigene Darstellung, Bildquellen vgl. links: Revers eines Solidus des Iulianus, RIC VII Antiochia 205, 361-363 n. Chr., Numismatische Sammlung der Universität Heidelberg, Inv.-Nr.: N83006, online: <https://pecunia.zaw.uni-heidelberg.de/ikmk/object?lang=en&id=ID1700>; rechts: Revers eines Aes 3 des Iulianus, RIC VII Lugdunum 221, 361-363 n. Chr., American Numismatic Society, Inv.-Nr.: 1970.165.91, online: <https://numismatics.org/collection/1970.165.91>.



Beim Multiclass-Modell werden demnach visuell ähnliche Darstellungen, wie hier *Kaiser Iulianus* in zwei gänzlich verschiedene Klassen aufgeteilt. Das liegt am letzten Layer des Modells, der sogenannten Softmax-Funktion, die aus allen möglichen Klassen die wahrscheinlichste ausgibt. Multiclass-Modelle sind darauf ausgelegt, einem Bild eine einzelne der vorgegebenen Klassen zuzuordnen, und können dementsprechend nicht damit umgehen, wenn mehrere oder keine Klassen auf einem Bild abgebildet sind. Das gelingt mit Multilabel-Modellen, die jede Klasse unabhängig voneinander betrachten. Beim Labeln von Bildern werden deshalb klassischerweise Multilabel-Modelle gewählt.<sup>76</sup> Sie erlauben die Vorhersage von beliebig vielen Klassen, die binär – also entweder vorhanden oder nicht vorhanden – klassifiziert werden. Um VGG-19 oder andere vortrainierte CNNs in Multilabel-Modelle umzuwandeln, muss nur die Softmax-Funktion durch eine sogenannte Sigmoid-Funktion ausgetauscht werden, die alle Klassen separat betrachtet.<sup>77</sup> Dadurch haben die beiden abgebildeten Münzen von Iulianus die gemeinsame Darstellung *Iulianus* und *Globus* und sich unterscheidende Elemente *Toga*, *Helm*, *Speer* und *Szepter*. Diese Klassen haben eine genaue visuelle Entsprechung auf der Münze und können somit leichter erkannt werden. Außerdem lernt das Modell mit dem Multilabel-Ansatz explizit Interdependenzen zwischen den Klassen.<sup>78</sup> So trägt Iulianus in Toga nie einen Speer und andersherum ist Iulianus mit Speer nie in Toga. Wenn das Modell dementsprechend einen Speer erkennt und die Münze so stark abgegriffen ist, dass eine Toga nicht zweifelsfrei festgestellt werden kann, kann es trotzdem vorhersagen, dass sie nicht abgebildet ist.

Overfitting tritt zwar auch bei Multilabel-Modellen auf, aber es beeinflusst die finale Klassifikation nicht im gleichen Maße wie bei Multiclass-Modellen. Multiclass-Algorithmen werden normalerweise nicht mit Münzen trainiert, die zu keiner Klasse gehören. Sie lernen nur

<sup>75</sup> Eigene Darstellung, Bildquellen vgl. Ebd.

<sup>76</sup> HERRERA u.a., Multilabel Classification, S. 14.

<sup>77</sup> Vgl. Kapitel 3.3 „Das Convolutional Neural Network“.

<sup>78</sup> Ebd., S. 26.

zwischen den Klassen, die dem Modell zur Verfügung gestellt werden, zu unterscheiden. Wenn also das Modell als Input eine unbekannte Münze bekommt, wird sie zwangsläufig einen falschen Output liefern. Mit Multilabel-Modellen verringert sich dieses Problem. Seltene Münzen teilen manche Darstellungen mit anderen Münzen. Zum Beispiel teilen Kaiser, die durch eine zu geringe Zahl ihnen zugeordneter Münzen nicht der KI zum Training gegeben werden können, nun das Label *Mann* und *Kaiser* mit anderen Kaiserdarstellungen. So kann eine – wenn auch verringerte – Aussagekraft der KI-Prognose auch bei seltenen Stücken erhalten bleiben. Unbekannte Münze mit gänzlich neuen Darstellungen werden so auch als unbekannt klassifiziert, was von zentralem Interesse ist.<sup>79</sup> Diese Frage kann in der Interpretation der Ergebnisse der KI wiederaufgegriffen werden und es kann überprüft werden, wie sich die KI verhält, wenn sie unerwarteten Input in Form beispielsweise einer modernen Münze erhält.

Um für die numismatische Forschung nutzbar zu sein, muss eine KI neben statistisch signifikanten Ergebnissen auch verständliche und erklärbare Ergebnisse liefern. Die Entscheidungen neuronaler Netze sind allerdings grundsätzlich schwer nachvollziehbar, weswegen sie auch oft als „Black Boxes“ bezeichnet werden. Wenn nicht klar nachvollzogen werden kann, nach welchen Kriterien das Modell entscheidet, kann der Mensch auch nicht gezielt eingreifen, um gegebenenfalls offensichtliche Fehler zu beheben. Das sorgt für weniger Vertrauen in den Automatisierungsansatz.<sup>80</sup> Um die Vertrauenswürdigkeit von KI-Modellen generell zu erhöhen, schlägt Fleisher zwei notwendige Kriterien vor. Erstens müssen die Ergebnisse interpretierbar sein.<sup>81</sup> Das bedeutet in Bezug auf die Münzerkennung, dass die KI ihren Input und Output offenlegen muss und dass der Output in einem vom Menschen lesbaren Format vorliegen muss. Zweitens soll die KI so angepasst werden, dass sie erklären kann, warum eine Entscheidung zustande kam.<sup>82</sup> Das funktioniert bei CNNs vergleichsweise gut. Hier kann eine Saliency Map genutzt werden.<sup>83</sup> Mit ihr können die Bildregionen des Input-Bilds visuell hervorgehoben werden, die einen großen Einfluss auf die Entscheidung des Modells hatten.

---

<sup>79</sup> ANWAR, Hafeez u.a.: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention, 2019, S. 1, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

<sup>80</sup> FLEISHER, Will: Understanding, Idealization, and Explainable AI, in: Episteme 19, 4 (2022), S. 534–560, hier S. 534f.

<sup>81</sup> FLEISHER, Will: Understanding, Idealization, and Explainable AI, in: Episteme 19, 4 (2022), S. 534–560, hier S. 546.

<sup>82</sup> Ebd., S. 540.

<sup>83</sup> GAMPE, Sebastian: Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen, Frankfurt, 2021, S. 11, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf) [25.08.2025] Die vorliegende Arbeit nutzt für die Implementierung den gleichen Algorithmus wie Gampe, das Gradient-weighted Class Activation Mapping (Grad-CAM).

Abbildung 3: Output des Modells



84

Mit dieser Kombination von einem Multilabel-CNN und der Nutzung der Saliency Map sind die Ergebnisse interpretierbar. Die für den Menschen sofort verständlichen Label werden eindeutig einer Bildregion zugeordnet. Dies zeigt beispielhaft der Output für die bestimmte Kategorie *Stier* bei der vorliegenden Beispielmünze (Abb. 3). Hier zeigt sich, dass der Altar, der in dieser Opferszene vor Kaiser Caligula abgebildet ist, auch auf Basis der entsprechenden Bildregion (hier in gelb und rot farblich hervorgehoben) von dem KI-Modell erkannt wird. Trotzdem kann die Saliency Map nur ein erster Schritt in Richtung erklärbarer Münz-KI sein, da sie keine Aussage darüber trifft, welche Features genau für das Resultat verantwortlich waren.<sup>85</sup> Wurde beispielsweise der Kopf des Stieres aufgrund seiner Formen erkannt, oder wurde allgemein von der Opferszene auf den Opferstier geschlossen.

## 2. Von Münzen zu Forschungsdaten

Um eine vollständig automatisierte Klassifikation von Münzen voranzutreiben, muss verstanden werden, welche ihrer Elemente sich eignen und notwendig sind. Im folgenden Kapitel wird deshalb dargestellt, welche Kategorien die Numismatik für eine Klassifikation standardisiert hat. Darüber hinaus wird zu prüfen sein, ob und gegebenenfalls wie sich diese Kategorien durch die Digitalisierung der Bestände verändert haben. Im Anschluss wird erklärt, warum in der vorliegenden Arbeit von diesen Kategorien die bildlichen Darstellungen ausgewählt wurden. Danach werden die für die weiteren Schritte notwendigen Daten heruntergeladen.

<sup>84</sup> Sesterz des Kaisers Caligula, RIC I<sup>2</sup> Gaius/Caligula 44, 39–40 n. Chr., American Numismatic Society, 1967.153.113, online: <https://numismatics.org/collection/1967.153.113>.

<sup>85</sup> FLEISHER, Understanding, Idealization, and Explainable AI, S. 544.

## 2.1. Bestandteile antiker Münzen

Münzen sind Teil des kulturellen Erbes ihrer Gesellschaften und Ausdruck ihrer Regierungsform. Die Numismatik untersucht dabei ihre materielle, bildliche und textuelle Dimension.<sup>86</sup> Alle drei Ebenen stehen in einem engen Zusammenhang mit der sozialen Funktion von Münzen als Zahlungsmittel, Herrschaftsinsignien und Kommunikationsmedien. Durch die Betrachtung all dieser Ebenen können römische Münzen systematisch typologisiert werden.

Die materielle Dimension von Münzen erschließt sich zunächst über das aufgewendete Material und dessen gesellschaftlich zugeschriebenen Wert.<sup>87</sup> In der römischen Zeit definiert sich dieser Wert über die Menge an Edelmetall. Es gab festgelegte Nominale (Nennwerte), die allerdings inflationsbedingt nicht immer der gleichen Edelmetallmenge entsprachen.<sup>88</sup> Für numismatisch-historische Fragestellungen liefert das Material häufig den Ausgangspunkt, um den Kreis vor-maliger Besitzer einzuschränken. Goldmünzen galten bereits zeitgenössisch als aristokratisch und zirkulierten nur in elitären Kreisen.<sup>89</sup> Aufgrund ihres Informationsgehalts sind Nominal und Material seit jeher die zentralen Kategorien für die Erfassung von Münzen und auch für die Interpretation der bildlichen und textuellen Ebene von Münzen zentral. Goldmünzen sind wesentlich kleiner als andere Münzen und folglich fehlen manche Motivkomponenten. Buntmetallmünzen sind häufiger korrodiert und verfärbt, was deren Klassifikation erschwert.

Aber auch die fotografische Dokumentation der Münzen wird durch ihre Materialität herausfordernd. Denn Farbe und Albedo (Rückstrahlvermögen) der Münzen hängen von ihrer Metallzusammensetzung und ihrer Aussetzung mit Umgebungsluft und Erde ab.<sup>90</sup> Oft werden im Erfassungsprozess Kamera- und Blitzlicheinstellung nicht für jede Münze neu angepasst. Goldmünzen wirken so sehr überbelichtet, während Bronzemünzen sehr dunkel erscheinen. Da die Münzenoberfläche ein Relief ist, verändert sich das Münzbild je nach Position der Lichtquelle signifikant. Dennoch sind Münzbilder im Vergleich mit den folgenden Kategorien in der numismatischen Erfassung weitgehend standardisiert. Bei römischen Münzen gibt es ein gesichertes Oben und Unten und die Münzen werden entsprechend in der gleichen Ausrichtung fotografiert. Für die vorliegende Arbeit bedeutet das, dass in einem technischen

---

<sup>86</sup> KEMMERS, Fleur/MYRBERG, Nanouschka: Rethinking Numismatics. The Archaeology of Coins, in: *Archaeological dialogues* 18 (2011), S. 87–108, hier S. 88f.

<sup>87</sup> Ebd., S. 93.

<sup>88</sup> In der Römischen Republik wurden die Nominale noch durch Inschriften auf den Münzen angezeigt, wie ein „X“ für den Denar. In der Kaiserzeit kann nur noch von den Maßen auf das Nominal geschlossen werden.

<sup>89</sup> Ebd., S. 95.

<sup>90</sup> ARANDJELOVIĆ, Automatic Attribution of Ancient Roman Imperial Coins, S. 1732.

Vorverarbeitungsschritt die Münzbilder lediglich hinsichtlich Belichtung und Kontrast vereinheitlicht werden müssen.<sup>91</sup>

Bei der Klassifikation von Münzen nehmen die direkt mit der Materialität verbundenen Merkmale Material, Nominal und Maße eine zentrale Rolle ein. Bei diesen Daten besteht allerdings nur ein geringer Bedarf an Automatisierung. Die Abmessungen der Münze und ihr Gewicht können durch für die Numismatik entwickelte Fotoapparate bereits automatisiert erfasst werden. Der manuelle Bestimmungsaufwand für das Materialfeld ist gering, da es nur drei Ausprägungen haben kann: Gold, Silber und Buntmetall. Eine automatisierte Materialerkennung auf Basis der Farbe der Münze ist allerdings denkbar und wohl Gegenstand künftiger Forschung. Das Nominal der Münzen kann wiederum aus der Kombination von ihrem Material und ihren Maßen abgeleitet werden.<sup>92</sup> Aus der Materialität ergeben sich weitere Forschungsfragen zu sogenannten Stempelgleichheiten, die jedoch aufgrund unzureichender Daten erst langfristig Gegenstand der Automatisierung sein können.<sup>93</sup>

Bei der bildlichen Dimension von römischen Münzen handelt es sich um auf ihnen dargestellte Personen, Tiere, mythische Kreaturen, Pflanzen oder sonstige Gegenstände. Das Bild des Averses trägt meistens eine Büste oder einen Kopf, aus dem sich der jeweilige Machthaber ableiten lässt. Während der Republik gibt es eine Autorität erst in der Zeit der Bürgerkriege, als sich etwa Marcus Antonius oder Julius Caesar auf Münzen abbilden ließen. Davor finden sich auf dem Avers oft mit Rom verbundene Gottheiten. In der Kaiserzeit handelt es sich bei den Dargestellten und der Autorität nicht immer um dieselbe Person. Manche Kaiser prägen auch nahe Familienangehörige auf dem Avers. Die systematische Inszenierung von Herrschaft galt als ein

---

<sup>91</sup> HOLT, Frank L.: When Money Talks, New York, 2021, S. 95.

<sup>92</sup> Eine Ausnahme bilden Dupondien, die durch die Strahlenkrone auf dem Kaiserportrait des Avers signalisiert werden.

<sup>93</sup> In sogenannten Stempelstudien können Münzen noch unterhalb der Kategorie Münztyp untersucht werden. Ausgangspunkt dieser Studien ist der Produktionsprozess der Münzen: Ein Münzrohling wurde auf einen Stempel gelegt, darüber der Gegenstempel platziert und mit einem Hammerschlag die jeweiligen Bilder eingeprägt. Auf diese Weise war prägen wesentlich effizienter als gießen. Vgl. HOWGEGO, Ancient history from coins, S. 26. Der Prozess führte allerdings dazu, dass sich die Münzen eines Münztyps untereinander mitunter stark unterschieden. Falls der Rohling verrutschte oder zu klein war, wurde das Bild nur teilweise auf die Münze eingeschlagen. Auch die Stempel selbst wurden von Hand gefertigt und weisen demnach Variationen auf. Für Stempelstudien ist hier relevant, dass sich die Avers- und Reversstempel unterschiedlich schnell abnutzten und dadurch nicht gleichzeitig ausgetauscht wurden. So kann die Reihenfolge der Stempel anhand von Münzen nachgezeichnet werden. Auf der Deutungsebene bedeutet das, dass Münzfunde mit vielen Stempelkopplungen vor der Verbergung wenig im Umlauf gewesen sein müssen. Der Fund steht dann üblicherweise in engem Zusammenhang mit unmittelbaren staatlichen Ausgaben wie beispielsweise der Bezahlung der Truppen. Um Stempelstudien durchzuführen, müssen allerdings ein großer Teil aller Münzen eines Typs vorliegen. Trotz fortschreitender Digitalisierung wird es wohl noch einige Zeit dauern, bis diese Voraussetzung erfüllt ist.



kohäsiver Faktor, der den inneren Zusammenhalt des Reiches stärken sollte.<sup>94</sup> Auch die Bilder des Reverses sind ikonografisch innerhalb der Herrscherinszenierung zu lesen und teilen ihre Bildsprache mit anderen Monumenten. Howgego erkennt auf Münzen Traians die gleiche Motivik wie auf dem Traiansbogen in Benevent. Dazu gehören der Triumph über die Daker und der neu gebaute Hafen in Ostia.<sup>95</sup> Oft wurde bei einzelnen Münztypen an bereits bekannte Darstellungen und Gottheiten erinnert.<sup>96</sup> Aber nicht nur die einzelne Münze, sondern auch eine Serie unterschiedlicher Münztypen kann ein bestimmtes Programm verfolgen. So inszenierte Philippus Arabs seine Säkularfeiern im Jahr 248 n. Chr. mit der Prägung von Serien mit verschiedenen exotischen Tieren, die bei Tierhetzen präsentiert wurden.<sup>97</sup> Über diese Bildsprache konnte auch die analphabetische Bevölkerung erreicht werden, die mit ihr vertraut war.<sup>98</sup>

Ikonographische Darstellungen auf Avers und Revers unterscheiden sich hinsichtlich ihrer Normierung und ihres Erhaltungszustandes. Auf den Portraits des Averses lassen sich die Gesichtszüge deutlich erkennen. Bei den Ganzkörperdarstellungen des Reverses ist das nicht der Fall. Das bedeutet keineswegs, dass die Bilder des Avers eindeutiger klassifizierbar sind. Es gibt starke Abweichungen in der Darstellung der gleichen Person, bei Caracalla gibt es beispielsweise unterschiedliche Bilder für Altersstufen seiner Jugend. Außerdem ähneln sich die Darstellungen von Personen der gleichen Dynastie. Größere Zäsuren gibt es bei umkämpften Machtwechseln, die durch eine große bildliche Abgrenzung vom Vorgänger ausgedrückt werden. Die Reverse zeigen eine große Vielfalt an Personen, oft Gottheiten oder Personifikationen, die bestimmte Tugenden oder Orte repräsentieren. Die Abbildung von Personifikationen und Gottheiten entstammt der griechischen Münzprägetradition.<sup>99</sup> Sie haben wiederkehrende Attribute, über die sie eindeutig wiedererkannt werden können. Diese können sich sehr ähneln, etwa bei Abundantia und Annona, die beide teilweise mit einem Füllhorn und Getreideähren über einem Modius (Getreidekorb) dargestellt werden.<sup>100</sup> Die dargestellten Figuren auf dem Revers sind oft zu schemenhaft, um sie eindeutig einer bestimmten Person zuzuordnen, sodass die Attribute für deren Bestimmung ausschlaggebend sind. Es gibt aber auch Fälle, wo die Identifikation der abgebildeten Person nur anhand der Legende der Münze, also über eine Inschrift auf

---

<sup>94</sup> Ebd., S. 43.

<sup>95</sup> Ebd., S. 76f.

<sup>96</sup> HIJMAN, Art and Economy: Images on Roman Coins, S. 525.

<sup>97</sup> Vgl. RIC IV Philip I 12–23.

<sup>98</sup> KEMMERS/MYRBERG, Rethinking Numismatics, S. 96.

<sup>99</sup> WOYTEK, Bernhard: Heads and Busts on Roman Coins. Some Remarks on the Morphology of Numismatic Portraiture, in: *Revue numismatique* 171 (2014), S. 45–71, hier S. 50.

<sup>100</sup> Vgl. die Darstellung der Annona bei RIC II 3<sup>2</sup> Hadrian 179 und die der Abundantia bei RIC II Traian 492.

dem Avers oder dem Revers, stattfindet. Wenn die Legende fehlt, lässt sich die Frage, um welche Person es sich handelt, teilweise nicht final beantworten.<sup>101</sup> Diese unsicheren Bestimmungen betreffen allerdings nur sehr wenige Münzen der römischen Numismatik. Obwohl die Portraiterkennung mit CNNs bisher noch keine durchschlagenden Erfolge erzielen konnte, handelt es sich um einen Bereich, der wohl vergleichsweise einfach automatisiert werden kann, da die Kaiserdarstellungen auf dem Avers häufiger im Datensatz vorkommen als die variationsreicheren Darstellungen des Reverses. Es wird zu untersuchen sein, ob die Kaiser von der vorliegenden KI besser erkannt werden können als andere Darstellungen.

Die textuelle Dimension von Münzen besteht aus Inschriften, sogenannten Legenden. Sie befinden sich auf kaiserzeitlichen Münzen am Rand, bei republikanischen in der Mitte. Auf dem Avers benennt die Legende meistens den Dargestellten samt Ehrentiteln. Auch die Legenden des Revers beschreiben häufig noch die abgebildete Person auf dem Avers. Neue Personifikationen, deren Attribute zu unbekannt waren, um von der Bevölkerung verstanden zu werden, wurden jedoch durchgehend auf die Reversseiten gesetzt und für die Herrschaftslegitimation genutzt.<sup>102</sup> Eine Erkennung von Legendenbestandteilen mit CNNs wurde bisher noch nicht umgesetzt und es gibt nur eine Studie, die sie mit analytischen Methoden versucht.<sup>103</sup> Sowohl die Legenden selbst als auch ihre Beschreibungen waren dafür wohl noch zu wenig standardisiert. Insbesondere die Legende ist am Rand von Korrosion und Stempelabnutzung betroffen. Auch bei einer dezentrierten Prägung ist bei kaiserzeitlichen Münzen die Legende schnell nicht mehr sichtbar. Allgemeine Optical-Character-Recognition-Systeme (OCR) funktioniert bei Legenden nicht, auch weil sie keinen eindeutigen Kontrast zum Münzhintergrund haben. Außerdem können die Ligaturen der Münzlegenden nicht erkannt werden.<sup>104</sup> Für eine vollständige Münzbestimmung ist allerdings auch eine Bestimmung ihrer Legenden notwendig, sodass deren Gelingen eine vordringliche Aufgabe für die künftige Forschung ist.

Die Legendenbestimmung geht über den Rahmen der vorliegenden Arbeit hinaus, auch weil die provinzialrömischen Prägungen durch ihre griechischen Legenden ein zusätzliches Maß an Komplexität ins Spiel bringen. Bei diesen Prägungen handelt es sich um Prägungen griechischer Städte, die auch noch nach der Expansion Roms vor allem in Bronze weitergeführt

---

<sup>101</sup> Hijmans verweist auf RPC I, 5203 und stellt den Diskurs, ob Augustus oder Nero auf dem Revers abgebildet ist. HIJMANS, *Art and Economy: Images on Roman Coins*, S. 522. Der RPC spiegelt diese Forschungskontroverse nicht wider und schreibt lediglich ‚Nero‘.

<sup>102</sup> HOWGEGO, *Ancient history from coins*, S. 75.

<sup>103</sup> Vgl. KAVELAR u.a., *Reading the Legends of Roman Republican Coins*.

<sup>104</sup> Ebd., S. 4.

wurden. Diese Prägung erklärt sich nach Hijmans nicht durch wirtschaftliche Notwendigkeit: Sie wurden zu selten geprägt, um für den Wirtschaftskreislauf relevant zu sein.<sup>105</sup> Sie wurden also vielmehr aufgrund eines spezifischen ideellen Anlasses und Adressaten, nämlich der Bevölkerung der jeweiligen Städte, geprägt.<sup>106</sup>

Die textuelle Ebene von Münzen erlaubt auch die Erfassung der Münzstätte (Prägeort) von Münzen. In republikanischer Zeit wurden Münzen nur in Rom geprägt. Bis in die Zeit der Flavier wurde außerdem in Lugdunum geprägt, erst in der Spätantike gibt es andere Münzprägstätten.<sup>107</sup> Bei den kaiserlichen Prägungen werden diese durch Münzstättenzeichen auf der Münze angegeben und können im Rahmen einer automatisierten Legendenerkennung extrahiert werden. Münzstätten provincialrömischer Münzen werden nicht über die Legende gekennzeichnet. Hier gibt es allerdings schon Vorarbeiten, die die Münzstätten mithilfe von CNNs vorhersagen.<sup>108</sup>

Auch für die Erkennung von Darstellungen auf Münzen sind die Provinzialprägungen eine Herausforderung, wenn auch eine wesentlich kleinere. Die provincialrömischen Prägungen weisen regionalspezifische Motive auf, die sich von der standardisierten Bildsprache der kaiserlichen Numismatik unterscheiden. Die Ausstellung der bürgerschaftlichen Eigenidentität geschah trotzdem im Rahmen einer Loyalität zum Kaiser, dessen Abbild wie bei Reichsprägungen auf dem Avers eingeprägt ist.<sup>109</sup> Visuell können sich provincialrömische Kaiserdarstellungen allerdings stark von denen der Reichsprägungen unterscheiden. Innerhalb der gleichen Region ähneln sich Provinzialprägungen. Manchmal wurden Münzstempel zwischen den Städten ausgetauscht.<sup>110</sup> Kleinere Städte ließen ihre Münzen teilweise von größeren Städten prägen, mit denen sie die gleichen Darstellungen teilten.<sup>111</sup> Aufgrund der zusätzlichen Komplexität ergibt es zugunsten der Generalisierbarkeit Sinn, auch provincialrömische Darstellungen für das KI-Training zu verwenden. Die vorliegende Arbeit ist die einzige Münzklassifikation, die diese Daten verwendet, wohl weil sie am wenigsten vereinheitlicht und nicht einfach herunterzuladen sind. Es gilt zu untersuchen, wie gut die KI provincialrömische Prägungen erkennen kann.

---

<sup>105</sup> HOWGEGO, *Ancient history from coins*, S. 10f.

<sup>106</sup> HIJMANS, *Art and Economy: Images on Roman Coins*, S. 526.

<sup>107</sup> HOWGEGO, *Ancient history from coins*, S. 11.

<sup>108</sup> GAMPE/TOLLE, *Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset*.

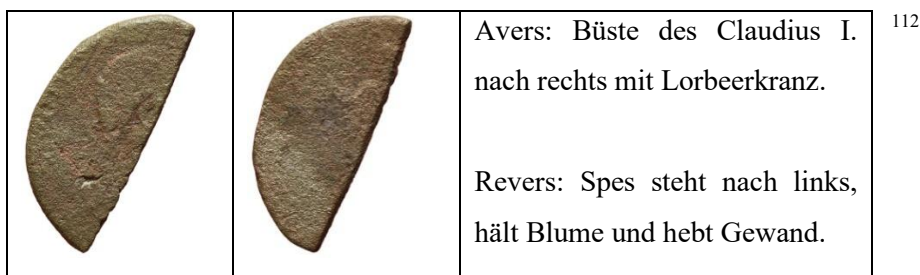
<sup>109</sup> KEMMERS/MYRBERG, *Rethinking Numismatics*, S. 97.

<sup>110</sup> HIJMANS, *Art and Economy: Images on Roman Coins*, S. 529.

<sup>111</sup> HOWGEGO, *Ancient history from coins*, S. 29.

Es zeigt sich, dass die viele Kategorien der Münzerfassung sich implizit oder aus der Logik anderer Kategorien bilden. Für die Münzbestimmung wurden die materiellen Elemente bereits weitgehend automatisiert oder sie erfordern keinen großen manuellen Aufwand. Zentral ist demgegenüber die Erkennung der Ikonographie und der Legenden von Münzen. Diese kann sich auch gegenseitig ergänzen. Interdependenzen erlauben Bestimmungen, die auf den ersten Blick unmöglich erscheinen. Die folgende Münze ist stark abgenutzt und zur Hälfte abgebrochen (Abb. 4).

Abbildung 4: beschädigter Sesterz des Claudius



Auf dem Avers ist noch ein Hinterkopf auszumachen und auf dem Revers ein Kopf und mit etwas Fantasie ein erhobener Arm als Teil einer Personendarstellung. Dennoch ist dem Kunsthistorischen Museum Wien eine Bestimmung dieser Münze gelungen. Hauptindiz war wohl der erhobene Arm, der ein eindeutiges Indiz für die Göttin Spes ist, die in ihrer erhobenen Hand eine Blume hält. Von diesen Spes-Darstellungen gibt es wiederum nur ein Dutzend mit Legende auf dem Revers. Die Kopfform auf dem Avers passt bei diesen wiederum nur zu den zwei angegebenen Münztypen des Claudius. Die vorliegende Erkennung von Münzbildern ist damit nur der erste Schritt. Für eine vollständige Münzklassifikation muss zumindest auch eine Legendenerkennung angestrebt werden.

## 2.2. Entwicklung ihrer Klassifizierung

Die frühe Numismatik hat bereits die Kategorien und das Vokabular erarbeitet, mit dem Münzen bis heute beschrieben werden. Mit Linked Open Data wurden diese Daten aus Katalogen strukturiert, erweitert und über Internetportale zugänglich gemacht. Dabei bestehen mehrere Systematiken parallel, von denen in der vorliegenden Arbeit nur die am häufigsten verwendeten genutzt werden.

<sup>112</sup> Für Bilder und Beschreibungen vgl. Sesterz des Claudius, RIC I<sup>2</sup> Claudius 99 oder 115, 41–54 n. Chr., Kunsthistorisches Museum Wien, Inv.-Nr.: FD 421, online: <https://www.ikmk.at/object?id=ID29082>.

Die Erfassung antiker Münzen hat eine lange Tradition, die maßgeblich von Numismatikern des 19. Jahrhunderts geprägt wurde. Henri Cohen arbeitete als einer der führenden Numismatiker seiner Zeit die Münzsammlung der Bibliothèque Nationale de France auf und legte so einen Grundstein für die systematische Katalogisierung römischer Münzen.<sup>113</sup> Er untersuchte dabei die vorgestellten Kriterien Kaiserportraits, Inschriften, Legenden und ikonographische Motive, um Münzen chronologisch und typologisch zu ordnen.<sup>114</sup> Wenn Münzen das gleiche Nominal, die gleiche Legende und die gleichen Objekte in der gleichen Darstellungsweise und Anordnung teilen, gehören sie zum gleichen Münztyp. Harold Mattingly erweiterte diese Typologisierung durch die Einbeziehung von Fundkontexten und stilistischen Analysen.<sup>115</sup> Aus seiner Feder stammt die Reihe *Coins of the Roman Empire in the British Museum* (BMCRE) und unter Einschluss von Münzen anderer großer Sammlungen die Reihe *Roman Imperial Coinage* (RIC).<sup>116</sup> Der RIC ist der zentrale Referenzkatalog für die kaiserliche Münzprägung. Dieser arbeitet strikt nach einer übersichtlichen Systematik. Münzen werden nach Kaiser und Nominal geordnet und dann wiederum für jede Averslegende eine Tabelle mit verknüpften Reversdarstellungen und -legenden angelegt.<sup>117</sup> Die Zuordnung zu einem Typ ist historisch gewachsen und stets im Wandel. Neuauflagen enthalten mehr Referenzstücke, sind meist feingliedriger und bilden dadurch eine geringere Varianz innerhalb desselben Typs ab.<sup>118</sup> Diese Darstellungsart hat den Nachteil, dass bei schlecht erhaltenem Avers nach den Reversdarstellungen über mehrere Seiten gesucht werden muss. Der RIC hat sich allerdings durch seinen Anspruch auf Vollständigkeit zum Vorbild für numismatische Kataloge anderer Zeiten entwickelt.

---

<sup>113</sup> HOLT, *When Money Talks*, S. 90.

<sup>114</sup> COHEN, Henry: *Description Historique Des Monnaies Frappées Sous L'Empire Romain Communément Appelées Médailles Impériales*, Paris, Londres, 1859.

<sup>115</sup> Von Cohen übernahm er wohl einige französische Begriffe zur Beschreibung der Münzen, etwa „exergue“, was im deutschen mit „im Abschnitt“ übersetzt wurde und den unteren Teil der Münze bezeichnet, wo sich häufig die Münzstättenzeichen befinden.

<sup>116</sup> SUTHERLAND, Carol Humphrey Vivian: *From 39 BC to AD 69 (The Roman Imperial Coinage, I)*, London, 1923.

<sup>117</sup> MATTINGLY, Harold/SYDENHAM, Edward Allen/SUTHERLAND, Carol Humphrey Vivian: *Macrinus to Pupienus (The Roman Imperial Coinage, IV, 2)*, London, 1938, S. 44 C. gibt den Grad der Seltenheit bedeutet und steht für Common, C.3 Pl. 12, 7 wiederum für den dritten Band der Cohen'schen Reihe zu Münzen des Altertums. Vgl. COHEN, *Description Historique Des Monnaies Frappées Sous L'Empire Romain Communément Appelées Médailles Impériales*.

<sup>118</sup> Eines der jüngsten Beispiele für eine überarbeitete Typisierung stellt die Neuauflage des RIC zur Prägung Hadrians. Vgl. HOWARD, Emma/AMANDRY, Michel/ABDY, Richard A. (Hrsg.): *From AD 117 to AD 138 - Hadrian (The Roman Imperial Coinage, II, 3)*, London, 2019.

Diese sind für die provinzialrömischen Münzen die Reihe Roman Provincial Coinage (RPC) und für die republikanischen der Roman Republican Coinage (RRC)<sup>119</sup>. Andere Systematiken wie die Reihen Sylloge Nummorum Romanorum legen den Fokus stärker auf Abbildungen von Münzen.<sup>120</sup> Das Dictionary of Roman Coins schlüsselt wiederum jede bekannte Darstellung und Legende auf. In mehreren Auflagen wurde dieses Wörterbuch erweitert.<sup>121</sup> Diese Werke werden bei der Bestimmung oft unterstützend herangezogen, klassifiziert wird aber nach dem RIC / RRC / RPC. Alle drei Bände können digitalisiert über Onlineportale abgerufen werden. Das sind Online Coins of the Roman Empire (OCRE),<sup>122</sup> Coins of the Roman Republic Online (CRRO),<sup>123</sup> und RPC-online.<sup>124</sup> Der RPC-online ist dabei den gedruckten Bänden voraus und enthält Daten zu noch nicht publizierten Münzen.

Die online-Kataloge CRRO / OCRE / RPC-online arbeiten mit Normdaten und nach dem Linked-Open-Data Prinzip. Das heißt, die Kataloge speichern die Merkmale des Münztyps entweder als Freitext, oder wo es Normdaten gibt, als Normdaten. Die Datenbanksysteme der GLAM-Institutionen liefern wiederum die Münzen zum jeweiligen Münztyp mit ihren eindeutigen Eigenschaften wie Aussehen (durch Fotografien), Maße und Gewicht. Bei Fundmünzen werden zusätzlich die Fundorte übermittelt. Diese Daten können bei CRRO und OCRE direkt quantitativ in Form von Balkendiagrammen und Karten ausgewertet werden.<sup>125</sup> Während der RIC und der RRC bereits auf einzelne, bestimmte Stücke in den großen Sammlungen verweisen, können durch Linked-Open-Data die mit dem Typ verknüpften Münzen direkt angesehen und verglichen werden. Dadurch verbinden die genannten Onlineportale die Vorteile von Wörterbüchern und gedruckten Vergleichsstücken mit den klassischen Münzkatalogen und haben sich gegenüber diesen für den Bestimmungsprozess durchgesetzt.

### 2.3. Technische Umsetzung – FAIRe Forschungsdaten

Der erste Schritt für die Erstellung eines Datensatzes ist der Download der Daten aus entsprechenden Datenbanken. Diese sind zunächst die digitalisierten Typenkataloge OCRE / CRRO /

---

<sup>119</sup> CRAWFORD, Michael H.: Roman Republican coinage, Cambridge, 1974.

<sup>120</sup> HOLT, When Money Talks, S. 94.

<sup>121</sup> STEVENSON, Seth William: A dictionary of Roman Coins, Republican and Imperial, London, 1889. Den Fokus auf die verbindenden Elemente im Sinne eines Wörterbuchs legt auch Schmidt-Dick mit ihrem mehrbändigen Typenatlas römischer Reichsprägungen. Vgl. SCHMIDT-DICK, Franziska: Typenatlas der römischen Reichsprägung von Augustus bis Aemilianus (Denkschriften / Österreichische Akademie der Wissenschaften, Philosophisch-Historische Klasse), Wien, 2003.

<sup>122</sup> <https://numismatics.org/ocre/>.

<sup>123</sup> <https://numismatics.org/crro/>.

<sup>124</sup> <https://rpc.ashmus.ox.ac.uk/>.

<sup>125</sup> Vgl. <https://numismatics.org/ocre/visualize/distribution>.

RPC-online samt den darin verknüpften Münzen aus GLAM-Institutionen. Außerdem wurden Münzen aus dem Portal Corpus Nummorum (CN) heruntergeladen, das provinzialrömische Münzen aus dem Raum Moesia inferior, Thracien, Mysien und Troas enthält.<sup>126</sup> So wurde die Zahl der Provinzialrömischen Münzen erhöht, bei denen eine größere Varianz der Darstellung erwartet wird, die eine größere Zahl Münzen für das KI-Training erfordert. Außerdem wurden Münzen aus dem Portal Interaktive Kataloge der Münzkabinette (IKMK) heruntergeladen, der größte Verbund von Münzen aus deutschsprachigen GLAM-Institutionen.<sup>127</sup>

Trotz der fortgeschrittenen Digitalisierung der Münzen, ist noch keine umfassende Zugänglichkeit dieser Daten gewährleistet. Nur die Daten in OCRE, CRRO und CN haben eine gemeinsame Schnittstelle, von der Daten im standardisierten RDF-Format aufgerufen werden können.<sup>128</sup> Mit einer einzigen Anfrage kann so ein Großteil der Daten heruntergeladen werden. Die anderen Datenbanken (IKMK und RPC-online) haben keine Schnittstellen und es können nur die Daten heruntergeladen werden, die über die Webseite abrufbar sind. Hier müssen Downloadskripte erstellt werden, die auf den Webseiten die richtigen Daten auswerten und konvertieren. Dabei muss beachtet werden, dass GLAM-Institutionen und deren Dateninfrastruktur nicht darauf ausgelegt sind, dass Daten in größerem Umfang heruntergeladen werden können. Für IKMK wurde bereits in einem früheren Projekt des Verfassers ein Skript erstellt, welches die relevanten Daten herunterlädt und vereinheitlicht.<sup>129</sup> Ein solches für RPC-online wurde im Rahmen der vorliegenden Arbeit angefertigt. Die als Resultat angelegte Datenbank ist modular aufgebaut, sodass künftig neue Datenquellen, etwa aus dem Münzhandel eingebunden werden können.

Nach dem Download werden die Daten zunächst zwischengespeichert und daraufhin mit einer zu zeigenden Systematik vereinheitlicht und in einer Datenbank abgespeichert. Jede Beschreibung der Vorder- und Rückseite wurde dabei mit ihrem jeweiligen Avers- und Reversbild,

---

<sup>126</sup> <https://www.corpus-nummorum.eu/>.

<sup>127</sup> <https://ikmk.net/>.

<sup>128</sup> Der API-Endpoint ist über <https://nomisma.org/sparql> erreichbar. Die Verknüpfung der Daten von CRRO, OCRE und CN kann gelingen, weil ihre Datenbanken zur Beschreibung der Münzdaten die Nomisma-Ontologie verwenden. Vgl. <https://nomisma.org/ontology>.

<sup>129</sup> Vgl. BÖRNER, Susanne: Bericht über die numismatischen Aktivitäten im Jahr 2023. Tätigkeitsbericht für die Numismatische Kommission der Länder, 2024, S. 5, online: [https://numismatische-kommission.de/fileadmin/user\\_upload/Nk/pdf/tatigkeitsberichte/2023/Boerner\\_Taetigkeitsbericht\\_fuer\\_23.pdf](https://numismatische-kommission.de/fileadmin/user_upload/Nk/pdf/tatigkeitsberichte/2023/Boerner_Taetigkeitsbericht_fuer_23.pdf) [12.08.2025]. Im Rahmen dieser Arbeit wurden auch unveröffentlichte Daten heruntergeladen, die in einer früheren Version der Datenbanken zugänglich waren. Es hat sich herausgestellt, dass diese Daten nicht qualitativ schlechter sind und trotzdem dem numismatischen Anspruch entsprechen. Aufgrund der geringen Mittel für Numismatik besitzen GLAM-Institutionen oft ein Vielfaches ihrer veröffentlichten Daten bereits digitalisiert und nahezu fertig bearbeitet.

sowie den URLs zur spezifischen Münze und ihrem Münztyp verknüpft.<sup>130</sup> So können die Beschreibungen des gleichen Münztyps schnell aufgerufen und miteinander verglichen werden. Die Bilder werden dabei an einer anderen Stelle durch ein Skript heruntergeladen und separat gespeichert. Die Dateinamen der Bilder entsprechen dabei der URL der Münze mit dem Zusatz „\_a“ für den Avers beziehungsweise „\_r“ für den Revers.<sup>131</sup> Die Bilder wurden außerdem beim Herunterladen in einer einheitlichen Größe von 448x448 Pixeln gespeichert. Dadurch konnte der benötigte Festplattenspeicher auf wenige Dutzend Gigabyte reduziert werden und die Bilder haben dennoch eine ausreichende Größe, um trotz auflösungsverringender Vorverarbeitungsschritte für die KI nutzbar zu sein.<sup>132</sup>

Abbildung 5: Beispieleintrag in der Datenbank

Coin	<a href="https://pecunia.zaw.uni-heidelberg.de/ikmk/object?id=ID3">https://pecunia.zaw.uni-heidelberg.de/ikmk/object?id=ID3</a>	133
Type	<a href="https://numismatics.org/ocre/id/ric.2.tr.570">https://numismatics.org/ocre/id/ric.2.tr.570</a>	
Obverse Image	<a href="https://pecunia.zaw.uni-heidelberg.de/ikmk/image/ID3/vs_opt.jpg">https://pecunia.zaw.uni-heidelberg.de/ikmk/image/ID3/vs_opt.jpg</a>	
Reverse Image	<a href="https://pecunia.zaw.uni-heidelberg.de/ikmk/image/ID3/rs_opt.jpg">https://pecunia.zaw.uni-heidelberg.de/ikmk/image/ID3/rs_opt.jpg</a>	
Obverse Description	An der linken Schulter drapierte Büste des Traianus mit Strahlenkrone in der Brustansicht nach rechts.	
Reverse Description	Steinerne Brücke mit Türmen an beiden Seiten über einen Fluss, darunter rechts ein Schiff.	

Die Abbildung zeigt einen Münzdatensatz in der Datenbank mit allen relevanten Informationen. Über den Münztyp können die Beschreibungen mit anderen Beschreibungen des gleichen Typs verglichen werden. Alle Informationen außer den Beschreibungstexten liegen nach dem Download als URL vor. Auf diese Weise sind sie zwar für den Menschen schlechter lesbar,

<sup>130</sup> Im Portal IKMK befinden sich zwei Datenfelder, die beide zum Münztyp verweisen. Das eine ist ein halbstrukturiertes Feld, das einen Literaturverweis zum RRC / RIC / RPC enthält, das andere ist eine direkte Verlinkung zu den Typdatenbanken. Diese Felder enthalten fast immer die gleichen Informationen, außer es handelt es sich um Stücke, die nicht in den Referenzkatalogen enthalten sind. Dann gibt es Literaturhinweise zu ähnlichen Stücken, die Verlinkung fehlt allerdings, um diese Stücke nicht zu falschen Typen in den Referenzportalen auszuspielen. Um möglichst korrekte Daten zu erhalten, wurden diese beiden Felder auf eine einheitliche Schreibweise gebracht und miteinander verglichen. Dabei hat sich herausgestellt, dass je nach Institution etwa 5% der Daten keine Übereinstimmung vorweisen. Um diese Daten zu korrigieren, wurde ein Qualitätskontrolltool entwickelt, dass an die Institutionen zur Nutzung gegeben wurde.

<sup>131</sup> Dadurch können die Bilder direkt über den Permalink der Münze aufgerufen und die ursprünglichen URLs zu den Originalbildern werden obsolet. Diese Bild-URLs enthalten zudem bei manchen Bildern Sonderzeichen, die im Windows-Filesystem nicht erlaubt sind und deswegen nicht als Dateiname genutzt werden können

<sup>132</sup> Vgl. Kapitel 3.2 Münzbildvorbereitung.

<sup>133</sup> Eigene Darstellung.



ermöglichen aber so Verfahren zur teilautomatisierten Qualitätskontrolle. Erstens können nämlich die Datensätze entfernt werden, deren URLs im Internet nicht aufrufbar sind, sei es durch Rechtschreibfehler in der URL, sei es, weil sie in der Zwischenzeit gelöscht wurden. Zweitens kann überprüft werden, ob funktionierenden Links auch dem erwarteten Schema entsprechen. Auch hier zeigten sich einige durch die manuelle Erfassung verursachte Fehler, die aus dem Datensatz entfernt werden konnten.<sup>134</sup> Durch eine weitere automatisierte Kontrolle konnte wiederum die Datenqualität verbessert werden. In den Daten befanden sich tausende veraltete URLs. Diese entstehen, wenn neue Bände der Typenkataloge erscheinen, die die Nummerierung des gleichen Münztyps verändern. In den online-Portalen wie OCRE werden diese neuen Nummerierungssysteme übernommen. Die Links zu den alten Datensätzen bleiben allerdings erhalten und werden nur auf den neuen Eintrag umgeleitet. Für GLAM-Institutionen ist es deshalb nicht notwendig, die Münzen nach der aktualisierten Systematik zu erfassen, da die alte URL direkt zur neuen führt. Diese Funktionsweise ist für numismatische Daten essenziell, da der Datenbestand über Jahre angewachsen ist und mit vielen Altsystemen aufgenommen wurde, die teilweise weiterbestehen, teilweise ersetzt wurden.<sup>135</sup> Um eine Vergleichbarkeit der Daten zu gewährleisten, müssen allerdings die alten Nummerierungssysteme zu den neuen hin übersetzt werden.

Nach dem Qualitätssicherungsschritt ergibt sich für den Download der Daten die folgende deskriptive Statistik. Die heruntergeladenen Daten von 500.000 Münzen gehören zu knapp 100.000 verschiedenen Münztypen. Einem Münztyp werden dementsprechend durchschnittlich fünf Münzen zugeordnet. Auch wenn es hier starke Abweichungen pro Typ gibt, ist die Datenmenge wohl ausreichend, um die KI zu trainieren. Für die Erkennung von Konzepten in der Ikonographie ist allerdings auch zentral, ob die deutschen und englischen Beschreibungen das Abgebildete in ausreichender Variation abbilden. Der englische Datensatz deckt alle Münztypen ab. Zu allen Münzen liegt jeweils eine englischsprachige Avers- und Reversbeschreibungen vor.<sup>136</sup> Allerdings hat jeder Münztyp auch nur eine einzige englische Beschreibung. Für

---

<sup>134</sup> Dazu gehören Fälle, in denen nicht ein bestimmter Münztyp verlinkt wurde, sondern die vorausgehende Suchanfrage in den Referenzportalen, die zu dem jeweiligen Münztyp führt. Die Suchanfrage in OCRE: [https://numismatics.org/ocre/results?q=Venus+AND+authority\\_facet%3A\"Amandus\"](https://numismatics.org/ocre/results?q=Venus+AND+authority_facet%3A\) hat zum Beispiel nur einen Treffer, nämlich: <https://numismatics.org/ocre/id/ric.5.ama.3>. Da die Eingabe der Referenzen händisch geschieht, verwundert die Verwechslung dieser URLs nicht.

<sup>135</sup> TOLLE, Karsten/WIGG-WOLF, David: Improving Data Quality by Rules: A Numismatic Example, in: CAA (2017), S. 193–201, hier S. 193.

<sup>136</sup> Nach dem Herunterladen befinden sich 13 doppelten Einträge im Datensatz zum RPC-online. Das lässt sich dadurch erklären, dass die Katalognummer des Datensatzes nicht mit der Nummer in der URL des Datensatzes übereinstimmt. Diese falsche Zuordnung aus dem Titel betrifft auch die Suchfunktion beim RPC. Wenn man

künftige Forschung liegt es nahe, noch weitere englischsprachige Beschreibungen herunterzuladen und so mehr englische Schreibweisen abzudecken. Hier böten sich die großen Datenbestände der American Numismatic Society oder des British Museum an. Da aber die englischsprachigen Daten bereits aus mehreren Quellen stammen, und sich die gleiche Beschreibung des Averses oder Reverses auf mehreren Münztypen findet, ist auch hier eine Varianz der Schreibweisen zu erwarten, wenn unterschiedliche Münzen das gleiche Motiv tragen. Für den deutschsprachigen Datensatz ist die Datengrundlage nicht in gleicher Weise umfassend. Nur etwa 28.000 oder ein Viertel aller Münztypen haben eine oder mehrere deutschsprachige Beschreibungen. Dabei konnten 4292 deutsche Beschreibungen, das heißt 2146 Münztypen, aufgrund veralteter oder fehlerhafter URLs nicht den Münztypen aus dem englischsprachigen Datensatz zugeordnet werden. Die hier beobachtete Fehlerquote von etwa 10% findet sich im Folgenden auch im inhaltlichen Vergleich der Münzbeschreibungen und scheint ein grundsätzliches Merkmal manueller Münzbeschreibungen in den GLAM-Institutionen zu sein. Trotz der geringeren Größe des deutschen Datensatzes, handelt es sich um den größten gesammelten Bestand deutschsprachiger Münzbeschreibungen. Zudem hat er häufiger als der englische Datensatz mehrere sich unterscheidende Beschreibungen pro Münztyp. Es kann davon ausgegangen werden, dass er deshalb keine größeren blinden Flecken im Vergleich zum englischsprachigen hat.

### 3. Von Forschungsdaten zur automatisierten Münzklassifikation

Der vorliegende Datensatz wird im Folgenden dazu genutzt, das vortrainierte CNN VGG-19 zu trainieren. Das Modell braucht dabei zwei Inputs, die Fotografien der Averse und Reverse als einheitlichen Vektor und die Darstellungen, die zur einzelnen Münze gehören. Wie aus den Münzbeschreibungen die Konzepte extrahiert werden können, und wie die Fotografien in ein einheitliches Format übersetzt werden können, wird in den folgenden Kapiteln erläutert.

#### 3.1. Münzbeschreibungen und Named Entity Recognition

Bisher wurde im Zusammenhang mit der Erstellung der Datenbank nur allgemein von Münzbeschreibungen gesprochen, sie unterscheiden sich jedoch in Form, Adressat und Inhalt

---

beispielsweise nach Artemis-Selene sucht, taucht der Datensatz <https://rpc.ashmus.ox.ac.uk/coins/6/30732> mit dem Zusatz (temporär) und der Beschreibung: „Artemis-Selene(?) standing facing in long chiton, head right, wearing radiate crown and crescent, holding bow and long torch“ auf. Wenn man allerdings auf die URL klickt, erscheint ein anderer Datensatz, der auf dem Revers die folgende Beschreibung hat: „Apollo (?) standing right, clasping hand of turreted Amazon, standing left and holding shield“ Das RPC-Team wurde auf die betreffenden 13 Duplikate aufmerksam gemacht. Da dieser Fehler nur 13 Datensätze betrifft, ist er größentechnisch für diese Arbeit nicht relevant und die betreffenden Datensätze wurden entfernt.

erheblich voneinander. Aus den Referenzbänden und - Portalen (OCRE / CRRO / RPC-online) stammen Typbeschreibungen, die sich als Hilfestellung für die Klassifikation an Numismatiker richten. Die Objektbeschreibungen der GLAM-Institutionen haben nicht den gleichen wissenschaftlichen Anspruch und dienen entweder dem internen Gebrauch oder richten sich teilweise an ein breiteres Publikum, etwa in Museen. Außerdem sind sie nicht so standardisiert wie die Typbeschreibungen und erlauben die Beschreibung von Sonderfällen und eigenen Deutungen. Die Objekt- und Typbeschreibungen stimmen in den meisten Fällen überein, sodass teilweise von einer „doppelten Buchführung“ gesprochen wird.<sup>137</sup> Die bisherige Forschung zu Automatisierung der Münzklassifikation problematisiert die Unterschiede zwischen Objekt- und Typbeschreibung nicht, obwohl sie erhebliche Auswirkungen auf die Vergleichbarkeit der Daten hat.

Die aus OCRE / CRRO / RPC-online heruntergeladenen Münzen haben jeweils nur eine englischsprachige Typbeschreibung. Für die Münzen aus CN und IKMK liegt jeweils eine deutsch- und englischsprachige Beschreibung vor. Münzen aus IKMK haben eine deutschsprachige Objektbeschreibung und eine zweite englische Typenbeschreibung aus OCRE / CRRO / RPC-online. Wenn also im Folgenden die deutschen und englischen Beschreibungen im heruntergeladenen Datensatz verglichen werden, geschieht das bei Daten aus IKMK zwischen Objekt- und Typbeschreibungen. Es ist davon auszugehen, dass die Beschreibungen deswegen nicht hundertprozentig übereinstimmen. CN hat als zweisprachiges Portal ihre ursprünglich deutschsprachigen Typbeschreibungen ins Englische übersetzt, die direkt verglichen werden können.<sup>138</sup>

Die Unterschiede zwischen Objekt- und Typbeschreibung ergeben sich aus der historisch gewachsenen Natur der Münzerfassungssysteme. Zu Beginn der elektronischen Erfassung gab es je nach Datenbank noch kein Internet und noch keine Referenzportale zu welchen verlinkt werden konnte, sodass eine lokale, eigene Objektbeschreibung wie bei IKMK notwendig war. Aus informatischer und numismatischer Sicht ist eine Zweiteilung der Objekt- und Typbeschreibungen bis heute sinnvoll. Zunächst ist es der Schnelligkeit von Datenbanken zuträglich, wenn alle Daten zentral in einem Datensatz gespeichert werden. Außerdem braucht es eine gesonderte Objektbeschreibung, wenn es sich bei einer Münze um eine Imitation oder Fehlprägungen

---

<sup>137</sup> Ebd., S. 195.

<sup>138</sup> Teilweise enthalten die englischen Beschreibungen Begriffe, die in den anderen Portalen nicht vorkommen und wohl auf eine direkte Übersetzung aus dem Deutschen zurückzuführen sind. So werden Pferdeprotome auf Englisch als *forepart of a prancing horse* bezeichnet und nicht als *horse protomes*. Vgl. [https://www.corpus-nummorum.eu/CN\\_30817](https://www.corpus-nummorum.eu/CN_30817). Auch die Verwendung des Partizip Präsens *sitting* weist auf eine automatisierte Übersetzung hin, die Referenzkataloge nutzen ausschließlich die Partizip Perfekt-Form *seated*.

handelt, die nicht in Typenkatalogen repräsentiert ist. Das gleiche gilt für Münzen, die derartig abgenutzt sind, dass sie nicht mehr eindeutig einem bestimmten Münztyp zugeordnet werden können.<sup>139</sup> Objektbeschreibungen bieten zudem die Möglichkeit, neuere Deutungen abzubilden, während die Typenkataloge oft bereits jahrzehntealten numismatischen Konsens abbilden. Sie verschweigen damit Darstellungen, deren Deutung besonders umstritten sind. So lässt beispielsweise Elagabal bei einigen Münzen ein längliches Objekt über dem Lorbeerkranz seiner Büste prägen. Die Typenkataloge ignorieren dieses Objekt, während es in den Objektbeschreibungen gemäß neueren Deutungen als Stierphallus beschrieben wird.<sup>140</sup>

Der Unterschied zwischen Objekt- und Typbeschreibungen bedeutet eine methodische Herausforderung. Soll die KI nur auf Darstellungen trainiert werden, die tatsächlich auf der Münze abgebildet sind, oder auch auf jene, die zwar dem jeweiligen Münztyp zugeordnet sind, aber möglicherweise auf der Münze fehlen? Streng genommen würde es zwei verschiedener KIs bedürfen. Die Typerkennung würde dabei die Münze ihrem wahrscheinlichsten Typ zuordnen, während die Objekterkennung vom Münztyp abweichende Darstellungen erkennt. Dieses Unterfangen würde jedoch an den im Datensatz vorgefundenen Tatsachen scheitern. Hier zeigt sich nämlich, dass Unterschiede zwischen Typ- und Objektbeschreibungen häufig nicht bedeuten, dass ein bestimmtes Objekt nicht abgebildet wurde, sondern dass es schlicht vergessen wurde. Um eine Objekterkennung anzuvisieren, müsste für jeden Einzelfall überprüft werden, ob eine abweichende Beschreibung auch tatsächlich ein abweichendes Münzbild bedeutet.

Deshalb muss eine Entscheidung getroffen werden. Die vorliegende Arbeit verwendet Typbeschreibungen, da es möglich ist, von der Objektbeschreibung auf sie zu schließen. Umgekehrt ist das nicht der Fall. Das eröffnet Möglichkeiten zur Überprüfung der NER, die sonst nur qualitativ überprüfbar wäre, was bei dem vorliegenden Datensatz aufgrund seiner Größe nicht möglich ist. Wenn allerdings die Heuristik zugrunde gelegt wird, dass deutsche und englische Beschreibungen inhaltlich übereinstimmen, sollten die Ergebnisse der NER ebenfalls übereinstimmen. Die Ergebnisse können so mittels der knapp 30.000 Münztypen aus CN und IKMK überprüft werden, zu denen sowohl englische als auch deutsche Beschreibungen vorliegen. Wenn

---

<sup>139</sup> Ebd.

<sup>140</sup> Vgl. Denar des Elagabal, RIC IV Elagabalus 46d, 221 n. Chr., Münzkabinett Berlin, Inv.-Nr.: 18200736, online: <https://ikmk.smb.museum/object?id=18200736>.

die daraus extrahierten Konzepte übereinstimmen, kann davon ausgegangen werden, dass die NER in beiden Sprachen zufriedenstellende Ergebnisse geliefert hat.<sup>141</sup>

Die Deduktion von Typenbeschreibungen aus Objektbeschreibungen ist möglich, da mit dem Leidener Klammersystem in den Typen vorhandene Darstellungen, die auf dem Objekt fehlen, mit eckigen Klammern markiert werden.<sup>142</sup> Mit der Sprache Regex, die spezifische Muster in Texten erkennt,<sup>143</sup> kann diese Regel mit einem einfachen und effektiven Schema ausgenutzt werden: Ein Wort sei eine beliebige Kombination von Buchstaben A-Z (inklusive Umlaute). Dadurch werden alle Sonderzeichen ignoriert und alle Zeichen zwischen Buchstaben, also Bindestriche und Satzzeichen als Leerzeichen interpretiert. Zur weiteren Vereinfachung wurden Großbuchstaben durch Kleinbuchstaben ersetzt. Umlaute ausgeschrieben und griechische Buchstaben durch ihre lateinische Entsprechung ersetzt.<sup>144</sup>

Panzerbüste des Anthemius [...] in der Frontalansicht. [Schild nicht sichtbar].<sup>145</sup>

panzerbueste des anthemius [...] in der frontalansicht schild nicht sichtbar.

Dieser Solidus des Anthemius stimmt beispielsweise in allem außer dem fehlenden Schild mit dem Münztyp RIC II, 1<sup>2</sup>, 750 überein. In der Konsequenz handelt es sich also streng genommen um einen neuen Münztyp, der aber aufgrund seiner Seltenheit nicht im RIC erfasst wurde. Durch den Regex-Befehl bleibt der Begriff *Schild* in der Beschreibung erhalten und kann mit

---

<sup>141</sup> Der so genannte  $\beta$ -Fehler, nämlich dass die Beschreibungen in Realität nicht übereinstimmen, durch das NLP aber fälschlicherweise angegeben werden, wird durch dieses Vorgehen nicht gesondert betrachtet. Hier ist allerdings davon auszugehen, dass wenn die allgemeine Genauigkeit und Übereinstimmung zunimmt, auch der  $\beta$ -Fehler zurückgeht. Das gilt allerdings nur, wenn die beiden NER-Algorithmen bereits viele Konzepte erkennen. Erkennen Sie auf den Beschreibungen nur ein übereinstimmendes Konzept, ist das kein ausreichender Indikator für die Güte der NER.

<sup>142</sup> Um unvollständige Legenden und Darstellungen widerzugeben, nutzt die Numismatik das ursprünglich für die Papyrologie entwickelte Leidener Klammersystem. Relevant sind in der Numismatik runde Klammern zum Ausschreiben von abgekürzten Wörtern und eckige Klammern, die nicht mehr erhaltene Buchstaben anzeigen.

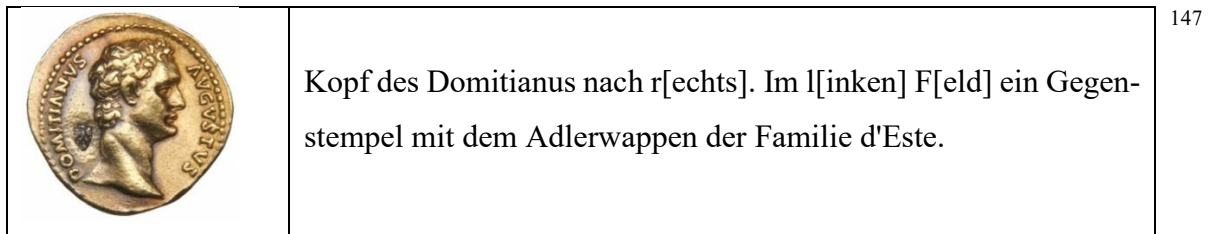
<sup>143</sup> JURAFSKY, Dan/MARTIN, James H.: Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Prentice Hall series in artificial intelligence), Upper Saddle River, 2009, S. 5.

<sup>144</sup> Im Englischen werden Sinnzusammenhänge mit Kommata und Semikola getrennt, im Deutschen werden normalerweise ganze Sätze für jede Sinneinheit geschrieben. Die Satzzeichen sind jedoch für eine Begriffsextraktion nicht relevant und können deshalb sicher entfernt werden. Wenn die Beschreibungen nur noch aus kleingeschriebenen Wörtern im Buchstabenspektrum von A-Z bestehen, sind die nächsten computerbasierten Schritte wesentlich einfacher. Auf diese Weise konnte die Ausführungszeit der Begriffssuche von ursprünglich 10 Minuten auf wenige Millisekunden reduziert werden. Hier muss allerdings darauf geachtet werden, dass keine Informationen verloren gehen. Es hat sich herausgestellt, dass manchmal im Englischen und Deutschen Akzente, wahrscheinlich aus Zufall, in den Daten auftauchen. Im Deutschen werden Umlaute und das Eszett ersetzt. Außerdem taucht manchmal „P“ als Symbol für das Christogramm auf und wird auch dementsprechend umgewandelt.

<sup>145</sup> Solidus des Anthemius, RIC II 1<sup>2</sup> 750, 467–472 n. Chr., Münzkabinett der Staatlichen Museen Berlin, 18309884, online: <https://ikmk.smb.museum/object?id=18309884>

NER extrahiert werden.<sup>146</sup> Nur in seltenen Fällen enthalten die Objektbeschreibungen Zusätze, die einen potenziellen Störfaktor bei der NER darstellen. Um in diesen nicht fälschlicherweise Darstellungen zu erkennen, müssen sie entfernt werden. Das betrifft Gegenstempel (Abb. 8), Beizeichen, Ritzungen, Graffiti und weitere Besonderheiten der einzelnen Münze.

Abbildung 6: Aureus mit Gegenstempel



Ohne eine klare Trennung dieser neuzeitlichen Markierung von den ursprünglichen Münztypinformationen könnte ein automatisiertes System das in der Abbildung erkennbare Adlerwappen fälschlicherweise als eine antike Darstellung eines Adlers interpretieren. Es wurde eine Liste von Gegenstempeln und weiteren Besonderheiten erstellt, auf dessen Grundlage Einzelfälle dieser Art entfernt werden können.<sup>148</sup>

### 3.1.1 Umsetzung durch den Knowledge Graph

Die NER wird in der vorliegenden Arbeit mittels eines zu erstellenden Knowledge Graph implementiert, in dem Konzepte, mit denen die KI trainiert werden soll, über semantisch sinnvolle Relationen (z. B. durch Verben) miteinander in einem Netzwerk verknüpft und gespeichert werden. Dabei können auch orthographische und morphologische Varianten erfasst und mit den Konzepten verknüpft werden. Wie im Folgenden erläutert wird, kann auf diese Weise nicht nur die Vieldeutigkeit einzelner Begriffe reduziert, sondern auch ihr kontextueller Zusammenhang besser nachvollzogen werden. Im Datensatz finden sich allerdings auch häufig flektierte Formen und Rechtschreibfehler, die nicht zwangsläufig aufwändig im Knowledge Graph abgelegt

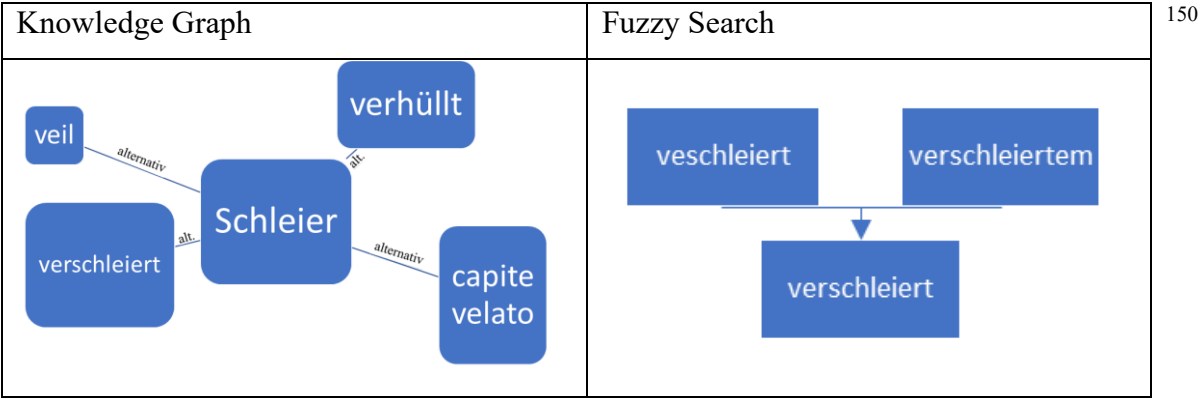
<sup>146</sup> Die NER erkennt nicht, dass Schild in einem Zusammenhang mit *nicht* steht. Vgl. Kapitel 3.1.3.

<sup>147</sup> Avers eines Aureus des Domitian, RIC II 1<sup>2</sup> 750, 92–94 n. Chr., Münzkabinett der Staatlichen Museen Berlin, 18205378, online: <https://ikmk.smb.museum/object?id=18205378>.

<sup>148</sup> Das betrifft nicht nur einzelne Begriffe, sondern auch teilweise die gesamte Beschreibungssystematik. In dem Beschreibungssystem der Universität Tübingen wird beispielsweise eine Datierung der Münze nach dem Kaiserjahr den Münzbeschreibungen nachgestellt. Die Beschreibung lautet demnach „Der Pharos (Leuchtturm) von Alexandria, darauf Schneckenhorn (buccinum) blasende Tritone und Kaiserstatue. Beiderseits L - IZ (= Jahr 17 des Hadrianus, d. h. 132/133 n. Chr.).“ Vgl. Hemidrachme des Hadrian, RPC III Nr. 5853, 132/133 n. Chr. Münzkabinett des Instituts für Klassische Archäologie der Eberhard-Karls-Universität Tübingen, Inv.-Nr.: SNG Tübingen 7, 271, online: <https://www.ikmk.uni-tuebingen.de/object?lang=de&id=ID9402>. Bei dieser Münze würde die NER Hadrian erkennen, auch wenn er nicht abgebildet ist. Diese Datierungen befinden sich immer innerhalb von „(=[...])“ und können einfach entfernt werden.

werden müssen. Es wäre zu mühsam, jede mögliche Wortendung in den Katalog mitaufzunehmen. Es sollten daher Wörter auf Basis ihrer lexikalischen Grundform (Lemma) erkannt werden. Um diese Lemmata in den Beschreibungen zu erkennen, bietet sich die sogenannte Fuzzy Search an, die bei der Stichwortsuche eine gewisse Abweichung toleriert. So können auch Rechtschreibfehler und flektierte Wortformen ihren jeweiligen Konzepten im Knowledge Graph zugeordnet und durch ihn identifizierbar gemacht werden.<sup>149</sup>

Abbildung 7: Funktionsweise von Fuzzy Search und Knowledge Graph



Im Knowledge Graph können alle alternativen Schreibweisen des Konzepts *Schleier* in den im Datensatz vorkommenden Sprachen Deutsch, Englisch, Latein und Altgriechisch gespeichert werden. Alle Begriffe, die einem Konzept zugeordnet werden, werden in der NER mittels der Fuzzy Search im Datensatz gesucht. Auf diese Weise kann der *Schleier* auch mithilfe anderer sinnbehafteter Verbindungen wie *tritt gemeinsam auf* oder *tritt ausschließlich gemeinsam auf* mit anderen Konzepten verknüpft werden. Wie sich noch zeigen wird, ist diese zusätzliche Funktionalität des Knowledge Graphen gegenüber einem klassischen Thesaurus notwendig, um die extrahierten Konzepte weiter zu vereinheitlichen.

Auf dieser Grundlage kann der Knowledge Graph mit den Begriffen aus den Münzbeschreibungen gefüllt werden. Zunächst wurde automatisiert eine Liste aller Wörter im Datensatz erstellt und damit geprüft, welche Wörter in den Knowledge Graph aufgenommen werden müssen. Diese Liste enthält 15.000 deutsche, englische, lateinische und altgriechische Wörter. Ein Wort kann dabei auf mehrere Darstellungen hinweisen und eine Darstellung kann durch

<sup>149</sup> Mit der Nutzung der Fuzzy Search kann der Aufwand reduziert werden, es entstehen allerdings neue Fehler, wenn falsche Wörter erkannt werden. Als geeignete Lösung hat sich herausgestellt, die Fuzzy-Search um eine gesonderte Liste zu ergänzen, die Begriffe enthält, nach denen nur exakt gesucht werden soll. Der Begriff „Gewand“ sollte beispielsweise nicht „gewandt“ finden, „Greif“ nicht „greifend“ und „Waage“ nicht „waagerecht“.

<sup>150</sup> Eigene Darstellung.

mehrere Wörter beschrieben werden. Ein Beispiel dafür ist das Kultbild der Diana Nemorensis. Dabei handelt es sich um eine archaische Darstellung dreier Göttinnen in einem heiligen Hain, die sich auf zahlreichen republikanischen Denaren wiederfindet (Abb. 9).<sup>151</sup>

Abbildung 8: Das Kultbild der Diana Nemorensis



Diese Statuengruppe mit den drei Göttinnen Frontalansicht ist auf römischen Münzen nur ein Mal zu finden und es ergibt Sinn, das Kultbild auch als eigenständige Darstellung von Diana Nemorensis, deren Büste auf dem Avers zu sehen ist, getrennt zu betrachten. Diana Nemorensis wird auf dem Avers der Münze als Büste dargestellt und gleicht in dieser Darstellung anderen Dianabüsten. *Büste der Diana Nemorensis* ist demnach eine alternative Schreibweise des Konzepts *Diana* sein, während die *Statuengruppe der Diana Nemorensis* ein eigenständiges Konzept im Knowledge Graph darstellt. Die Statuengruppe wird jedoch auf unterschiedlichste Weise beschrieben. In den Beschreibungen finden sich für das Kultbild unter anderem die folgenden Varianten.

Drei Göttinnen (Diana, Hekate, Selene) stehen in der Vorderansicht [...].<sup>153</sup>

Statuengruppe der Diana Nemorensis mit Diana, Hekate und Selene [...].<sup>154</sup>

Dreigestaltige Kultstatue der Diana Nemorensis in der Frontalansicht [...].<sup>155</sup>

<sup>151</sup> ALFÖLDI, Andreas: Diana Nemorensis, in: *American Journal of Archaeology* 64, 2 (1960), S. 137–144, hier S. 137.

<sup>152</sup> Münzkabinett der Staatlichen Museen Berlin, Inv.-Nr.: 18202033, RRC Nr. 486,1, online: <https://ikmk.smb.museum/object?id=18202033>.

<sup>153</sup> Reversbeschreibung RRC Nr. 486,1, 43 v. Chr., Münzkabinett der Staatlichen Museen Berlin, Inv.-Nr.: 18202033, online: <https://ikmk.smb.museum/object?id=18202033>.

<sup>154</sup> Reversbeschreibung RRC Nr. 486,1, 43 v. Chr., Münzkabinett der Universität Wien, Inv.-Nr.: S\_00475, online: <https://ikmk-ing.univie.ac.at/object?id=ID192>.

<sup>155</sup> Reversbeschreibung RRC Nr. 486,1, 43 v. Chr., Münzkabinett des Instituts für klassische Archäologie der Universität Tübingen, Inv.-Nr.: III 31/2 (B1), online: <https://www.ikmk.uni-tuebingen.de/object?id=ID1437>.



Dreifiguriges Kultbild der Diana Nemorensis (Diana, Hekate, Selene) [...].<sup>156</sup>

Die Eigenständigkeit dieser Darstellung definiert sich demnach über die Begriffe *Kultbild* / *Kultstatue* / *Statuengruppe der Diana Nemorensis*. Die erste Beschreibung „Drei Göttinnen (Diana, Hekate, Selene) [...]“ enthält allerdings keinen expliziten Bezug zum Kultbild der Diana Nemorensis und nur durch manuelle Kontrolle kann die Begriffskombination *Drei Göttinnen Diana Hekate Selene* als eindeutiges Zeichen für das Kultbild der Diana Nemorensis erkannt werden. Dieses Beispiel zeigt, dass jedes Wort in all seinen Kontexten betrachtet werden muss, bevor es in den Knowledge Graphen eingeordnet werden kann.

Neben komplexen Ikonographien wie der eben genannten können auch Homonyme durch den Abgleich mit dem Knowledge Graph erkannt werden. Um Homonyme unterscheiden zu können, muss durch die NER der unmittelbare Kontext des jeweiligen Wortes untersucht werden. Beispiel: Ein häufiges Homonym im Datensatz ist das Wort *Tympanum* (beziehungsweise gr. *Tympanon*), das einerseits zur Beschreibung einer Trommel, andererseits der dreieckigen Giebelfläche eines Tempels herangezogen wird.<sup>157</sup> Bei genauer Betrachtung zeigt sich hier, dass bereits in unmittelbarer Nähe liegende Präpositionen (im Tympanon) oder Komposita (Tympanonfeld) ausreichen, um den Giebel von der Trommel zu unterscheiden. Homonyme fallen zwar nur bei manueller Durchsicht auf, sind aber sprachspezifisch und können deshalb über den Vergleich der Beschreibungen ausfindig gemacht werden.<sup>158</sup> Homonyme können um die jeweiligen Kontexte bereichert in bereits vorgestellter Weise als alternative Schreibweisen im Knowledge Graph gespeichert werden.

Der Knowledge Graph sollte möglichst kurz sein und gleichzeitig aus gut definierbaren, literaturgestützten Begriffen bestehen.<sup>159</sup> Das heißt, es sollten nur diejenigen Begriffe aufgenommen werden, die eindeutig auf eine bestimmte Darstellung hinweisen. Ansonsten würden möglicherweise Darstellungen erkannt, die in der Ikonographie überhaupt nicht impliziert sind. Um die Eindeutigkeit zu gewährleisten, wurden die folgenden Schritte durchgeführt: Zunächst wurde

---

<sup>156</sup> Reversbeschreibung RRC Nr. 486,1, 43 v. Chr., Gorny & Mosch Giessener Münzhandlung, Auktion 311, Losnummer: 279, online: <https://www.coinarchives.com/a/lotviewer.php?LotID=2541665&AucID=6391&Lot=279&Val=dc5c32556513ff67303e2da00b2acfe2>.

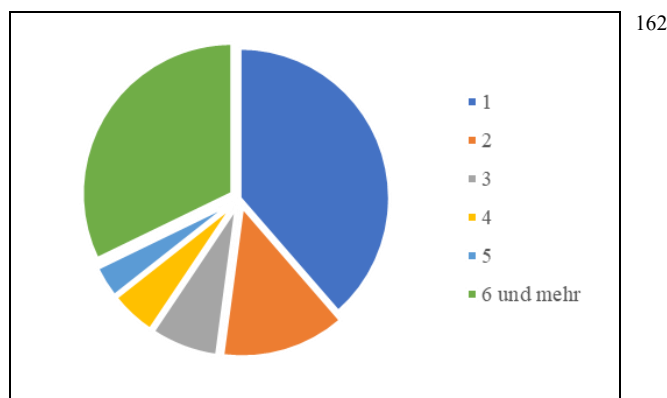
<sup>157</sup> Weitere Beispiele finden sich im Englischen: ‚veil‘ für Segel und Schleier und ‚palm‘ für Palme beziehungsweise Handfläche.

<sup>158</sup> Die unterschiedlichen Schreibweisen von Tympanum als numismatisch fachsprachlicher Begriff fallen etwa auf, da er Synonym mit den deutschen Begriffen Giebel beziehungsweise Trommel und den englischen Begriffen Pediment beziehungsweise drum verwendet wird. Im Vergleich der englischen und deutschen Beschreibungen fällt dann auf, wenn sie sich in unerwarteten Begriffen unterscheiden, in diesem Fall also die Begriffe Trommel und Giebel.

<sup>159</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 212.

ein Katalog an ikonographischen Darstellungen auf Basis der bestehenden Normdatensysteme erstellt. Die von Deligio und Gencer auf Basis ihrer NER-KI extrahierten Begriffe, eine Liste von 300 Objekten, bildete hierbei den Ausgangspunkt für die Suche nach weiteren Begriffen.<sup>160</sup> Alle Variationen von Schreibweisen wurden aus den Normdatensystemen übernommen, wenn sie im Datensatz auftauchen. Griechische Gottheiten wurden, falls sie eine römische Entsprechung haben, unter dem römischen Namen einsortiert. Dieser so erstellte Katalog umfasst etwa 1.500 normdatenverknüpfte Begriffe.<sup>161</sup> Allerdings ist die dann noch verbleibende Zahl von 13.500 Wörter zu umfangreich für eine manuelle Kontrolle, sodass vorab noch seltene und damit statistisch insignifikante Wörter aussortiert wurden. Diese finden sich im Datensatz häufig. Die Mehrheit der Wörter im gesamten Datensatz kommt seltener als drei Mal vor (Abb. 10).

Abbildung 9: Verteilung der Wörter nach Wortwiederholungen



Dieser Befund ist bei natürlicher Sprache – insbesondere im Deutschen – wenig überraschend. Unter den seltenen Wörtern finden sich häufig seltene Bilddarstellungen (*Bienenwabe*), flektierte Formen (*angedeuteten*) oder zusammengesetzte Substantive (*Ausfallschrittstellung*). Für die NER sind solche Einzelvorkommen weitgehend unproblematisch, da relevante Begriffe in der Regel mittels Fuzzy Search erkannt werden. Gleichzeitig gibt es eine kleinere Gruppe von Wörtern, die sehr häufig im Datensatz auftreten: Etwa tausend Wörter kommen mehr als hundertmal vor, knapp zweihundert davon sogar häufiger als tausendmal. Bei diesen Wörtern handelt es sich größtenteils um numismatische Fachbegriffe wie etwa *Strahlenkrone*, sowie um häufige Verben und Konjunktionen. Um eine systematische Konzeptzuordnung für die

<sup>160</sup> Vgl. [https://github.com/Frankfurt-BigDataLab/NLP-on-multilingual-coin-datasets/blob/main/lists/csv/nlp\\_list\\_obj.csv](https://github.com/Frankfurt-BigDataLab/NLP-on-multilingual-coin-datasets/blob/main/lists/csv/nlp_list_obj.csv).

<sup>161</sup> Für die folgenden manuell extrahierten Konzepte existieren demnach noch keine numismatischen Normdaten. Um dennoch die Normdatenbindung zu gewährleisten, werden Normdaten aus Wikidata oder VIAF verknüpft.

<sup>162</sup> Eigene Darstellung.

Münzbeschreibungen zu ermöglichen, wurden deshalb alle etwa 4000 Begriffe, die mindestens fünfmal vorkommen, manuell überprüft. Konkret wurden zunächst diejenigen Beschreibungen untersucht, bei denen die Differenz der vorhandenen Wörter gegenüber den extrahierten Konzepten am größten ist. Danach wurden die aus deutschen und englischen Beschreibungen des gleichen Münztyps extrahierten Konzepte verglichen und bei fehlender Übereinstimmung die noch nicht extrahierten Konzepte im Knowledge Graph nachgetragen. Nach jedem Schritt wurde überprüft, ob durch das Hinzufügen eines Begriffs möglicherweise an anderer Stelle ungewollte Darstellungen extrahiert wurden. Dieses Vorgehen stellt sicher, dass die Begriffsextraktion möglichst vollständig bleibt und keine thematisch relevanten Begriffe übersehen werden – also keine „blinden Flecken“ entstehen. Dieser manuelle Prozess dauerte etwa zwei Monate. Von den untersuchten 4.000 Begriffen wurden 3.524 Begriffe als direkt auf eine Darstellung auf Münzen hinweisend identifiziert. Bei den Übrigen handelt es sich um Füllwörter oder Verben, die keine Darstellung implizieren, sondern das Verhältnis der Darstellungen zu einander beschreiben. Die 3.524 ausgewählten Begriffe gehören als unterschiedliche Schreibweisen zu 801 identifizierten, normdatenverknüpften Konzepten. Diese bilden das Zentrum des Knowledge Graphs. Der Prozess, die Begriffe in den Knowledge Graph einzufügen, ist wie bereits diskutiert hermeneutisch. Das heißt, dass die im folgenden beschriebenen Herausforderungen erst im Laufe der Arbeit aufgetaucht sind.

### 3.1.2 Umgang mit Mehrdeutigkeit

Nachdem mit unterschiedlichen Schreibweisen für gleiche Konzepte mithilfe von Knowledge Graph und Fuzzy Search umgegangen wurde, stellt sich heraus, dass viele Beschreibungen noch nicht vereinheitlicht werden konnten. Diese mehrdeutige und unsichere Beschreibungen erfordern eine andere Vorgehensweise. Damit knüpft die vorliegende Arbeit an aktuelle interdisziplinäre Forschungsarbeiten der Archäologie und Informatik an, die damit befasst sind, Vagheiten und Unsicherheiten in Daten zu modellieren.<sup>163</sup> Daraus ergibt sich, dass für die Modellierung eine Data Augmentation notwendig ist, das heißt die Anreicherung der Datengrundlage durch weitere Daten. Auch hier bedarf es numismatischen Fachwissen und der Erkennung von Mustern in den Beschreibungen.<sup>164</sup> Manche Darstellungen enthalten implizite Informationen, die

---

<sup>163</sup> Vgl. die NFDI-Temporary Working Group „Community Standards for Modelling Fuzziness & Wobbliness in research data“ seit Januar 2025. online: <https://www.nfdi4objects.net/portal/twgs/community-standards-for-modeling-fuzziness-wobbliness-in-research-data-using-semantic-web-technologies-and-formalisms-fuzzy-wobblysw>.

<sup>164</sup> Aussagen auf Basis von Beobachtungen im Datensatz geschehen allerdings unter Vorbehalt. Es ist denkbar, dass neue Münzen gefunden werden, die nicht in dieses Schema fallen. Dann müsste der Knowledge Graph angepasst werden.

zur Erzeugung von Redundanz genutzt werden können. Tolle und Wigg-Wolf nutzen im Besonderen zwei Varianten der Data Augmentation auf Münzdaten. Als Erstes setzen sie die Daten in einen logischen Zusammenhang. Beispiel: eine Münze eines bestimmten Kaisers muss in seiner Regierungszeit geprägt worden sein. Im zweiten Schritte werden redundante, implizite Informationen wie zum Beispiel das Material, das im Nominal steckt, hinzugefügt.<sup>165</sup> Wenn redundante Informationen impliziten genannten Objekten hinzugefügt werden, können diese mit expliziten beschriebenen Objekten verglichen werden. Diese Methoden werden in der vorliegenden Arbeit auf die Münzbeschreibungen angewendet. Der logische Zusammenhang der Ikonographie besagt, dass bestimmte Darstellungen nur mit bestimmten anderen auftauchen können. Diese beiden Prinzipien, Kontextlogik und Redundanz erzeugung, schlagen auch Pavlek und Winters vor, um seltene Konzepte mit ähnlichen Konzepten zu verbinden und damit häufigere, allgemeinere Klassen zu generieren.<sup>166</sup> Data Augmentation auf Basis von Kontextlogik und Redundanz lässt sich anhand der folgenden Beschreibung veranschaulichen:

Dionysos auf einem nach rechts laufenden Panther nach rechts sitzend, langen Thyrsos in der Linken haltend und den rechten Arm auf den Panther gestützt.<sup>167</sup>

Zentral ist der Begriff *Thyrsos*, der den Stab des Dionysos bezeichnet.<sup>168</sup> *Thyrsos* ist damit durch die Logik der Begriffe *Stab* und *Dionysos* bereits hinreichend beschrieben. Der Panther ist allgemein ein zentrales Attribut des Dionysos und tritt nur mit ihm auf Münzen auf. Das bedeutet, dass der Begriff *Panther* durch das allgemeinere, redundant in ihm implizierte Konzept *Großkatze* ersetzt werden kann. Aus der Kombination von *Dionysos* und *Großkatze* kann *Panther* logisch abgeleitet werden. Die vorliegende Beschreibung lässt sich dementsprechend durch die Konzepte *Stab*, *Bacchus* (lat. für Dionysos) und *Großkatze* komplett beschreiben. Für das Modelltraining werden in der vorliegenden Systematik zusätzlich die redundanten Konzepte *Mann* und *Gottheit* mit der Beschreibung verbunden. Die so generierten Konzepte können leichter mit anderen bildlichen Darstellungen verglichen werden und unnötige Komplexität kann entfernt werden. Dieses Vorgehen entspricht den Anforderungen an eine KI-gestützte

---

<sup>165</sup> TOLLE/WIGG-WOLF, Improving Data Quality by Rules: A Numismatic Example, S. 196.

<sup>166</sup> PAVLEK u.a., Standards and Quantification of Coin Iconography, S. 209.

<sup>167</sup> Vgl. CN Coin 343, online: <https://www.corpus-nummorum.eu/coins/343>.

<sup>168</sup> Der Thyrsus, also der Stab des Dionysos ist sein zentrales Attribut und er trägt ihn auf allen römischen Münzen, außer bei Prägungen, wo er als Kind abgebildet ist. Auf diesen Münzen, wo Dionysos als Kind abgebildet ist, wird er allerdings immer mit einer anderen Person dargestellt, sodass aus deren Kombination abgeleitet werden kann, dass kein Thyrsos abgebildet ist. Vgl. etwa RPC III 2408A, online: <https://rpc.ashmus.ox.ac.uk/coins/3/2408A>.

Auswertung, da es mit wenigen, präzise definierten, aber ausreichend breit gefassten und inklusiven Kategorien arbeitet.

Die Anreicherung der Daten auf Basis von Redundanz lässt sich im Knowledge Graph leichter umsetzen als auf Grundlage von Logik, da für letztere ein genaues Verständnis der Quellen notwendig ist. In der vorliegenden Arbeit werden für die Data Augmentation verschiedene Vorgehensweisen für Tiere, Personen und Objekte vorgeschlagen.

Bei Tieren ergibt es Sinn, die Systematik der Biologie zu übernehmen, die mit den Normdaten schon teilweise verknüpft ist.<sup>169</sup> Ein Beispiel sind die Münzen, die Philippus Arabs anlässlich seiner Säkularfeier zum 1000-jährigen Bestehen Roms prägte. Die hier geprägten Münzen zeigen auf dem Revers Tierdarstellungen: Löwen, Wölfe, Gazellen, Antilopen, Nilpferde. Da für die Feier auch exotische Tiere nach Rom gebracht wurden, die den Stempelschneidern wohl unbekannt waren, sind die Darstellungen nicht eindeutig den Tieren zuzuordnen. Bei zwei Darstellungen lässt der RIC sogar explizit zwei Deutungen zu: einerseits „Stag (elk?), walking right.“<sup>170</sup>, andererseits „Stag (elk?), walking left.“<sup>171</sup> Hier lohnt sich eine genauere Betrachtung der drei Münzen diesen Typs, die im Münzkabinett Berlin liegen. Sie haben jeweils unterschiedliche Beschreibungen. Die eine hält die Deutung als Elch für falsch und schreibt „Hirsch nach r.“<sup>172</sup> Die zweite übernimmt die Vagheit aus dem RIC mit „Hirsch (Elch?) nach l.“<sup>173</sup>. Die dritte Münze zeigt dagegen eine Darstellung, die richtigerweise und ohne Vagheit mit „Antilope nach l[inks]“ beschrieben ist.<sup>174</sup> Damit ist sie allerdings mit dem falschen Münztyp verknüpft. Dieses Beispiel zeigt, dass eine unsichere Typologisierung im Englischen auch bei deutschen Beschreibungen zu unsicheren Forschungsdaten führt. Außerdem treten menschliche Fehler in Form von Verwechslungen dort häufiger, wo sich die Konzepte ähneln, wie es bei Hirschen, Elchen und Antilopen der Fall ist.<sup>175</sup>

---

<sup>169</sup> Auch Wikidata und das Deutsche Archäologische Institut verwenden diesen Ansatz.

<sup>170</sup> Vgl. RIC IV Philip I 19, online: [https://numismatics.org/ocre/id/ric.4.ph\\_i.19](https://numismatics.org/ocre/id/ric.4.ph_i.19).

<sup>171</sup> Vgl. RIC IV Philip I 20, online: [https://numismatics.org/ocre/id/ric.4.ph\\_i.20](https://numismatics.org/ocre/id/ric.4.ph_i.20).

<sup>172</sup> RIC I<sup>2</sup> Nero 581, Münzkabinett Berlin, Inv.-Nr.: 18209968, online: <https://ikmk.smb.museum/object?id=18209968>.

<sup>173</sup> RIC VI Philip I 20, Münzkabinett Berlin, Inv.-Nr.: 18209969, online: <https://ikmk.smb.museum/object?id=18209969>.

<sup>174</sup> RIC VI Philip I 20, Münzkabinett Berlin, Inv.-Nr.: 18203965, online: <https://ikmk.smb.museum/object?id=18203965>.

<sup>175</sup> In diesem Beispiel könnte die unklare Benennung dadurch umschifft werden, dass die Konzepte mit Fragezeichen innerhalb der Klammern gelöscht werden. Sobald aber Unsicherheiten wie bei „Goat or Hind walking, left.“ (vgl. RIC IV Philip I 23) mit einem „oder“ getrennt werden, ist dieser Ansatz nutzlos, ganz abgesehen davon, dass durch die Entfernung aller Elemente in runden Klammern sicher auch relevante Informationen verloren gingen.

Diese Fehler können durch den Knowledge Graph gelöst werden, indem die nächsthöhere, implizierte Ebene betrachtet wird. Diese ist im Beispiel von *Hirschen* und *Elchen* die Familie der *Stirnwaffenträger*. Während die genaue Klassifikation in dem oben genannten Beispiel unsicher ist, ist auf den ersten Blick zu erkennen, dass ein Tier mit zwei Hörnern zum biologischen Taxon der *Stirnwaffenträger* gehört, zu der dann auch Rehe, (Dorcas-)Gazellen, Antilopen und Oryx zählen. Die Systematik aus der Biologie hat allerdings auch Grenzen, wenn man sie auf die Münzdarstellungen anwendet, die das reale Tier unterschiedlich genau treffen. Im Kontext von antiken Münzen ergibt es zum Beispiel Sinn, Ziegen und Widder von den restlichen Stirnwaffenträgern abzugrenzen. Die Ziege wird nämlich mit Ziegenbart dargestellt<sup>176</sup> und der Widder hat ein spezifisches, rundes Horn.<sup>177</sup> Dieses Horn findet sich auf Münzen jedoch auch losgelöst von einer Widderdarstellung, oft als Kopfschmuck wie bei der Büste des Serapis Pantheos aus Alexandria.<sup>178</sup> Es muss sich also um ein separates Objekt handeln, das aber wiederum auch in der Darstellung eines Widders zu sehen und damit mit dem Tier verknüpft ist. Demgegenüber ergibt es jedoch keinen Sinn, innerhalb der *Stirnwaffenträger* die Untergruppe der Hornträger von den Hirschen, die ein zapfenförmiges Geweih haben, abzugrenzen, weil diese zu selten auf Münzen geprägt wurden, und damit für die Bilderkennung uninteressant sind. Für die Konzepte *Stirnwaffenträger*, *Ziegen* und *Widder* ergibt sich demnach der Knowledge Graph wie in folgender Abbildung.

Abbildung 10: Knowledge Graph zu Stirnwaffenträgern

Oberkonzept	Tier				Objekt	179
Unterkonzept	↳ Stirnwaffenträger	↳ Ziege	↳ Widder	→	↳ Widderhorn	
Damit verknüpft	↳ Reh ↳ Elch ↳ Hirsch [...]					

Redundante Informationen zu Personen können ähnlich wie die bei Tieren hinzugefügt werden. Zunächst können die Personen im Knowledge Graph mit den Konzepten *männlich* respektive *weiblich*, sowie bestimmten Gruppenbezeichnungen wie *Götter* oder *Kaiser* verbunden werden.

<sup>176</sup> Vgl. RIC IV Philip I 23.

<sup>177</sup> Vgl. RIC II 1<sup>2</sup> Vespasian 1082.

<sup>178</sup> Vgl. RPC III 5840/6, Münzkabinett Berlin, Inv.-Nr.: 18201159, online: <https://ikmk.smb.museum/object?id=18201159>.

<sup>179</sup> Eigene Darstellung.

Das ist für die meisten Personen selbsterklärend. Einige wenige Personifikationen sind allerdings nicht eindeutig weiblich oder männlich. So zeigen Münzen mit der Legende Oriens, die zuerst unter Hadrian auftauchen, auf dem Revers den Gott Sol.<sup>180</sup> Oriens trägt hier die Bedeutung der aufgehenden Sonne. Gallienus prägt demgegenüber Oriens als eine Frau mit Mauerkrone.<sup>181</sup> Manche Personifikationen ändern demnach über die Zeit ihre Attribute. Diese komplexen Zusammenhänge müssen mithilfe von logischer Data Augmentation modelliert werden, indem sie zunächst mit sowohl *männlich* als auch *weiblich* verknüpft werden. In einem anschließenden Schritt wird die Auswahl des Geschlechts auf Basis der dargestellten Attribute getroffen werden. Im Fall von Oriens ist die Strahlenkrone Sols ein eindeutiges Zeichen für *männlich*, während die Mauerkrone ein Zeichen für *weiblich* ist.

Für diese komplexere logische Data Augmentation bedient sich die vorliegende Arbeit eines anderen Teilbereichs der NLP, der Relation Extraction (RE), wie sie von Deligio und Gencer vorgestellt.<sup>182</sup> Nach dieser Systematik funktionierte auch die Systematik des mittlerweile offline genommenen Ikonografie-Portals DIANA.<sup>183</sup> Allerdings werden diese Methoden aufgrund ihrer Ungenauigkeit für die hier vorgenommene Arbeit mit dem Knowledge Graph angepasst. Die vorliegende Arbeit begreift die Ikonographie von Münzen als Ansammlung von Darstellungen unabhängig von ihrer Position. Die RE geht demgegenüber davon aus, dass die zentralen Elemente in Münzbeschreibungen Personen sind und die Tiere und Objekte sich mit einem bestimmten Prädikat auf Sie beziehen.<sup>184</sup> Aus dem bereits oben genannten Beispiel des Dionysos würde die RE demnach die folgenden Relationen extrahieren:

$$\text{Dionysos} \xrightarrow{\text{sitzt}} \text{Panther}, \text{Dionysos} \xrightarrow{\text{hält}} \text{Thyrus.}^{185}$$

Oft findet sich diese Logik auch in den Münzbeschreibungen, die zunächst die abgebildete Person, danach ihre Kleidung und anschließend ihre Gegenstände in der rechten und linken Hand nennen. Im Englischen wird dazu entweder das Gerundium nach dem Schema „Victory

<sup>180</sup> Vgl. RIC II 3<sup>2</sup> Hadrian 49–51, online: [{49-51}](https://numismatics.org/ocre/id/ric.2_3(2).hdn).

<sup>181</sup> Vgl. RIC V Gallienus 448, online: [https://numismatics.org/ocre/id/ric.5.gall\(1\).448](https://numismatics.org/ocre/id/ric.5.gall(1).448).

<sup>182</sup> DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

<sup>183</sup> CALTABIANO u.a., DIANA, S. 415.

<sup>184</sup> TOLLE/WIGG-WOLF, Improving Data Quality by Rules: A Numismatic Example, S. 74.

<sup>185</sup> DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, S. 56, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

standing front, holding shield“<sup>186</sup>, oder Partizipialkonstruktionen wie „laureate, draped and cuirassed male bust“<sup>187</sup> verwendet. Im Deutschen ist eine Schreibweise mit flektierten Aktivformen vorherrschend. Allerdings gibt es, wohl durch die Bedeutung der englischen Typenreferenz, auch einige Partizipialkonstruktionen. In den deutschen Objektbeschreibungen ist die Varianz hier größer, da es in keiner Datenbank eine einheitliche Beschreibungsrichtlinie gibt, wie sie OCRE / CRRO / RPC-online zugrunde liegt. Obwohl Münzbeschreibungen allgemein bereits der RE ähneln, sind die Beziehungen zwischen den Darstellungen häufig nicht eindeutig.<sup>188</sup> Wenn in einer Münzbeschreibung etwa nur *mit Schleier* steht, bleibt unklar, ob die Person ein verschleiertes Haupt hat oder den Schleier nur über dem Arm hält. Und selbst wenn die Prädikate eindeutig sind, unterliegen sie noch einer subjektiven Interpretation. Das betrifft auch ein Beispiel, das Deligio und Gencer zur Erklärung ihres RE-Modells heranziehen:<sup>189</sup> „Hygieia standing right, feeding serpent held in right arm [...]“. <sup>190</sup> Die Schlange ist demnach als Objekt und über das Prädikat *wird gefüttert* mit Hygieia verbunden. Auf einer anderen Münze der Hygieia (lat. Salus) fehlt diese Verbindung und die Schlange wird selbst zum Subjekt: „Salus stehend nach l[inks], r[echts] Schlange umwindet Altar, l. [...]“, obwohl die Darstellung auch auf ein Füttern hinweist.<sup>191</sup>

Diese Beispiele zeigen, dass die RE an der inneren Varianz des vorliegenden Datensatzes scheitert, weil die Relationen nicht immer eindeutig sind. Außerdem nimmt die RE keine, jedoch notwendige, Qualitätskontrolle vor. Diese ist mit dem Knowledge Graph möglich, indem das Wissen der NER global und nicht nur für jede Beschreibung separat gespeichert wird. Bei dem Beispiel des Dionysos kann zum Beispiel *Dionysos* bereits im Knowledge Graph über die Relation *hält* mit dem *Stab* verbunden werden, da die Konzepte immer über dieses Prädikat verbunden sind. Der Knowledge Graph bietet zudem die Möglichkeit, nur die für die KI-Erkennung relevanten Relationen zu speichern. Denn bei manchen Darstellungen ergibt die genauere und detailliertere Systematik der RE Sinn, bei anderen führt sie wahrscheinlich zu zusätzlichen

---

<sup>186</sup> RIC I<sup>2</sup> Augustus 49, online: [https://numismatics.org/ocre/id/ric.1\(2\).aug.49](https://numismatics.org/ocre/id/ric.1(2).aug.49).

<sup>187</sup> RIC II 3<sup>2</sup> Hadrian 990–992, online: [https://numismatics.org/ocre/id/ric.2\\_3\(2\).hdn.{990-992}](https://numismatics.org/ocre/id/ric.2_3(2).hdn.{990-992}).

<sup>188</sup> DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, S. 76, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

<sup>189</sup> DIES.: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, S. 4, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

<sup>190</sup> CN Coin 824, online: <https://www.corpus-nummorum.eu/coins/824>.

<sup>191</sup> RIC III Pius 636a, 140–144 n. Chr., Münzsammlung der Universität Frankfurt am Main, Inv.-Nr.: T0505, online: <https://www.mkfrankfurt.uni-frankfurt.de/object?id=ID2365>.



Problemen. Sinnvoll ist beispielsweise eine feingliedrigere Unterscheidung des Objekts „Lorbeerkranz“, je nachdem, ob er auf einer Kaiserbüste<sup>192</sup> oder in der Hand einer Person<sup>193</sup> abgebildet wird. Diese Darstellungen unterscheiden sich visuell stark voneinander. Würden aus dem Lorbeerkranz verschiedene Klassen je nach Darstellung gebildet, würde sich die innere Varianz der Klasse verringern und damit möglicherweise die Genauigkeit der KI verbessern. Es gibt allerdings Objekte, die immer gleich abgebildet werden. Die Cornucopia (dt. Füllhorn) hat beispielsweise unabhängig davon, von welcher Personifikation oder Gottheit sie gehalten wird, die gleiche bildliche Darstellung, was die RE jedoch nicht erkennt, sodass das Konzept *Cornucopia* in mehrere Kategorien aufgespalten wird. Der KI würde es dann schwerfallen, diese neuen Klassen voneinander zu unterscheiden.

Von besonderem Informationsgehalt sind demnach die Klassen, die immer zusammen auftreten. Beispiele hierfür sind die Göttin Vesta und die Personifikation Africa. Vesta ist grundsätzlich verschleiert<sup>194</sup> und Africa trägt grundsätzlich eine Elefantenhaube.<sup>195</sup> Der griechisch-römische Gott Serapis trägt immer einen Kalathos (Getreidekorb) als Zeichen der Fruchtbarkeit. In der Numismatik sind diese Attribute bekannt und werden häufig, wie im vorliegenden Beispiel, nicht erwähnt. In der deutschen Beschreibung „Vesta steht in der Vorderansicht“<sup>196</sup> fehlt der explizite Hinweis auf den Schleier, während er im Englischen „Vesta, veiled, draped, standing“<sup>197</sup> ausdrücklich genannt wird. Das Phänomen betrifft, wie in dem vorliegenden Beispiel, vor allem deutsche Beschreibungen, während im Englischen auch die ikonografisch gesicherten Attribute explizit genannt werden. Diese Unterschiede zwischen den Sprachen können dann in der NER korrigiert werden, wenn numismatisches Fachwissen umfassend im Knowledge Graph berücksichtigt und gespeichert wird.

Personen auf römischen Münzen treten grundsätzlich in drei Positionen auf: entweder sie stehen, dazu gehören auch diverse Formen von Ausfallschritten, die eine Dynamik andeuten, oder sie sitzen, oder sie liegen. Diese drei Varianten können nicht voneinander unterschieden werden, wenn nur die Personen an sich durch die NER erkannt werden. Mit einer RE, die auch die Verben im Text berücksichtigt, wäre das möglich, allerdings geben die in den Datenbanken

---

<sup>192</sup> Vgl. etwa RIC II 3<sup>2</sup> Hadrian 1, online: [https://numismatics.org/ocre/id/ric.2\\_3\(2\).hdn.1](https://numismatics.org/ocre/id/ric.2_3(2).hdn.1).

<sup>193</sup> Vgl. etwa RIC I<sup>2</sup> Augustus 254A, online: [https://numismatics.org/ocre/id/ric.1\(2\).aug.254A](https://numismatics.org/ocre/id/ric.1(2).aug.254A).

<sup>194</sup> STEVENSON, A dictionary of Roman Coins, Republican and Imperial, S. 854f.

<sup>195</sup> Ebd., S. 25f.

<sup>196</sup> Denar des Vespasian RIC II 1<sup>2</sup> Vespasian 360, 72–73 n. Chr., Münzkabinett Berlin, Inv.-Nr.: 18219454, online: <https://ikmk.smb.museum/object?id=18219454>.

<sup>197</sup> RIC II 1<sup>2</sup> Vespasian 360, online: [http://numismatics.org/ocre/id/ric.2\\_1\(2\).ves.360](http://numismatics.org/ocre/id/ric.2_1(2).ves.360).

vorhandenen Daten eine RE wie besprochen nicht her. Die unterschiedlichen Positionen der Personen implizieren allerdings auch Objekte, die nur sehr selten in Beschreibungen explizit genannt werden. Wenn eine Person beispielsweise *sitzt*, dann kann aus dem Verb “sitzen” auf die Präsenz eines Stuhls geschlossen werden. Ausgeschlossen sind in diesem Fall Personen, die auf dem Boden sitzen, denn diese werden im Deutschen mit dem Prädikat *lagernd*<sup>198</sup>, im Englischen mit *reclining*<sup>199</sup> verbunden. Aus Verben kann nicht nur auf Objekte, sondern auch auf Tiere geschlossen werden. „Nero reitet [...] nach r[echts]“<sup>200</sup> impliziert beispielsweise ein Pferd. Erst wenn Personen auf einem anderen Tier reitend dargestellt sind, wie etwa in „der kindliche Iupiter reitet auf einer Ziege nach r[echts]“<sup>201</sup>, wird das Tier eindeutig genannt. “Reiten/riding“ impliziert demnach genau dann die Präsenz eines Pferdes, wenn kein anderes Tier angegeben wird.

Zu den auf Münzen dargestellten Objekten gibt es noch keine Vorarbeiten zur Data Augmentation. Hier sind jedoch Formen der logischen Anreicherung denkbar. Ein Beispiel ist das *Labarum*, eine von Kaiser Konstantin eingeführte Standarte, die oben das Christogramm trägt.<sup>202</sup> In den Münzbeschreibungen findet sich neben dem Wort “*Labarum*” allerdings auch „Standarte mit Chi-Rho-Zeichen“.<sup>203</sup> Während die NER in diesem Beispiel *Standarte* und *Christogramm* als separate Konzepte erkennt, handelt es sich doch um deren Kombination in Form des *Labarums*. Die Definition des *Labarums* sollte demnach so erweitert werden, dass sie alle Schreibweisen umfasst, bei denen Chi-Rho gemeinsam mit einem Feldzeichen auftaucht.<sup>204</sup> Mit dieser Methode können auch andere Konzepte, die sich aus der Kombination von Konzepten ergeben, definiert werden.

### 3.1.3 Fehlerquellen und Analyseprobleme

Nach der logischen und redundanten Data Augmentation bei Tieren, Personen und Objekten kann statistisch untersucht werden, wie sich die Überschneidung der Beschreibungen im Datensatz erhöht hat. Hierzu lohnt zunächst ein Blick in den Datensatz der deutschen

---

<sup>198</sup> Vgl. CN Coin 17553, online: [https://www.corpus-nummorum.eu/CN\\_17553](https://www.corpus-nummorum.eu/CN_17553).

<sup>199</sup> STEVENSON, A dictionary of Roman Coins, Republican and Imperial, S. 25f.

<sup>200</sup> Vgl. RIC P Nero 581, um 67 n. Chr., Münzkabinett der Staatlichen Museen Berlin, Inv.-Nr.: 18220873, online: <https://ikmk.smb.museum/object?id=18220873>.

<sup>201</sup> Vgl. RIC V Valentinian II 3, 257–258 n. Chr., Münzkabinett der Staatlichen Museen Berlin, Inv.-Nr.: 18200704, online: <https://ikmk.smb.museum/object?id=18200704>.

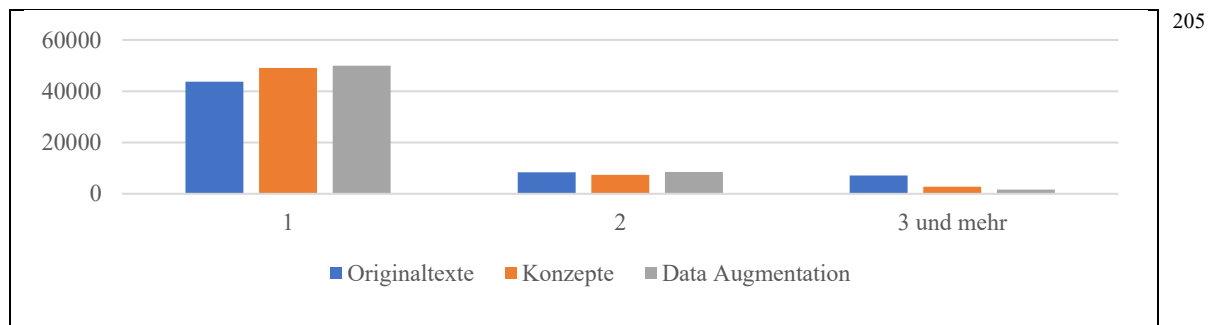
<sup>202</sup> Ebd., S. 500f.

<sup>203</sup> Vgl. RIC VIII Sirmium 29, 351 – 355 n. Chr., Bischöfliches Dom- und Diözesanmuseums Mainz, Inv.-Nr.: 1017, online: <https://numid.uni-mainz.de/object?id=ID94>.

<sup>204</sup> Weitere Schreibweisen für das *Labarum* sind „Feldzeichen mit Christusmonogramm“, „Standarte mit Chi-Rho“, oder im Englischen „standard with cross / Chi-Rho.“

Beschreibungen, der im Gegensatz zum Englischen mehrere Beschreibungen pro Münztyp aufweist. Und es gibt hier auch in den bereinigten Volltexten noch in erheblichem Umfang sich unterscheidende Beschreibungen des gleichen Typs. Jeweils etwa 8.000 Typen haben noch zwei beziehungsweise mehrere Münzbeschreibungen. Es gilt zu untersuchen, wie sich die Überschneidung durch die Konzeptextraktion beziehungsweise die Data Augmentation erhöht hat.

**Abbildung 11: Anzahl unterschiedlicher deutscher Beschreibungen pro Münztyp**



Bei knapp über 42.000 Beschreibungen liegt nur eine Beschreibung pro Münztyp vor, sei es, weil nur eine Münze des Typs heruntergeladen wurde, sei es, weil die Münzen eine übereinstimmende Beschreibung haben (Abb. 12). Mit der NER wurden aus diesen Beschreibungen die Konzepte extrahiert, sodass die Beschreibungen nur noch eine Kombination aus Konzepten sind. Dadurch erhöht sich die Zahl der übereinstimmenden Münztypen auf knapp unter 50.000. Mit der NER konnte auch die Zahl der Münztypen, die drei oder mehr Beschreibungen haben, erheblich reduziert werden. Das stützt die These, dass die Varianz innerhalb der deutschen Beschreibungen ursprünglich hoch ist. Schon die simple Abbildung „Blitzbündel auf Thron“<sup>206</sup> findet sich im Deutschen in elf Varianten, im Englischen sind es nur zwei.<sup>207</sup> All diese Beschreibungen sind nach der NER als Liste der Konzepte *Blitzbündel*, *Thron* vereinheitlicht.

Die Data Augmentation, das heißt die logische und redundante Anreicherung dieser Konzepte, kann die Zahl der Münztypen mit drei oder mehr Beschreibungen weiter verringern. Die verbleibende Zahl von etwa 800 Münztypen mit drei oder mehr unterschiedlichen Beschreibungen pro Münztyp erklärt sich hier nach stichprobenartiger Durchsicht vor allem durch falsche

<sup>205</sup> Eigene Darstellung.

<sup>206</sup> Vgl. RIC II Trajan 829, online: <https://numismatics.org/ocre/id/ric.2.tr.829>.

<sup>207</sup> Einige der Beschreibungen sind „geflügelter Blitz auf Thron liegend“, „geflügeltes Blitzbündel auf Thron“, „frontaler Thron darauf geflügeltes Blitzbündel“ oder schlicht „Thron mit Blitzbündel“.

Zuordnung und Sonderfälle in den Daten. Allerdings wurde auch die Zahl der Münztypen mit zwei unterschiedlichen Konzeptkombinationen nicht weiter verringert.

In den folgenden Fällen stehen Fehler im Vordergrund. Ein Beispiel hierfür sind zwei Münztypen, die Trajan Decius für Trajan prägen ließ.<sup>208</sup> Beide Münzen haben die gleiche Motivik und unterscheiden sich nur darin, dass die eine auf dem Avers eine Büste des Trajan mit Gewandresten, die andere einen Kopf des Trajan zeigt. Die Münztypen unterscheiden sich demnach in den Konzepten *Büste*, *Gewand* beziehungsweise *Kopf*. Zum ersten Typ liegt die Beschreibung „Kopf des Divus Trajan mit Strahlenkrone nach r., Gewandspuren l.“<sup>209</sup> vor. Die Kombination *Kopf*, *Gewand* ist allerdings laut RIC nicht möglich. Hier wurde also entweder der Kopf mit der Büste verwechselt oder die Gewandspuren wurden fälschlicherweise hinzugefügt. Zum gleichen Typ gibt es auch die Beschreibung „Büste des verstorbenen Trajan nach rechts mit Strahlenkrone“<sup>210</sup>. Hier fehlen in den Beschreibungen wohl die Gewandreste. Beide Beschreibungen sind vom RIC hier nicht vorgesehen. Ohne Betrachtung der Münzfotografien lässt sich auch nicht beantworten, ob die Münzen zu den richtigen Typen zugeordnet wurden oder zu der jeweils anderen Variante gehören. Die Varianz der Münzbeschreibungen bei diesem Typ kann ohne Überprüfung der Fotografien nicht vereinheitlicht werden.

So zeigt sich auch im Vergleich mit den englischen Beschreibungen, dass die Unterschiede zwischen deutschen und englischen Beschreibungen nicht vollständig ausgeglichen werden können. Die Extraktion der Konzepte hat in drei Vierteln (das heißt bei knapp 30.000, Abb. 13) der Münztypen dazu geführt, dass jeder dieser Münztypen genau eine spezifische Beschreibung hat. Durch Data Augmentation erhöht sich dieser Wert leicht auf rund 32.000.

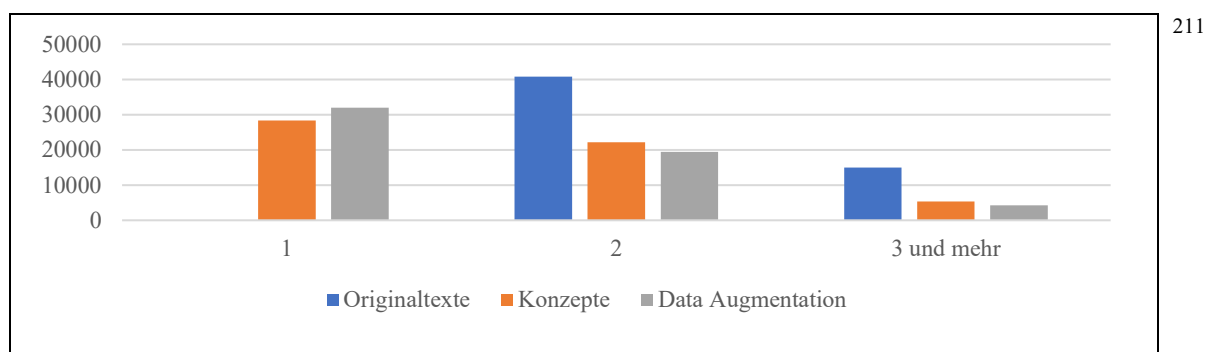
---

<sup>208</sup> RIC IV Trajan Decius 86A / 86B, online: [https://numismatics.org/ocre/id/ric.4.tr\\_d.86](https://numismatics.org/ocre/id/ric.4.tr_d.86).

<sup>209</sup> Doppeldenar des Traianus Decius, 249-251 n. Chr, Münzkabinett der Staatlichen Museen Berlin, 18273996, online: <https://ikmk.smb.museum/object?id=18273996>.

<sup>210</sup> Vgl. Doppeldenar des Traianus Decius, 249-251 n. Chr, Münzkabinett des Kunsthistorischen Museums Wien, RÖ 19196, online: <https://www.ikmk.at/object?id=ID71306>.

Abbildung 12: Anzahl unterschiedlicher Konzepte pro Münztyp im Deutschen und Englischen



211

Gleichzeitig bleibt aber ein erheblicher Anteil (über 20.000) von Münztypen mit zwei unterschiedlichen Konzeptkombinationen bestehen, sowohl nach der NER als auch nach der Data Augmentation. Der Grad der Vereinheitlichung ist erheblich niedriger als bei den ausschließlich deutschen Beschreibungen (Abb. 12). Eine reine NER samt Anreicherung reicht anscheinend nicht aus, um deutsche und englische Beschreibungen zu vereinheitlichen.

Um die Gründe zu verstehen, ist ein erkundender Blick auf die Daten erforderlich. Dazu kann das bereits genannte Beispiel der Münzen des Traianus Decius herangezogen werden, das nicht nur Fragen der falschen Zuordnung aufwirft. Die Beschreibungen unterscheiden sich dergestalt, dass die NER in einen Fall das Konzept *Kopf*, im anderen *Büste* extrahiert. Die Darstellungen wurden demnach unterschiedlich interpretiert, obwohl der ikonografische Unterschied in der Numismatik eigentlich genau geregelt ist. Eine Büste zeigt im Gegensatz zum Kopf noch einen Teil der Schulter.<sup>212</sup> Wie in diesem Beispiel wird bei den sich unterscheidenden Begriffen oft unmittelbar deutlich, ob es sich um einen Fehler der NER, oder der Beschreibungen handelt. Von den 20.000 Münztypen, deren Beschreibungen noch nicht vereinheitlicht werden konnten, sind in einem Drittel der Fälle die sich unterscheidenden Begriffe „Büste/Kopf“, „Gewand/kein Gewand“, „Kürass/keine Kürass“, „Kopfbinde/Lorbeerkranz“, „nackt/nicht nackt“.<sup>213</sup> Es zeigt sich nach stichprobenartiger Kontrolle, dass diese Unterschiede hauptsächlich auf Fehler der Beschreibungen zurückzuführen sind.

Diese Fehler sind nicht nur auf die Erfassenden zurückzuführen, sondern haben auch inhaltliche Gründe. Manchmal inkludiert etwa eine Typologie mehrere Darstellungen. Solche Fälle werden im RIC / RPC / RRC nicht einheitlich gehandhabt. Ein Beispiel aus dem RIC sieht so aus: „Bust

<sup>211</sup> Eigene Darstellung.


<sup>212</sup> STEVENSON, A dictionary of Roman Coins, Republican and Imperial, S. 147.

<sup>213</sup> Auffallend ist die Anzahl der Fehler bei den Kategorien Gewand und Kürass, obwohl die Personen, die immer Gewand oder Kürass tragen, bereits im Knowledge Graph mit den Konzepten verbunden wurden.

of Trajan, laureate, right (sometimes draped).”<sup>214</sup> Ob der betreffende Avers demnach drapiert ist oder nicht, lässt sich aus dieser Beschreibung nicht ableiten. Dieses Problem ist quantitativ signifikant. 990 RIC-Typen enthalten das Wort „sometimes“, 290 das Wort „usually“. Über tausend Münztypen lassen sich demnach nicht abschließend vereinheitlichen, wenn die Typologie nicht zuvor präzisiert wird.

Andere Unterschiede ergeben sich aus dem Zweck der einzelnen Beschreibung. Münzbeschreibungen bewegen sich grundsätzlich im Spannungsfeld zwischen knapper, wissenschaftlicher Beschreibung und ausführlicheren, museumsdidaktischen Beschreibungen. Das betrifft eher die deutschen Objektbeschreibungen als die Typbeschreibungen, die sich primär an wissenschaftliches Publikum richten. Beim abgebildeten Denar des Augustus (Abb. 14) entspricht der Beschreibungstext in der Datenbank dem Objekttext, wie er in der zur Zeit geschlossenen Ausstellung im Alten Museum Berlin zu finden ist und richtet sich demnach auch an numismatische Laien.

#### Abbildung 13: Denar mit museumsdidaktischer Beschreibung

	<p>Im Schritt fahrendes Viergespann (Triumphalquadriga) nach r., die Seitenpaneele sind mit geometrischen Mustern dekoriert, das Frontpaneel mit einer Victoria. Das Viergespann (quadriga) wird gekrönt von vier Pferden im Galopp.</p> <p>Slow quadriga right, with four miniature horses.</p>
---	--

<sup>215</sup>

In diesem Fall kann aber die Extraktion der Konzepte von den expliziten und impliziten Beschreibungen profitieren. Es zeigt sich, dass die Begriffe „slow quadriga“ bei OCRE wenig intuitiv immer eine *Triumphalquadriga* beschreiben. Diese unterscheidet sich von klassischen Quadrigen durch den höheren Aufbau und erinnert an Wagen wie das *Carpentum*.<sup>216</sup> Es ergibt also Sinn, die *Triumphalquadriga* mit dem klassischen *Wagen* zu verknüpfen und nicht mit dem Konzept *Quadriga*. Das gilt insbesondere, weil eine *Quadriga* auf Münzen immer von

<sup>214</sup> Typbeschreibung: RIC II Trajan 99, online: <https://numismatics.org/ocre/id/ric.2.tr.99>.

<sup>215</sup> Bild und deutschsprachige Beschreibung: RIC I<sup>2</sup> Nr. 96, um 18 v. Chr., Münzkabinett Berlin, Inv.-Nr.: 18202562, online: <https://ikmk.smb.museum/object?id=18202562>, englischsprachige Beschreibung: RIC I<sup>2</sup> Nr. 96, online: [https://numismatics.org/ocre/id/ric.1\(2\).aug.96](https://numismatics.org/ocre/id/ric.1(2).aug.96).

<sup>216</sup> Vgl. der Leichenwagen Agrippinas auf dem Revers des Sesterz des Kaisers Caligula [https://numismatics.org/ocre/id/ric.1\(2\).gai.55](https://numismatics.org/ocre/id/ric.1(2).gai.55)

bestimmten Tieren gezogen wird, die durch logische Data Augmentation hinzugefügt werden können. *Triumphalquadrigen* werden dagegen auch ohne Tiere abgebildet.<sup>217</sup>

Wie ausschmückend die Beschreibungen sind, unterscheidet sich von Institution zu Institution und wohl auch je nach Geschmack des Bearbeitenden. Viele Unterschiede sind wohl eher zufällig, als dass sie wie im vorherigen Beispiel eindeutig auf einen musealen Kontext zurückzuführen sind. So gibt es zum Beispiel Münzbeschreibungen, die im Englischen wesentlich ausführlicher als im Deutschen gehalten sind. Der RPC beschreibt etwa eine Schiffdarstellung mit „galley left with steersman and nine rowers; curved stern above steersman, standard and vexillum; on prow, Triton, left, blowing horn.“<sup>218</sup> Beim Münzkabinett in Wien heißt es zu diesem Typ schlicht „Schiff nach links“.<sup>219</sup> Der Begriff „Schiff“ kann hier nicht weiter durch Data Augmentation angereichert werden, da nicht alle auf Münzen dargestellten Schiffe Ruder-schiffe sind, geschweige denn auf dem Schiffsbug den Gott Triton zeigen.

Manche Unterschiede zwischen expliziten und impliziten Beschreibungen führen auch dazu, dass die Beschreibungen keine Konzepte teilen. Die englische Beschreibung „Regalianus and Dryantilla, standing facing each other; between, altar“<sup>220</sup> hat beispielsweise keine gemeinsamen Konzepte mit der deutschen Beschreibung „Das Kaiserpaar, die Kaiserin l[inks], der Kaiser r[echts], im Handschlag (dextrarum iunctio)“. Hier werden die Kaisernamen nicht explizit benannt und der Altar ignoriert. Die englische Beschreibung wiederum macht explizit, dass sich das Paar die Hände im symbolischen Handschlag gibt. In Anbetracht dieser Unterschiede befindet sich die Übereinstimmung der Beschreibungen pro Münztyp von über 90 % wohl bereits im oberen Spektrum dessen, was die Datengrundlage zulässt. Zur weiteren Vereinheitlichung müssten die Beschreibungen manuell korrigiert werden.

Bei diesen Beispielen, bei denen die Beschreibungen des Deutschen und Englischen durch den Knowledge Graph nicht vereinheitlicht werden konnten, wurde den Münzen die englische Typbeschreibung zugeteilt, da sie standardisierter und wahrscheinlich häufiger korrekt ist als die Objektbeschreibungen. So konnten die letzten 18.000 Münztypen mit mehreren Beschreibungen vereinheitlicht werden. Dieses Vorgehen kann nicht auf das Portal CN angewendet werden, da es dort sowohl eine englische als auch eine deutsche Typbeschreibung gibt. Hier stellt sich

---

<sup>217</sup> Vgl. RIC I<sup>2</sup> Augustus 120, online: [https://numismatics.org/ocre/id/ric.1\(2\).aug.120](https://numismatics.org/ocre/id/ric.1(2).aug.120).

<sup>218</sup> Vgl. RPC III 1518, online: [rpc.ashmus.ox.ac.uk/coins/3/1518](http://rpc.ashmus.ox.ac.uk/coins/3/1518).

<sup>219</sup> Vgl. Medaillon des Hadrian, RPC III 1518, 129–138 n. Chr., Inv.-Nr.: GR 16146, <https://www.ikmk.at/object?id=ID103391>.

<sup>220</sup> Vgl. RIC V Regalianus I, online: <https://numismatics.org/ocre/id/ric.5.reg.1>.

die Frage, welche Einträge zutreffender sind. Die englischen scheinen in CN grundsätzlich eine Übersetzung aus dem Deutschen zu sein. Diese These ergibt sich daraus, dass manche Beschreibungen im Englischen sprachlich auffällig sind. Der „Ährenkranz“ stellt sich zum Beispiel im Englischen etwas unnatürlich klingenden als „grain ear wreath“ dar, was auf eine wörtliche Übersetzung hindeutet. In den Beschreibungen aus dem RRC/RIC/RPC findet man dagegen ausschließlich den korrekten Ausdruck „wreath of corn ears“.<sup>221</sup> Andere Unterschiede scheinen dagegen einen anderen Hintergrund zu haben. So steht dem deutschen „Doppelaxt von Lorbeerkrantz umrandet“ das englische „Double axe within olive wreath“ entgegen.<sup>222</sup> Tatsächlich scheinen in diesem Fall die länglichen Blätter des Kranzes eher einem Olivenkranz zu entsprechen. Es ist schwer vorstellbar, dass hier einem Übersetzungsprogramm oder einer menschlichen Übersetzung ein Fehler unterlaufen ist. Eine mögliche Erklärung ist dagegen, dass nach der Übersetzung aus dem Deutschen jede englische Beschreibung händisch geprüft wurde und Ungenauigkeiten im Englischen, nicht jedoch im Deutschen, berichtigt wurden. Demnach wären die englischen Beschreibungen neuer und es gibt Anlass, auch bei den CN-Daten im Zweifel die englischen Beschreibungen vorzuziehen.

### 3.2. Münzbildvorbereitung

Die informatische Vorverarbeitung der Münzbilder ist aus mehreren Gründen notwendig. Das CNN benötigt als Input ein Bild in Form eines fest definierten Vektors von 112 x 224 x 3 (Höhe, Breite und Farbdimension). Um dies zu erreichen, wurden die Münzbilder bereits beim Download in einheitlicher Auflösung gespeichert. Wie bei der Aufbereitung der Münzbeschreibungen für den Knowledge Graph geht es auch bei der visuellen Erfassung der Bildinhalte darum, die Varianz der Daten zu vereinheitlichen. Dazu gehört bei Münzfotografien die Position der Münze im Bild und die Größe des Hintergrunds. Münzbilder von GLAM-Institutionen wurden bereits im korrekten Winkel aufgenommen, sodass die KI keine rotierten Münzen erkennen muss, was die Komplexität des Modells reduziert und die Genauigkeit erhöht.

Die Position der Münze sollte demnach so vereinheitlicht werden, dass das Bild voll ausgefüllt ist, also kein Hintergrund sichtbar ist. Dadurch sind Darstellungen, die immer auf der gleichen Münzposition vorkommen, auch immer in der gleichen Region der Input-Bilder lokalisiert. Dadurch fällt es dem Modell leichter, die semantischen Zusammenhänge auf Basis der Position

---

<sup>221</sup> Vgl. Bronze aus Priapos, CN Coin 53631, 1. Jh. v. Chr., online: <https://www.corpus-nummorum.eu/coins/53631>.

<sup>222</sup> Vgl. CN Type 19380, online: <https://www.corpus-nummorum.eu/coins/56016>.



zu erkennen. Außerdem könnte ein Verzicht auf Vorverarbeitung der Bilder dazu führen, dass das sogenannte „Shortcut Learning“ eintritt. Dabei lernt das Modell nicht den realen Zusammenhang zwischen Input und Output, erkennt also nicht die Darstellungen auf Basis ihrer realen Position auf der Münze. Stattdessen fokussiert es sich auf Daten, die zufällig mit dem Output korrelieren und lediglich einfacher zu erkennen sind als andere. Bei anderen Modellen hat sich gezeigt, dass die Änderung von Hintergründen zu komplett anderen Vorhersagen des Modells führt. Das tritt zum Beispiel bei dem ImageNet-Datensatz auf, mit dem die CNNs wie VGG-19 vortrainiert werden. Hier wird zum Beispiel eine Graslandschaft als Schafsherde erkannt, da Gras im ImageNet-Datensatz ein guter Indikator für die Präsenz von Schafen ist.<sup>223</sup> Beim vorliegenden Quellenbestand kann dieses Problem auftreten, weil das Erscheinungsbild der Münzbilder nicht zufällig ist, sondern jeweils die Handschrift ihrer Fotografen und Institutionen trägt.<sup>224</sup> Das bedeutet, dass die Aufnahmeart auf unbekannte Weise mit bestimmten Darstellungen korrelieren könnte, wodurch Shortcut Learning auftreten kann.

Das betrifft nicht nur den Aufnahmewinkel, die Belichtung und den Hintergrund der dargestellten Münze, sondern auch den Abstand zwischen Münze und Bildrand. Dieser Umstand hat Loyal dazu veranlasst, für das Training ihres CNNs nur die mit gleicher Einstellung fotografierten Münzen der ANS zu verwenden.<sup>225</sup> Es kann nicht quantifiziert werden, welchen Einfluss diese unbereinigten Daten auf die Qualität der KI-Entscheidung haben. Wenn allerdings die Vorhersagequalität des Modells auf unbekannten Daten stark von der auf dem Trainingsdatensatz abweicht, könnte dafür Shortcut Learning die Ursache sein. Shortcut-Learning kann mit zwei Methoden minimiert werden. Eine Möglichkeit ist es, die KI auf Basis einer größeren, generalisierbaren Datengrundlage lernen zu lassen. Dazu würden Ränder oder Ausschnitte bei Münzbildern zufällig generiert. Das würde allerdings einen erhöhten Trainingsaufwand bedeuten, da mehr Daten untersucht würden.

---

<sup>223</sup> GEIRHOS, Robert u.a.: Shortcut Learning in Deep Neural Networks, in: Nature Machine Intelligence 2, 11 (2020), S. 665–673, hier S. 665.

<sup>224</sup> So wurden beispielsweise für das Projekt Daten aus dem Heidelberger Graffiti-Portal heruntergeladen. Dieses Datenbankprojekt als Resultat des SFBs „Materiale Textkulturen“ sammelt römische Goldmünzen, die ein Graffito tragen. Die Datenbank basiert zu einem großen Teil auf dem Trierer Goldmünzhort. Da dort, wie das Projekt gezeigt hat, vor allem Münzen des Nero ein Graffito aufweisen, enthält die Datenbank zu einem großen Anteil diese Münzen. Diese Münzen haben durch das in Trier verwendete Fotografiergerät einen spezifischen Schatten am Rand. Das Modell könnte daraus ableiten, dass dieser spezifische Schatten einen großen Einfluss darauf, ob auf der Münze Nero abgebildet ist.

<sup>225</sup> LOYAL, Antje: Maschinelles Lernen angewendet auf Bilder antiker Münzen, Frankfurt, 2018, S. 38, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit\\_Antje\\_Loyal.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit_Antje_Loyal.pdf) [25.08.2025].

In der vorliegenden Arbeit werden stattdessen die Bilder mit analytischen Methoden vereinheitlicht, da dieses Vorgehen zum einen vorhersagbar ist und zum anderen weniger Aufwand als die Nutzung von KI erfordert. Holt zeigt eine Vielzahl von Algorithmen auf, die bei der Vereinheitlichung von Bildern helfen könnten. Auf einer großen Varianz an Münzen erzielen diese allerdings noch keine ausreichenden Ergebnisse.<sup>226</sup> Um die Münzen von ihrem Hintergrund zu trennen, werden deshalb in dieser Arbeit zunächst die Münzbilder insgesamt angeglichen. Da Münzbilder Fotografien eines Reliefs sind, ist die Position der hellen und dunklen Stellen von der Belichtungsquelle abhängig. Zu wenig oder zu viel Belichtung führt dabei zwangsläufig zu Datenverlust.<sup>227</sup> Um dem entgegenzuwirken müssen also die lokalen Kontraste erhöht und so die Merkmale der Münze ausreichend sichtbar gemacht werden. Das geschieht durch Contrast Limited Adaptive Histogram Equalization (CLAHE). Mit diesem Algorithmus wird der Kontrast der Bildregionen erhöht und die Helligkeit im Bild insgesamt angeglichen. Danach können mittels Kantendetektion (Canny Edge Detector) die Ränder der einzelnen Bildelemente kalkuliert werden. Wenn man davon ausgeht, dass die längste zusammenhängende Kante die Münze darstellt, kann das Bild auf den Umriss der Münze zugeschnitten werden. Falls die Fläche außerhalb des Umrisses einfarbig ist, kann davon ausgegangen werden, dass es sich um den Hintergrund handelt und somit entfernt werden kann.

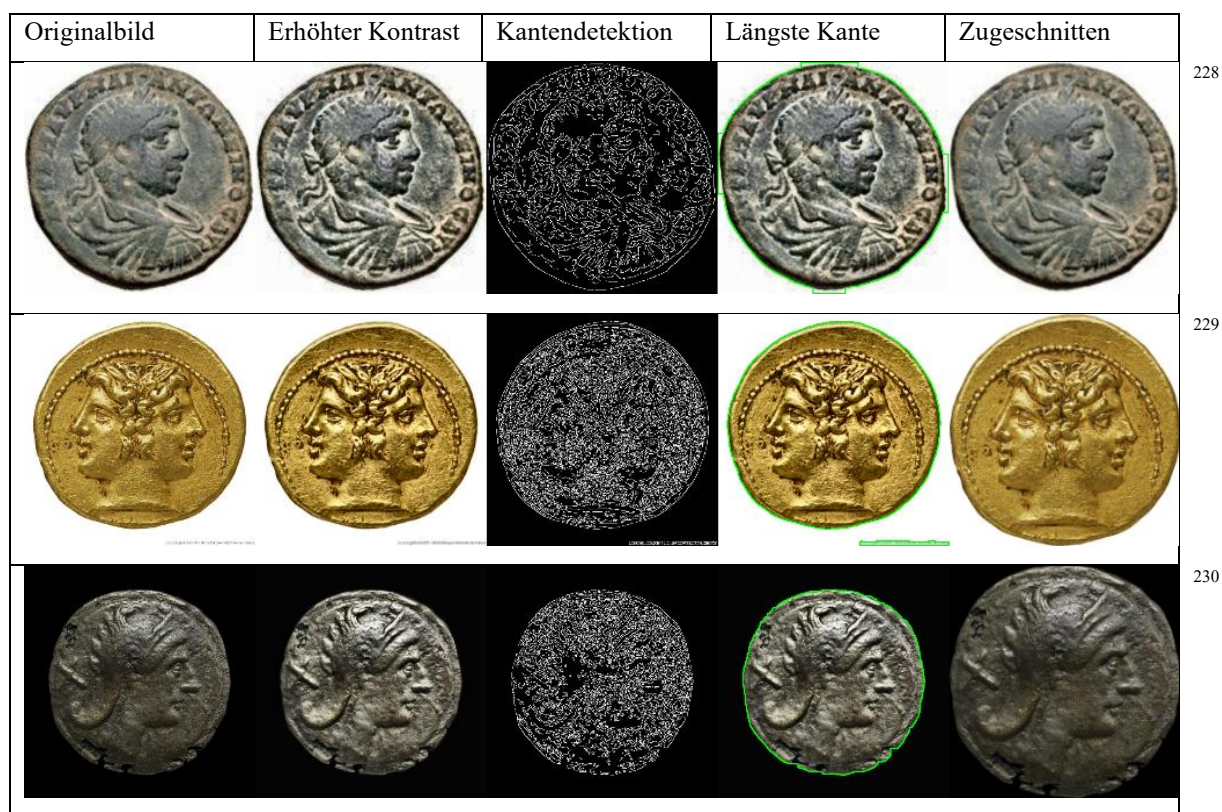
Die folgenden drei Bilder, eine Bronze-, eine Gold- und eine Silbermünze, zeigen, dass die verwendeten Algorithmen mit unterschiedlichen Materialien umgehen können. Der Kontrast lokaler Bildregionen konnte vereinheitlicht werden, wodurch die Kantendetektion den genauen Umriss der Münze erkennt. Die schwarzen und weißen Hintergründe können entfernt werden. Die Goldmünze zeigt zudem, wie eine institutionelle Besonderheit ausgeglichen werden konnte. In diesem Fall ist es die Bibliothèque Nationale de France, die den Permalink der Münzen in das Bild selbst setzt, sodass die Münze bei Bildern dieser Provenienz immer nach oben versetzt ist.

---

<sup>226</sup> HOLT, When Money Talks, S. 95.

<sup>227</sup> ZAMBANINI, Sebastian/KAMPEL, Martin/SCHLAPKE, Mario: On the Use of Computer Vision for Numismatic Research, in: The 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (2008), hier S. 2.

Abbildung 14: Vereinheitlichung von Münzbildern



Die genannten Beispiele sind allerdings dahingehend nicht repräsentativ, als dass sie einen einheitlichen Hintergrund haben. Dieser wird von modernen Kamerasystemen hinzugefügt, nachdem die Kamera selbst die Münze von ihrem tatsächlich fotografierten Hintergrund trennt. Viele Münzbilder enthalten allerdings den natürlichen Hintergrund, auf welchem sie fotografiert wurden. Diese Hintergründe sind aufgrund der unterschiedlichen Position der Lichtquelle, welche die Reflexionen und Schatten auf der Münze bestimmt, nicht einheitlich. Zudem finden sich im Datensatz die unterschiedlichsten Hintergrundfarben. Die folgenden Beispiele zeigen solche Münzen, bei denen die Entfernung des Hintergrunds notwendig, aber nicht trivial ist.

Abbildung 15: Münzen mit natürlichem Hintergrund

Originalbild	Erhöhter Kontrast	Kantendetektion	Längste Kante	Zugeschnitten
--------------	-------------------	-----------------	---------------	---------------

<sup>228</sup> Eigene Darstellung, Bildquelle: Avers eines Vierers aus Perinthos, CN Coin 14747, 218–222 n. Chr., online: <https://www.corpus-nummorum.eu/coins/14747>.

<sup>229</sup> Eigene Darstellung, Bildquelle: Avers eines republikanischen Hemistaters, RRC 28/2, 225–212 v. Chr., Bibliothèque Nationale de France, Inv.-Nr.: REP-21349, online: <https://gallica.bnf.fr/ark:/12148/btv1b10453404n>.

<sup>230</sup> Eigene Darstellung, Bildquelle: Avers eines republikanischen Decussis, RRC 41/1, 215–212 v. Chr., British Museum, Inv.-Nr.: 1877,0604.4, online: [https://www.britishmuseum.org/collection/object/C\\_1877-0604-4](https://www.britishmuseum.org/collection/object/C_1877-0604-4).



Auffallend bei dem ersten Bild ist, dass der grüne Hintergrund eine leichte, wie Leder wirkende Musterung aufweist, die durch CLAHE hervorgehoben wird. Bei der Detektion der Kanten und dem Zuschneiden der Bilder spielt das jedoch keine Rolle. Zu sehen ist außerdem, dass der Lichteinfall die untere Seite der Münze unterbeleuchtet erscheinen lässt. Dies konnte durch CLAHE ausgeglichen werden. Bei der zweiten Münze geht der Schatten der Münze auf dem Hintergrund in die Münze über, da diese die gleiche graue Farbe hat. Auch hier kann CLAHE genutzt werden, um zwar keinen kompletten Umriss zu erstellen, aber einen, der ausreicht, das Bild auf die Münze zuzuschneiden. Hier bietet es sich an, zwei Ergebnissoptionen des Vorverarbeitungsschrittes zu erlauben. Wenn die längste, äußere Kante eine ovale Form hat, handelt es sich wohl um die Münze und der Hintergrund kann entfernt werden. Wenn sie eine andere Form hat, würde die Entfernung des Hintergrundes auch Teile der Münze entfernen, sodass nur ein Zuschneiden vorgenommen werden kann. Hier bedarf es weiterer Forschung, die geeignete Methoden dafür findet, alle Hintergründe zu entfernen.

<sup>231</sup> Eigene Darstellung, Bildquelle: Avers eines republikanischen Denars, RRC 414/1, 63 v. Chr., Numismatische Sammlung der Universität Graz, Inv.-Nr.: 82, online: <http://gams.uni-graz.at/o:numis.909>.

<sup>232</sup> Eigene Darstellung, Bildquelle: Avers einer republikanischen Semuncia, RRC 42/5, 214–212 v. Chr., Rutgers University / Badian Collection, Inv.-Nr.: BAD0095, online: <http://dx.doi.org/doi:10.7282/T3TH8Q0F>.

<sup>233</sup> Eigene Darstellung, Bildquelle: Revers einer republikanischen Litra, RRC 13/2, 280–276 v. Chr., Numismatica Ars Classica Auktion 72, Los 396, online: <https://numismatics.org/sitnam/id/a8c06f6a>.

Die vorgestellten Methoden funktionieren nicht beim dritten Beispiel, das zur SITNAM-Collection der American Numismatic Society gehört.<sup>234</sup> Diese Münzbilder sind Scans alter Auktionskataloge und haben einen sehr uneinheitlichen Hintergrund und haben damit zunächst einmal eine schlechte Qualität. Diese Münzen wurden gerade aufgrund dieser Tatsache im Datensatz belassen, damit das Modell auch bei schlechteren Münzbildern funktioniert. Die SITNAM-Collection enthält 40.000 Münzen und hat damit einen signifikanten Anteil am vorliegenden Datensatz. All diese Münzen gehören in den Zeitraum der römischen Republik mit ihrer spezifischen Bildsprache, was Probleme des Shortcut-Learning verursacht. Nach dem Training der KI muss also quantitativ untersucht werden, ob sich die Qualität der Vorhersagen je nach Institution signifikant voneinander unterscheidet. Werden zum Beispiel Münzen kleinerer Institutionen, die Münzen etwas anders fotografieren als die größeren, tatsächlich genauso gut erkannt?

### 3.3. Das Convolutional Neural Network

In den vorhergehenden Kapiteln wurden die heruntergeladenen Daten für das Training des Modells vorbereitet. Dabei konnten aus den Münzbeschreibungen die relevanten Konzepte extrahiert und diese wiederum mit der Data Augmentation angereichert werden. Die Bilder konnten außerdem zugeschnitten und wenn möglich von ihrem Hintergrund getrennt werden. Mit diesen Daten kann im Folgenden das CNN trainiert werden.

#### 3.3.1 Training und Gütekriterien

Von den durch NER und Data Augmentation extrahierten 800 Konzepten kommen viele nur so selten im Datensatz vor, dass sie von der KI nicht erkannt würden. Für das Training der KI wurden deswegen nur die häufigsten 295 Klassen mit mehr als tausend Einträgen im Datensatz ausgewählt. Die seltenen Klassen wurden entfernt, da für sie weitere Vorverarbeitungsschritte notwendig geworden wären. Trotzdem ist der reduzierte Datensatz immer noch unausgeglich, das heißt, zu viele Darstellungen kommen zu selten im Datensatz vor. Ein *Mann* befindet sich zum Beispiel auf 80% der Münzen, *Silenus* als eine der vielen seltenen Personen kommt nur auf 0,08% der Münzen vor. Durch die Erstellung der vorliegenden KI als Multilabel-Modell werden die Klassen unabhängig voneinander betrachtet. Auf diese Weise wird die Genauigkeit der KI nicht durch die Verringerung der Klassen beeinträchtigt. Für künftige Forschungen müssen allerdings Methoden gefunden werden, die auch seltene Darstellungen durch KI erkennbar machen. Das geht allerdings über den Rahmen der vorliegenden Arbeit hinaus.

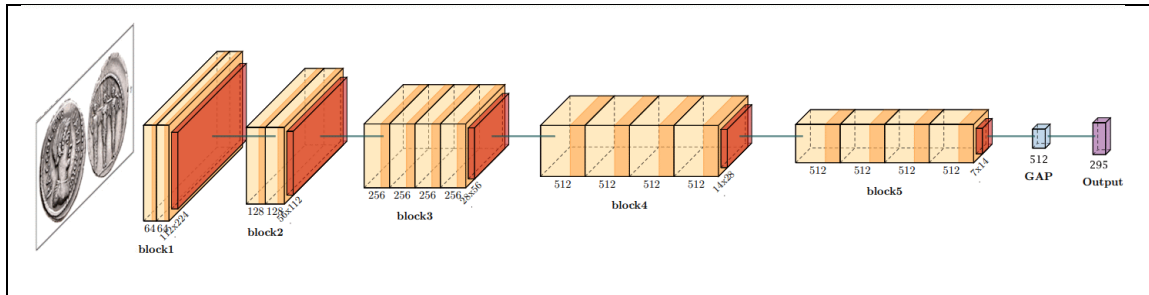
---

<sup>234</sup> Vgl. SITNAM: standalone numismatic database of Public Domain, Fair Use objects from private collections or archival auction catalogs published by the American Numismatic Society, online: <https://numismatics.org/sitnam>.



Das bereits vortrainierte VGG-19-Modell konnte mit den ausgewählten 295 Klassen und den Münzbildern weitertrainiert (Finetuning) werden. Die Architektur von VGG-19 wurde dafür in der vorliegenden Arbeit leicht angepasst (Abb. 15).

Abbildung 16: Funktionsweise des CNNs



235

Die wesentliche Anpassung betraf den Input: Anstelle eines quadratischen Bildes wurde ein zusammengesetztes Bild aus Avers und Revers der Münze verwendet. Daraus ergibt sich für die Bilder ein Vektor von (112x224x3: Höhe, Breite, Farbdimension). Die Farbdimension wurde als Indikator des Materials beibehalten, dessen Einfluss auf die Darstellungen vom Modell gelernt werden soll. Um kleine Abweichungen durch manuelle Bildaufnahme auszugleichen, wurde eine zufällige Rotation der Bilder von bis zu 5° integriert. Während dieses Bild VGG-19 durchläuft, erkennen die Convolutional Layers (beige) auf Basis lokaler Pixel Informationen wie Kanten und Farben. Mittels der Max Pooling Layer (rot) werden aus diesen Informationen diejenigen mit den höchsten Werten ausgewählt, sodass die Dimension reduziert wird, aber gleichzeitig die wichtigsten Features zu zunehmend größeren Bildregionen erhalten bleiben.<sup>236</sup> In einem letzten Schritt werden diese Informationen in ein Global Average Pooling (GAP) zur Reduktion der Komplexität auf 512 Features verringert, auf deren Basis das neuronale Netz die Entscheidung zu den ausgewählten 295 Klassen trifft.

Der Datensatz wurde für das Training im Verhältnis 60% Training, 20% Validierung und 20% Test aufgeteilt.<sup>237</sup> Es wurde ein automatischer Stopp des Trainings implementiert, wenn die auf dem Testdatensatz erzielten Genauigkeiten über mehrere Epochen unter denen des Trainingsdatensatzes liegen. Das wäre ein eindeutiges Zeichen für Overfitting durch Shortcut

<sup>235</sup> Eigene Darstellung.

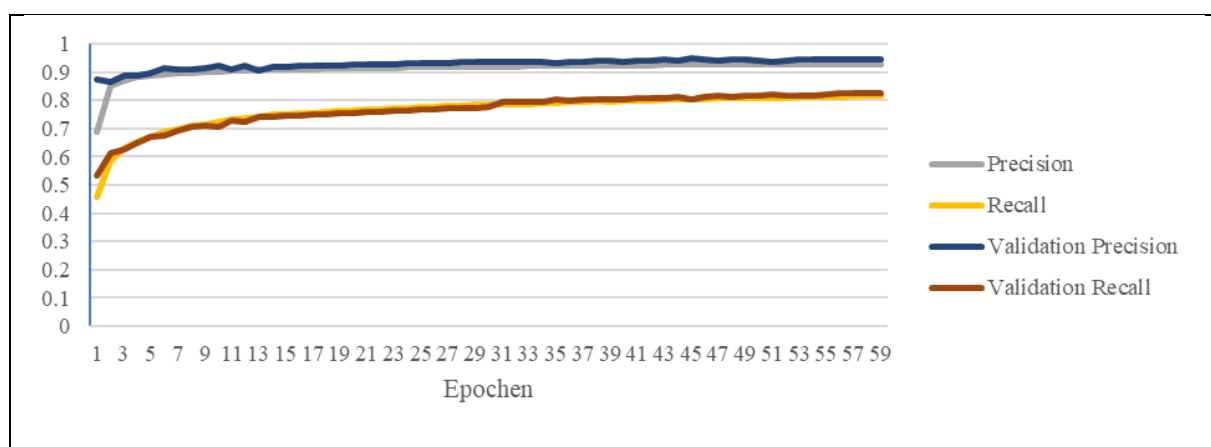
<sup>236</sup> SIMONYAN, Karen/ZISSERMAN, Andrew: Very Deep Convolutional Networks for Large-Scale Image Recognition, in: 3rd International Conference on Learning Representations, ICLR 2015, San Diego, 7-9 May 2015 (2015), hier S. 2.

<sup>237</sup> Das Modell wurde anhand des Trainingsdatensatz optimiert, während der Validierungssplit zur Hyperparametersteuerung und Early Stopping diente. Der Testsatz wurde ausschließlich für die Endbewertung des Modells genutzt.

Learning.<sup>238</sup> Außerdem wurde während des Trainings ein Dropout-Layer hinzugefügt, das die 512 Features zufällig um die Hälfte auf 256 reduziert. Dadurch kann sich das Modell im Training nicht auf die gleichen Features konzentrieren, wodurch ebenfalls Overfitting verringert wird. Trainiert wurden ausschließlich die letzten beiden Blöcke (Block 4/5), während die Gewichte der vorangehenden Layer aus dem auf ImageNet vortrainierten Modell übernommen wurden. Das Training wurde auf dem bwUniCluster 2.0 (Hochleistungsrechner Baden-Württemberg) durchgeführt. Eingesetzt wurde eine Partition mit vier Nvidia Tesla V100 GPUs.<sup>239</sup> Für die Implementierung eines CNN wurde die TensorFlow-Bibliothek für Python verwendet, da diese für Operationen mit Hochleistungsrechnern ausgelegt ist.<sup>240</sup> Durch die Reduktion der Komplexität des Modells dauerte das Training über sechzig Epochen etwa 24 Stunden (Abb. 18).

Das CNN erzielt trotz heterogener Datenquellen und unvollständig harmonisierter Münzbilder sowie -beschreibungen die bislang besten Gütemaße für Münzdaten. Nach 60 Trainingsepochen erreichte es eine Präzision (Precision) von 97,08 % (Trainingsdatensatz) bzw. 97,57 % (Validierungsdatensatz). Diese Werte zeigen, dass das Modell sehr zuverlässig in seinen Vorhersagen ist. Für eine vollständige Evaluation der Modellgüte bedarf es neben der Precision auch der Sensitivität des Modells (Recall). Precision beschreibt dabei den Anteil der korrekt als positiv erkannten Fälle (echte Positive) an allen vom Modell als positiv eingestuften Fällen (echte Positive + falsche Positive). Recall hingegen misst den Anteil der korrekt als positiv erkannten Fälle (echte Positive) an allen tatsächlich positiven Fällen (echte Positive + falsche Negative).

Abbildung 17: Trainingsprozess des Modells



241

<sup>238</sup> GEIRHOS u.a., Shortcut Learning in Deep Neural Networks, S. 670.

<sup>239</sup> Vgl. <https://wiki.bwhpc.de/e/BwUniCluster2.0>

<sup>240</sup> Vgl. <https://www.tensorflow.org>.

<sup>241</sup> Eigene Darstellung.

Die Kombination von Precision und Recall ist besonders bei unausgeglichene Datensätzen notwendig, da Modelle auf Datensätzen mit sehr seltenen Klassen durch systematische „Nicht-vorhanden“-Vorhersagen scheinbar gute Ergebnisse erzielen können. Diese Strategie würde auf dem vorliegenden Datensatz bei 295 Klassen und durchschnittlich 10 Klassen pro Münze eine Precision von 96%, allerdings einen Recall von 0% erreichen. Der Recall des vorliegenden Modells ist nach sechzig Epochen mit 81,6% und 82,5% geringer als die Precision. Auffällig ist, dass die Präzision von Beginn an sehr hoch ist und sich sowohl im Training als auch in der Validierung schnell stabilisiert. Der Recall hingegen startet deutlich niedriger und verbessert sich erst über die Epochen kontinuierlich. Dies deutet darauf hin, dass das Modell zunächst sehr konservativ vorgeht: Es sagt nur selten falsche Labels voraus (hohe Präzision), erkennt jedoch zu Beginn viele korrekte Labels nicht (niedriger Recall). Das Dropout trägt vermutlich dazu bei, dass Trainings- und Validierungskurven eng beieinanderliegen und das Modell kaum überanpasst, gleichzeitig aber die Verbesserung des Recalls nur langsam erfolgt. Auf geringes Overfitting weist auch hin, dass die Genauigkeit des Modells auf dem Validierungsdatensatz durchgehend höher als auf dem Trainingsdatensatz liegt. Das ist ungewöhnlich und liegt ebenfalls am Dropout-Layer. Es wurde nur im Training angewendet, sodass beim Validierungsdatensatz doppelt so viele visuelle Features für die Entscheidung zur Verfügung standen. Angesichts des Kurvenverlaufs ist eine weitere Steigerung des Recalls bei längerer Trainingsdauer möglich, das Training wurde jedoch nach sechzig Epochen beendet, um Ressourcen zu sparen und da die Ergebnisse bereits für die folgenden inhaltlichen Untersuchungen des Modells ausreichen.

Die KI erkennt trotz der hohen Genauigkeit pro Klasse nur in 37 % der Fälle alle Darstellungen auf der Münze vollständig. Das Hauptproblem für die Anwendung der KI liegt somit nicht in falschen, sondern in unvollständigen Vorhersagen. Korrekte, unvollständige Vorhersagen ermöglichen es jedoch, den Nutzenden Varianten auszugeben, die teilweise zutreffen. Dadurch bleibt die KI trotz der eingeschränkten Vollständigkeit besser nutzbar als bei falschen Vorhersagen. Deswegen wurde der Trainingsprozess so gestaltet, dass der Fokus auf Genauigkeit beibehalten wird.<sup>242</sup> Es soll untersucht werden, welche der Darstellungen nur unvollständig erkannt werden können. Dafür wurden Precision und Recall für jede Darstellung einzeln

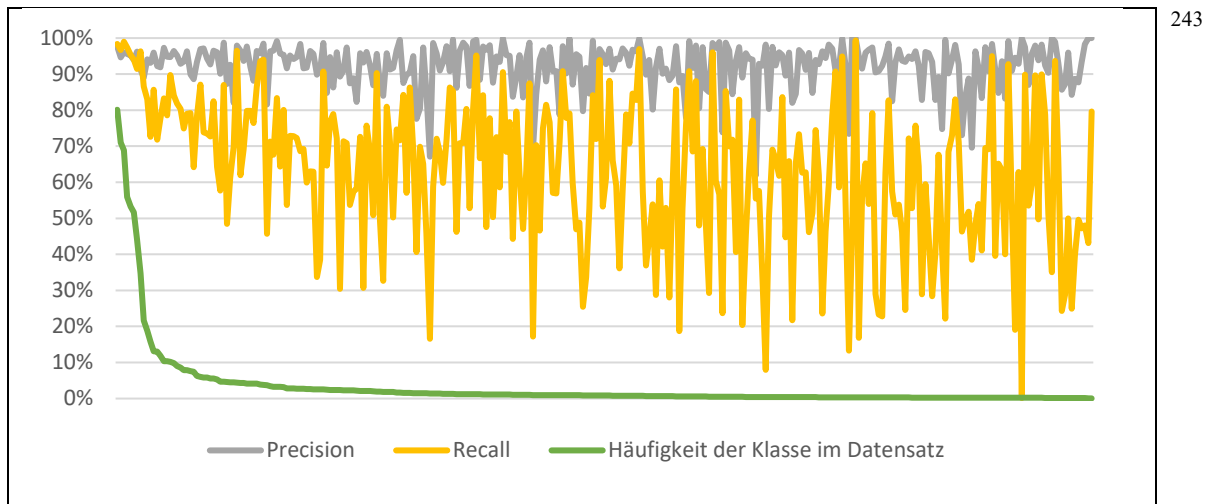
---

<sup>242</sup> Die Sigmoid-Ausgabe mit der Binary-Crossentropy-Loss-Funktion führt in hochunausgebalancierten Multi-Label-Daten dazu, dass False Positives stärker ins Gewicht fallen als False Negatives. Verstärkt wird dieser Effekt durch den eingesetzten Dropout (0,5), der dazu beiträgt, dass nur eindeutig erkennbare Muster positiv klassifiziert werden.



überprüft und angezeigt (Abb. 19). Es kann davon ausgegangen werden, dass das Modell durch Overfitting seltenere Klassen schlechter erkennt als häufige. Der Befund bestätigt dieses Bild.

Abbildung 18: Qualitätsmaße pro Klasse nach Häufigkeit im Datensatz



Es zeigt sich eindeutig, dass Precision und Recall abnehmen, je seltener die Darstellung im Datensatz auftaucht. Während die Precision bis auf wenige zu untersuchende Ausreiser nicht unter 90% fällt, ist die Varianz beim Recall sehr hoch. Im Mittel sinkt sie allerdings bei den selteneren Klassen auf etwa 60%. Hier gibt es zu viele Ausreiser, um sie einzeln zu untersuchen. Sie lassen sich aber wahrscheinlich dadurch erklären, dass beim Aufteilen in den Trainings-, Test- und Validationsdatensatz nur sehr wenige Datensätze im Trainingsdatensatz gelandet sind, die dadurch weniger präzise erkannt wurden.

Die häufigsten Klassen, die künstlich generierten Oberkategorien, haben eine nahezu perfekte Erkennungsrate. Hier ist der Recall (R) sogar höher als die Precision (P). Das heißt fast alle Kaiser werden erkannt (99,0%), mit einer Erfolgsquote von 96,0%. Ähnlich verhält es sich mit der Kategorie *männlich* (P: 98%, R: 98%) und *Büste* (P: 96%, R: 97%). Frauen (P: 94%, R: 94%) werden nur leicht schlechter erkannt als Männer. Das liegt möglicherweise daran, dass sie häufiger auf den schlechter erhaltenen Reversen abgebildet sind. Diese hohen Genauigkeiten wurden dementsprechend auch bei schlecht erhaltenen Münzen erzielt. Das zeigt, wie sinnvoll die Erweiterung der extrahierten Konzepte durch Oberkonzepte in der Data Augmentation ist. Bei unbekannten Personen oder sehr schlecht erhaltenen Münzen kann die Erkennung von diesen groben Kategorien die weitere Klassifikation verlässlich eingrenzen.

<sup>243</sup> Eigene Darstellung.

Die Darstellungen mit der geringsten Precision sind *Konstantin II* (P: 62 %, R: 55%), *Jugendlicher* (P: 67 %, R: 17%) und *Valerian* (P: 69 %, R: 17 %). Allein aufgrund dieser Begriffe lässt sich noch kein Muster erkennen. Unter ihnen sind mit Konstantin II und Valerian zwei Kaiser, die mit 1.500 respektive 3.000 ihnen im Datensatz zugeordneten Münzen zu den seltenen Kategorien gehören. Allerdings ist der Recall bei *Valerian* sehr überraschend und erklärungsbedürftig. Beide Kaiser haben wiedererkennbare Gesichtszüge, die auch von der KI erkannt werden könnten. Sie gehören auch nicht zu der Gruppe von Kaisern, die sehr ähnlich dargestellt werden, wie es etwa bei Valentinian und Valens der Fall ist.<sup>244</sup> Es bleibt unklar, warum die KI bei *Konstantin II* und *Valerian* so schlecht abschneidet. Ebenfalls einen sehr niedrigen Recall hat die Klasse *Jugendlicher*, obwohl sie über 5.000 Mal im Datensatz vorkommt. Das liegt wahrscheinlich daran, dass sich jugendliche Darstellungen nicht besonders von Erwachsenen unterscheiden, bei Männern häufig nur durch den fehlenden Bart. Bei dem *Jugendlichen* handelt es sich außerdem um eine Beschreibung, die von den unterschiedlichen Institutionen sehr unterschiedlich gehandhabt wird und von vielen überhaupt nicht verwendet wird.

Da die Ergebnisse der KI nicht nachvollziehend erklärbar sind, kann für jede einzelne Darstellung nur gemutmaßt werden, weshalb sie schlechter oder besser erkannt wurde. Fasst man hingegen die einzelnen Klassen zu Gruppen zusammen, eröffnet sich nunmehr die Möglichkeit, deren durchschnittliche Genauigkeit zu vergleichen. Zunächst wird geprüft, ob Personen oder Tiere und Objekte besser erkannt werden. Zweitens stellt sich die Frage, welcher Teilbereich der Numismatik, ob provinzialrömische, republikanische oder kaiserliche Prägungen, die höchste Genauigkeit erzielt. Daraus leitet sich drittens ab, ob die Genauigkeit der Münzerkennung von der katalogisierenden Institution abhängt. Dabei kann untersucht werden, ob die Unterschiede aus der Varianz der Münzfotografie herrühren oder ob es eher ausschlaggebend ist, ob die Institution gewöhnliche Münzen oder seltene Darstellungen besitzt. Alle drei Fragen konnten bisher in der automatisierten Münzklassifikation aufgrund fehlender Daten noch nicht beantwortet werden.

Bei der Art der Darstellung zeigt sich, dass Personen gegenüber Tieren und Objekten die schlechteste Precision und Recall haben (1 Prozentpunkt schlechter bei Precision, 4 Prozentpunkte schlechter bei Recall). Allerdings ist bei den Personen die Standardabweichung höher. Das heißt, es gibt eine Reihe von Personen wie die bereits genannten *Valerian* und *Konstantin*

---

<sup>244</sup> GAMPE/TOLLE, Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, S. 51.

II, die sehr schlecht erkannt werden und dadurch die durchschnittliche Precision und Recall der Personen stark verschlechtern. Werden diese herausgerechnet, sind Precision und Recall zwischen Tieren, Objekten und Personen sehr ähnlich. Dadurch kann die Vermutung, dass Personen auf Basis der Objekte und damit schlechter als Objekte erkannt werden, nicht bestätigt werden. Zur Überprüfung, ob das Modell seine Entscheidungen eher aufgrund der Gesichtszüge einer Person trifft, oder ob die Attribute der Person wesentlich für die Prognose sind, ist dementsprechend eine qualitative Überprüfung notwendig.

Die Ergebnisse für die Teilbereiche der Numismatik sind sehr deutlich. Es fällt sofort auf, dass die provinzialrömischen Münzen am schlechtesten erkannt wurden. Bei den Münzen von CN wird die geringste Genauigkeit (P: 87%, R 69%) erreicht, während die mit dem RPC-online verknüpften Münzen besser, aber trotzdem unterdurchschnittlich abschneiden (P: 92%, R: 75%). Hier bleibt zunächst unklar, ob provinzialrömische Münzen aufgrund ihrer speziellen Ikonographie oder ihres durchschnittlich schlechteren Erhaltungszustands und ihrer niedrigeren Fotografieauflösung seltener richtig erkannt werden. Die schlechtere Vorhersagequalität bei provinzialrömischen Münzen ist erklärungsbedürftig. Ausschließlich Provinzialrömischen Motive werden nämlich gut erkannt. Die *Preiskrone* wird zum Beispiel durchschnittlich erkannt, (P: 88%, R: 82%), während der Berg *Argaios* sogar zu den am besten erkannten Klassen gehört (P: 99%, R: 85%). Der Grund für das schlechte Abschneiden der Provinzialrömischen Münzen muss dementsprechend einen anderen Grund haben, der im Folgenden besprochen wird.

Die republikanischen Münzen, die die geringste Varianz der Motive haben, erzielen die höchsten Genauigkeiten. Die bereits zitierten Scans von Auktionskatalogen in geringer Auflösung der SITNAM-Collection erzielen die mit weitem Abstand höchste Genauigkeit mit einer Precision von 99,8% und einem Recall von 99,7%. Diese nahezu perfekte Erkennungsrate überrascht angesichts der Datengrundlage. Die Erkennung von Mustern scheint nur sehr wenig von schlechten Münzbildern beeinflusst zu werden. Stattdessen konnte das Modell dank der unzähligen Bilder der SITNAM-Collection, die nur zu wenigen Münztypen gehören, die Eigenheiten der republikanischen Prägungen sehr gut lernen. Ausschlaggebend muss also sein, dass die KI diejenigen Epochen besser erkennt, bei denen die Datengrundlage gut und die Varianz der Darstellungen niedrig ist.

Am besten erkennt das Modell Münzen der großen englischsprachigen Sammlungen. Dazu gehören die Bestände der American Numismatic Society (P: 97%, R: 92%), des Fitzwilliam Museum in Cambridge (P: 97%, R: 89%) und des Ashmolean Museum in Oxford (P: 96%, R:

88%). Alle diese Werte liegen weit über dem Durchschnitt. Dies verwundert nicht vor dem Hintergrund, dass die dort verwahrten Stücke gut erhalten und beschrieben sind sowie als einheitliche und hochauflösende Bilder vorliegen. Die größte deutsche Sammlung aus dem Münzkabinett Berlin liegt abgeschlagen (P: 94%, R: 82%) unmittelbar vor der Sammlung der Bibliothèque Nationale de France (P: 94%, R: 82%) im Mittelfeld. Diese ebenfalls großen Sammlungen bilden wie die genannten die Numismatik in gleichem Umfang wie die englischsprachigen ab. Außerdem entfernen beide Institutionen die Hintergründe von den Münzen. Die geringere Genauigkeit lässt sich nur durch unterschiedliche Sprachen und Forschungstraditionen ausgelöste Beschreibungsfehler erklären. Das zeigt, dass es von zentraler Bedeutung ist, falsche Beschreibungen zu korrigieren. Die am besten erkannte deutschsprachige Sammlung ist die der Universität Graz (P: 94%, R: 85%) und damit gerade die, die Münzen auf grünem Hintergrund fotografiert. Das deckt sich mit der Vermutung, dass die trotz Entfernung und Zuschneiden im Bild verbliebenen grünen Pixel dazu geführt haben, dass das Modell die spezifischen Eigenheiten der Grazer Sammlung durch Shortcut Learning lernt. Dadurch ist das Modell schlechter generalisierbar. Das heißt, wenn das Modell auf einem anderen Münzdatensatz mit grünem Hintergrund angewendet wird, das Münzen aus anderen Jahrhunderten als die Grazer Sammlung aufweist, werden die Ergebnisse vermutlich schlechter ausfallen.

Um annäherungsweise zu verstehen, wie die KI ihre Entscheidungen trifft, kann überprüft werden, welche Darstellungen gemeinsam auftauchen (Kovarianz). Dabei kann die Kovarianz der Ursprungsdaten mit der der KI-Prognose verglichen werden. Die Stärke dieses Zusammenhangs ist quantifizierbar. Als Beispiel können die Konzepte *Libertas* und *Pileus* (Filzkappe) betrachtet werden. Der *Pileus* galt als Zeichen freigelassener Sklaven<sup>245</sup> und *Libertas* trägt ihn häufig in ihrer linken Hand.<sup>246</sup> Wenn alle Datensätze betrachtet werden, die entweder *Libertas* oder *Pileus* enthalten, dann kommen diese Darstellungen in 80% der Fälle zusammen vor. Dass der Wert nicht höher ist, ist nicht weiter verwunderlich, da der *Pileus* auch alleine abgebildet wurde, etwa in der Münze der Caesarmörder als Zeichen der wiedererlangten Freiheit.<sup>247</sup> Außerdem haben auch andere Gottheiten wie die Dioskuren als Attribut den *Pileus*.<sup>248</sup> In den KI-Vorhersagen stimmen *Libertas* und *Pileus* sogar in 90% der Fälle überein. Wenn ein *Pileus* vorhanden ist, ist im Originaldatensatz nur zu 76% auch *Pietas* abgebildet, im Modell sind es

---

<sup>245</sup> STEVENSON, A dictionary of Roman Coins, Republican and Imperial, S. 629.

<sup>246</sup> Ebd., S. 518.

<sup>247</sup> Vgl. Denar des Brutus, RRC 508/3, 43 - 42 v. Chr., Münzkabinett Berlin, Obj.-Nr.: 18202198, online: <https://ikmk.smb.museum/object?id=18202198>.

<sup>248</sup> Ebd., S. 331.

88%. Wenn Libertas vorkommt, hält sie zu 86% eine Filzkappe, im Modell sind es 92%. Die Daten der KI zeigen, dass das Modell weit öfter diese Zusammenhänge in den Daten herstellt, als es sie in der Realität gibt. Das bedeutet, dass zu sehr Overfitting auf häufigen Darstellungen betrieben wurde, während seltenere Kombinationen schlechter erkannt werden und als die häufigen fehlklassifiziert werden. Im vorliegenden Beispiel erkennt das Modell häufiger den Pileus gemeinsam mit Pietas. Im Originaldatensatz ist die Kombination der Darstellungen weit seltener. Möglicherweise wurden hier andere Personen im Umfeld des Pileus von der KI für Libertas gehalten. Dieser Befund kann auch positiv gedeutet werden. Es zeigt nämlich, dass das Modell semantische Zusammenhänge auf Münzen eindeutig erkennt und Personen aufgrund ihrer Attribute vorhersagt. Auch andere Attribute bestätigen diesen Zusammenhang. Aclepius kommt nur etwa in 93% der Fälle mit seinem Schlangentab vor, die KI verbindet sie zu 97%.

Dennoch stellt sich die Frage, welche Auswirkungen diese erhöhten Zusammenhänge haben. Hier zeigt sich, dass vor allem provinzialrömische Münzen betroffen sind. Bei einer Münze aus Pergamon ist beispielsweise der eponyme Heros Pergamos auf dem Avers abgebildet, während den Revers ein Schlangentab schmückt.<sup>249</sup> Die KI erkennt auf dieser Münze Asclepius. Hier denkt die KI also Pergamos sei Asklepios. Das tritt bei provinzialrömischen Münzen häufiger auf, da Pergamos wie viele andere lokale mythische Figuren zu selten im Datensatz auftaucht, sodass er nicht für das Training verwendet wurde. Hier ist wohl eine Erkennung der Legenden notwendig, um noch weitere Kontextinformationen liefern. Pergamos zeichnet sich nicht durch besondere Attribute aus, sondern kann über die Münzlegende „ΠΕΡΓΑΜΟC“<sup>250</sup> identifiziert werden. Das zeigt, dass provinzialrömische Prägungen nicht per se schlechter erkannt werden. Sie weisen allerdings häufig Personen auf, die zu selten vorkommen und damit nicht für das Training verwendet wurden. Abgebildete seltene hält die KI dann für andere, häufigere Personen.

### 3.3.2 Qualitative Überprüfung

Mithilfe einer Saliency Map können die Vorhersagen des Modells auf Basis der ausschlaggebenden Regionen auf der Münzfotografie qualitativ überprüft werden. Das kann jedoch immer nur für eine Einzelmünze geschehen. Da aber jede Münze für sich ein Einzelfall ist, können wiederkehrende Muster nur im Vergleich verschiedener Saliency Maps miteinander verstanden werden. Dabei lässt sich zunächst an einem einfachen Beispiel, dem Vergleich von Büsten und

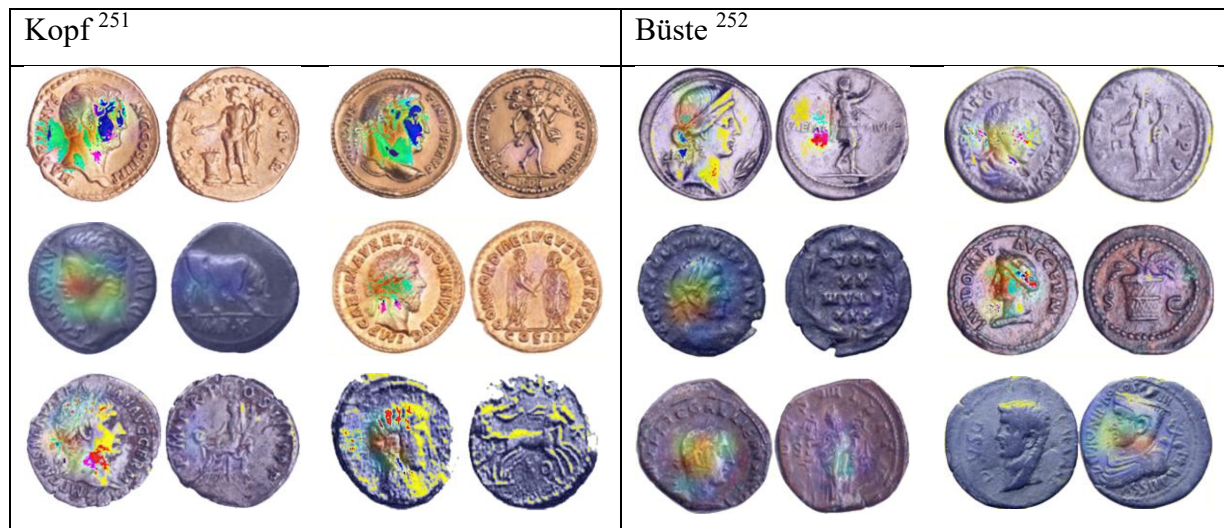
---

<sup>249</sup> RPC II 924, online: <https://rpc.ashmus.ox.ac.uk/coins/2/924>.

<sup>250</sup> STEVENSON, A dictionary of Roman Coins, Republican and Imperial, S. 331.

Köpfen, prüfen, ob die Saliency Map und das CNN verständlichen Output liefern (Abb. 20). Wenn dies der Fall ist, dann sollte eine Büste anhand der Halspartie erkannt werden, während der Kopf anhand wiederkehrender Merkmale wie Haare oder Nasen erkannt werden sollte.

Abbildung 19: Vorhersagen zu Büste und Kopf



Zunächst ist festzustellen, dass das Farbschema der Saliency Maps nicht auf allen Münzmateri-  
 alien in gleicher Weise nachvollziehbare Ergebnisse liefert. Dennoch lassen sich allgemeine  
 Beobachtungen anstellen. Sowohl *Büste* als auch *Kopf* werden auf der richtigen Seite der Münze  
 und innerhalb des Münzbildes, also nicht auf der Legende oder dem Hintergrund erkannt. Auf  
 der rechten unteren Münze wird die Büste außerdem korrekterweise auf dem Revers erkannt,  
 während auf dem Avers ein Kopf abgebildet ist. Es ist auffällig, dass die Erkennung des Kopfes  
 meistens über die Gesichtspartie läuft, wie zu erwarten war. Die Büste wird allerdings nicht  
 über die Schulterpartie erkannt, sondern häufig anhand ähnlicher Kopfregionen. Hier stellt sich  
 die Frage, wie die KI die beiden Kategorien mit so einer guten Genauigkeit auseinanderhalten  
 kann, wenn sie doch nicht eindeutig erkennt, dass sich eine Büste über die abgebildete Schulter

<sup>251</sup> Vgl oben links: American Numismatic Society, Inv.-Nr.: 1960.175.22, online: <http://numismatics.org/collection/1960.175.22>; Mitte links: Bibliothèque Nationale de France, Inv.-Nr.: IMP-5072, online: <https://gallica.bnf.fr/ark:/12148/btv1b10443959j>; unten links: American Numismatic Society, Inv.-Nr.: 1911.23.19, online: <http://numismatics.org/collection/1911.23.19>; oben rechts: American Numismatic Society, Inv.-Nr.: 1944.100.6008, online: <http://numismatics.org/collection/1944.100.6008>; Mitte rechts: American Numismatic Society, Inv.-Nr.: 1973.16.13, online: <http://numismatics.org/collection/1973.16.13>; unten rechts: Archäologisches Museum Arta, Inv.-Nr.: unbekannt, online: <https://rpc.ashmus.ox.ac.uk/coin/287004>.

<sup>252</sup> Vgl oben links: American Numismatic Society, Inv.-Nr.: 1937.158.438, online: <http://numismatics.org/collection/1937.158.438>; Mitte links: Münzkabinett Berlin, Inv.-Nr.: 18252027, online: <https://ikmk.smb.museum/object?id=18252027>; unten links: American Numismatic Society, Inv.-Nr.: 1911.23.19, online: <http://numismatics.org/collection/1911.23.19>; oben rechts: American Numismatic Society, Inv.-Nr.: 1944.100.30647, online: <http://numismatics.org/collection/1944.100.30647>; Mitte rechts: American Numismatic Society, Inv.-Nr.: 1944.100.54618, online: <http://numismatics.org/collection/1944.100.54618>; unten rechts: Azktion Roma Numismatics E-Live 6, 25 Mar. 2023, Los 86, online: <https://rpc.ashmus.ox.ac.uk/coin/450838>.

definiert. Auch bei anderen Klassen sind die Ergebnisse der Saliency Map wechselhaft, sodass im Vergleich mehrerer Darstellungen nicht eindeutig identifiziert werden kann, ob die KI die Klassen auch aufgrund der richtigen Bildregion erkennt.

Aussagekräftiger ist die Saliency Map bei der Fragestellung, wo Konzepte erkannt werden, die aufgrund des Erhaltungszustands nicht mehr sichtbar sind. Hier zeigen Beispiele, dass die KI dazu in der Lage ist. Bei der folgenden Münze werden eine *Blume* und *Spes* erkannt, obwohl der erhobene Arm der *Spes* nur angedeutet ist. Hier zeigt sich, dass *Blume* und *Spes* die exakt gleiche Saliency Map haben (Abb. 21). Das heißt, die Entscheidung für diese Labels fand über die Stellung der Person statt, die, wie besprochen, einzigartig für *Spes* ist. Die KI stößt bei dem zuvor genannten Beispiel der stark abgegriffenen Münze (Abb. 4) allerdings an ihre Grenzen und erkennt weder *Spes* noch *Blume*.

Abbildung 20: Saliency Map der *Spes* mit *Blume*



Die einzelnen Darstellungen werden dementsprechend in der ungefähr richtigen Bildregion erkannt, aber nicht genau dort, wo sie erwartet werden. Damit trifft das Modell seine Entscheidungen nie über die Legende, wie es vorherige CNNs gemacht haben. Das ist zum Beispiel beim Portraiterkennungs-CNN von Loyal der Fall, welches seine Entscheidungen zu mehr als 50% durch die Legende trifft.<sup>254</sup> Das lässt sich wohl durch den genutzten Datensatz erklären. Loyal strebt eine Kaisererkennung bei kaiserlichen Prägungen an. Bei diesen werden die Kaiser häufig in der lateinischen Legende genannt. Dies ist bei den provinzialrömischen und republikanischen Münzen des vorliegenden Datensatz nicht der Fall. Die griechischen Inschriften der provinzialrömischen Münzen haben wohl auch dazu geführt, dass das Modell keinen Fokus auf die Legende legt.

---

<sup>253</sup> Eigene Darstellung, Bildquelle: As des Domitian, RIC II-1<sup>2</sup> Domitian 1290, 77–78 n. Chr., Münzsammlung der Universität Heidelberg, online: <https://pecunia.zaw.uni-heidelberg.de/ikmk/object?id=ID1094>.

<sup>254</sup> Dieser Umstand ist erst Gampe bei seinen Nachprüfungen aufgefallen. Um eine Erkennung nur auf Basis der Bildelemente zu erreichen, trainiert er das Modell mit auf das Portrait reduzierte Münzen, bei denen die Legende fehlt, vgl. GAMPE, Sebastian: Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen, Frankfurt, 2021, S. 2, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf) [25.08.2025].



Bei der qualitativen Überprüfung kann auch überprüft werden, wie die KI mit gänzlich unbekannten Münzfotografien umgeht. In diesem Sinne werden die Avers- und Reversfotografien getauscht. Bei einer stichprobenartigen Kontrolle von Münzen zeigt sich allerdings, dass es keinen Unterschied macht, ob das Bild der KI in der richtigen oder falschen Anordnung gegeben wird. Die KI fokussiert ihre Entscheidungen auf Bildregionen, unabhängig davon, ob diese sich auf dem Avers oder Revers befinden. Die Kombination von unmöglichen Avers- und Reverskombinationen liefert allerdings falsche Ergebnisse.

Exemplarisch kann auch untersucht werden, welche Ergebnisse die KI bei Münzen anderer Zeiten liefert. Dabei zeigt sich, dass das Modell allgemeine Konzepte wie *Mann* und *Büste* sogar auf Geldscheinen erkennt (Abb. 22). Das spricht für eine hohe Robustheit der KI und ermöglicht Überlegungen zur Anwendung auf weiteren Teilbereichen der Numismatik.

Abbildung 21: KI-Vorhersagen auf Objekten anderer Epochen



## 4. Schluss

Die vorliegende Arbeit hatte zum Ziel, die Möglichkeiten und Grenzen einer automatisierten Klassifikation antiker Münzen aufzuzeigen. Dazu wurde eine KI mit den häufigsten Darstellungen auf Münzen trainiert. Es hat sich gezeigt, dass die Digitalisierung und Standardisierung von Münzdaten eine unverzichtbare Grundlage für jede Form automatisierter Klassifikation bildet. Deshalb wurden in erster Linie Daten von GLAM-Institutionen verwendet. Englischsprachige Daten wurden aus den Portalen OCRE / CRRO / RPC-online und CN heruntergeladen, während deutschsprachige Daten aus den Portalen IKMK und CN kamen. Auf diese Weise wurde ein Datensatz von über 500.000 Münzen erstellt.

Als Untersuchungsgegenstand wurde die Ikonographie der Münzen ausgewählt, da in ihr der größte Anteil der für die Klassifikation notwendigen Informationen enthalten ist. Dabei haben sich die Beschreibungsdaten als unstandardisiert und variationsreich herausgestellt. In der

<sup>255</sup> links: Medaille Reichsschatzsekretär Dr. Karl Helfferich, Münzkabinett Berlin, Inv.-Nr.: 18236103, online: <https://ikmk.smb.museum/object?id=18236103>; rechts: 100.000 kroatische Dinar 1993, Münzkabinett Berlin, Inv.-Nr.: 18246988, online: <https://ikmk.smb.museum/object?id=18246988>.



Numismatik, wo Beschreibungen oft durch jahrzehntelange Traditionen und institutionelle Eigenheiten geprägt sind, bleibt die Vereinheitlichung eine langfristige Aufgabe. Die durch die vorliegende Arbeit erkannten Fehler wurden den Münzsammlungen zur Verfügung gestellt und dadurch ein Beitrag dazu geleistet, diese Aufgabe zu beschleunigen.

Die Varianz der Münzbeschreibungen konnte durch den Einsatz eines Knowledge Graph zu großen Teilen aufgefangen werden. Dazu wurden zunächst die relevanten Darstellungen samt alternativer Schreibweisen im Knowledge Graph gespeichert. Als herausfordernd stellte sich hierbei heraus, dass einzelne Objekte und Personen je nach Kontext auf eine gänzlich andere Darstellung hinweisen können. Deshalb war dieser Arbeitsschritt mit erheblichem manuellen Aufwand verbunden, obwohl bereits bestehende Normdatensysteme als Ausgangspunkt des Knowledge Graphs genutzt werden konnten. Die Konzeptextraktion wurde mittels eines Vergleichs der Objekt- mit den Typbeschreibungen durchgeführt. Dazu wurden die Objektbeschreibungen unter Ausnutzung des Leidener Klammersystems mittels Regex so gut wie möglich in Typbeschreibungen umgewandelt. Dabei stellte sich heraus, dass Unterschiede zwischen den Sprachen noch eine zusätzliche Herausforderung für die Vereinheitlichung darstellen. Diese Unterschiede blieben bestehen, obwohl durch die Speicherung von impliziten Zusammenhängen in der Data Augmentation knappe und ausführliche Münzbeschreibungen vergleichbar gemacht wurden. Auch in den Prognosen des Modells hat sich der Befund bestätigt, dass die deutschen und englischen Beschreibungen manuell auf Fehler überprüft werden müssen. Die KI erzielt auf Münzen deutscher Institutionen grundsätzlich schlechtere Ergebnisse als auf denen von englischsprachigen.

Auf der bildanalytischen Ebene konnte durch die Vorbereitung der Münzbilder und das Training eines Convolutional Neural Networks eine beachtliche Erkennungsleistung erzielt werden. In den meisten Fällen konnten die Bilder von ihrem Hintergrund getrennt werden und in einheitlichem Format der KI zum Training gegeben werden. Hier gibt es allerdings für die künftige Forschung noch Möglichkeiten, die Algorithmen zu verbessern, um auch bei unebenen Hintergründen eine Erkennung der Münzumrisse zu gewährleisten.

Die Gütekriterien des Modells verdeutlichten, dass die vorgelegte KI die bisher zuverlässigsten Ergebnisse in der Erkennung von Darstellungen auf römischen Münzen erzielt. Dabei liegt der Fokus stärker auf der Precision als auf dem Recall. Unter allen Klassen wurden gerade die in der Data Augmentation hinzugefügten Oberkategorien am besten erkannt. Je seltener eine Darstellung auftaucht, desto schlechter wird sie erkannt. Die Quantitativen Analysen haben gezeigt,

dass zwar Personen, Objekte und Tiere ähnlich erkannt werden. Allerdings wurde Overfitting bei den Personen mit ihren Attributen beobachtet. Künftige KIs sollten häufiger seltene Kombinationen und Darstellungsweisen von Personen gegeben werden, um Overfitting zu verhindern. Außerdem zeigen sich Übersetzungsprobleme in den Daten in der durchweg schlechteren Erkennung der deutschen Münzinstitutionen.

Für die Qualität der KI-Vorhersage ist es weniger relevant als erwartet, welche Qualität das Münzbild hat. Allerdings lösen möglicherweise manche Hintergründe Shortcut Learning aus. Für die Qualität der Vorhersage ist ausschlaggebend, wie typisch eine Darstellung ist. Deswegen ist die provinzialrömische Numismatik der Teilbereich der am schlechtesten erkannt wird, während die republikanischen Prägungen überdurchschnittlich vorhergesagt werden. Hier bietet es sich an, für die provinzialrömischen Prägungen gezielt neue Oberklassen zu bilden, die häufiger und damit besser erkannt werden können. Dadurch kann der Output des Modells besser eingeordnet werden.

Für künftige Forschungen ergeben sich mehrere Möglichkeiten. Für eine vollautomatisierte Münzklassifikation ist die Verknüpfung der vorliegenden KI mit anderen Bereichen notwendig. Die Erkennung von bildlichen Darstellungen muss zumindest um eine Legendenerkennung, wahrscheinlich auch um eine Münzstättenenerkennung erweitert werden. In der Praxis sollte diesen wiederum eine KI vorangestellt werden, die erkennt, ob es sich um eine römische Münze handelt oder nicht. Dadurch können andere Münzen zum Beispiel aus der griechischen Numismatik aussortiert werden und an eine noch zu entwickelnde KI zur Erkennung griechischer Münzen übergeben werden. Das würde die Nutzbarkeit der KI erhöhen, wenn in einem zu klassifizierenden Bestand nicht nur römische Münzen vorhanden sind.

Die vorliegende KI kann aber auch direkter genutzt werden. Dadurch, dass sie die bisher besten Ergebnisse auf der gesamten römischen Numismatik erzielt, kann sie als Vergleichsmaßstab für künftige Forschungen verwendet werden. Dabei können die bereits auf die Numismatik abgestimmten Gewichte des Modells für künftige Modelle weiterverwendet werden. Das von Gampe auf den Portraits trainierte Modell erzielt bessere Ergebnisse auf dem OCRE-Datenset als das nur mit dem ImageNet trainierte Datenset.<sup>256</sup> Das vorliegende Modell könnte

---

<sup>256</sup> GAMPE/TOLLE, Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, S. 54.

beispielsweise auf einem Datensatz der griechischen Numismatik angewendet werden, die viele Motive mit der Provinzialrömischen Numismatik teilt.

Die Erkennung von Darstellungen kann methodisch weiterentwickelt werden, wenn der Knowledge-Graph weiter verfeinert wird. Dazu könnte in der numismatischen Forschungsliteratur nach weiteren Zusammenhängen zwischen Darstellungen gesucht werden. Es würde außerdem helfen, wenn Personen noch kleinteiliger erkannt würden, wenn zum Beispiel die Personen desselben Kaiserhauses gemeinsam zu einer Dynastie zusammengefasst würden. Das würde auch bei den vielen Kaisern der Reichskrise des dritten Jahrhunderts helfen, die häufig zu wenige Münzen für ein KI-Training prägten. Diese hätten dann ein gemeinsames verbindendes Konzept mit dem sie erkannt werden könnten.

## 5. Abkürzungsverzeichnis:

CN	Corpus Nummorum
GLAM	Galleries, Libraries, Museums, Archives
IKMK	Interaktive Kataloge der Münzkabinette
OCRE	Online Coins of the Roman Empire
CRRO	Coinage of the Roman Republic Online
RIC	Roman Imperial Coinage
RRC	Roman Republican Coinage
RPC	Roman Provincial Coinage
RPC-online	Roman Provincial Coinage Online
SITNAM	A standalone numismatic database of Public Domain, Fair Use objects from private collections or archival auction catalogs published by the American Numismatic Society

## 6. Literaturverzeichnis

### 6.1. Quellenverzeichnis

AMANDRY, Michel/BURNETT, Andrew M.: From Nerva to Hadrian. AD 96–138 (Roman Provincial Coinage, III), 2015.

BÖRNER, Susanne: Bericht über die numismatischen Aktivitäten im Jahr 2023. Tätigkeitsbericht für die Numismatische Kommission der Länder, 2024, online: [https://numismatische-kommission.de/fileadmin/user\\_upload/Nk/pdf/taetigkeitsberichte/2023/Boerner\\_Taetigkeitsbericht\\_fuer\\_23.pdf](https://numismatische-kommission.de/fileadmin/user_upload/Nk/pdf/taetigkeitsberichte/2023/Boerner_Taetigkeitsbericht_fuer_23.pdf) [12.08.2025].

BRUUN, Patrick: Constantine and Licinius A.D. 313 - 337. (The Roman Imperial Coinage, VII), London, 1966.

BURNETT, Andrew M./AMANDRY, Michel/CARRADICE, Ian: From Vespasian to Domitian. AD 69–96 (Roman Provincial Coinage, II), 1999.

BURNETT, Andrew M./AMANDRY, Michel/RIPOLLÈS, Pere.: From the death of Caesar to the death of Vitellius. 44 BC–AD 69 (Roman Provincial Coinage, I), 1992.

BUTCHER, Marguerite: De Gordien Ier à Gordien III. 238–244. Province d'Asie (Roman Provincial Coinage, VII.1), 2006.

CARRADICE, Ian/BUTTREY, Theodore: From AD 69 to 96 (The Roman Imperial Coinage, II, 3), 2007.

CARSON, Robert/SUTHERLAND, Carol Humphrey Vivian/KENT, John Philip Cozens: The Family of Constantine I (The Roman Imperial Coinage, VIII), London, 1981.

COHEN, Henry: Description Historique Des Monnaies Frappées Sous L'Empire Romain Communément Appelées Médailles Impériales, Paris, Londres, 1859.

CRAWFORD, Michael H.: Roman Republican Coinage, London, 1974.

HOSTEIN, Antony/MAIRAT, Jerome: From Trajan Decius to Uranius Antoninus. AD 249–254 (Roman Provincial Coinage, IX), 2016.

HOWARD, Emma/AMANDRY, Michel/ABDY, Richard A. (Hrsg.): From AD 117 to AD 138 - Hadrian (The Roman Imperial Coinage, II, 3), London, 2019.

HOWGEGO, Christopher: From Antoninus Pius to Commodus. AD 138–193: Egypt (Roman Provincial Coinage, IV.4), 2023.

KENT, John Philip Cozens: The Divided Empire and the Fall of the Western Parts, AD 395 - 491 (The Roman Imperial Coinage, X), London, 1994.

MAIRAT, Jerome/BUTCHER, Marguerite: From Gordian I to Gordian III. 238–244. All provinces except Asia (Roman Provincial Coinage, VII.2), 2022.

MATTINGLY, Harold/SYDENHAM, Edward Allen: Antoninus Pius to Commodus (The Roman Imperial Coinage, III), London, 1930.

MATTINGLY, Harold/SYDENHAM, Edward Allen/SUTHERLAND, Carol Humphrey Vivian: Pertinax to Geta (The Roman Imperial Coinage, IV, 1), London, 1936.

MATTINGLY, Harold/SYDENHAM, Edward Allen/SUTHERLAND, Carol Humphrey Vivian: Macrinus to Pupienus (The Roman Imperial Coinage, IV, 2), London, 1938.

MATTINGLY, Harold/SYDENHAM, Edward Allen/SUTHERLAND, Carol Humphrey Vivian: Gordian III to Uranius Antoninus (The Roman Imperial Coinage, IV, 3), London, 1949.

MATTINGLY, Harold/SYDENHAM, Edward Allen/WEBB, Percy: Valerian to Florian (The Roman Imperial Coinage, V, 1), London, 1927.

MATTINGLY, Harold/SYDENHAM, Edward Allen/WEBB, Percy: Probus to Amandus (The Roman Imperial Coinage, V, 2), London, 1933.

PEARCE, John William Ernest: Valentinian I to Theodosius I (The Roman Imperial Coinage, IX), London, 1933.

STEVENSON, Seth William: A dictionary of Roman Coins, Republican and Imperial, London, 1889.

SUTHERLAND, Carol Humphrey Vivian: From 39 BC to AD 69 (The Roman Imperial Coinage, I), London, 1923.

SUTHERLAND, Carol Humphrey Vivian/CARSON, Robert Andrew Glindinning: From Diocletian's Reform (A.D. 294) to the Death of Maximinus (A.D. 313) (The Roman Imperial Coinage, VI), London, 1967.

## 6.2. Literaturverzeichnis

ALFÖLDI, Andreas: Diana Nemorensis, in: *American Journal of Archaeology*, 2 (1960), S. 137–144.

ANWAR, Hafeez u.a.: Deep Ancient Roman Republican Coin Classification via Feature Fusion and Attention, 2019, online: <http://arxiv.org/pdf/1908.09428v2> [25.08.2025].

ANWAR, Hafeez/ZAMBANINI, Sebastian/KAMPEL, Martin: A Bag of Visual Words Approach for Symbols-Based Coarse-Grained Ancient Coin Classification, 2013, online: <http://arxiv.org/pdf/1304.6192v1> [25.08.2025].

ARANDJELOVIĆ, Ognjen: Automatic Attribution of Ancient Roman Imperial Coins, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (2010), S. 1728–1734.

CALTABIANO, Maria u.a.: DIANA. An Approach to Coin Iconography According to Time and Space Through Digital Maps, in: 2013 Digital Heritage International Congress (DigitalHeritage) (2013), S. 413–416.

COOPER, Jessica/ARANDJELOVIĆ, Ognjen: Understanding Ancient Coin Images, 2019, online: <http://arxiv.org/pdf/1903.02665v2> [25.08.2025].

CRAWFORD, Michael H.: *Roman Republican coinage*, Cambridge, 1974.

DELIGIO, Chrisowalandis/GENCER, Kerim: Natural Language Processing auf mehrsprachigen Münzdatensätzen. Untersuchung der Qualität, Datenqualität und Übertragbarkeit auf andere Datensätze, 2021, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis\\_Deligio\\_Gencer\\_DBISOnline.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterthesis_Deligio_Gencer_DBISOnline.pdf) [29.01.2025].

DELIGIO, Chrisowalandis/TOLLE, Karsten/WIGG-WOLF, David (Hrsg.): Supporting the Analysis of a Large Coin Hoard with AI-Based Methods, 2024.

DIETZ, Laura/KOTOV, Alexander/MEIJ, Edgar: Utilizing Knowledge Graphs for Text-Centric Information Retrieval, in: *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, hrsg. v. Kevyn COLLINS-THOMPSON u.a. 2018, S. 1387–1390.

FLEISHER, Will: Understanding, Idealization, and Explainable AI, in: *Episteme*, 4 (2022), S. 534–560.

GAMPE, Sebastian: Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen, Frankfurt, 2021, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf) [25.08.2025].

GAMPE, Sebastian/TOLLE, Karsten (Hrsg.): Creating an Additional Class Layer with Machine Learning to counter Overfitting in an Unbalanced Ancient Coin Dataset, Frankfurt, 2024.

GEIRHOS, Robert u.a.: Shortcut Learning in Deep Neural Networks, in: *Nature Machine Intelligence*, 11 (2020), S. 665–673.

GUO, Zhongliang u.a.: A Siamese Transformer Network for Zero-Shot Ancient Coin Classification, in: *Journal of Imaging*, 6 (2023).

HEKSTER, Olivier/MANDERS, Erika: Coinage in the Roman Provinces: the RPC and CHRE Projects, in: *Journal of Roman Studies* (2022), S. 199–213.

HERRERA, Francisco u.a.: *Multilabel Classification*, Cham, 2016.

HIJMANS, Steven: Art and Economy: Images on Roman Coins, in: *Sessions 4–5, Single Contributions* (2023).

HOLT, Frank L.: *When Money Talks*, New York, 2021.

HOWGEGO, Christopher: *Ancient history from coins (Approaching the ancient world)*, London, New York, 1995.

JURAFSKY, Dan/MARTIN, James H.: *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (Prentice Hall series in artificial intelligence), Upper Saddle River, 2009.

KAMPEL, Martin/HUBER-MÖRK, Reinhold/ZAHARIEVA, Maia: Image-Based Retrieval and Identification of Ancient Coins, in: *IEEE Intelligent Systems*, 2 (2009), S. 26–34.

KAVELAR, Albert/ZAMBANINI, Sebastian/KAMPEL, Martin: Reading the Legends of Roman Republican Coins, in: *Journal on Computing and Cultural Heritage*, 1 (2014), S. 1–20.

KEMMERS, Fleur/MYRBERG, Nanouschka: Rethinking Numismatics. The Archaeology of Coins, in: *Archaeological dialogues* (2011), S. 87–108.

KIM, Jongpil/PAVLOVIC, Vladimir: Discovering Characteristic Landmarks on Ancient Coins using Convolutional Networks, in: *2016 23rd International Conference on Pattern Recognition (ICPR)*, Cancun, Mexico (2016), S. 1595–1600.



LOYAL, Antje: Maschinelles Lernen angewendet auf Bilder antiker Münzen, Frankfurt, 2018, online: [http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit\\_Antje\\_Loyal.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2021/11/Masterarbeit_Antje_Loyal.pdf) [25.08.2025].

MELNYKOV, Oleksandr/KAPELESHCHUK, Andrei: A CNN-Assisted Decision Support System for Ancient World Coin Classification, in: *Computer systems and information technologies*, 1 (2024), S. 99–107.

NÖLLE, Michael u.a.: Dagobert - A New Coin Recognition and Sorting System, 2003.

PAVLEK, Barbara/WINTERS, James/MORIN, Olivier: Standards and Quantification of Coin Iconography. Possibilities and Challenges, in: *Digital Scholarship in the Humanities*, 1 (2022), S. 202–217.

SCHLAG, Imanol/ARANDJELOVIĆ, Ognjen: Ancient Roman Coin Recognition in the Wild Using Deep Learning Based Recognition of Artistically Depicted Face Profiles, in: *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)* (2017), S. 2898–2906.

SCHMIDT-DICK, Franziska: Typenatlas der römischen Reichsprägung von Augustus bis Aemilianus (Denkschriften / Österreichische Akademie der Wissenschaften, Philosophisch-Historische Klasse), Wien, 2003.

SIMONYAN, Karen/ZISSERMAN, Andrew: Very Deep Convolutional Networks for Large-Scale Image Recognition, in: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, 7-9 May 2015* (2015).

STOLL, Heinrich Wilhelm: Amalthea, in: *Ausführliches Lexikon der Griechischen und Römischen Mythologie*, hrsg. v. Wilhelm Heinrich ROSCHER, Leipzig 1884, S. 262–266.

TOLLE, Karsten u.a.: Semantic Search Based on Natural Language Processing. A Numismatic Example, in: *Journal Of Ancient History And Archaeology*, 3 (2018), S. 68–79.

TOLLE, Karsten/WIGG-WOLF, David: Improving Data Quality by Rules: A Numismatic Example, in: *CAA* (2017), S. 193–201.

WEISSER, Bernhard: Numismatiker in Deutschland. Zur Situation im Jahre 2019, in: *Geldgeschichtliche Nachrichten*, hrsg. v. Gesellschaft für Internationale Geldgeschichte, S. 264–277.

WOYTEK, Bernhard: Heads and Busts on Roman Coins. Some Remarks on the Morphology of Numismatic Portraiture, in: *Revue numismatique* (2014), S. 45–71.

ZAMBANINI, Sebastian/KAMPEL, Martin/SCHLAPKE, Mario: On the Use of Computer Vision for Numismatic Research, in: The 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (2008).

## 7. Abbildungsverzeichnis

Abbildung 1: Forschungsstand zu CNNs in der antiken Numismatik .....	14
Abbildung 2: Multiclass und Multilabel Modelle .....	17
Abbildung 3: Output des Modells .....	20
Abbildung 4: beschädigter Sesterz des Claudius .....	26
Abbildung 5: Beispieleintrag in der Datenbank .....	30
Abbildung 6: Aureus mit Gegenstempel .....	36
Abbildung 7: Funktionsweise von Fuzzy Search und Knowledge Graph .....	37
Abbildung 8: Das Kultbild der Diana Nemorensis .....	38
Abbildung 9: Verteilung der Wörter nach Wortwiederholungen .....	40
Abbildung 10: Knowledge Graph zu Stirnwaffenträgern .....	44
Abbildung 11: Anzahl unterschiedlicher deutscher Beschreibungen pro Münztyp .....	49
Abbildung 12: Anzahl unterschiedlicher Konzepte pro Münztyp im Deutschen und Englischen .....	51
Abbildung 13: Denar mit museumsdidaktischer Beschreibung .....	52
Abbildung 14: Vereinheitlichung von Münzbildern .....	57
Abbildung 15: Münzen mit natürlichem Hintergrund .....	57
Abbildung 16: Funktionsweise des CNNs .....	60
Abbildung 17: Trainingsprozess des Modells .....	61
Abbildung 18: Qualitätsmaße pro Klasse nach Häufigkeit im Datensatz .....	63
Abbildung 19: Vorhersagen zu Büste und Kopf .....	68
Abbildung 20: Saliency Map der Spes mit Blume .....	69
Abbildung 21: KI-Vorhersagen auf Objekten anderer Epochen .....	70