



TECHNISCHE HOCHSCHULE MITTELHESSEN

THM

**CAMPUS
GIESSEN**

MNI

Mathematik, Naturwissenschaften
und Informatik

Master Thesis

Einsatzmöglichkeiten der Bildverarbeitung in der klassischen
Archäologie

Ein Proof of Concept am Beispiel der Untersuchungen von Meleager-Sarkophagen

zur Erlangung des akademischen Grades

Master of Science

eingereicht im Fachbereich Mathematik, Naturwissenschaften und Informatik an der
Technischen Hochschule Mittelhessen

von

Manuel Groh

21. August 2019

Referent: Prof. Dr. Peter Kneisel

Ko-Referentin: Prof. Dr. Katharina Lorenz

Erklärung der Selbstständigkeit

Hiermit versichere ich, die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die Zitate deutlich kenntlich gemacht zu haben.

Gießen, den 14. April 2021

Manuel Groh

Danksagung

Mein größter Dank geht an Katharina, Peter und Andreas, die mir eine neue Welt gezeigt und mich herzlich darin aufgenommen haben.

Die Objekterkennung durch künstlich neuronale Netze ist in der Informatik ein aktuelles Forschungsfeld. In dieser Arbeit wird untersucht, ob sich die Ergebnisse jener Forschungen auch in die Domäne der Klassischen Archäologie übertragen lassen. Im Ergebnis zeigt sich, dass Netze, die initial zur Erkennung von Menschen auf Kamerabildern ausgelegt sind, auch auf Personen in Abbildungen von Meleager-Sarkophagen anwendbar sind. Mit Erkennungsraten von bis zu ca. 60% bei Personen und über 80% bei deren Posen wird deutlich, dass ein grundsätzlicher Mehrwert für die Klassische Archäologie gegeben ist. Ohne technische Neuentwicklungen bietet sich die Nutzung künstlich neuronaler Netze als Assistenz zur bisherigen Forschungsarbeit an. Mit ihrer Hilfe können in kurzer Zeit größere Bilddatenbanken exploriert und voruntersucht werden um z.B. Darstellungen szenisch zu unterteilen oder Personen und Personengruppen bestimmter Eigenschaften zu finden.

Inhaltsverzeichnis

1	Einführung	1
1.1	Motivation und Problemstellung	1
1.1.1	Bezug auf das Themenfeld	2
1.2	Forschungsstand	4
1.3	Methodik	6
2	Material	9
2.1	Netze	9
2.1.1	Funktionsweise	10
2.1.2	Frameworkauswahl	14
2.2	Bilder	15
2.2.1	Inhaltliche Auswahl	16
2.2.2	Technische Auswahl	18
3	Vorgehen	21
3.1	Personenerkennung	22
3.2	Posenerkennung	29
4	Ergebnis und Zusammenfassung	35
4.1	Anwendbarkeit	35
4.2	Fazit	37
4.3	Weiterer Forschungsbedarf	38
	Literaturverzeichnis	39
	Glossar	45
	Akronyme	53
	Abbildungsverzeichnis	55
	Tabellenverzeichnis	57
	Anhang	59
A	Bilder	61
B	Tabellen	71

1 Einführung

1.1 Motivation und Problemstellung

Mit dem Erfolg künstlich neuronaler Netze zur Bildverarbeitung (Convolutional Neural Networks (CNNs)) in Wettbewerben der Objektklassifizierung und -detektion ab dem Jahr 2012 ist die Forschung in diesem Bereich rasant vorangeschritten.¹ Entwickler wie Alex Krizhevsky, Christian Szegedy und Ross Girshick haben auf diesem Feld erfolgversprechende Resultate erzielt.² Insbesondere durch aktuelle Themen der Automobilindustrie, wie dem autonomen Fahren, wird die Entwicklung der benötigten Hard- und Software stark vorangetrieben.³ Gerade durch die so entstehende, kontinuierliche Verbesserung von Grafikprozessoren, können diese Fortschritte erreicht werden. Die zudem benötigten Daten u.a. von Straßenverläufen und Fahrweisen zur Entwicklung dieser Systeme, werden weltweit von verschiedenen Institutionen und Unternehmen auf Teststrecken erhoben.⁴ Dabei ist die Bildverarbeitung eine zentrale Komponente der Objekterkennung neben den anderen sensorischen Datenquellen wie Lidar und Ultraschall. Von diesen bietet allerdings derzeit nur die automatische Bildverarbeitung Erkenntnisse über die Art bzw. Kategorie des detektierten Objektes. In Wettbewerben wie der ImageNet Large Scale Visual Recognition Challenge (ILSVRC) müssen die CNNs eine große Anzahl Bilder in 1000 verschiedene Objektklassen einteilen. Die aktuellen Netze erreichen dabei erstaunlich präzise Ergebnisse, die je nach Aufgabenfeld die menschliche Fehlerquote schon unterschreiten.⁵

1 Siehe z.B. [Rus15]. Während [Kri12] noch Fehlerquoten über 33% (Detektion) und 15% (Klassifizierung) erreichte, so liegen aktuelle Teilnehmer bei ca. 6% und 3%

2 Basierend auf dem neuen Entwurf von [Kri12] haben sich andere Entwickler wie in [Sze13] und [Gir14] die Grundlegende Architektur zu eigen gemacht und stetig fortentwickelt.

3 In ihren Onlineartikeln schreiben [Han17] und [Gre18], dass künstlich neuronale Netze bzw. Machine Intelligence für Abläufe im Auto wie Motorensteuerung oder Menüführung eingesetzt werden können und nicht ausschließlich der Umgebungsüberwachung dienen. Für einen detaillierten Überblick der Terminologien und Einsatzmöglichkeiten im Volkswagen Konzern siehe auch [Hof17]

4 Ein bekanntes Beispiel in Deutschland ist der KITTI-Datensatz des Karlsruhe Institute of Technology in Zusammenarbeit mit Toyota. Siehe [Gei13]

5 Für die ILSVRC siehe [Rus15]. Andrej Karpathy schreibt in seinem Blog über die menschliche Fehlerquote in diesem Wettbewerb und über die Fehlergründe. In einem Update aus 2015 spricht er schon von maschinellen Leistungen, welche die menschliche hierbei übertreffen. Siehe [Kar14]

Dieser Stand der Forschung und diese Art der Technologie hat bereits in vielen Fachdisziplinen Einzug gehalten, die mit optischen Daten arbeiten, z.B. in der Medizin.¹ In geisteswissenschaftlichen Fächern wie der Kunstgeschichte oder den Altertumswissenschaften, die ebenfalls primär oder sekundär mit Fotografien oder Abbildungen arbeiten, ist bisweilen nur wenig zu Methoden künstlich neuronaler Netze geforscht und gearbeitet worden. Elliot Crowley und Andrew Zisserman haben in ihrem Paper „The State of the Art: Object Retrieval in Paintings using Discriminative Regions“ bereits gezeigt, dass sich die Techniken der Objektklassifizierung und -detektion ebenfalls auch auf Gemälde anwenden lassen, wobei sie in dieser Arbeit jedoch noch auf traditionellere Methoden als CNNs zurückgriffen.² Ein Proof of Concept der Verwendungsmöglichkeiten von CNNs in den bildorientierten Geisteswissenschaften, speziell aber der Altertumswissenschaften, wie sie es die Klassische Archäologie ist, hat noch nicht in der Breite stattgefunden. Eine erste Adaption der Objekterkennung mit Hilfe künstlich neuronaler Netze konnte Crowley mit der Arbeit „The Art of Detection“ anwenden, wobei die Technik im Bereich der Multi-Objekt-Erkennung schon weiter fortgeschritten ist.³

1.1.1 Bezug auf das Themenfeld

Aus wissenschaftlicher Sicht und speziell in ikonographischen und semiotischen Untersuchungen, ist der Fokus auf Interaktionen zwischen Personen oder Objekten von hoher Relevanz. Dabei werden speziell die dargestellten Szenen mit dem überlieferten Mythos verglichen und Diskrepanzen interpretiert.⁴ Durch Untersuchungen von Posen und z.B. der Mimik können Rückschlüsse auf die dargestellten Szenen gezogen werden. Von besonderer Bedeutung sind dabei die Attribute, die eine Person identifizierbar machen. Bei diesen körperlichen Merkmalen oder persönlichen Gegenständen kann es sich um kontextuelle oder kontextfreie Attribute handeln. Einerseits kann eine Segmentierung einzelner Personen aus einer szenischen Darstellung also nur bedingt von Nutzen für den Betrachter sein, da ggf. wichtige Identifikationsmerkmale verloren gehen. Andererseits dient eine genaue Segmentierung der räumlichen Beschreibung eines Werkes und kann als vorbereitender Arbeitsschritt zur Bestimmung von Interaktionen zwischen Personen oder Objekten genutzt werden. Weiterführende Interpretationen sind zumeist abhängig von den Kontexten u.a. der Szene, des ursprünglich umgebenden Raums sowie der kulturellen und gesellschaftlichen Hintergründe.

Ohne diese Kontextualisierung kann eine automatisierte Auswertung also nur ein erstes Hilfsmittel als Ausgangsbasis für weitere Untersuchungen sein. Dennoch bietet ein

1 Siehe z.B. in der Erkennung von Hautkrebs [Nas16]

2 Siehe [Cro14]

3 Siehe [Cro16] im Vergleich zu [Ren15]

4 Siehe [Lor16] Seite 37

solcher Automatismus gerade durch seine Unwissenheit über den Kontext auch einen qualitativen Vorteil: Bei der Betrachtung einzelner Bildelemente kann er nicht quantitative Erhebungen interpretatorisch beeinflussen, da er über diese Daten, anders als ein lernender Mensch, nicht verfügt.

Bezogen auf Gruppen- und Szenendarstellungen ist auch zu untersuchen, ob durch eine initiale quantitative Datenerhebung es nicht auch möglich ist, kontextuelle Gruppierungen vorzunehmen. Durch erfassbare Daten wie Posen und Blickrichtungen können Richtungsvektoren berechnet werden, die als Interaktionsrelationen zwischen Objekten interpretierbar sind.¹ So könnten durch ein Clustering der detektierten Objekte mit ihren Relationen, Abbildungen in einzelne Szenarien aufgeteilt werden.

Technologisch handelt es sich dabei um eine Adaption von Objektsegmentierungen, Human Pose Estimations und Face Pose Estimations, welche allesamt auf CNNs basieren können.² Als Datengrundlage werden Abbildungen von römischen Sarkophagen der Meleager-Sage verwendet. Für die klassische Archäologie stellt dieses Konvolut mit ca. 200 bekannten Artefakten eine umfangreiche und gut untersuchte Gruppe dar. Sie umfasst etwa 800 Personen, deren Untersuchungsergebnisse mit denen der bestehenden Forschung verglichen werden können.³ Zudem ist auch die Zugänglichkeit der Daten ein ausschlaggebender Grund für die Auswahl. Sowohl in den Katalogen von Carl Robert „Einzelmythen: Hippolytos - Meleager. Die antiken Sarkophagreliefs III 2.“ und Guntram Koch „Die mythologischen Sarkophage: Meleager. Die antiken Sarkophagreliefs XII 6.“ als auch in online verfügbaren Datenbanken wie Arachne der Uni Köln und Foto Marburg sind viele Abbildungen vorhanden.⁴

Untersucht werden soll daher, ob es möglich ist bestehende Technologien direkt auf antike Reliefs anzuwenden oder ob ergänzenden Vorarbeiten zu leisten sind. Die Ergebnisse sollen zum einen stochastisch ausgewertet werden, um z.B. wiederkehrende Muster zu identifizieren oder Aussagen über räumliche Relationen zu treffen. Zum anderen ist ein qualitativer Vergleich der anfallenden Daten mit dem Stand der Forschung anzustellen, um den Nutzen der digitalen Bildverarbeitung in den Altertumswissenschaften einschätzen zu können. Resultate der Arbeit können dabei in die weiteren Untersuchungen des

¹ Zu verschiedenen Möglichkeiten die Kopfposition zu bestimmen siehe [Kaz14], [Liu16b] und [Rui18]

² Alle erwähnten Modelle basieren auf CNNs. Unter Methodik wird weiterführend beschrieben, warum sich diese Arbeit darauf fokussiert.

³ Siehe auch die Begründung zur Materialwahl im DFG Projektantrag „ReGroup: Graph-based analysis of group depictions in the Roman imperial period and its implications for art-historical knowledge production in the digital age“ von Katharina Lorenz und Andreas Kuczera

⁴ Unter <https://arachne.uni-koeln.de> werden mit dem Suchbegriff „Meleager Sarkophag“ über 800 Ergebnisse geliefert.

Projektes „ReGroup“ einfließen, welches die räumliche Anordnung und die Interaktionen zwischen Personen graphbasiert erfassen möchte.¹

Ebenfalls ist zu prüfen, ob die oben genannten Techniken alleinstehend oder in Kombination bessere Resultate liefern können. Es wird vorab die Hypothese aufgestellt, dass die Posenanalyse weniger falsch-positive Ergebnisse liefert, wenn die Untersuchung auf Personen eingeschränkt werden kann. Im Kern laufen diese Analysen auch in diesem Stil ab, jedoch gibt es für jeden Anwendungsfall auch spezialisierte Netze oder Algorithmen, welche unter Umständen erst in Kombination einen deutlichen Mehrwert für die Untersuchung liefern.

1.2 Forschungsstand

Elliot Crowley ist mit seiner Adaption der Objekterkennung in der Kunstgeschichte in diesem Bereich am weitesten fortgeschritten. Diese Untersuchungen sind für den konkreten Anwendungsfall der römischen Sarkophage nur sehr begrenzt nutzbar, da sich die Domäne der Gemälde von den Sarkophagreliefs hinsichtlich Material und Farbe deutlich unterscheidet. Im Fokus der folgenden Arbeit steht die Detektion von Personen, welche in den trainierten Datensätzen nur einen kleinen Teil ausmachen.²

Ähnlich der Arbeit Crowleys hat auch Hongping Cai in „Beyond Photo-Domain Object Recognition“ an dem Problem der „Cross-Depiction“ gearbeitet.³ Inhaltlich basieren diese Arbeiten auf sehr frühen CNNs oder Support Vector Machines (SVMs), deren Technik in den letzten Jahren weit fortgeschritten ist.⁴ Damit kann ein früherer Ansatz, wie in den oben genannten Arbeiten also nur als Ausgangspunkt weiterer Forschungen zu sehen sein.

1 Siehe auch die Zielsetzung des DFG Projektantrags „ReGroup: Graph-based analysis of group depictions in the Roman imperial period and its implications for art-historical knowledge production in the digital age“ von Katharina Lorenz und Andreas Kuczera

2 Im ImageNet Datensatz befinden sich im 2011 Fall Release ca. 32.000 Bilder, von denen etwa 7.000 auf die Kategorie „person, individual, someone, somebody, mortal, soul“ entfallen. In dem von Crowley genutzten Datensatz (siehe [Cro16]) befinden sich nur zehn Klassen, von denen keine Personen klassifizieren.

3 Siehe [Cai15a]. In [Cai15b] beschreibt er das Cross-Depiction Problem als „The cross-depiction problem is that of recognising visual objects regardless of whether they are photographed, painted, drawn, etc.“

4 Sie Kapitel Abschnitt 2.1 für weitere technische Ausführungen zur Funktionsweise der Objekterkennung.

Schon mit den von Crowley genutzten Faster R-CNN ist die Erkennung multipler Objekte möglich. Weiterentwickelt wurde die Idee zur Segmentierung einzelner Objekte mit Mask R-CNN.¹ Andere Arbeiten wie You Only Look Once (YOLO) oder Single Shot Detector (SSD) sind ebenfalls dafür geeignet.² In dieser Thesis soll untersucht werden, welche der genannten Netze sich direkt, oder nach einiger Anpassung, für die Personenerkennung in antiken Reliefs eignen und im Ergebnis ggf. die Genauigkeit der darauffolgenden Untersuchungen wie der Posen- und Gesichtserkennung verbessern.

Im Bereich der Posenerkennung gibt es mit den Arbeiten von Bao, Cao, G. Wei, S. Wei und Zhang einige recht aktuelle Forschungsstände, deren Adaptionenmöglichkeit und Genauigkeit für stochastische Untersuchungen im konkreten Anwendungsfall zu prüfen sind.³ Mit dem aktuellsten Paper aus diesem Jahr „View Invariant 3D Human Pose Estimation“, wird die räumliche Dimension in den Vordergrund gestellt. Ob sich diese Dreidimensionalität auch perspektivisch aus reinen Tiefendaten generieren lässt, wäre weiterführend ein interessanter Aspekt.⁴

Die Gesichtserkennung bzw. die daraus resultierende Kopfposition basiert zumeist auf sogenannten Keypoints, die im Wesentlichen die körperlichen Gesichtsmarkale wie Augen, Ohren, Nase und Mund repräsentieren. Der Erfolg der Anwendung auf steinerne Abbildungen hängt also vornehmlich von der Erkennungsleistung und dem Vorhandensein dieser Merkmale ab. Die Toleranz dieser Algorithmen in Hinblick auf fragmentierte Teile der Reliefs ist daher von besonderer Bedeutung. Arbeiten von Nathaniel Ruiz und Xiabing Liu versprechen einen hohen Detailgrad ohne die typischen Grundvoraussetzungen.⁵ Es ist zu prüfen, ob sich ähnliche Leistungen auch in diese Domäne übertragen lassen. Von deutlich höherer Komplexität sind unvollendete Sarkophage, die für den künftigen Besitzer bzw. dessen Angehörige teilweise die Gesichter unbehauen ließen, um sie persönlich anpassbar zu machen.⁶ Diese zeichnen sich dementsprechend durch das Fehlen jeglicher Merkmale aus. Nur die ungefähre Blickrichtung lässt sich aus der Körperhaltung und Grundform des Gesichts schließen.

1 Siehe [He17]

2 Für YOLO siehe [Red18] und für SSD [Liu16a]

3 Basierend auf Step Deep CNNs siehe [Bao18], basierend auf Part Affinity Fields siehe [Cao18]. Andere Ansätze mit dem Fokus auf technischen Aspekten wie Invarianz gegenüber der Aufnahme oder Ausführungen in Echtzeit siehe [Wei19], [Wei16] und [Zha19]

4 Siehe [Wei19]

5 Siehe [Rui18] und [Liu16b]. Beide Netze extrahieren nicht die Gesichtsmarkale um anschließend eine 2D zu 3D Regression vorzunehmen, sondern berechnen direkt die Eulerschen Winkel zur Drehlage im Raum.

6 Über das Problem der Personenköpfe schreibt [Koc93]: „Häufig sind die Porträts sehr klein oder nur flüchtig ausgearbeitet“. In [And84] wird zudem über das Phänomen der bossierten, also nicht ausgearbeiteten Porträts diskutiert.

1.3 Methodik

Im zeitlich begrenzten Rahmen der Masterthesis kann die Eigen- und Neuentwicklung von Netzarchitekturen, basierend auf den genannten Arbeiten nicht im Fokus stehen. Vielmehr ist zu prüfen, welche konkreten und im besten Fall vortrainierten Netze verfügbar sind, um sie im Kontext dieses Anwendungsfalls zu testen. Mit Googles Tensorflow gibt es ein Framework, welches im Bereich der künstlich neuronalen Netze sehr breite Anwendung findet.¹ In Kombination mit Bibliotheken, die über viele Jahre führend im Bereich der Bildverarbeitung sind, wie z.B. OpenCV ergibt sich ein Projektsetting, welches für den generellen Proof-of-Concept geeignet scheint.² Sprachlich ist zum jetzigen Zeitpunkt die Programmiersprache Python die Fachdomäne für die Entwicklung von CNNs, wobei auch C++ durch seine Nähe zur Grafikkhardware in Frage kommt. Beide Varianten bleiben potentiell möglich, andere Sprachen wie Java werden ebenfalls bereits unterstützt, jedoch werden hier die Ressourcen an vorhandenen Netzen aufgrund des Verbreitungsgrades deutlich kleiner sein.

Mit der Wahl von CNNs ist andererseits die Abwahl anderer Ansätze wie klassischer Bildverarbeitung oder SVMs verbunden. Diese Entscheidung resultiert auf der Grundlage des Arbeitsmaterials. Bei antiken Reliefs handelt es sich um Personendarstellungen in verschiedensten Positionen bis hin zu unnatürlichen Körperhaltungen und materialbedingten Fehl- oder Schadstellen. Zudem können einzelne körperliche Merkmale künstlerisch nicht exakt anatomisch abgebildet worden sein. CNNs kommen mit dieser großen Variation aufgrund ihrer hochkomplexen Struktur besser zurecht wie die klassische mathematischen Verfahren, die eine gute, wohldefinierte Ausgangsbasis benötigten. Zudem zeigen die Resultate der verschiedenen Wettbewerbe, dass die Präzision der neuronalen Lösungen signifikant über den klassischen Ansätzen liegt.³

Die Einsatzmöglichkeiten der Bildverarbeitung werden im Verlauf dieser Arbeit an den Methoden der Personen- und Posenerkennung gezeigt. Im Kontext der mythologischen Sarkophage sind auch weitere Methoden, wie die Untersuchung der Veränderungen von Reproduktionen über die Zeit möglich. Diese haben ihren Ursprung häufig in der Kunstwissenschaft oder beziehen sich auf andere Ressourcen, weswegen sie für den Einsatz in der klassischen Archäologie nur begrenzt von Interesse sind.⁴ Natürlich

1 Siehe <https://www.tensorflow.org/> für weiterführende Informationen.

2 Siehe <https://opencv.org/> für weiterführende Informationen.

3 Siehe die Ergebnisse der ILSVRC seit 2012

4 In [Aub14] wurden Gemälde von Objekten aus verschiedenen Blickrichtungen dreidimensional rekonstruiert. In [Cro15] wird die Personenzuordnung per Gesichtserkennung auf Gemälde übertragen. [Joh08] behandelt die Künstlererkennung auf Grundlage anderer Werke. In [Mon14] werden Reproduktionen auf Ihre Veränderungen untersucht. [Res14] vergleicht Bilder und Bildausschnitte in Archiven um Zusammenhänge darzustellen und in [Sch11] werden ähnlich der hier zu betrachtenden Untersuchungen Handgesten in Bildern erkannt und klassifiziert.

lassen sich die hier zu untersuchenden Methoden auch auf andere Ressourcen mit figürlichen Darstellungen anwenden, dies gilt neben den Gemälden auch für Malereien auf Gegenständen oder noch räumlichere Objekte wie Statuen. Dennoch zeichnen sich die Mythendarstellungen auf Sarkophagen im Gegensatz zu vielen anderen Materialgruppen dadurch aus, dass die Darstellungen szenisch unterteilt sind und vergleichsweise viele Personen erscheinen. Diese Unterteilung durch automatisierte Verfahren zu ermitteln, steht bei der Posen- und Blickrichtungserkennung im Vordergrund und unterscheidet sich damit gänzlich von anderen Techniken der Bildverarbeitung oder der Datenerhebung in Gemälden.¹

Der Zugang zu den benötigten Ressourcen, also Abbildungen der Sarkophage sind über die Datenbanken Arachne der Universität Köln, Foto Marburg der Phillips Universität Marburg und Census der Humboldt-Universität zu Berlin möglich.²

-
- 1 Quantitative Erhebungen von Farbwerten und deren Verteilung wie z.B. in [vP15] sind aus kunstwissenschaftlicher Sicht interessant, stellen aber keine Fragestellungen im Bereich dieser Domäne der Sarkophage dar.
 - 2 Siehe dazu die Onlinezugänge unter <https://arachne.dainst.org/>, <https://www.bildindex.de/> und census.bbaw.de.

2 Material

Wie bereits beschrieben, gibt es zum jetzigen Stand der Forschung eine Vielzahl an Netzarchitekturen und -varianten. Für den angestrebten Proof-of-Concept, welcher zum Ziel hat die Anwendung der künstlich neuronalen Netze für die steinernen Reliefs der Sarkophage zu testen, ist es jedoch nicht nötig eigene Trainings durchzuführen. Um das Potential der Generalisierung von Strukturen und Mustern noch deutlicher zu zeigen, ist es sogar sinnvoll, auf natürliche Aufnahmen vortrainierte Netze zu nutzen. Ein Netz, welches darauf trainiert wurde, Menschen in einer realen Umgebung zu erkennen und im Einsatz sogar steinerne Abbildungen als Menschen identifiziert, verdeutlicht bereits die mögliche Übertragbarkeit in andere Domänen. Durch ein erweitertes Training ist es wahrscheinlich, dass die Erkennungsrate in diesem neuen Bereich noch steigt.¹ Aufgrund dieser Prämisse und dem Verzicht auf ein eigenes Training bieten sich einige Frameworks zur Nutzung besonders an, da sie über einen sogenannten Model Zoo, also einer Sammlung vortrainierter Netze verfügen, deren einfache Anwendbarkeit sichergestellt ist.

2.1 Netze

Die Objekterkennung in Abbildungen durch CNNs lässt sich gut in drei Stufen beschreiben, welche in der Abb. 2.1 visualisiert werden. Technisch am einfachsten ist die Bildklassifizierung, bei der das Hauptmotiv im Bild erkannt wird. Dabei ist es noch nicht von Interesse, an welcher Position oder wie groß das Objekt ist.² In einem weiteren Schritt können die einzelnen Merkmale, die zur Erkennung geführt haben genutzt werden, um das Motiv im Bild einzugrenzen. Man spricht hierbei von einer Lokalisierung des Objekts. Auch diese Netzarchitekturen sind vergleichsweise simpel im Aufbau, da ihre Resultate weiterhin nur aus einer Objektklasse und zwei Punkten für den Rahmen der sogenannten Bounding Box besteht. Problematisch wird es die

1 In der Domäne der Gemälde konnten Crowley et al. zeigen, dass die Genauigkeit einzelne Klassen zu Erkennen durch erweitertes Training im Mittel um 13% steigen kann. Siehe dazu [Cro16].

2 In der von [Rus15] beschriebenen „ImageNet Large Scale Visual Recognition Challenge“ ist die ursprüngliche Aufgabe nur das Hauptmotiv zu erkennen. Zur besseren Vergleichbarkeit aufgrund falscher Abschätzungen oder multipler Objekte können dabei fünf Klassenvorschläge pro Bild abgegeben werden.

Auswahl eines Hauptmotivs voraussagen, wenn in dem Bild verschiedene Objekte gleich präsent sind.

Mit der Erweiterung auf multiple Objekte werden die Netze komplexer. Entweder in einem vorgelagerten Schritt oder auch im Laufe des Prozesses werden sogenannte Region of Interests (ROIs) ermittelt und klassifiziert. Diese können sehr vielzählig sein, mit dem Ziel, jedes Objekt unabhängig von Größe und Bildverhältnis zu treffen. Während in früheren Ansätzen klassische Verfahren der Bildverarbeitung genutzt wurden um einzelne Bildsegmente zu ermitteln, so nutzen neuere Netzarchitekturen zur Bestimmung der ROIs die gleichen Daten, die auch zur Klassifizierung genutzt werden. Zusammen mit einer entsprechenden Klassenwahrscheinlichkeit der erkannten Klasse und ggf. weiteren Informationen wird die Region in Form einer Bounding Box dann als Object Proposal vom Netz ausgegeben.¹

Im letzten Schritt können die so genau eingefassten Objekte weiter verfeinert werden. Bei der Instanzsegmentierung zum Beispiel wird ein Objekt pixelgenau maskiert, was nützlich ist um einzelne Bildelemente freizustellen. Konkret auf Personen bezogen ist auch eine Keypointdetection möglich, genutzt werden kann, um Körperposen genau festzustellen und zu vergleichen.

2.1.1 Funktionsweise

Convolutional Neural Networks arbeiten nach einem recht simplen Prinzip. Das originale Bild wird als zweidimensionale Matrix von Farbwerten verstanden. Bei einem Grauwertbild handelt es sich um eine Matrix, bei Farbbildern mit den Farbkanälen Rot, Grün

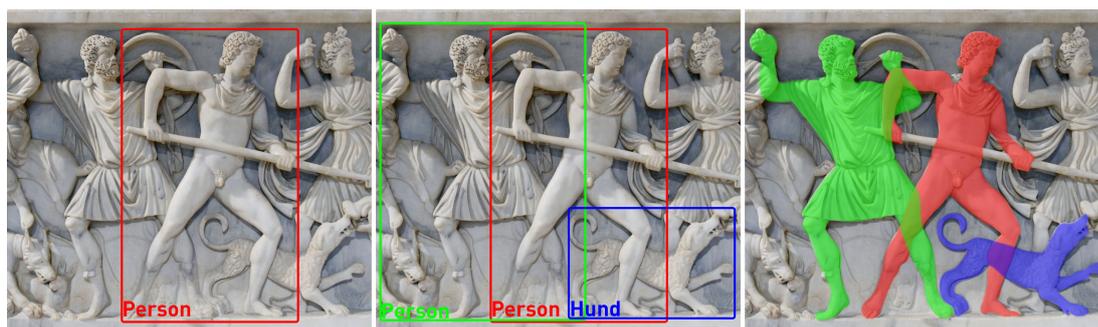


Abbildung 2.1: Übersicht über die Stufen der Objekterkennung auf Grundlage der Abbildung 3.1. Von links: Klassifizierung und Lokalisierung, Objekterkennung, Instanzsegmentierung

¹ In einem Blogpost von Piotr Dollár, einem der Autoren von Mask R-CNN (siehe [He17]), wird die historische Herleitung von Object Proposals beschrieben. Dabei sind es Algorithmen zur Bildsegmentierung, die aus einzelnen Bestandteilen durch z.B. Ähnlichkeiten in Farbe und Textur komplexe Objekte generieren. Siehe [Dol13].

und Blau dementsprechend um drei. In der originalen Architektur des AlexNet werden die Bilder vorab auf 224 mal 224 Pixel umgerechnet, was bei regulären Fotografien einer deutlichen Verkleinerung entspricht.¹ Dieser Umstand erklärt auch, warum Netze, die diese Architektur nutzen kleine Objekte in Relation zur Bildgröße nicht erkennen können. In moderneren Architekturen, wie der R-CNN Familie wird nicht mehr das gesamte Bild, sondern die einzelnen ROIs betrachtet, weswegen dieser Umstand nur noch bedingt aktuell ist. Eine ähnliche Problematik gilt für sehr kleine Objekte, die um weiter verarbeitet werden zu können, im ersten Schritt als mögliches Objekt erkannt werden.

Im weiteren Verlauf werden verschiedene Kernels über die Matrizen geschoben. Ein solcher Kernel, beispielsweise als Matrix der Größe 11 mal 11 Pixel, repräsentiert eine spezielle Form wie etwa eine senkrechte Linie oder einen Punkt. Welche Formen genau darin hinterlegt sind, ist nicht vorprogrammiert, sondern ein Resultat des Lernprozesses. Sie sind daher je nach Training unterschiedlich und nicht vorhersehbar. Beim Traversieren eines Kernels über die Matrix wird in jedem Schritt die Form mit dem Ausschnitt verglichen und die Ähnlichkeit als Wert ausgegeben. Das Resultat ist somit eine neue Matrix, welche in ihrer Größe der originalen fast entspricht, sofern der Kernel in Einzelschritten verschoben wurde. Bei größeren Schrittweiten (Strides) ist die Ergebnismatrix dementsprechend kleiner.²

Jeder Kernel erzeugt aufgrund der Farbkanäle drei neue Matrizen, die nun abermals von weiteren Kernels traversiert werden. Aufgrund der Kombinatorik entstehen in wenigen Schritten sehr viele dieser Ergebnismatrizen. Man spricht hierbei auch von Feature Maps, da die einzelnen Zellenwerte auf die Präsenz der jeweiligen Formen hinweisen. Mit jeder Stufe werden diese Formen dabei konkreter, von geometrischen Strukturen über farbige Muster hin zu genauen Texturen. Beispielhaft wurden diese von Fabio Grätz in der Abb. 2.2 visualisiert.³ Zu erkennen ist, dass tiefere Ebenen komplexere Strukturen verarbeiten können, wie in diesem Fall Ketten in der 40. Schicht. Im letzten Schritt vor der Klassifizierung kann eine solche Feature-Map hochdimensional mehrere zehntausend Werte beinhalten.⁴

Dieses Ergebnis dient in den folgenden, vollständig verknüpften Schichten, auch dichte Ebenen genannt, als Eingabe. Sie sind die klassischen Strukturen künstlich neuronaler

1 Siehe [Kri12] in der Figur 2. Trotz einer solchen Beschränkung handelt es sich in der Eingabe um 150.528 einzelne Zahlenwerte.

2 In der klassischen Bildverarbeitung wird auch von Filtern oder Faltungsmatrizen gesprochen. Bekannte Vertreter sind z.B. der Gauß-Filter als diskrete Glockenkurve zur Weichzeichnung von Bildern oder die Sobel-Operatoren zur Kantendetektion.

3 Eine detaillierte Anleitung zur Visualisierung von trainierten Konvolutionsschichten und weitere Beispiele für die Abstraktionsstufen der Ebenen findet sich bei [Grä19].

4 Im klassischen AlexNet in [Kri12] werden zuletzt 43.264 Werte generiert, unter Mask R-CNN in [He17] sind es bis zu 100.352 je nach Einstellung.

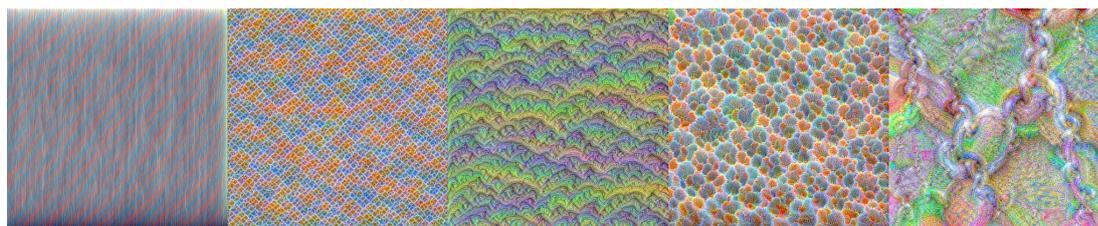


Abbildung 2.2: Visualisierung von Featuremaps. Zu sehen sind exemplarisch Elemente der Ebenen 7, 14, 20, 30 und 40. ([Grä19])

Netze. Jedes Element einer Ebene ist mit allen Elementen der vorherigen und der nachfolgenden Ebene verknüpft. Ein einzelnes Element, auch Neuron genannt, nimmt dabei alle Ausgabewerte der Vorgänger, multipliziert sie jeweils mit einem antrainierten Faktor und summiert sie auf. Abhängig von einem ebenfalls gelernten Schwellenwert, wird das Ergebnis an die nachfolgenden Elemente weitergegeben. Der Ablauf ist dabei dem in natürlichen neuronalen Strukturen sehr ähnlich.¹ Am Ende der tiefen Schichten kommen als Resultat bei der AlexNet-Architektur 4096 Werte heraus, die eindimensional als Feature-Vektor dargestellt werden.²

Zur Klassifizierung kann letztendlich dieser Vektor von einer SVM verarbeitet werden. Der Vektor wird dabei als Punkt im multidimensionalen Raum verstanden, welcher durch eine Hyperebene unterteilt ist. Je nachdem, ob der Punkt sich vor oder hinter der Ebene befindet, handelt es sich um das gesuchte Objekt oder nicht. Das Problem der SVMs ist somit die Beschränkung auf binäre Klassifikationen, da für jede Objektklasse eine separate SVM benötigt wird. Viele Netzarchitekturen ersetzen daher die SVMs durch eine weitere tiefe Schicht, welche für jede Klasse ein Neuron besitzt. Durch Multiplikation und Addition des Featurevektors in jedem Neuron lässt sich so eine Klassenwahrscheinlichkeit für jede Objektklasse ermitteln. Im Ergebnis ist das detektierte Objekt also von der Klasse mit der höchsten Wahrscheinlichkeit.³

Die Funktionsweise der R-CNN-Familie hat sich stetig optimiert, der generelle Ablauf ist allerdings gleich geblieben. Zu Beginn nutzte R-CNN einen Algorithmus namens Selective Search um ROIs zu generieren. Dieser fragmentiert dazu das Bild in viele kleinstmögliche Teile, welche dann aufgrund von Ähnlichkeiten in Farbe, Textur und Größe paarweise zu größeren Fragmenten zusammengesetzt werden. Im letzten Schritt ist somit das gesamte

1 Der Ursprung künstlich neuronaler Netze liegt in den Arbeiten von Warren McCulloch und William Pitts, die Neuronen mathematisch modelliert haben. Siehe dazu [McC43].

2 Siehe [Kri12], andere Netzwerke wie Mask R-CNN in [He17] sind mit 1024 bis 2048 Werte im letzten Schritt geringer aufgestellt.

3 Schon [Kri12] nutzten als letzte Schicht eine Softmax-Regration, welche die Summe der Ausgabewerte aller Neuronen auf 100% normalisiert, um eine Wahrscheinlichkeitsverteilung zu verdeutlichen. SVMs sind im Vergleich mathematisch berechenbar und benötigen daher kein Training. Durch die Entkopplung von Beschreibung und Klassifizierung wird allerdings das Training der Klassifizierer komplexer.

Bild ein Fragment, wodurch sich eine hierarchische Baumstruktur ergibt.¹ R-CNN nimmt nun die oberen 2.000 Elemente des Baumes als ROIs, skaliert sie auf eine fest definierte Größe und untersucht sie einzeln. Die Kernels in den Konvolutionsschichten verarbeiten jeden Ausschnitt zu einem Feature-Vektor von 4096 Elementen, die dann nachträglich von SVMs klassifiziert werden.²

Die Probleme bei diesem Vorgehen liegen hauptsächlich in der Zeit und Effizienz. Sowohl die Generierung der ROIs, als auch die weitere Verarbeitung sind sehr langsam und bei einer Verarbeitungszeit von mehr als 40 Sekunden pro Bild auch nicht für Echtzeitanwendungen geeignet. Bei einer genauen Betrachtung des Ablauf stellt man fest, dass die Berechnung der Features sehr redundant ist. Ein komplettes Bild kann in eine Featuremap konvertiert werden. Diese hat zwar andere Abmessungen, jeder Punkt im originalen Bild lässt sich aber auf diese Matrix abbilden. Mit dieser Erkenntnis ist klar, dass eine zweitausend-fache Feature-Generierung nicht zweckmäßig ist. Das zweite Netz der Familie, Fast R-CNN, optimiert sich eben durch jene Feststellung. Das Bild wird nur einmalig in seine Features konvertiert und die ROIs nur noch darauf abgebildet. Einfach beschrieben nutzt man zur Klassifizierung nicht einen Ausschnitt aus dem Bild, sondern aus der resultierenden Matrix. Auch die Klassifizierung selbst wurde in das Netz hinein verlagert. Anstelle der SVMs nutzen die Entwickler hier tiefe Schichten.³ Die Struktur des Fast R-CNN ist somit kompakter als sein Vorgänger, wodurch auch das Training effizienter durchgeführt werden kann. In der Durchführung stellt sich eine Zeitersparnis um den Faktor 20 ein. Mit etwa zwei Sekunden ist es ebenfalls noch nicht im Bereich der Echtzeit, für die Exploration großer Datenmengen jedoch geeigneter als zuvor.

Mit dieser Optimierung bleibt weiterhin das Problem des zeitintensiven Selective Search. In Faster R-CNN wurde dieser initialer Schritt weggelassen und durch ein spezielles Subnetz ersetzt. Diese neue Teilkomponente nutzt die Featuremap und nicht das originale Bild um daraus ROIs zu generieren. Der Vorteil liegt nicht nur in der Zeitersparnis, als neuronales Netz ist diese Komponente, im Gegensatz zu dem statischen Selective Search Algorithmus, ebenfalls trainierbar. Während der Selective Search ganz generell Strukturen in einem Bild gruppiert, kann ein separates Training bei dieser Architektur auch dazu führen, dass domänenspezifische Objekte besser erkannt werden.⁴ Im Vergleich zu Fast R-CNN stellt sich abermals ein Faktor 10 in der Zeitersparnis ein. Mit etwa 0,2 Sekunden ist der Faster R-CNN nun auch für Echtzeitanwendungen geeignet.

1 Für den Selective Search Algorithmus siehe [Uij13].

2 Für das R-CNN Paper siehe [Gir14].

3 Für das Fast R-CNN Paper siehe [Gir15].

4 Für das Faster R-CNN Paper siehe [Ren15].

Als aktuell jüngster Vertreter der Familie ist Mask R-CNN zu nennen. Die Generierung der ROIs ist hier weiter optimiert worden, der im Ergebnis sichtbare Unterschied liegt allerdings im letzten Schritt. Neben der Objektklasse und der entsprechenden Bounding Box wird zudem noch eine Objektmaske ausgegeben. Diese Architektur wurde bereits von den Entwicklern ebenfalls darauf ausgelegt, um Personenposen zu analysieren. Dazu wird für jeden Keypoint eine separate Maske mit ausgegeben.¹ Eine solche Objektmaske ist eine Matrix in der Größe der Bounding Box, welche für jeden Bildpunkt eine Wahrscheinlichkeit angibt, dass dieser zu dem Objekt gehört. Bezogen auf die Keypoints wird so nur der Bildpunkt mit der höchsten Wahrscheinlichkeit gewählt.²

Im Gegensatz zur R-CNN-Familie nutzen Netze wie RetinaNet, YOLO und SSD statt zwei getrennte Schritte zur Erkennung und Klassifizierung nur einen. Bei diesen Verfahren wird das Bild in ein grobes Raster eingeteilt und einzelne Zellen ausgewählt, welche wahrscheinlich Bestandteil eines Objektes sind. Diese Auswahl basiert wie bei CNNs generell auf den Resultaten der Featuremaps. Auf diese so ausgewählten Zellen, auch Anker genannt, werden fest definierte Rahmen unterschiedlicher Größe gelegt und geprüft, welcher von diesen am wahrscheinlichsten dem Objekt entspricht.³ Durch diese festgelegten Rahmen sind die Netze mit einem Durchgang weniger flexibel bei der Objektgröße als andere Netztypen, jedoch sind sie so auch schneller. Je nach Anwendungsfall kann so zwischen Geschwindigkeit und Genauigkeit gewählt werden.

2.1.2 Frameworkauswahl

Zur Auswahl stehen insbesondere die Sammlungen der Frameworks Tensorflow von Google und Facebooks Detectron. Beide Frameworks bieten rund 30 bis 40 verschiedene Modelle an, die hauptsächlich auf state-of-the-art Architekturen basieren.⁴ Um den zeitlichen Aufwand eigener Programmierleistungen stark reduziert und zugleich die korrekte Visualisierung der Ergebnisse sicherzustellen, ist im Umfang des Detectron-Projekts bereits ein einfaches Programm zur direkten Ausführung der Netze enthalten. Da zudem die neueren Entwicklungen im Bereich der CNN-Netzarchitekturen im Rahmen des Facebook AI Research Programms entstanden sind, wurde sich gegen die weitere Nutzung des Tensorflow Frameworks und für den Facebook eigenen Model Zoo entschieden.⁵

1 Für das Mask R-CNN Paper siehe [He17].

2 Eine weitere, nicht rein statistische Auswahl ist beispielsweise im Kontext aller Keypoints möglich. Über die räumliche Ordnung der Körperteile und deren relative Größe zueinander kann potentiell eine genauere Festlegung erfolgen.

3 Für das RetinaNet Paper siehe [Lin17].

4 Für das Detectron-Framework siehe [Gir18], für Tensorflow siehe [Aba15].

5 Die Publikationen der Facebook Research Gruppe sind online zugänglich unter <https://research.fb.com/publications/?cat=13>.

Dieser besteht aus insgesamt 75 vortrainierten Netzen, welche sich initial in zwei Kategorien aufteilen. Um eine Klassifizierung vorzunehmen, nutzt etwa die Hälfte vorab generiert ROIs. Die verbleibenden 37 Netze führen die ROI-Ermittlung und Klassifizierung gemeinsam aus, weswegen sie auch „Ende-zu-Ende trainierte“ Netze genannt werden. Sie lassen sich ohne weitere Abhängigkeiten leichter ausführen. Von diesen gehören zehn zur Netzarchitektur der Faster R-CNN und acht basieren auf RetinaNet. Die restlichen 19 sind Mask R-CNN Netze, von denen wiederum acht für die Erkennung von Keypoints geeignet sind.

Ein großes Problem der neuronalen Netzwerke ist, das mit zunehmender Tiefe oder Breite der Ebenen die Qualität der Erkennung nicht mehr steigt, sondern stagniert oder gar rückläufig wird. Werden zu viele Konvolutionsschichten aneinander gereiht, sind die Ergebnisse der Ebenen zu spezifisch und nicht mehr allgemein genug, um Variationen von Objekten erkennen zu können. Um auch in tiefen Schichten weiterhin einen Zugriff auf die abstrakten Ergebnisse ermöglichen, wurden mit der ResNet-Architektur „Kurzschlüsse“ zwischen den Ebenen eingeführt.¹ Die Detectron-Netze greifen auf vier Variationen der ResNet-Architektur zurück um entsprechende ROIs zu generieren. Diese unterscheiden sich in den angegebenen Testergebnissen sowohl in Qualität als auch in der Ausführungszeit. Des weiteren stellt Facebook von jeder möglichen Kombinationsmöglichkeit zwei Varianten zur Verfügung, die sich in der Art des Trainings unterscheiden.

Ziel der weiteren Untersuchungen ist es daher festzustellen, welche der Netzkombinationen im vorliegenden Fall am erfolgreichsten ist. Dabei spielen sowohl die Generierung der ROIs als auch deren Klassifizierung eine Rolle.

2.2 Bilder

Für die zu untersuchenden Abbildungen gelten zwei Maßstäbe: Sie müssen sowohl technisch als auch inhaltlich für diese Aufgabe relevant sein. Viele Resultate der verfügbaren Datenbanken eignen sich aufgrund dieser Einschränkungen nicht. Im weiteren Verlauf wurden sechs Abbildungen untersucht, die in ihrer Art diesen Eigenschaften entsprechen. Eine größere Anzahl an Untersuchungsobjekten wäre möglich gewesen, erschien allerdings in diesem Fall nicht notwendig. Insgesamt sind auf den Abbildungen fünf Sarkophage zu sehen, einer davon aus verschiedenen Perspektiven. Darauf sind in Summe 78 Personen dargestellt, von denen 22 nicht auf einem Korpus, sondern auf dessen Deckel erscheinen. Eine größere Anzahl an Testobjekten hätte die teils manuelle

¹ Für das Problem der tiefen Schichten und dessen Lösung siehe [He16].

Validierung in der Zeit dieser Arbeit nicht möglich gemacht, daher wurde der Fokus auf eine möglichst breite Verteilung der Eigenschaften gelegt.

Viele der in den Datenbanken abgelegten Ressourcen zeigten sich von geringer technischer Qualität. Wie exemplarisch in der Abb. 2.5 sind es meist Digitalisate älterer schwarz-weiß Fotografien, die aufgrund der geringen Größe und Auflösung eine genaue Identifizierung aller Personen erschweren oder unmöglich machen. Auch wenn das Ergebnis nicht direkt ohne nötige Anpassungen auf andere Objekte übertragbar ist, so ist eine Einschränkung des Materials an dieser Stelle sinnvoll, da diese Arbeit die grundsätzliche Einsetzbarkeit künstlich neuronaler Netze auf Steinreliefs untersuchen soll und nicht dazu dient, bereits vorhandene Datensätze zu explorieren. Natürlich ist es aber auch von großer Bedeutung, mit vermeintlich schlechterem Material umgehen zu können.¹ Um einen generellen Nutzen zu demonstrieren, wird das Resultat einer möglichst objektiven Untersuchung im Anschluss auch auf weitere Sarkophage angewandt, deren Abbildungen nicht den hier beschriebenen Kriterien entsprechen.

2.2.1 Inhaltliche Auswahl

Der Sarkophag auf Abb. 3.1 steht in den Kapitolinischen Museen in Rom und zeigt sieben Personen, die ohne größere Überlappungen deutlich erkennbar sind. Alleine der Held Meleager ist gänzlich unverdeckt, die anderen Personen sind nur leicht durch andere Bildelemente verdeckt. Zwei der Figuren sind sitzend auf einem Pferd dargestellt, die restlichen stehen. Auffallend ist, dass nur Meleager selbst fast völlig unbekleidet ist, während alle weiteren Personen komplett bekleidet sind. Im Schnitt nehmen alle Figuren 80% der Bildhöhe ein, was im Vergleich zu den anderen Bildern sehr groß ist.

Der auf Abb. 3.3 dargestellte Sarkophag befindet sich im archäologischen Museum Eleusis und zeigt elf Personen. Neben einem sitzenden Reiter liegen zwei Teilnehmer der Szene auf dem Boden. Im Vergleich zu dem ersten Bild sind hier fast die Hälfte der Personen nur teilweise oder gar nicht bekleidet. Durch die größere Anzahl an Personen in der Szene sind auch mehr davon verdeckt. Von zweien sieht man nur den Kopf und je eine Hand, fünf weitere sind durch andere Bildelemente nur leicht verdeckt. Aufgrund des weiten Bildausschnitts, der noch den Sockel und die Abdeckung des Sarkophags zeigt, sind die Figuren mit im Schnitt 20% der Bildhöhe deutlich kleiner als im ersten Beispiel.

¹ Es ist bereits eine Feldstudie in Planung, bei der die genannten Bildkorpora gänzlich untersucht werden sollen. Diese Studie hat die Anwendung der hier beschriebenen Konzepte zum Ziel und soll die im Fazit beschriebene potentielle Anwendbarkeit nachweisen.

Die Abb. 3.4 zeigt nur das Fragment eines Reliefs aus dem archäologischen Museum Saloniki. Drei Personen sind dabei nur in Teilen vorhanden, wobei dies von einem einzelnen Kopf über einen Torso mit Arm und Bein bis zu einer Person reicht, der nur ein Arm fehlt. Eine Figur steht soweit im Hintergrund, dass sie fast vollständig verdeckt ist, die verbleibenden vier sind nur leicht beeinträchtigt. Mit 38% mittlerer Größe (ohne Einbeziehung des einzelnen Kopfes) sind die Darstellungen trotz des weiten Bildausschnittes vergleichsweise groß.

Der vierte Sarkophag auf Abb. 2.3, welcher in der Villa Doria Pamphili steht, zeichnet sich in der Untersuchung durch seinen Deckel aus. 15 Personen sind auf dem Korpus, weitere 22 auf dem Deckel dargestellt. Dadurch ergibt sich auch eine große Differenz der Personengrößen von 39% zu 16%. Die Szenerie ist zudem die umfangreichste auf den ausgewählten Abbildungen. Etwa die Hälfte der Personen sind hinter anderen verdeckt, neben Meleager sind dafür zwei weitere komplett im Vordergrund abgebildet. Ebenfalls knapp die Hälfte trägt Kleidung, die andere ist weitestgehend unbekleidet. In dieser Abbildung gibt es keine Reiter, nur eine Person liegt am Boden. Die Szenerie auf dem Deckel zeigt nicht den Kampf, sondern die Heimkehr des verletzten Helden. Neben den zwei Figuren an den Rändern sind noch zwei weitere auf ihren Pferden sitzend dargestellt. Der proportional übergroße Meleager wird liegend von einer Gruppe getragen. Auch auf dem Deckel ist ein Teil der Figuren neben Meleager unbekleidet, die Mehrheit jedoch wird in Kleidung dargestellt.

Die letzten beiden Abbildungen 4.2 und 2.4, welche für die Untersuchung ausgewählt wurden, zeigen einen Sarkophag, der in der Woburn Abbey in Bedfordshire ausgestellt ist. Unter den zehn dargestellten Personen sind ein Kind und ein Reiter zu sehen, von welchem abgesehen niemand sitzt oder liegt. Auffällig ist, dass neben dem Kind nur zwei Personen an den beiden Seiten unbekleidet dargestellt werden. Durch den fließenden Übergang zweier Szenen während und nach der Jagd, welche sich an dem Eber bzw. seinem Kopf ausmachen lassen, handelt es sich bei beiden Personen um Meleager selbst. Durch die Darstellungen auf der gesamten Höhe des Sarkophags und des recht genauen Bildausschnittes ergibt sich eine mittlere Figurengröße von 59%, die nur aufgrund des Kindes und der teilweise im Bodenbereich verdeckten Personen nicht noch größer ist. Auf der letzten Abbildung sind die fünf Figuren rechts des Reiters nochmal in einem deutlich flacheren Winkel aufgenommen, um die hohe räumliche Tiefe dieses Sarkophages zu zeigen. Im Vergleich zu einigen sehr flachen Reliefs sind die Figuren auf diesem weitestgehend freigestellt vom Hintergrund. Durch den kleineren Bildausschnitt haben die Personen eine Größe von durchschnittlich 74%.

Abbildungsrechte nicht vorhanden

Abbildung 2.3: Meleager-Sarkophag mit Deckel, Villa Doria Pamphili - Casino Belrespiro (arachne.dainst.org/entity/2888823)

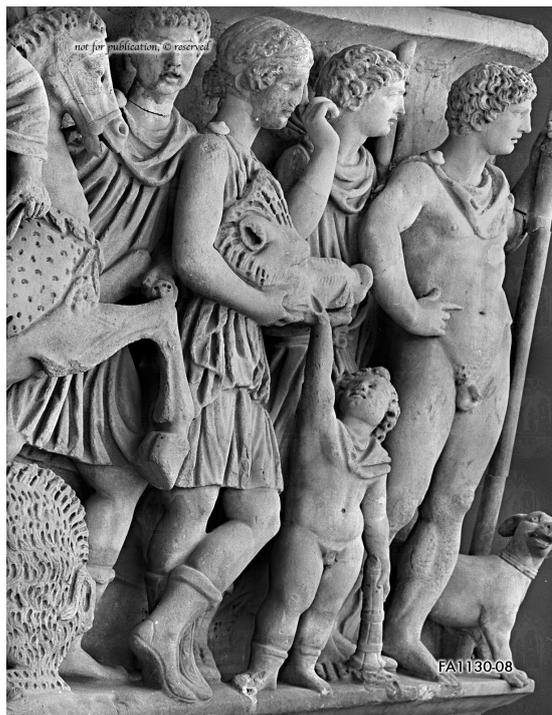


Abbildung 2.4: Meleager-Sarkophag, Bedfordshire, Woburn Abbey (arachne.dainst.org/entity/405092)

Die ausgewählten Abbildungen enthalten im Schnitt neun Personen mit unterschiedlichen Eigenschaften.¹ Neben den stehenden oder kämpfenden Akteuren sind vier in liegender Position dargestellt, ein Kind zeichnet sich auf zwei Abbildungen durch eine große Differenz in der Körpergröße aus, das letzte Relief ist nicht frontal sondern stark seitlich aufgenommen worden. Einige Personen sind sehr flach im Hintergrund dargestellt, wodurch nur einzelne Körperteile erkennbar sind. Ähnliche Einschränkungen finden sich auch bei einem Fragment, auf welchem zwei Personen nur noch in geringen Teilen zu erkennen sind. Auf einem Sarkophagdeckel sind zudem nochmal über 20 Personen deutlich kleinerer Größe dargestellt, deren Erkennungsrate möglicherweise aufgrund des geringen Verhältnisses zur Bildhöhe geringer ist. Mit dieser breiten Auswahl an Posen, Größen, Kontrasten und Bildwinkeln lässt sich die Qualität der Ausgabewerte gut einordnen.

2.2.2 Technische Auswahl

Um sich ausschließlich auf die formulierten Probleme der fehlenden Farben und übergreifenden Textur zu konzentrieren, müssen weitere potentielle Fehlerquellen und Hemmnisse

¹ In der Tab. 2.1 ist die Personenanzahl und deren durchschnittliche Größe in Relation zur Höhe der jeweiligen Abbildungen aufgelistet.

Abbildung	Anzahl Personen	Relative Größe
3.1	7	80%
3.3	11	20%
3.4	8	34%
2.3	78	25%
4.2	10	58%
2.4	5	74%

Tabelle 2.1: Übersicht über die Personenanzahl und deren Größe in Relation zur Höhe der Abbildung

vermieden werden. Abbildungen, auf denen die Personen kaum manuell zu identifizieren sind, stellen ein solches Hemmnis dar. Um dieses auszuschließen und eine genaue Lokalisierung zuzulassen, müssen die Abbildungen eine Mindestgröße und -auflösung besitzen. Die zuvor ausgewählten Abbildungen haben eine Mindestbreite von 1200 und eine Mindesthöhe von 700 Pixeln. Da die Auflösung allerdings zwischen 72 und 300 Dots per Inch (dpi) schwankt, ist eine Vergleichbarkeit der Größe in Pixeln nicht gegeben. Aus diesem Grund werden die Maße zunächst auf 100 dpi und dann auf metrische Einheiten umgerechnet. Dabei zeigt sich, dass die Abbildungen mindestens etwa 30 cm breit und 20 cm hoch sind. Die kleinste Abbildung ist 4.2 mit 766 cm² im Vergleich zur Abb. 3.1, welche 3532 cm² groß ist. Im Vergleich zu weiteren erhobenen Abbildungen aus der Census-Datenbank in Tabelle B.28 zeigt sich, dass die durchschnittliche Größe dort bei gerade einmal bei 181 cm² liegt, was einer Seitenlänge von etwa 18 mal 10 cm entspricht.¹ Exemplarisch ist hier die Abb. 2.5 zu sehen.

Neben der Größe wird eine weitere Vermutung untersucht, welche sich auf die syntaktische Beschaffenheit der Abbildungen bezieht. Werden Aufnahmen manuell begutachtet, so müssen diese ausreichend gut belichtet und der Kontrast so gewählt sein, dass sich die Figuren optisch abheben. Gleiche Grundvoraussetzungen könnte man demnach auch für eine maschinelle Auswertung annehmen. Eine generelle Argumentation über die Helligkeits- und Kontrastwerte der Bilder lässt sich jedoch nicht führen. Bei dem Vergleich der ausgewählten mit weiteren Bildern der Census-Datenbank fällt auf, dass sich diese Werte nicht signifikant unterscheiden. Die Helligkeit liegt im Schnitt bei 43%, was nur auf eine leichte Unterbelichtung hinweist. Der Kontrastumfang wird als Betrag berechnet, welcher sich aus den gemittelten Helligkeitswerten der hellen und dunklen Bildteile ergibt. Dabei wird das Histogramm an seinem Mittelwert, der Helligkeit, getrennt. Der Kontrastumfang liegt mit 39% allerdings ebenfalls auf einem Niveau, das regulären Fotografien entspricht und ist somit auch kein entscheidendes Maß für die technische Qualität. Zudem werden die Pixelwerte der Bilder für die Verarbeitung in neuronalen Netzen vorverarbeitet. Dabei wird eine Normalisierung des Histogramms

¹ Die Größen der Abbildungen können in den Tabellen B.27 und B.28 verglichen werden.

durchgeführt, welche die Bildhelligkeit mittelt und den Kontrast erhöht.¹ In der Folge sind die vorab gemessenen Werte der Bilder nicht mehr aussagekräftig.



Abbildung 2.5: Meleager-Sarkophag, Rom, Kapitolinische Museen (<http://census.bbaw.de/easydb/censusID=159869>)

¹ Für einen Vergleich von Methoden der Histogrammnormalisierung siehe [Vad11].

3 Vorgehen

Die Untersuchung ist von der Forschungsfrage motiviert, ob es mit Hilfe der CNNs möglich ist, Personen in Reliefs zu identifizieren und falls ja, ob sich die Posen der so ermittelten Personen ebenfalls auslesen und statistisch untersuchen lassen. Dazu werden im folgenden die ausgewählten Abbildungen von allen Ende-zu-Ende trainierten Netzen des Detectron Model Zoo verarbeitet und im ersten Schritt die Qualität der Object Proposals bzw. deren Bounding Boxes evaluiert. Für die anschließende Auswertung der Keypoint-Erkennungsleistung, werden im weiteren Verfahren diese Resultate von den spezialisierte Netzen weiterverarbeitet.

Initial wird grundsätzlich ein Farbbild vorausgesetzt, dessen drei Farbkanäle auch in den ersten Schritten separat voneinander untersucht werden. Gerade bei alten Fotografien und deren Digitalisaten ist jedoch aufgrund der monochromatischen Aufnahme kein Unterschied in den Kanälen vorhanden. Bei neueren Aufnahmen zeigt sich allerdings, dass auch hier die zusätzliche Farbinformationen nur bedingt hilfreich sind, da alle untersuchten Reliefs höchstens noch vereinzelte Reste der ursprünglichen Bemalung aufweisen. Die sichtbaren Farben sind daher abhängig vom verarbeiteten Material, also zum Beispiel dem entsprechenden Marmor und seiner Textur. Ein Kernel, welcher in den ersten Schichten auf einfache geometrische Formen spezialisiert ist, erkennt somit Strukturen im Material über einzelne Bildelemente und den Hintergrund hinweg. Es ist eine der Untersuchung vorausgehende Vermutung, dass dieser Umstand die Erkennungsleistung negativ beeinflussen könnte. In der Abb. 3.1 ist eine solche gleichbleibende Textur zu erkennen. Die dargestellten Figuren setzen sich hauptsächlich durch den Schattenwurf vom Hintergrund ab.

Ein weiteres mögliches Hemmnis ergibt sich aus dem Grundtraining, welches einer Nutzung oder einem spezialisierten Training vorausgeht. Um zum einen eine Vergleichbarkeit der Erkennungsrate zu schaffen und um die ursprünglich unstrukturierten Kernels in den Konvolutionsschichten auf eben jene Formen und Strukturen anzupassen, werden hierbei die Netze auf dem ImageNet- oder COCO-Datensatz trainiert.¹ Die Kernel sind in den ersten Schichten allgemein genug um daraus in tieferen Ebenen komplexe Elemente wie

¹ Der ImageNet-Datensatz ist unter [Den09] beschrieben, für den COCO-Datensatz siehe [Lin14]. Die Verwendung der Abbildungen zum Training der Detectron-Netze wird bei [Gir18] beschrieben.



Abbildung 3.1: Sarkophag mit der Jagd auf den calydonischen Eber, Rom, Kapitolinische Museen (Marie-Lan Nguyen / Wikimedia Commons)

Personen oder Kraftfahrzeuge zu erkennen. In der Art des Training liegt damit auch das zweite Problem: Die im Datensatz enthaltenen Personen sind grundsätzlich in eine heutige, natürliche Umgebung eingebettet.¹ Alle Netze sind daher auf weitestgehend bekleidete Personen mit natürlicher Hautfarbe angelernt. Ob sie auch in der Lage sind, eine Person auch als solche zu identifizieren, wenn es weder aufgrund der Farbgebung von Kleidung, Haut und Haaren dem Gelernten entspricht, ist zu prüfen.

Ein dritter Umstand, der jedoch nicht zu größeren Abweichungen in der Erkennungsleistung führen sollte, ist das häufige Fehlen von Körperteilen. In der künstlerischen Darstellung einiger Reliefs sind einige Personen stark überlappend im Hintergrund angeordnet, sodass teilweise nur ihr Kopf zu sehen ist. Bei einigen weiteren Abbildungen sind einzelne Elemente abgebrochen oder das ursprüngliche Motiv nur noch in Fragmenten erhalten. Allerdings verdecken sich auch auf den Aufnahmen von Straßenszenen viele Personen gegenseitig oder befinden sich räumlich hinter einem anderen Objekt. Die Netze sind daher auf Teilerkennungen bis zu einem gewissen Grad ausgelegt, der in den Reliefs nicht stärker zum tragen kommen sollte als auf natürlichen Aufnahmen.²

3.1 Personenerkennung

Der Ablauf sieht vor, zuerst mit allen Netzen Personen zu lokalisieren um sie dann nur mit den speziellen Netzen auf ihre Keypoints zu untersuchen. Diese Aufteilung entspringt der Grundannahme, dass die Netze, welche auf viele Klassen trainiert sind genauer zwischen verschiedenen Objekten unterscheiden können als solche, welche nur

¹ Für beispielhafte Abbildungen der Datensätze siehe [Den09] und [Lin14].

² In [He17] sind Abbildungen dargestellt, auf welchen ebenfalls größere und sich überlappende Personengruppen gezeigt werden.

auf die Erkennung von einzelnen Körperteilen ausgelegt sind.¹ Die so im ersten Schritt erkannten Personen können dann separat genauer betrachtet werden. Der Ablauf ist zweigeteilt, da vermutet wird, dass die Erkennungsleistungen von Personen und die der Keypoints eines Netzes variieren können. Ein mögliches Ergebnis kann daher die Nutzung von zwei unterschiedlichen Netzen für die jeweiligen Aufgaben sein.

Die sechs zu testenden Abbildungen werden von allen Netzen verarbeitet und die erkannten Objekte nach der Klasse „Person“ gefiltert. Im Ergebnis liefern die Netze eine Menge von Object Proposals, welche neben einer Klassenbezeichnung und der entsprechenden Bounding Box des Objekts auch einen Wahrscheinlichkeitswert enthalten. Diese Klassenwahrscheinlichkeit gibt an, wie sicher die Aussage des Netzes ist. Resultate dieser Ausführungen werden nach ihren Wahrscheinlichkeiten gruppiert und validiert. Diese ersten Daten geben zunächst nur die Anzahl der jeweiligen Object Proposals und deren Klassenwahrscheinlichkeiten an. Im generellen Überblick zeigt sich, dass rein aufgrund der Klassenwahrscheinlichkeit kein indirekter Rückschluss auf die Korrektheit möglich ist. Die Verteilung der Klassenwahrscheinlichkeiten in Abb. 3.2 zeigt zunächst eine deutliche Häufung im unteren Viertel, verläuft dann aber annähernd kontinuierlich bis zu einer kleinen Spitze im sehr sicheren Bereich über 90%. Insgesamt wurden von 37 Netzen etwa 5.400 Object Proposals für 78 dargestellte Personen generiert, was einer Quote von 187% entspricht. Umgekehrt müsste man die Ergebnisse unterhalb von 15% ausschließen, um eine exakte Deckungsquote zu erhalten.

Sehr deutlich in Tabelle B.1 ist zu erkennen, dass die Netze vom Typ RetinaNet und Keypoint über 50% mehr Object Proposals generieren, als die vom Typ Faster R-CNN und Mask R-CNN. Jedoch zeigt sich in der Übersicht auch, dass die RetinaNet-Netze mit ihren Ausgaben sehr unsicher sind. Der Schwerpunkt liegt hier bei unter 30% Wahrscheinlichkeit, im oberen Bereich über 80% sind gar keine Object Proposals

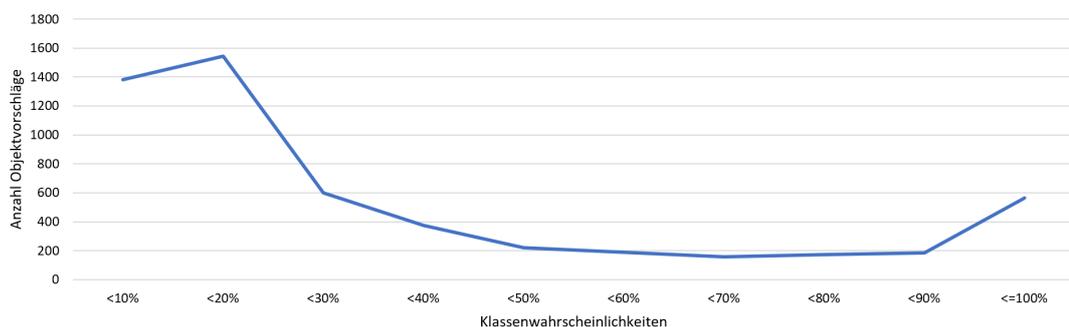


Abbildung 3.2: Übersicht der Klassenwahrscheinlichkeiten in zehn-Prozent-Schritten als Visualisierung der Daten aus Tab. B.1.

¹ Die Begriffe „Klasse“ und „Objekt“ sind angelehnt an die Bezeichnungen aus dem objektorientierten Programmierparadigma. Eine Klasse entspricht damit einer Kategorie von Dingen, hier z.B. „Person“, „Car“ oder „Trafficlight“. Das Objekt hingegen bezeichnet eine konkrete Instanz der Klasse.

vorhanden. Bei allen anderen Netzen zeigen sich in Relation sehr ähnliche Kurven, die zwischen den exponierten Rändern im Mittelfeld bei 60% zu einer Senke zulaufen. Aus dieser ersten Betrachtung lässt sich daher nur ein erstes Indiz auf die schlechtere Qualität der RetinaNet-Netze ziehen.

Alleine über die Anzahl der Object Proposals und deren Klassenwahrscheinlichkeiten lassen sich allerdings keine weiteren qualitativen Schlüsse ziehen. Um die Netzkategorien bzw. deren konkrete Implementierungen besser quantifizieren zu können, müssen die Ergebnisse validiert werden. In der Objekterkennung nutzt man dazu den Intersection over Union (IOU)-Wert, welcher einen berechneten Bildausschnitt mit einem vorher definierten vergleicht.¹ Dazu wird der Quotient aus den Flächeninhalten der Vereinigungs- und Schnittmenge gezogen. Stimmt die Bounding Box des Object Proposals mit der sogenannten Groundtruth überein, so sind beide Mengen gleichgroß und der Quotient dementsprechend bei 100%. Je nach Domäne wird ein IOU größer 50% als ausreichend angesehen, um präzisere Ergebnisse zu erhalten ist bei stark überlappenden Entitäten wie im Fall der Sarkophag eine strengere Auslegung mit 70% sinnvoll. Um diese Werte zu ermitteln, müssen alle Testbilder annotiert werden. Jede sichtbare Person wird dazu möglichst exakt eingerahmt und die Koordinaten der Boxen ausgelesen. Diese Annotation wie in Abb. 3.3 zu sehen, wird als Goldstandard der Validierung zu Grunde gelegt und bezeichnet sich in der Bildverarbeitung als Groundtruth.

Um die Resultate besser interpretieren zu können, werden im nächsten Schritt so alle Object Proposals unabhängig von ihrer Wahrscheinlichkeit gegen die Groundtruths validiert. Das statistische Maß des Recall gibt dabei die korrekten Object Proposals im Verhältnis zur Menge der vorhandenen Personen an. Die Precicion ist ein ähnlicher Wert, welcher allerdings in Relation zur Gesamtmenge der Object Proposals gemessen wird. Auf der anderen Seite setzt die False Discovery Rate (FDR) die inkorrekten Object Proposals (Fehler 1. Art) ebenfalls in Relation zur Gesamtmenge. In der Abb. 3.4 sind alle generierten Bounding Boxes zu sehen.² Deutlich zu erkennen ist, dass sehr viele Vorschläge gemacht werden, jedoch nur wenige eine erhöhte Klassenwahrscheinlichkeit besitzen.

Ein generelles Problem ist, dass sich Recall und False Discovery Rate gegenseitig beeinflussen. Allein durch eine große Anzahl Object Proposals, wie im Fall der RetinaNet-Netze, können viele Personen korrekt erkannt werden. Je mehr Object Proposals ein Netz produziert, desto höher ist aber auch die Quote der Fehler 1. Art, wodurch in der Gesamtbetrachtung die Qualität wiederum abnimmt. Die Precicion bringt durch

1 Bei Wettbewerben wie dem PASCAL VOC wird der IOU-Wert gemittelt als Qualitätsmaß der Einreichungen genutzt. Siehe dazu [Eve10] und [Eve15].

2 Die Annotationen entsprechen den Ergebnissen des Netzes 37697946. Der für die hervorgehobenen Bounding Boxes gewählte Schwellenwert von 60% ergibt sich aus den weiteren Untersuchungen.



Abbildung 3.3: Sarkophag mit Jagd auf den kalydonischen Eber, Eleusis, Archäologisches Museum (©Bildarchiv Foto Marburg / Konrad Helbig). Hier annotiert mit der Groundtruth.

den Bezug auf die Gesamtmenge etwas Abhilfe, vernachlässigt dabei aber die Anzahl der potentiell vorhandenen Personen. Mit dem F1-Score wird durch das harmonische Mittel aus Precision und Recall eine verlässliche Abwägung geschaffen, auf die sich alle weiteren Beurteilungen stützen. Da sowohl Precision als auch Recall möglichst hoch sein sollten, wird auch die Beurteilung besser, je höher der F1-Score ist.

Alle Maße werden durch einen Schwellenwert beeinflusst, der angibt ob ein Object Proposal angenommen wird oder nicht. Durch eine Grenze der Klassenwahrscheinlichkeiten verringert sich die Gesamtmenge der Object Proposals und dadurch auch die zu untersuchenden Metriken. Wie Abb. 3.2 zu sehen, würde schon eine Schwelle bei etwa 10% ausreichen, um die Gesamtmenge um ein Drittel zu reduzieren. Ohne jede Schwelle weisen die Keypoint-Netze in der Abb. 3.5 im Mittel den höchsten F1-Score von 29%, dicht gefolgt von den Mask R-CNN-Netzen mit 28% aus. Bei einer Erhöhung



Abbildung 3.4: Sarkophag-Fragment mit der Eberjagd des Meleager, Saloniki, Archäologisches Museum (©Bildarchiv Foto Marburg / Jan Gloc; Paul Haag). Hier annotiert mit generierten Object Proposals (Dicke Linie: Über 60% Klassenwahrscheinlichkeit).

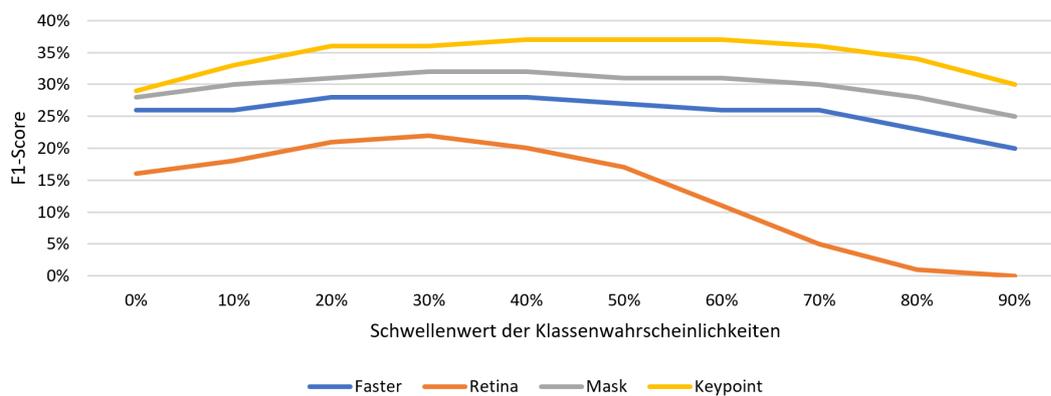


Abbildung 3.5: Übersicht der F1-Scores nach Netzarchitektur als Visualisierung der Tab. B.22

des Schwellenwertes in zehn Prozent Schritten zeigt sich bei allen Netzkategorien eine parabelförmige Verteilung des F1-Scores. Dabei steigen die Werte der Keypoint-Netze mit maximal 37% stärker an, als die der Kategorien Mask R-CNN und FASTER R-CNN mit 32% bzw. 28% im Maximum. Gut zu erkennen ist, dass die F1-Scores der RetinaNet-Netze sowohl deutlich niedriger starten, als auch früher und stärker abfallen. Das zeigt, dass trotz der vielen Object Proposals die Qualität dieser Netzkategorie insgesamt deutlich niedriger ist. Im Mittel aller Netze ist der F1-Score bei einem Schwellenwert von 30% am höchsten, die Keypoint-Netze erreichen ihr Maximum bei 40% Schwellenwert.

Die einzelne Betrachtung des Recalls zeigt in Tab. 3.1 ein ähnliches Bild. Im Verlauf der Werte über die verschiedenen Schwellen wird deutlich, dass die Netze der Keypoint-Architektur alle anderen übersteigen. Während die Mask R-CNN und FASTER R-CNN-Netze fast gleichauf liegen, fallen die Werte der RetinaNet-Netze schneller ab. Die FDRs der Architekturen von FASTER R-CNN, Mask R-CNN und Keypoint in Tab. B.23 sind annähernd identisch. Auffällig ist zunächst der Verlauf der RetinaNet-Kurve, welche deutlich höher startet, aber die anderen Kurven schon bei 20% Schwellenwert schneidet und weiter bei 70% gegen null fällt. Eine Erklärung findet sich hier in der andersartigen Verteilung der Klassenwahrscheinlichkeiten dieses Netztyps. Die vielen Object Proposals waren zu großen Teilen von geringer Wahrscheinlichkeit, was bei einer Erhöhung des Schwellenwerts zu einer stärkeren Reduzierung der Gesamtmenge führt, als bei den anderen Netzen. Die verbleibenden Object Proposals sind dann zwar insgesamt korrekter, allerdings auch von geringerem Umfang, was die Diskrepanzen zwischen Recall und FDR erklärt.

Bei der Betrachtung aller Recall-Werte ohne Schwellenwert in der Tab. B.2 fällt auf, dass von allen Netzen nur zwei in der Lage sind, etwas mehr als 50% der Personen zu erkennen. Der Grundannahme nach liegt das wahrscheinlich an den farblichen und textuellen Herausforderungen der Bilddomäne, jedoch wurde ein möglicher Einflussfaktor in der Untersuchung noch nicht ausgeschlossen. In der Abb. 2.3 ist neben dem großen Relief auf dem Sarkophagkorpus auch der Deckel zu sehen, auf welchem 22, in Relation zur Bildgröße sehr kleine Personen dargestellt sind. Durch die deutlich geringere Anzahl an Bildpunkten auf der Oberfläche sind diese möglicherweise schwieriger zu erkennen, als die großen Hauptmotive. Zur Untersuchung dieses Einflussfaktors wurde eine zweite

Typ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Faster	31%	27%	25%	23%	21%	19%	18%	17%	15%	12%
Retina	28%	24%	17%	15%	12%	9%	6%	3%	0%	0%
Mask	35%	32%	29%	27%	26%	24%	23%	21%	19%	16%
Keypoint	48%	43%	39%	34%	32%	30%	28%	27%	24%	20%

Tabelle 3.1: Recall nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11

Evaluierung der Daten unter Ausschluss dieser Figuren ausgeführt und in die weiteren Betrachtungen mit einbezogen.¹

Alles in Allem zeigt sich, dass die Netze der Keypoint-Architektur aufgrund der Metriken aus Tab. B.22 am besten für die Aufgabe der Personenerkennung geeignet sind. Daher ist eine detailliertere Untersuchung der konkreten Netze diesen Typs sinnvoll. Mit einem maximalen und mittleren F1-Score von 42% bzw. 38% hebt sich das Netz mit der ID 37732415 in der Tab. 3.2 leicht von den anderen ab. Während im Mittel der F1-Score bei den Keypoint-Netzen erst ab 40% Schwellenwert seinen Höhepunkt erreicht, zeigt sich das Maximum hier schon bei 20%. Der Recall in Tab. B.25 wird dadurch nur um elf Prozent verringert, während er bei einem höheren Schwellenwert um weitere zehn Prozent fallen würde. Initial konnte dieses Netz mit 53% etwa die Hälfte aller Personen auf den Sarkophagen identifizieren, nach Begrenzung durch den Schwellenwert sind es noch 42%.

Weniger klar wird die Übersicht in Tab. 3.3, wenn man die Personenmenge um die kleinen Figuren des Sarkophagdeckels verringert. Zwar haben die Keypointnetze laut Tab. B.24

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	30%	34%	34%	31%	34%	33%	33%	29%	27%	24%
37697714	28%	33%	36%	36%	39%	39%	37%	38%	35%	33%
37697946	26%	29%	35%	37%	37%	40%	41%	38%	33%	32%
37698009	26%	29%	34%	36%	36%	37%	38%	37%	35%	32%
37732318	32%	37%	40%	39%	40%	40%	37%	38%	36%	29%
37732355	27%	32%	34%	36%	36%	37%	34%	35%	31%	26%
37732415	35%	36%	42%	40%	40%	38%	39%	39%	39%	33%
37792158	29%	31%	36%	35%	36%	37%	36%	35%	34%	29%

Tabelle 3.2: F1-Score der Keypointnetze nach Schwellenwert als Ausschnitt aus den Tabellen B.2 bis B.11

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	25%	30%	32%	28%	31%	32%	32%	30%	27%	25%
37697714	23%	28%	30%	35%	39%	39%	39%	40%	39%	39%
37697946	24%	28%	35%	40%	41%	46%	48%	45%	42%	41%
37698009	24%	28%	33%	38%	40%	40%	42%	41%	42%	41%
37732318	27%	32%	36%	38%	39%	38%	39%	40%	40%	38%
37732355	22%	28%	31%	32%	35%	37%	36%	37%	37%	33%
37732415	31%	34%	40%	39%	42%	42%	44%	44%	47%	40%
37792158	26%	29%	35%	35%	38%	39%	40%	39%	38%	35%

Tabelle 3.3: F1-Score der Keypointnetze nach Schwellenwert bei gekürzter Personenzahl als Ausschnitt aus den Tabellen B.12 bis B.21

¹ Im weiteren Verlauf wird für einige Tabellen die Bezeichnung „gekürzte Personenzahl“ verwendet. Darunter sind die Untersuchungen bei Ausschluss der Figuren auf dem Sarkophagdeckel zu verstehen.

weiterhin den höchsten F1-Score, im Detail ergeben sich aber wichtige Verschiebungen. Das Netz mit der ID 37732415 hat wie zuvor den höchsten mittleren F1-Score von 40%, also sogar zwei Prozentpunkten mehr als bei der Betrachtung aller Figuren, jedoch erreicht ein anderes Netz das absolute Maximum. Bei einem Schwellenwert von 60% liegt der F1-Score des Netzes 37697946 bei 48%. Dieses hat ebenfalls bei dem angegebenen Schwellenwert in Tab. B.26 den höchsten mittleren Recall von 43%.

Die unterschiedlichen Ergebnisse der Netze in Abhängigkeit der Personengruppen zeigt, dass es nicht per se das eine beste Netz zur Personenerkennung gibt, sondern es vom konkreten Einzelfall, hier also der relativen Figurengröße abhängt. Genauso unterscheiden sich die Ergebnisse im Detail je nachdem, ob für die angewandte Forschungsfrage eine hohe Trefferquote oder eine geringe Fehlerquote, jeweils zu Lasten des anderen Wertes, gewünscht ist. Die hier getroffene Auswahl gewichtet mit dem F1-Score beide Aspekte gleichgewichtig.

3.2 Posenerkennung

Der zweite Teil der Untersuchung besteht aus der Extraktion von Keypoints aus den Object Proposals. Diese sind ebenfalls Teil der Netzresultate, sofern die Netze darauf ausgelegt sind. Die schon genauer betrachteten Keypoint-Netze sind Derivate der Mask R-CNN-Architektur, welche neben den Personen selbst auch deren wichtigsten Merkmale zur Posenbestimmung lokalisieren. Es ist nicht grundsätzlich davon auszugehen, dass ein Netz sowohl die Personen als auch deren Keypoints gleichermaßen gut erkennt, daher wurden im weiteren Vorgehen die Object Proposals des Netzes 37697946 mit der ermittelten Schwelle von 60% als Ausschnitte der Abbildungen separat weiter untersucht.

Insgesamt können die Keypoint-Netze 17 Punkte aus einer Person extrahieren. Neben der Nase sind dies die Ohren, Augen, Schultern, Ellenbogen, Handgelenke, Knie und Fußgelenke sowie beide Seiten der Hüfte. Zur Verdeutlichung sind in Abb. 3.6 alle Punkte auf einer Darstellung von Atalante aus Abb. 2.3 annotiert und entsprechend verbunden. Um qualitative Rückschlüsse auf die Erkennungsleistung der Netze zu ziehen, müssen diese Punkte einzeln validiert werden. Ein erkannter Punkt kann in der Abbildung an der richtigen oder falschen Stelle lokalisiert werden. Ein nicht erkannter Punkt kann auf der Abbildung nicht vorhanden sein oder wurde fälschlich nicht erkannt. Zudem ist es möglich, dass ein falsch lokalisierter Punkt an der falschen Stelle liegt oder auf der Abbildung nicht vorhanden ist. Ähnlich wie es für die Object Proposals die Klassenwahrscheinlichkeiten gibt, so werden auch die Keypoints mit Wahrscheinlichkeiten ausgegeben. In Abhängigkeit eines Schwellenwertes kann es daher zu veränderten Ergebnissen kommen. Ein korrekt erkannter Punkt wird unterhalb der Schwelle zu einem inkorrekt nicht erkannten Punkt. Ein falsch lokalisierter Punkt wird ebenfalls

zu einem nicht erkannten Punkt, jedoch kann dieser in der Abbildung vorhanden oder nicht vorhanden sein.

Eine Evaluierung unterscheidet sich durch diese vier Klassen auch von der Evaluierung der Object Proposals. Hier konnte nur unterschieden werden, ob deren Bounding Box korrekt oder inkorrekt ist bzw. welche Personen in den Abbildungen nicht lokalisiert wurden. Es gab allerdings keine Klasse „korrekt nicht erkannt“, wie es sie bei den Keypoints gibt. Insgesamt hat das ausgewählte Netz etwa 40 Bildausschnitte generiert. Durch die Untersuchung von acht Netzen führt dies zu über 5.400 Punkten, die validiert werden müssen. Zum aktuellen Zeitpunkt gibt es für diese Punkte noch keine Groundtruth-Annotationen, weswegen eine Validierung manuell erfolgen muss. Um die Gesamtmenge auf eine Größe zu reduzieren, deren Ergebnisse weiterhin aussagekräftig, aber in der Zeit realisierbar sind, wird die Menge der Bildausschnitte reduziert. Die etwa 40 Ausschnitte teilen sich etwa zu gleichen Teilen auf korrekte und inkorrekte Ergebnisse auf. Auch wenn es bei einer angewandten Untersuchung nicht zu diesem Filter kommen kann, werden nur die korrekten Ausschnitte weiter betrachtet. Mit diesen 24 Elementen, welche in den Abbildungen A.1 bis A.24 dargestellt sind, kommt man auf 3.264 zu untersuchende Punkte.

Diese werden zunächst ohne Einschränkungen, also bei einem Schwellenwert von 0% manuell validiert. Technisch werden dafür die Zahlenwerte 0 bis 3 genutzt, wobei 0 für korrekt nicht erkannte, 1 für korrekt erkannte, 2 für inkorrekt erkannte und 3 für inkorrekt nicht erkannte Punkte steht. Die inkorrekt erkannten Punkte (2) bezeichnen sich daher als Fehler erster Art, die inkorrekt nicht erkannten Punkte (3) als Fehler zweiter Art. Diese Erhebungen sind in den Tabellen B.38 bis B.45 für jedes Netz separat dargestellt. Für die Validierung der jeweiligen Schwellenwerte in zehn Prozent Schritten werden die Kategorien in Abhängigkeit der Wahrscheinlichkeiten fortgeschrieben. Diese sind als direkte Netzresultate in den Tabellen B.30 bis B.37 aufgelistet und dienen so als Grundlage für die Erhebung der Metriken bei höheren Schwellenwerten. Nur die Fehler erster Art müssen manuell geprüft werden, da sie sich je nach Präsenz des Punktes in ein korrektes Ergebnis oder einen Fehler zweiter Art verändern. Die Werte werden nach Netz und Schwellenwert zusammengefasst, wobei sich dann je 408 Punkte pro Netz in die vier Kategorien aufteilen. Der Recall ist somit der Quotient aus richtig erkannten Keypoints zu allen in der Abbildung vorhandenen. Die Precision im Unterschied dazu, nimmt als Relation alle vom Netz erkannten Punkte, also die korrekten und die Fehler erster Art. Beide Metriken berücksichtigen jedoch nicht die korrekt nicht erkannten Punkte, die genauso wie die korrekt erkannten Punkte eine positive Leistung des Netzes sind. Somit ist auch der F1-Score nicht genauso aussagekräftig, wie er noch für die Validierung der Object Proposals war. Um beide korrekten Kategorien in Relation zu allen 408 Punkten zu setzen wird die Accuracy als Metrik benötigt. In ihr werden die beiden positiven Werte gleichgewichtig betrachtet. Für einen späteren Einsatz kann

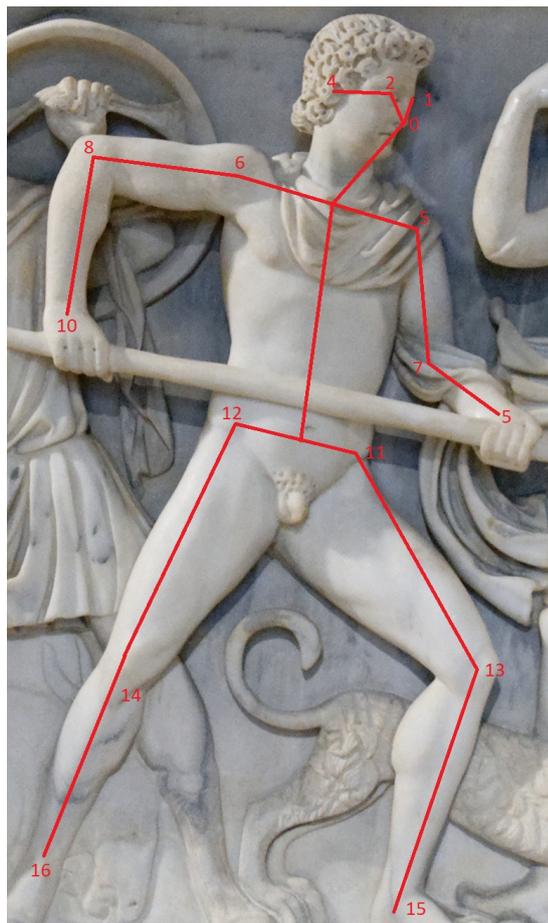


Abbildung 3.6: Mit Keypoints annotierte Abb. A.1

allerdings auch ein hoher Recall sinnvoller als eine hohe Accuracy sein, je nachdem ob durch eine manuelle oder automatische Weiterverarbeitung die Fehler erster oder zweiter Art kritischer zu betrachten sind.

Auffällig in Tab. B.47 ist, dass sich durch einen hohen Schwellenwert der Fehler erster Art fast gänzlich auflöst, was schon ab 60% zu einer Precision von über 90% führt. Die Netze 37697547 und 37792158 erreichen beide eine mittlere Precision von 86%, jedoch mit deutlichen Unterschieden in der Erkennungsleistung. Hier erreicht das erste Netz zuletzt nur noch 17 korrekt lokalisierte Punkte, während das zweite noch auf 45 kommt.¹ Natürlich ist die Menge der korrekt erkannten Punkte ohne jeden Schwellenwert am höchsten. Maximal werden in Tab. B.46 88% Recall vom Netz 37732355 erreicht. Es gehört zusammen mit den Netzen 37732318, 37792158 und 37732415 auch zu denen, die auch unter der höchsten Schwelle die besten Ergebnisse liefern. Das in diesem Szenario

¹ Die Werte sind berechnete Ergebnisse aus den Tabellen B.38 und B.30 für das Netz 37697547 bzw. B.45 und B.37 für das Netz 37792158.

der F1-Score nicht genauso aussagekräftig ist zeigt sich auch in der Tab. B.48. Anders als bei den Bounding Boxes zeigt sich hier nur eine monoton fallende Werte.

Die Accuracy in Tab. 3.4 zeigt durch die Berücksichtigung der korrekt nicht erkannten Punkte einen leichten Bogen, wenn auch nicht bei allen Netzen. Begründet wird dies durch die Anzahl der korrekt nicht erkannten Punkte, die im Mittel von 44 auf 103 steigen, während die korrekt erkannten Punkte von 212 um das dreifache auf 35 fallen.¹ Ein möglichst hohe Accuracy wird somit bei einem relativ niedrigen Schwellenwert von 10%-20% erreicht. Das Maximum liegt mit 68% im Netz 37792158 bei einer Schwelle von 20%, was bedeutet, dass zwei Drittel aller Punkte korrekt lokalisiert werden. Mit dem Netz 37697946, welches für die Bildausschnitte genutzt wird, ist ein etwas geringerer Wert von 66% sowohl bei 10% als auch bei 20% Schwellenwert messbar. Dadurch wird deutlich, dass wie vermutet das beste Netz zur Personenerkennung nicht auch das beste Netz zur Posenbestimmung ist. Der geringe Unterschied von nur 2% rechtfertigt aber weiterführende Überlegungen.

Bisher werden die Personen und deren Keypoints in zwei separaten Schritten erkannt. Durch die sehr ähnlichen Erkennungsleistungen der jeweils stärksten Netze stellt sich die Frage, ob die Keypointerkennung auch mit geringerem technischen Aufwand bei gleichbleibender Qualität in einem Schritt möglich ist. Dazu werden die Resultate des ersten Durchlaufs der Netze 37697946, 37732415 und 37792158 validiert.² Bei dieser dritten Untersuchung steht nur die Precicion im Vordergrund, da eine weiterführende Betrachtung vollständig annotierte Abbildungen benötigt. So ist die Groundtruth nicht gänzlich gegeben, weswegen nur die erkannten Punkte in Relation gesetzt werden konnten. Die Schwellenwerte werden für jedes Netz individuell eingestellt. Die Schwelle der Klassenwahrscheinlichkeiten wird in Abhängigkeit der jeweils besten F1-Scores

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	63%	66%	61%	59%	53%	48%	42%	37%	34%	30%
37697714	63%	65%	63%	62%	60%	54%	48%	41%	37%	34%
37697946	65%	66%	66%	62%	58%	50%	44%	37%	33%	30%
37698009	59%	62%	64%	59%	57%	53%	44%	39%	34%	31%
37732318	65%	67%	67%	65%	59%	55%	50%	45%	40%	36%
37732355	66%	67%	67%	65%	61%	55%	51%	46%	41%	36%
37732415	66%	65%	65%	65%	63%	59%	53%	47%	41%	38%
37792158	66%	68%	68%	65%	61%	56%	49%	43%	39%	37%

Tabelle 3.4: Accuracy der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.

- 1 Für diese Aussage wurden die Punkte aus den Tabellen B.38 bis B.45 mit ihren dazugehörigen Wahrscheinlichkeiten aus den Tabellen B.30 bis B.37 bei einem Schwellenwert von 0% bzw. 90% gezählt und gemittelt.
- 2 Die gewählten Netze zeigten in den bisherigen Untersuchungen der Personenerkennung bei voller und reduzierter Anzahl, sowie bei der Keypointerkennung die jeweils besten Resultate.

gewählt, für die Keypoints entsprechend der Accuracy. Im Resultat zeigt sich jedoch, dass die Precision in Tab. 3.5 signifikant niedriger ist als in dem zweistufigen Verfahren in Tab. B.47. Eine Erklärung für dieses Verhalten bzw. die Entwicklung eines optimierten einstufigen Verfahren ist somit weiterer Forschungsbedarf.

Abbildung	37697946	37732415	37792158
3.1	81%	88%	85%
3.3	50%	50%	58%
3.4	63%	48%	66%
2.3	65%	63%	77%
4.2	70%	66%	69%
2.4	57%	64%	69%

Tabelle 3.5: Precision im Ein-Schritt-Verfahren je Netz und Abbildung

4 Ergebnis und Zusammenfassung

Schlussendlich stellt sich die Frage, ob die getätigten Untersuchungen und somit das Proof-of-Concept einen Mehrwert für die Klassische Archäologie bietet. Sogar unter der Prämisse, ohne domänenspezifisches Training nur eine Auswahl aussichtsreicher Netze zu treffen, lässt sich die Frage positiv beantworten. Die finalen Ergebnisse zeigen deutlich, dass die moderne Bildverarbeitung ein gutes Werkzeug für die Arbeit der Forscherinnen und Forscher sein kann. Mit Hilfe der Keypoints können z.B. umfangreiche Bildkorpora schnell und effizient durchsucht werden um ähnlich abgebildete Personen oder Personengruppen zu identifizieren. Dies könnte bei der Zusammenführung von Datenbanken mit unterschiedlichen Abbildungen des gleichen Objekts genauso hilfreich sein wie bei der Materialrecherche für andere Forschungsarbeiten.

Es zeigt sich deutlich, dass das Bildmaterial vorab nicht aufbereitet werden muss. Damit beantwortet sich auch die Frage, ob vorhandene Korpora direkt nutzbar sind. Für einen produktiven Einsatz künstlich neuronaler Netze im Bereich der Personenerkennung ist jedoch ein weiterführendes Training nötig, um die Erkennungsraten weiter zu steigern. Dafür bedarf es einiger Abbildungen, die mit der Groundtruth annotiert werden.

4.1 Anwendbarkeit

Die Einschränkung auf die untersuchten sechs Testbilder galt, da diese aufgrund von Größe, Auflösung und Inhalt als besonders geeignet erschienen. Nichts desto trotz besteht ein weiterführender Nutzen der Technologie nur dann, wenn sie sich auch auf reguläre Bildressourcen anwenden lässt. Aus diesem Grund wird mit den Erkenntnissen über die passenden Netze und Schwellenwerte Bilder der Census-Datenbank ebenfalls ausgewertet. Die Resultate sind auch hier im ersten Schritt sehr vielversprechend. Trotz der im Durchschnitt geringen Größe der 52 Abbildungen können 192 Object Proposals erhoben werden, von denen auch ohne genaue Validierung viele korrekt erscheinen.¹ Im zweiten Durchgang ist die Erkennungsleistung der Keypoints jedoch deutlich geringer als erwartet. Spätestens hier scheint die Größe der Ausschnitte von den ursprünglich

¹ Beispielhaft sind in der Abb. 4.1 die Bounding Boxes annotiert dargestellt. Nur zwei der 15 Personen werden bei dem angelegten Schwellenwert nicht korrekt erkannt.

schon kleinen Bildern zu gering zu sein, um die Personenposes noch bestimmen zu können. Möglicherweise erlaubt es auch hier die Bildverarbeitung, zum Beispiel durch interpolierte Vergrößerungen, die Abbildungen vorab aufzubereiten.

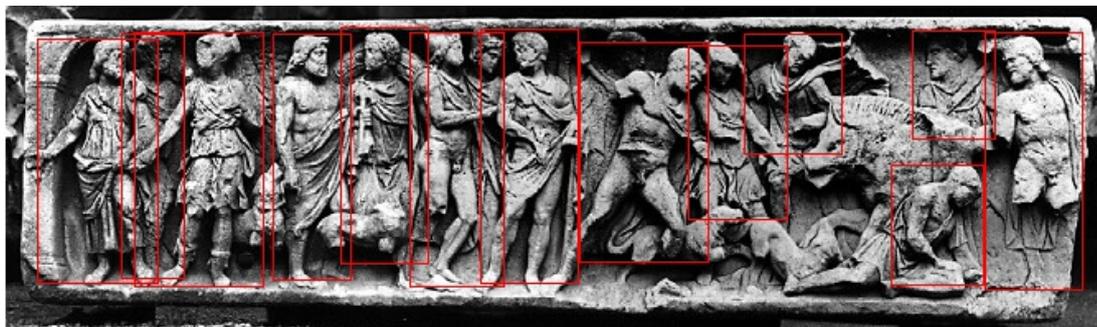


Abbildung 4.1: Mit generierten Object Proposals annotierte Abb. 2.5

4.2 Fazit

Im Ergebnis zeigt sich, dass es deutliche Unterschiede in der Leistung der verschiedenen Netzarchitekturen gibt. Schon von Beginn an verdeutlichte sich, dass die Netze vom Typ RetinaNet eine sehr unterdurchschnittliche Leistung erbrachten. Diese Erkenntnis bestätigt sich mit den validierten Ergebnissen der Personenerkennung, bei denen diese Netze zwischen 10% und 30% schlechtere F1-Scores erreichen als die Keypoint-Netze.

Die weiteren Unterschiede zwischen den Netztypen sind weniger signifikant, was an der sehr ähnlichen Funktionsweise liegt. Bei fast allen Schwellenwerten sind die Resultate der Keypoint-Netze am besten. Metriken von 42% und 48% F1-Score für die gesamte bzw. gekürzte Personenmenge weisen sehr deutlich auf ein großes Potential im praktischen Nutzen hin. Mit bis zu 59% Recall werden bereits über die Hälfte aller auf den Abbildungen der Reliefs dargestellten Personen korrekt erkannt. Bei der weiterführenden Bestimmung der Keypoints werden mit 68% Accuracy sogar zwei Drittel aller Punkte korrekt zugeordnet. Mit diesem Ergebnis lassen sich Posen bereits effektiv vergleichen. Zudem wird dadurch auch klar, dass die eingangs vermuteten Hemmnisse das Proof-of-Concept nicht verwarfen. Die von Facebook angegebenen Leistungsmerkmale von rund 55% Precision bei der Generierung von Object Proposals und 65% bei der Keypoint-Erkennung werden hier ebenfalls erreicht, wenngleich sich die Metriken aufgrund unterschiedlicher Bemessungsgrundlage nicht exakt vergleichen lassen. Die vermuteten negativen Einflüsse der gleichmäßigen Textur und Farblosigkeit der Reliefs auf den Abbildungen werden demnach durch die hohe Abstraktionsfähigkeit der Netze kompensiert. Einzig die hohe Dichte an sich überlappenden Personen führt häufiger zu fehlerhaften Zuordnungen von Körperteilen. Gerade bei Fragmenten können so die Körperhälften zweier Personen nicht korrekt unterschieden werden.

Weiterhin ist es ein Resultat, dass das zweistufige Vorgehen der getrennten Erkennung von Bounding Boxes und Keypoints sehr sinnvoll ist. Im Überblick aller Metriken ist deutlich geworden, dass die gleiche Netzarchitektur für beide Anforderungen am besten geeignet ist, dennoch unterscheiden sich die Werte der Precision zwischen dem ein- und zweistufigen Verfahren klar. In einem einzelnen Durchgang waren maximal nur 71% erreichbar, während bei der Anwendung gleicher Schwellenwerte im getrennten Durchlauf 81% möglich sind. Die initiale Vermutung, dass es Unterschiede in der maximalen Erkennungsleistung von Bounding Boxes und Keypoints je Netz gibt wurde somit bestätigt.

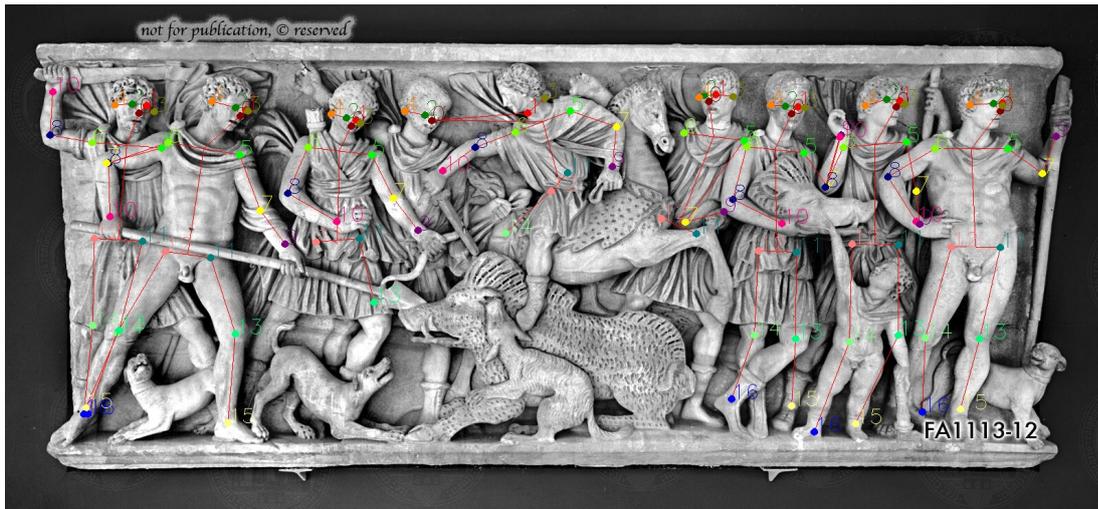


Abbildung 4.2: Meleager-Sarkophag, Bedfordshire, Woburn Abbey (arachne.dainst.org/entity/405082). Hier annotiert mit generierten Keypoints.

4.3 Weiterer Forschungsbedarf

Im Rahmen dieser Masterthesis konnten final nicht alle Fragestellungen geklärt werden. Nach der Erhebung aller Keypoints steht deren weitere Analyse noch aus. Über die räumliche Relation der Punkte zueinander können in einem nächsten Schritt die Personenposen untereinander verglichen werden. Eine detailliertere Untersuchung der Gesichter ist ebenfalls noch offen. Mit der Lokalisierung von Nase, Augen und Ohren allein ist kein Rückschluss auf weitere persönliche Merkmale wie Geschlecht, Alter oder Emotionen möglich. Für deren Auswertung bedarf es weit mehr Punkte. Allerdings kann schon allein über die Gerade von Ohr zu Nase die Kopfposition und damit eine ungefähre Blickrichtung festgestellt werden. In Kombination mit der zumeist korrekten Erkennung von linker und rechter Körperhälfte ist somit eine szenische Gruppierung der Personen denkbar.

Literaturverzeichnis

- [Aba15] ABADI, Martín; AGARWAL, Ashish; BARHAM, Paul; BREVDO, Eugene; CHEN, Zhifeng; CITRO, Craig; CORRADO, Greg S.; DAVIS, Andy; DEAN, Jeffrey; DEVIN, Matthieu; GHEMAWAT, Sanjay; GOODFELLOW, Ian; HARP, Andrew; IRVING, Geoffrey; ISARD, Michael; JIA, Yangqing; JOZEFOWICZ, Rafal; KAISER, Lukasz; KUDLUR, Manjunath; LEVENBERG, Josh; MANÉ, Dandelion; MONGA, Rajat; MOORE, Sherry; MURRAY, Derek; OLAH, Chris; SCHUSTER, Mike; SHLENS, Jonathon; STEINER, Benoit; SUTSKEVER, Ilya; TALWAR, Kunal; TUCKER, Paul; VANHOUCHE, Vincent; VASUDEVAN, Vijay; VIÉGAS, Fernanda; VINYALS, Oriol; WARDEN, Pete; WATTENBERG, Martin; WICKE, Martin; YU, Yuan und ZHENG, Xiaoqiang: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (2015), URL <https://www.tensorflow.org/>, software available from tensorflow.org
- [And84] ANDREAE, Bernard.: Bossierte Porträts auf römischen Sarkophagen. *Marburger Winckelmann-Programm, 1984.1984* (1984): S. 109–128, URL <https://zenon.dainst.org/Record/000444642>
- [Aub14] AUBRY, Mathieu; RUSSELL, Bryan C und SIVIC, Josef: Painting-to-3D model alignment via discriminative visual elements. *ACM Transactions on Graphics (ToG)* (2014), Bd. 33(2): S. 1–14
- [Bao18] BAO, W.; YANG, Y.; LIANG, D. und ZHU, M.: Human Pose Estimation Based on Step Deep Convolution Neural Network, in: *11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, S. 1–6
- [Cai15a] CAI, H.; WU, Q. und HALL, P.: Beyond Photo-Domain Object Recognition: Benchmarks for the Cross-Depiction Problem, in: *IEEE International Conference on Computer Vision Workshop (ICCVW)*, S. 74–79
- [Cai15b] CAI, Hongping; WU, Qi; CORRADI, Tadeo und HALL, Peter: The Cross-Depiction Problem: Computer Vision Algorithms for Recognising Objects in Artwork and in Photographs. *arXiv preprint arXiv:1505.00110* (2015), URL <http://arxiv.org/abs/1505.00110>
- [Cao18] CAO, Zhe; HIDALGO, Gines; SIMON, Tomas; WEI, Shih-En und SHEIKH, Yaser: OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE transactions on pattern analysis and machine intelligence*

- (2018), URL <http://arxiv.org/abs/1812.08008>
- [Cro14] CROWLEY, Elliot J und ZISSERMAN, Andrew: The state of the art: Object retrieval in paintings using discriminative regions (2014)
- [Cro15] CROWLEY, E.; PARKHI, O. und ZISSERMAN, A.: Face Painting: querying art with photos, in: *British Machine Vision Conference 2015*, S. 651–6513
- [Cro16] CROWLEY, E. und ZISSERMAN, A.: The Art of Detection, in: Gang Hua und Hervé Jégou (Herausgeber) *Computer Vision – ECCV 2016 Workshops*, Lecture Notes in Computer Science, Springer International Publishing, S. 721–737
- [Den09] DENG, J.; DONG, W.; SOCHER, R.; LI, L.-J.; LI, K. und FEI-FEI, L.: ImageNet: A Large-Scale Hierarchical Image Database, in: *IEEE conference on computer vision and pattern recognition*, Ieee, S. 248–255
- [Dol13] DOLLÁR, Piotr: Generating Object Proposals, Webseite (2013), online erhältlich unter <https://pdollar.wordpress.com/2013/12/22/generating-object-proposals/>; abgerufen am 30. November 2020.
- [Eve10] EVERINGHAM, M.; VAN GOOL, L.; WILLIAMS, C. K. I.; WINN, J. und ZISSERMAN, A.: The Pascal Visual Object Classes (VOC) Challenge. *International Journal of Computer Vision* (2010), Bd. 88(2): S. 303–338
- [Eve15] EVERINGHAM, M.; ESLAMI, S. M. A.; VAN GOOL, L.; WILLIAMS, C. K. I.; WINN, J. und ZISSERMAN, A.: The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision* (2015), Bd. 111(1): S. 98–136
- [FFL17] FEI-FEI LI, Serena Yeung, Justin Jones: CS231n: Convolutional Neural Networks for Visual Recognition, Lecture 11: Detection and Segmentation, Webseite (2017), URL http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf, online erhältlich unter http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture11.pdf; abgerufen am 08. November 2020.
- [Gei13] GEIGER, Andreas; LENZ, Philip; STILLER, Christoph und URTASUN, Raquel: Vision meets Robotics: The KITTI Dataset. *International Journal of Robotics Research (IJRR)* (2013)
- [Gir14] GIRSHICK, Ross; DONAHUE, Jeff; DARRELL, Trevor und MALIK, Jitendra: Rich feature hierarchies for accurate object detection and semantic segmentation, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, S. 580–587
- [Gir15] GIRSHICK, Ross: Fast R-CNN, in: *Proceedings of the IEEE international conference on computer vision*, S. 1440–1448
- [Gir18] GIRSHICK, Ross; RADOSAVOVIC, Ilija; GKIOXARI, Georgia; DOLLÁR, Piotr und HE, Kaiming: Detectron, <https://github.com/facebookresearch/detectron> (2018)

- [Grä19] GRÄTZ, Fabio M.: How to visualize convolutional features in 40 lines of code, Webseite (2019), online erhältlich unter <https://towardsdatascience.com/how-to-visualize-convolutional-features-in-40-lines-of-code-70b7d87b0030>; abgerufen am 19. August 2019.
- [Gre18] GREIS, Friedhelm: Wie Studenten autonome Autos schlau machen, Webseite (2018), online erhältlich unter <https://www.golem.de/news/neuronale-netze-wie-studenten-autonome-autos-schlau-machen-1807-135579.html>; abgerufen am 13. März 2019.
- [Han17] HANSEN, Sven: Mercedes: Was neuronale Netze im Auto leisten können und warum KI vielleicht scheitert, Webseite (2017), online erhältlich unter <https://heise.de/-3588654>; abgerufen am 13. März 2019.
- [He16] HE, Kaiming; ZHANG, Xiangyu; REN, Shaoqing und SUN, Jian: Deep residual learning for image recognition, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, S. 770–778
- [He17] HE, Kaiming; GKIOXARI, Georgia; DOLLÁR, Piotr und GIRSHICK, Ross: Mask r-cnn, in: *Proceedings of the IEEE international conference on computer vision*, S. 2961–2969
- [Hof17] HOFMANN, Martin; NEUKART, Florian und BÄCK, Thomas: Artificial intelligence and data science in the automotive industry. *arXiv preprint arXiv:1709.01989* (2017)
- [Joh08] JOHNSON, C. R.; HENDRIKS, E.; BEREZHNOY, I. J.; BREVDO, E.; HUGHES, S. M.; DAUBECHIES, I.; LI, J.; POSTMA, E. und WANG, J. Z.: Image processing for artist identification. *IEEE Signal Processing Magazine* (2008), Bd. 25(4): S. 37–48
- [Kar14] KARPATY, Andrej: What I learned from competing against a ConvNet on ImageNet, Webseite (2014), online erhältlich unter <http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet>; abgerufen am 13. März 2019.
- [Kaz14] KAZEMI, Vahid und SULLIVAN, Josephine: One Millisecond Face Alignment with an Ensemble of Regression Trees, in: *Computer Vision and Pattern Recognition*
- [Koc93] KOCH, Guntram: *Sarkophage der römischen Kaiserzeit*, Wissenschaftliche Buchgesellschaft (1993)
- [Kri12] KRIZHEVSKY, Alex; SUTSKEVER, Ilya und HINTON, Geoffrey E.: ImageNet Classification with Deep Convolutional Neural Networks, in: *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, NIPS’12, Curran Associates Inc., USA, S. 1097–1105, URL <http://dl.acm.org/citation.cfm?id=2999134.2999257>

- [Lin14] LIN, Tsung-Yi; MAIRE, Michael; BELONGIE, Serge; HAYS, James; PERONA, Pietro; RAMANAN, Deva; DOLLÁR, Piotr und ZITNICK, C Lawrence: Microsoft coco: Common objects in context, in: *European conference on computer vision*, Springer, S. 740–755
- [Lin17] LIN, Tsung-Yi; GOYAL, Priya; GIRSHICK, Ross; HE, Kaiming und DOLLÁR, Piotr: Focal loss for dense object detection, in: *Proceedings of the IEEE international conference on computer vision*, S. 2980–2988
- [Liu16a] LIU, Wei; ANGUELOV, Dragomir; ERHAN, Dumitru; SZEGEDY, Christian; REED, Scott; FU, Cheng-Yang und BERG, Alexander C: Ssd: Single shot multibox detector, in: *European conference on computer vision*, Springer, S. 21–37
- [Liu16b] LIU, X.; LIANG, W.; WANG, Y.; LI, S. und PEI, M.: 3D head pose estimation with convolutional neural network trained on synthetic images, in: *IEEE International Conference on Image Processing (ICIP)*, S. 1289–1293
- [Lor16] LORENZ, Katharina: *Ancient Mythological Images and Their Interpretation*, Cambridge University Press (2016)
- [McC43] MCCULLOCH, Warren S. und PITTS, Walter: A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* (1943), Bd. 5(4): S. 115–133, URL <https://doi.org/10.1007/BF02478259>
- [Mon14] MONROY, Antonio; BELL, Peter und OMMER, Björn: Morphological analysis for investigating artistic images. *Image and Vision Computing* (2014), Bd. 32(6): S. 414–423, URL <http://www.sciencedirect.com/science/article/pii/S026288561400064X>
- [Nas16] NASR-ESFAHANI, E.; SAMAVI, S.; KARIMI, N.; SOROUSHMEHR, S. M. R.; JAFARI, M. H.; WARD, K. und NAJARIAN, K.: Melanoma detection by analysis of clinical images using convolutional neural network, in: *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, S. 1373–1376
- [Red18] REDMON, Joseph und FARHADI, Ali: Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* (2018)
- [Ren15] REN, Shaoqing; HE, Kaiming; GIRSHICK, Ross und SUN, Jian: Faster r-cnn: Towards real-time object detection with region proposal networks, in: *Advances in neural information processing systems*, S. 91–99
- [Res14] RESIG, John: Using Computer Vision to Increase the Research Potential of Photo Archives, in: *Journal of Digital Humanities*
- [Rui18] RUIZ, Nataniel; CHONG, Eunji und REHG, James M: Fine-grained head pose estimation without keypoints, in: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, S. 2074–2083
- [Rus15] RUSSAKOVSKY, Olga; DENG, Jia; SU, Hao; KRAUSE, Jonathan; SATHEESH, Sanjeev; MA, Sean; HUANG, Zhiheng; KARPATHY, Andrej; KHOSLA, Aditya; BERNSTEIN, Michael; BERG, Alexander C. und FEI-FEI, Li: ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer*

- Vision (IJCV)* (2015), Bd. 115(3): S. 211–252
- [Sch11] SCHLECHT, J.; CARQUÉ, B. und OMMER, B.: Detecting gestures in medieval images, in: *18th IEEE International Conference on Image Processing*, S. 1285–1288
- [Sze13] SZEGEDY, Christian; TOSHEV, Alexander und ERHAN, Dumitru: Deep Neural Networks for Object Detection, in: C. J. C. Burges; L. Bottou; M. Welling; Z. Ghahramani und K. Q. Weinberger (Herausgeber) *Advances in Neural Information Processing Systems 26*, Curran Associates, Inc. (2013), S. 2553–2561, URL <http://papers.nips.cc/paper/5207-deep-neural-networks-for-object-detection.pdf>
- [Uij13] UIJLINGS, J.R.R.; VAN DE SANDE, K.E.A.; GEVERS, T. und SMEULDERS, A.W.M.: Selective Search for Object Recognition. *International Journal of Computer Vision* (2013), URL <http://www.huppelen.nl/publications/selectiveSearchDraft.pdf>
- [Vad11] VADDI, R. S.; BOGGAVARAPU, L. N. P.; VANKAYALAPATI, H. D. und ANNE, K. R.: Comparative analysis of contrast enhancement techniques between histogram equalization and CNN, in: *Third International Conference on Advanced Computing*, S. 106–110
- [vP15] VON PIPPICH, Waltraud: Rot rechnen, in: *Zeitschrift für digitale Geisteswissenschaften, Sonderband 1: Grenzen und Möglichkeiten der Digital Humanities*
- [Wei16] WEI, Shih-En; RAMAKRISHNA, Varun; KANADE, Takeo und SHEIKH, Yaser: Convolutional pose machines, in: *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, S. 4724–4732
- [Wei19] WEI, Guoqiang; LAN, Cuiling; ZENG, Wenjun und CHEN, Zhibo: View Invariant 3D Human Pose Estimation. *IEEE Transactions on Circuits and Systems for Video Technology* (2019)
- [Zha19] ZHANG, Feng; ZHU, Xiatian und YE, Mao: Fast human pose estimation, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, S. 3517–3526

Glossar

Accuracy

Dt.: Korrektklassifikationsrate, auch Treffergenauigkeit. Bezeichnet das statistische Maß der korrekt klassifizierten Werte (positiv und negativ) in Relation zu allen Werten.

AlexNet

CNN zur Objekterkennung. Beschrieben in [Kri12] gilt es als Vorlage für viele weitere Entwicklungen und deren generelle Funktionsweise.

Arachne

„Zentrale Objektdatenbank des Deutschen Archäologischen Instituts (DAI) und des Archäologischen Instituts der Universität zu Köln“ mit etwa 4,3 Mio. Einträgen. Verfügbar unter <https://arachne.dainst.org/>.

Bounding Box

Eingrenzung eines Objekts. Dabei handelt es sich um die Regression der Konturen eines Elements, um bei Berechnungen von Überlappungen oder Kollisionen die Komplexität des Systems zu reduzieren

C++

Eine genormte und maschinennahe Programmiersprache. Die Vertreter der C-Sprachfamilie gehören immer unter die am weitesten verbreiteten Programmiersprachen weltweit. Durch die Nähe der Sprache zur Hardware ermöglicht sie eine sehr effiziente Nutzung der Systemressourcen für die Berechnung komplexer Matrizen im Anwendungsfall der künstlich neuronalen Netze.

Census

Bilddatenbank des Instituts für Kunst- und Bildgeschichte der Humboldt-Universität zu Berlin. „Die Datenbank des Census dokumentiert die in der Renaissance (1400-

1600) bekannten antiken Bau- und Bildwerke.“ Verfügbar unter `census.bbaw.de`.

COCO

Common Objects in Context. Bilddatenbank für großformatige, hierarchisch organisierte Objekte. Insgesamt sind etwa 1,5 Mio. Objekte aus 80 Kategorien in 330.000 Abbildungen vorhanden. Siehe [Lin14].

Convolutional Neural Network

Ist eine Kategorie künstlich neuronaler Netzwerke, welche speziell für die Bildverarbeitung geeignet sind. Die Eingangsdaten liegen zweidimensional vor, da einzelne Bildelemente im Kontext der umgebenden Fläche stehen und werden daher nicht vorab auf eine Dimension geglättet.

Dots per Inch

Maß der Bildauflösung. Dt.: Bildpunkte pro Zoll. Die Maßeinheit gibt die räumliche Größe eines Pixels im digitalen Bild an und kann im Umkehrschluss genutzt werden um Bildgrößen in Relation zu setzen.

F1-Score

Dt.: F-Wert. Bezeichnet das kombinierte statistische Maß aus Recall und Precision. Dabei wird das harmonische Mittel aus beiden berechnet.

False Discovery Rate

Dt.: Falscherkennungsrate. Bezeichnet das statistische Maß der falsch positiv klassifizierten Werte (Fehler 1. Art) in Relation zu allen als positiv klassifizierten Werte. Die Falscherkennungsrate ist somit der Kehrwert der Precision.

Faster R-CNN

Eine Weiterentwicklung von R-CNN. Hier werden keine ROIs vorab festgelegt und bearbeitet, sondern die extrahierten Features genutzt um daraus entsprechende Bounding Boxes zu generieren. Siehe [Ren15].

Feature

Dt.: Eigenschaft. Hier: Die mathematische Beschreibung von Bildeigenschaften durch ein- oder zweidimensionale Vektoren.

Foto Marburg

Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Foto Marburg. „Der Bildindex ist eine Verbunddatenbank, in der rund 80 Kultur- und Wissenschaftseinrichtungen – Museen, Denkmalpflegeämter, Bibliotheken, Universitäts- und Forschungsinstitute – ihre Bild- und Datenbestände zur Kunst und Architektur gemeinsam im Internet publizieren.“ Verfügbar unter <https://www.bildindex.de/>.

Framework

Sammlung von Programmen und Daten. Ähnlich wie Bibliotheken oder Application Programming Interface (APIs) zeichnen sich Frameworks durch den Fokus auf die Erweiterbarkeit aus.

Groundtruth

Hier: Referenzwerte zur Ergebnisvalidierung. Als Groundtruth werden annotierte Daten bezeichnet, mit denen ein künstlich neuronales Netz trainiert wird. Die Annotationen werden daher als vollständig korrekt angenommen.

ImageNet

Bilddatensatz zum Training und zur Evaluation von Objekterkennungstechniken. Die Objektklassen sind dabei hierarchisch kategorisiert, um abstraktere Klassifizierungen zuzulassen. Siehe [Den09].

Intersection over Union

Kennzahl für die Ähnlichkeit von Mengen. Auch Jaccard-Koeffizient nach Paul Jaccard. Bezeichnet den Quotienten aus Schnitt- und Vereinigungsmenge.

Java

Eine plattformunabhängige, objektorientierte Programmiersprache. Durch die einfache Syntax, die starke Typisierung und die Plattformunabhängigkeit zählt Java zu einer der am weitesten verbreiteten Sprachen weltweit.

Kernel

Werkzeug in der Bildverarbeitung. Als Kernel werden spezielle Matrizen bezeichnet, welche dazu geeignet sind Merkmale eines Bildes zu extrahieren. Durch

Matritzenmultiplikation können so Kanten hervorgehoben oder Darstellungen geglättet werden.

Keypoint

Dt.: Schlüsselpunkt. Hier: Die Menge der wichtigsten körperlichen Merkmale um die Pose einer Person zu bestimmen.

Klassenwahrscheinlichkeit

Gibt die Wahrscheinlichkeit an, mit der ein Objekt einer Klasse angehört. Ein neuronales Netz zur Bildklassifizierung gibt für jede angelernte Klasse eine Wahrscheinlichkeit aus, welche meist in Summe 100% ergeben. Im Kontext wird häufig nur die höchste Wahrscheinlichkeit bzw. die dazu gehörige Klasse gemeint.

Mask R-CNN

Eine Weiterentwicklung von R-CNN und Faster R-CNN. Der Architektur wurde ein Element zur Erstellung einer binären Objektmaske hinzugefügt, welches das Hauptobjekt eines Bildausschnittes pixelgenau segmentiert. Gleichzeitig wurde die Bestimmung der Bounding Boxes optimiert. Siehe [He17].

Meleager

Held der griechischen Antike. Meleager (auch Meleagros) war Mitglied der Gruppe um Jason auf der Suche nach dem goldenen Fließ und Jäger des kalydonischen Ebers.

Model Zoo

Eine Sammlung vortrainierter Netze. Unter Angabe der Parameter, die für das Training gewählt wurden und der daraus resultierenden Ergebnismetriken kann eine Auswahl getroffen werden, welches Netz für die weitere Nutzung gewählt wird.

Netzarchitektur

Die grundlegende Struktur künstlich neuronaler Netze. Durch die Anordnungen und Verbindungen unterschiedlicher in Ebenen zusammengefasster Neuronen und anderer Algorithmen entstehen entstehen Netze, welche sich in ihrer Domäne oder Effektivität und Geschwindigkeit unterscheiden.

Neuron

Kleinster Bestandteil eines künstlich neuronalen Netzes. Mathematisch gesehen ist ein Neuron ein Element, welches beliebig viele Eingangswerte mit unterschiedlichen Faktoren multipliziert (gewichtet) und aufsummiert. Wenn die Summe einen Schwellenwert überschreitet, wird sie an die nachfolgenden Neuronen als neuer Eingabewert weitergegeben.

Object Proposal

Hier: Ausgabe eines CNNs. Unter Object Proposal werden der Bildausschnitt in Form einer Bounding Box, ein Klassenvorschlag bzw. dessen Klassenwahrscheinlichkeit und mögliche weitere Daten, wie die Keypoints zusammengefasst.

OpenCV

Eine Sammlung von Funktionen und Klassen zur Bildverarbeitung. Seit 2000 eine weit verbreiteter Standard bei der Auswertung von Kamerabildern.

Precision

Dt.: Genauigkeit. Bezeichnet das statistische Maß der korrekt positiven Werte in Relation zu allen als positiv klassifizierten Werte. Hier also die Anzahl der korrekten Object Proposals im Vergleich zu allen.

Python

Eine dynamisch typisierte, interpretierte Programmiersprache. Benannt nach Monty Python hat sich die Sprache als Einstieg in die Domäne der künstlich neuronalen Netze etabliert.

R-CNN

Regions with CNN Features. Bezeichnet eine der ersten Lösungen um multiple Objekte in einem Bild per CNNs zu erkennen. Siehe [Gir14].

Recall

Dt.: Sensitivität. Bezeichnet das statistische Maß der korrekt positiven Werte in Relation zu allen korrekten Werten. Hier also die Anzahl erkannter Personen zu allen vorhandenen Personen bzw. die Anzahl korrekt platzierter Keypoints zu allen erkennbaren Keypoints.

Region of Interest

Beschreibt einen Bildausschnitt mit besonderen Merkmalen. Durch die Gruppierung vorab erhobener Bildeigenschaften können Ausschnitte definiert werden, in denen ein Objekt vermutet wird. Diese Bildausschnitte werden zu Bounding Boxes verarbeitet.

ResNet

Grundlegende Netzarchitektur für CNNs. Durch die Optimierung des Datenflusses innerhalb der Schichten können deutlich tiefere Netze erzeugt werden, ohne zu spezifische Ergebnisse zu erhalten. Siehe [He16].

RetinaNet

Netzarchitektur zur Lokalisierung multipler Objekte. Ähnlich wie You Only Look Once und Single Shot Detector verläuft die Erkennung in einem Schritt. Siehe [Lin17].

Single Shot Detector

Netzwerkarchitektur zur Lokalisierung multipler Objekte. Nutzt ähnlich wie YOLO vordefinierte Bildbereiche zur Objekterkennung, verarbeitet diese aber nicht zu genaueren Bounding Boxes weiter. Die Lokalisierungsgenauigkeit ist dadurch deutlich geringer, der Ablauf jedoch auch für Echtzeitsysteme geeignet. Siehe [Liu16a].

Support Vector Machine

Unterteilt Punkte im Vektorraum in zwei Mengen. Die Trennung erfolgt durch eine (Hyper-) Ebene, welche die beiden Mengen mit möglichst breitem Rand voneinander separiert um auch bei weiteren Entscheidungen treffsicher zu klassifizieren.

Training

Optimierung von Klassifikatoren zur Steigerung der Genauigkeit. Durch die Evaluierung der Resultate eines Netz mit annotierten Daten können nachträglich die Schwellen und Gewichte der Neuronen und damit der Kernels angepasst werden um die Anzahl korrekter Ergebnisse zu steigern.

You Only Look Once

Netzwerkarchitektur zur Lokalisierung multipler Objekte. Im Unterschied zu den R-CNN-Entwicklungen werden zuerst für vordefinierte Bildbereiche Klassenwahr-

scheinlichkeiten berechnet, aus welchen anschließend die Bounding Boxes ermittelt werden. Siehe [Red18].

Akronyme

CNN

Convolutional Neural Network

dpi

Dots per Inch

FDR

False Discovery Rate

ILSVRC

ImageNet Large Scale Visual Recognition Challenge

IOU

Intersection over Union

ROI

Region of Interest

SSD

Single Shot Detector

SVM

Support Vector Machine

YOLO

You Only Look Once

Abbildungsverzeichnis

2.1	Übersicht über die Stufen der Objekterkennung auf Grundlage der Abbildung 3.1. Von links: Klassifizierung und Lokalisierung, Objekterkennung, Instanzsegmentierung	10
2.2	Visualisierung von Featuremaps. Zu sehen sind exemplarisch Elemente der Ebenen 7, 14, 20, 30 und 40. ([Grä19])	12
2.3	Meleager-Sarkophag mit Deckel, Villa Doria Pamphili - Casino Belrespiro (arachne.dainst.org/entity/2888823)	17
2.4	Meleager-Sarkophag, Bedfordshire, Woburn Abbey (arachne.dainst.org/entity/405092)	18
2.5	Meleager-Sarkophag, Rom, Kapitolinische Museen (http://census.bbaw.de/easydb/censusID=159869)	20
3.1	Sarkophag mit der Jagd auf den calydonischen Eber, Rom, Kapitolinische Museen (Marie-Lan Nguyen / Wikimedia Commons)	22
3.2	Übersicht der Klassenwahrscheinlichkeiten in zehn-Prozent-Schritten als Visualisierung der Daten aus Tab. B.1.	23
3.3	Sarkophag mit Jagd auf den kalydonischen Eber, Eleusis, Archäologisches Museum (©Bildarchiv Foto Marburg / Konrad Helbig). Hier annotiert mit der Groundtruth.	25
3.4	Sarkophag-Fragment mit der Eberjagd des Meleager, Saloniki, Archäologisches Museum (©Bildarchiv Foto Marburg / Jan Gloc; Paul Haag). Hier annotiert mit generierten Object Proposals (Dicke Linie: Über 60% Klassenwahrscheinlichkeit).	26
3.5	Übersicht der F1-Scores nach Netzarchitektur als Visualisierung der Tab. B.22	26
3.6	Mit Keypoints annotierte Abb. A.1	31
4.1	Mit generierten Object Proposals annotierte Abb. 2.5	36
4.2	Meleager-Sarkophag, Bedfordshire, Woburn Abbey (arachne.dainst.org/entity/405082). Hier annotiert mit generierten Keypoints.	38
A.1	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.1	61
A.2	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.3	62
A.3	Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 3.3	62

A.4	Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 3.3	62
A.5	Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 3.3	63
A.6	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.4	63
A.7	Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 3.4	64
A.8	Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 3.4	64
A.9	Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 3.4	65
A.10	Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 3.4	65
A.11	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 2.3	65
A.12	Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 2.3	65
A.13	Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 2.3	65
A.14	Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 2.3	65
A.15	Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 2.3	65
A.16	Positives Ergebnis Nr. 6 des Netzes 37697946 aus Abbildung 2.3	65
A.17	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 4.2	66
A.18	Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 4.2	66
A.19	Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 4.2	67
A.20	Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 4.2	67
A.21	Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 4.2	68
A.22	Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 4.2	68
A.23	Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 4.2	69
A.24	Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 2.4	69

Tabellenverzeichnis

2.1	Übersicht über die Personenanzahl und deren Größe in Relation zur Höhe der Abbildung	19
3.1	Recall nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11	27
3.2	F1-Score der Keypointnetze nach Schwellenwert als Ausschnitt aus den Tabellen B.2 bis B.11	28
3.3	F1-Score der Keypointnetze nach Schwellenwert bei gekürzter Personenanzahl als Ausschnitt aus den Tabellen B.12 bis B.21	28
3.4	Accuracy der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.	32
3.5	Precision im Ein-Schritt-Verfahren je Netz und Abbildung	33
B.1	Object Proposals und deren Klassenwahrscheinlichkeiten der Netze . . .	72
B.2	Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten	74
B.3	Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten	76
B.4	Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten	78
B.5	Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten	80
B.6	Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten	82
B.7	Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten	84
B.8	Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten	86
B.9	Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten	88
B.10	Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten	90

B.11 Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten	92
B.12 Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	94
B.13 Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	96
B.14 Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	98
B.15 Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	100
B.16 Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	102
B.17 Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	104
B.18 Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	106
B.19 Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	108
B.20 Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	110
B.21 Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl	112
B.22 F1-Scores nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11	114
B.23 False Discovery Rate nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11	114
B.24 F1-Score nach Netzarchitektur und Schwellenwert bei gekürzter Personenzahl als Zusammenfassung der Tabellen B.12 bis B.21	114
B.25 Recall der Keypointnetze nach Schwellenwert als Ausschnitt aus den Tabellen B.2 bis B.11	114
B.26 Recall der Keypointnetze nach Schwellenwert bei gekürzter Personenzahl als Ausschnitt aus den Tabellen B.12 bis B.21	114
B.27 Bildabmessungen der untersuchten Abbildungen	115
B.28 Abmessungen der Abbildungen der Census-Datenbank	115
B.30 Wahrscheinlichkeiten der Keypointerkennung für Netz 37697547	117
B.31 Wahrscheinlichkeiten der Keypointerkennung für Netz 37697714	119
B.32 Wahrscheinlichkeiten der Keypointerkennung für Netz 37697946	121
B.33 Wahrscheinlichkeiten der Keypointerkennung für Netz 37698009	123
B.34 Wahrscheinlichkeiten der Keypointerkennung für Netz 37732318	125
B.35 Wahrscheinlichkeiten der Keypointerkennung für Netz 37732355	127
B.36 Wahrscheinlichkeiten der Keypointerkennung für Netz 37732415	129

B.37	Wahrscheinlichkeiten der Keypointerkennung für Netz 37792158	131
B.38	Ergebnisse der Keypointerkennung für Netz 37697547	133
B.39	Ergebnisse der Keypointerkennung für Netz 37697714	135
B.40	Ergebnisse der Keypointerkennung für Netz 37697946	137
B.41	Ergebnisse der Keypointerkennung für Netz 37698009	139
B.42	Ergebnisse der Keypointerkennung für Netz 37732318	141
B.43	Ergebnisse der Keypointerkennung für Netz 37732355	143
B.44	Ergebnisse der Keypointerkennung für Netz 37732415	145
B.45	Ergebnisse der Keypointerkennung für Netz 37792158	147
B.46	Recall der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.	149
B.47	Precision der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.	149
B.48	F1-Score der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.	149

A Bilder



Abbildung A.1: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.1

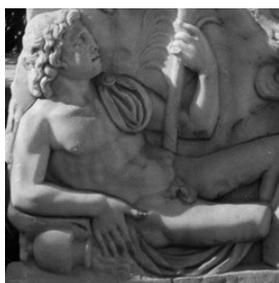


Abbildung A.2: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.3



Abbildung A.3: Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 3.3

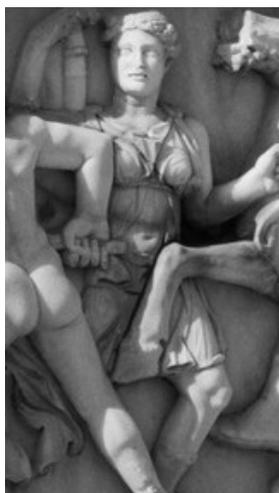


Abbildung A.4: Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 3.3

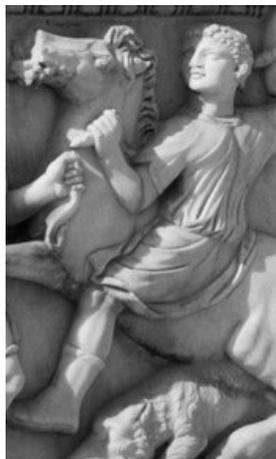


Abbildung A.5: Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 3.3

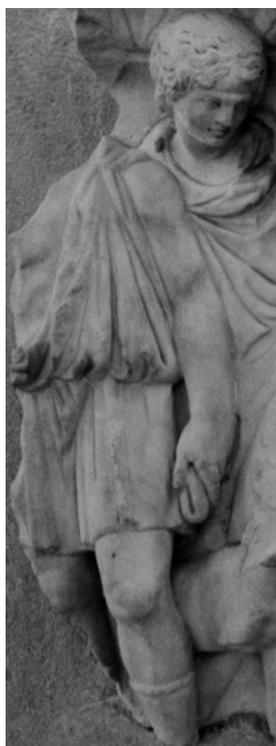


Abbildung A.6: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 3.4



Abbildung A.7: Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 3.4

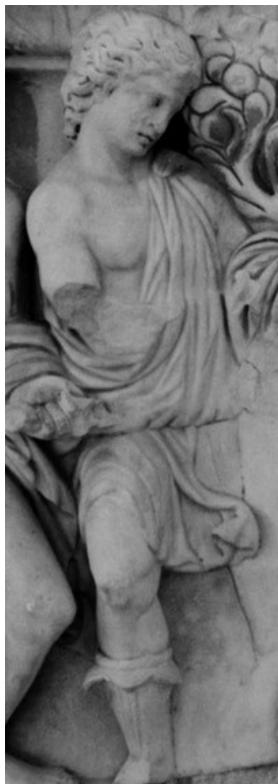


Abbildung A.8: Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 3.4



Abbildung A.9: Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 3.4



Abbildung A.10: Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 3.4

Abbildungsrechte nicht vorhanden

Abbildung A.11: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 2.3

Abbildungsrechte nicht vorhanden

Abbildung A.12: Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 2.3

Abbildungsrechte nicht vorhanden

Abbildung A.13: Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 2.3

Abbildungsrechte nicht vorhanden

Abbildung A.14: Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 2.3

Abbildungsrechte nicht vorhanden

Abbildung A.15: Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 2.3

Abbildungsrechte nicht vorhanden

Abbildung A.16: Positives Ergebnis Nr. 6 des Netzes 37697946 aus Abbildung 2.3



Abbildung A.17: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.18: Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.19: Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.20: Positives Ergebnis Nr. 4 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.21: Positives Ergebnis Nr. 5 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.22: Positives Ergebnis Nr. 1 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.23: Positives Ergebnis Nr. 2 des Netzes 37697946 aus Abbildung 4.2



Abbildung A.24: Positives Ergebnis Nr. 3 des Netzes 37697946 aus Abbildung 2.4

B Tabellen

Netz-ID	Netztyp	Vorschläge	<10%	<20%	<30%	<40%	<50%	<60%	<70%	<80%	<90%	<100%
35857197	Faster	117	17%	20%	10%	9%	4%	7%	6%	8%	6%	14%
35857281	Faster	92	13%	17%	8%	10%	3%	8%	3%	13%	12%	13%
35857345	Faster	129	29%	23%	9%	5%	7%	3%	5%	5%	3%	9%
35857389	Faster	127	32%	27%	6%	6%	7%	4%	4%	5%	0%	9%
35857890	Faster	122	22%	26%	11%	10%	5%	3%	4%	4%	2%	12%
35857952	Faster	79	19%	20%	20%	8%	3%	4%	0%	6%	6%	14%
35858015	Faster	104	25%	21%	11%	7%	3%	7%	2%	5%	3%	17%
35858198	Faster	85	20%	27%	11%	5%	6%	4%	2%	0%	1%	25%
36761737	Faster	139	20%	24%	12%	10%	6%	5%	3%	2%	2%	16%
36761786	Faster	83	22%	23%	10%	5%	6%	2%	4%	2%	7%	19%
35858791	Mask	102	27%	10%	8%	6%	2%	4%	7%	13%	9%	15%
35858828	Mask	105	20%	10%	8%	9%	9%	7%	7%	8%	10%	13%
35858933	Mask	149	19%	27%	13%	8%	6%	4%	5%	3%	3%	12%
35859007	Mask	108	21%	28%	12%	6%	5%	6%	4%	3%	6%	11%
35859745	Mask	80	23%	21%	9%	6%	4%	1%	4%	8%	0%	25%
35861795	Mask	123	28%	19%	11%	9%	3%	3%	2%	2%	8%	14%
35861858	Mask	79	27%	16%	11%	8%	3%	3%	5%	3%	5%	20%
36494496	Mask	119	24%	29%	12%	4%	4%	2%	5%	0%	4%	17%
36761843	Mask	138	22%	25%	12%	6%	7%	4%	3%	1%	4%	15%
36762092	Mask	86	21%	16%	9%	10%	5%	2%	1%	5%	6%	24%
37129812	Mask	223	11%	17%	12%	13%	4%	8%	6%	5%	8%	16%

Tabella B.1: Object Proposals und deren Klassenwahrscheinlichkeiten der Netze

Netz-ID	Netztyp	Vorschläge	<10%	<20%	<30%	<40%	<50%	<60%	<70%	<80%	<90%	<100%
36768636	Retina	208	20%	51%	12%	9%	4%	2%	0%	2%	0%	0%
36768677	Retina	190	28%	47%	11%	7%	2%	1%	2%	2%	1%	0%
36768744	Retina	233	27%	50%	13%	3%	4%	2%	2%	0%	0%	0%
36768840	Retina	200	39%	38%	11%	5%	2%	2%	3%	2%	0%	0%
36768875	Retina	185	30%	48%	9%	6%	2%	3%	2%	1%	0%	0%
36768907	Retina	167	28%	43%	16%	4%	4%	3%	2%	1%	0%	0%
36769563	Retina	223	33%	44%	13%	5%	3%	1%	1%	0%	0%	0%
36769641	Retina	157	50%	29%	5%	5%	3%	3%	3%	1%	1%	0%
37697547	Keypoint	195	22%	26%	11%	7%	7%	4%	3%	2%	6%	12%
37697714	Keypoint	200	25%	24%	13%	10%	3%	5%	1%	4%	3%	13%
37697946	Keypoint	204	31%	23%	12%	3%	7%	2%	2%	4%	4%	11%
37698009	Keypoint	190	26%	23%	14%	6%	3%	2%	3%	4%	4%	15%
37732318	Keypoint	140	26%	24%	6%	6%	4%	4%	2%	4%	6%	18%
37732355	Keypoint	170	25%	23%	8%	9%	3%	7%	2%	4%	5%	13%
37732415	Keypoint	154	24%	24%	11%	10%	3%	2%	3%	3%	1%	18%
37792158	Keypoint	185	29%	24%	15%	8%	2%	2%	1%	2%	4%	14%

Object Proposals und deren Klassenwahrscheinlichkeiten der Netze (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	117	25	32%	79%	21%	26%
35857281	Faster	92	26	33%	72%	28%	31%
35857345	Faster	129	26	33%	80%	20%	25%
35857389	Faster	127	29	37%	77%	23%	28%
35857890	Faster	122	26	33%	79%	21%	26%
35857952	Faster	79	19	24%	76%	24%	24%
35858015	Faster	104	23	29%	78%	22%	25%
35858198	Faster	85	15	19%	82%	18%	18%
36761737	Faster	139	29	37%	79%	21%	27%
36761786	Faster	83	20	26%	76%	24%	25%
35858791	Mask	102	26	33%	75%	25%	29%
35858828	Mask	105	28	36%	73%	27%	31%
35858933	Mask	149	32	41%	79%	21%	28%
35859007	Mask	108	23	29%	79%	21%	25%
35859745	Mask	80	23	29%	71%	29%	29%
35861795	Mask	123	31	40%	75%	25%	31%
35861858	Mask	79	22	28%	72%	28%	28%
36494496	Mask	119	29	37%	76%	24%	29%
36761843	Mask	138	28	36%	80%	20%	26%
36762092	Mask	86	25	32%	71%	29%	30%
37129812	Mask	223	35	45%	84%	16%	23%

Tabelle B.2: Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	208	28	36%	87%	13%	20%
36768677	Retina	190	21	27%	89%	11%	16%
36768744	Retina	233	26	33%	89%	11%	17%
36768840	Retina	200	22	28%	89%	11%	16%
36768875	Retina	185	21	27%	89%	11%	16%
36768907	Retina	167	17	22%	90%	10%	14%
36769563	Retina	223	21	27%	91%	9%	14%
36769641	Retina	157	17	22%	89%	11%	14%
37697547	Keypoint	195	41	53%	79%	21%	30%
37697714	Keypoint	200	39	50%	81%	20%	28%
37697946	Keypoint	204	37	47%	82%	18%	26%
37698009	Keypoint	190	35	45%	82%	18%	26%
37732318	Keypoint	140	35	45%	75%	25%	32%
37732355	Keypoint	170	33	42%	81%	19%	27%
37732415	Keypoint	154	41	53%	73%	27%	35%
37792158	Keypoint	185	38	49%	79%	21%	29%

Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	97	25	32%	74%	26%	29%
35857281	Faster	79	25	32%	68%	32%	32%
35857345	Faster	91	23	29%	75%	25%	27%
35857389	Faster	84	23	29%	73%	27%	28%
35857890	Faster	93	24	31%	74%	26%	28%
35857952	Faster	63	16	21%	75%	25%	23%
35858015	Faster	77	19	24%	75%	25%	25%
35858198	Faster	68	14	18%	79%	21%	19%
36761737	Faster	108	28	36%	74%	26%	30%
36761786	Faster	64	17	22%	73%	27%	24%
35858791	Mask	73	25	32%	66%	34%	33%
35858828	Mask	83	25	32%	70%	30%	31%
35858933	Mask	118	28	36%	76%	24%	29%
35859007	Mask	81	21	27%	74%	26%	26%
35859745	Mask	59	20	26%	66%	34%	29%
35861795	Mask	86	29	37%	66%	34%	35%
35861858	Mask	58	21	27%	64%	36%	31%
36494496	Mask	87	26	33%	70%	30%	32%
36761843	Mask	104	27	35%	74%	26%	30%
36762092	Mask	67	23	29%	66%	34%	32%
37129812	Mask	196	33	42%	83%	17%	24%

Tabelle B.3: Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	150	22	28%	85%	15%	19%
36768677	Retina	127	18	23%	86%	14%	18%
36768744	Retina	166	25	32%	85%	15%	20%
36768840	Retina	111	19	24%	83%	17%	20%
36768875	Retina	116	17	22%	85%	15%	18%
36768907	Retina	110	16	21%	85%	15%	17%
36769563	Retina	135	18	23%	87%	13%	17%
36769641	Retina	71	14	18%	80%	20%	19%
37697547	Keypoint	145	38	49%	74%	26%	34%
37697714	Keypoint	146	37	47%	75%	25%	33%
37697946	Keypoint	135	31	40%	77%	23%	29%
37698009	Keypoint	139	32	41%	77%	23%	29%
37732318	Keypoint	102	33	42%	68%	32%	37%
37732355	Keypoint	124	32	41%	74%	26%	32%
37732415	Keypoint	114	35	45%	69%	31%	36%
37792158	Keypoint	132	33	42%	75%	25%	31%

Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	73	23	29%	68%	32%	30%
35857281	Faster	63	23	29%	63%	37%	33%
35857345	Faster	61	20	26%	67%	33%	29%
35857389	Faster	51	19	24%	63%	37%	29%
35857890	Faster	63	21	27%	67%	33%	30%
35857952	Faster	47	15	19%	68%	32%	24%
35858015	Faster	55	18	23%	67%	33%	27%
35858198	Faster	44	14	18%	68%	32%	23%
36761737	Faster	77	24	31%	69%	31%	31%
36761786	Faster	45	15	19%	67%	33%	24%
35858791	Mask	64	22	28%	66%	34%	31%
35858828	Mask	72	24	31%	67%	33%	32%
35858933	Mask	81	25	32%	69%	31%	31%
35859007	Mask	55	18	23%	67%	33%	27%
35859745	Mask	45	18	23%	60%	40%	29%
35861795	Mask	64	27	35%	58%	42%	38%
35861858	Mask	44	19	24%	57%	43%	31%
36494496	Mask	57	17	22%	70%	30%	25%
36761843	Mask	72	26	33%	64%	36%	35%
36762092	Mask	54	21	27%	61%	39%	32%
37129812	Mask	160	33	42%	79%	21%	28%

Tabelle B.4: Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	59	16	21%	73%	27%	23%
36768677	Retina	47	16	21%	66%	34%	26%
36768744	Retina	53	15	19%	72%	28%	23%
36768840	Retina	46	12	15%	74%	26%	19%
36768875	Retina	39	11	14%	72%	28%	19%
36768907	Retina	46	12	15%	74%	26%	19%
36769563	Retina	51	11	14%	78%	22%	17%
36769641	Retina	33	14	18%	58%	42%	25%
37697547	Keypoint	100	30	38%	70%	30%	34%
37697714	Keypoint	102	32	41%	69%	31%	36%
37697946	Keypoint	92	30	38%	67%	33%	35%
37698009	Keypoint	94	29	37%	69%	31%	34%
37732318	Keypoint	71	30	38%	58%	42%	40%
37732355	Keypoint	86	28	36%	67%	33%	34%
37732415	Keypoint	79	33	42%	58%	42%	42%
37792158	Keypoint	87	30	38%	66%	34%	36%

Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	61	22	28%	64%	36%	32%
35857281	Faster	56	22	28%	61%	39%	33%
35857345	Faster	48	16	21%	67%	33%	25%
35857389	Faster	44	17	22%	61%	39%	28%
35857890	Faster	45	18	23%	60%	40%	29%
35857952	Faster	32	14	18%	56%	44%	25%
35858015	Faster	45	17	22%	62%	38%	28%
35858198	Faster	35	14	18%	60%	40%	25%
36761737	Faster	59	23	29%	61%	39%	34%
36761786	Faster	38	15	19%	61%	39%	26%
35858791	Mask	54	22	28%	59%	41%	33%
35858828	Mask	65	24	31%	63%	37%	34%
35858933	Mask	61	23	29%	62%	38%	33%
35859007	Mask	42	16	21%	62%	38%	27%
35859745	Mask	38	16	21%	58%	42%	28%
35861795	Mask	50	24	31%	52%	48%	38%
35861858	Mask	36	18	23%	50%	50%	32%
36494496	Mask	43	16	21%	63%	37%	26%
36761843	Mask	55	22	28%	60%	40%	33%
36762092	Mask	45	20	26%	56%	44%	33%
37129812	Mask	132	33	42%	75%	25%	31%

Tabelle B.5: Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	36	14	18%	61%	39%	25%
36768677	Retina	26	11	14%	58%	42%	21%
36768744	Retina	24	14	18%	42%	58%	27%
36768840	Retina	25	10	13%	60%	40%	19%
36768875	Retina	23	10	13%	57%	43%	20%
36768907	Retina	23	10	13%	57%	43%	20%
36769563	Retina	24	11	14%	54%	46%	22%
36769641	Retina	24	12	15%	50%	50%	24%
37697547	Keypoint	78	24	31%	69%	31%	31%
37697714	Keypoint	76	28	36%	63%	37%	36%
37697946	Keypoint	68	27	35%	60%	40%	37%
37698009	Keypoint	70	27	35%	61%	39%	36%
37732318	Keypoint	61	27	35%	56%	44%	39%
37732355	Keypoint	74	27	35%	64%	36%	36%
37732415	Keypoint	63	28	36%	56%	44%	40%
37792158	Keypoint	59	24	31%	59%	41%	35%

Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	51	20	26%	61%	39%	31%
35857281	Faster	48	22	28%	54%	46%	35%
35857345	Faster	42	15	19%	64%	36%	25%
35857389	Faster	37	16	21%	57%	43%	28%
35857890	Faster	37	15	19%	59%	41%	26%
35857952	Faster	26	13	17%	50%	50%	25%
35858015	Faster	38	16	21%	58%	42%	28%
35858198	Faster	32	14	18%	56%	44%	25%
36761737	Faster	46	19	24%	59%	41%	31%
36761786	Faster	34	15	19%	56%	44%	27%
35858791	Mask	50	22	28%	56%	44%	34%
35858828	Mask	54	21	27%	61%	39%	32%
35858933	Mask	49	23	29%	53%	47%	36%
35859007	Mask	35	14	18%	60%	40%	25%
35859745	Mask	33	15	19%	55%	45%	27%
35861795	Mask	40	22	28%	45%	55%	37%
35861858	Mask	30	17	22%	43%	57%	31%
36494496	Mask	38	15	19%	61%	39%	26%
36761843	Mask	46	22	28%	52%	48%	35%
36762092	Mask	37	19	24%	49%	51%	33%
37129812	Mask	105	29	37%	72%	28%	32%

Tabelle B.6: Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	17	10	13%	41%	59%	21%
36768677	Retina	12	7	9%	42%	58%	16%
36768744	Retina	17	12	15%	29%	71%	25%
36768840	Retina	16	9	12%	44%	56%	19%
36768875	Retina	13	8	10%	38%	62%	18%
36768907	Retina	15	9	12%	40%	60%	19%
36769563	Retina	12	8	10%	33%	67%	18%
36769641	Retina	17	12	15%	29%	71%	25%
37697547	Keypoint	65	24	31%	63%	37%	34%
37697714	Keypoint	56	26	33%	54%	46%	39%
37697946	Keypoint	61	26	33%	57%	43%	37%
37698009	Keypoint	59	25	32%	58%	42%	36%
37732318	Keypoint	53	26	33%	51%	49%	40%
37732355	Keypoint	57	24	31%	58%	42%	36%
37732415	Keypoint	48	25	32%	48%	52%	40%
37792158	Keypoint	45	22	28%	51%	49%	36%

Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	47	17	22%	64%	36%	27%
35857281	Faster	45	19	24%	58%	42%	31%
35857345	Faster	32	14	18%	56%	44%	25%
35857389	Faster	28	10	13%	64%	36%	19%
35857890	Faster	32	15	19%	53%	47%	27%
35857952	Faster	24	13	17%	46%	54%	25%
35858015	Faster	34	16	21%	53%	47%	29%
35858198	Faster	27	13	17%	52%	48%	25%
36761737	Faster	39	18	23%	54%	46%	31%
36761786	Faster	29	15	19%	48%	52%	28%
35858791	Mask	48	21	27%	56%	44%	33%
35858828	Mask	47	20	26%	57%	43%	32%
35858933	Mask	41	20	26%	51%	49%	34%
35859007	Mask	30	12	15%	60%	40%	22%
35859745	Mask	30	15	19%	50%	50%	28%
35861795	Mask	36	20	26%	44%	56%	35%
35861858	Mask	28	17	22%	39%	61%	32%
36494496	Mask	33	15	19%	55%	45%	27%
36761843	Mask	38	20	26%	47%	53%	34%
36762092	Mask	32	17	22%	47%	53%	31%
37129812	Mask	93	28	36%	70%	30%	33%

Tabelle B.7: Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	9	8	10%	11%	89%	18%
36768677	Retina	9	7	9%	22%	78%	16%
36768744	Retina	9	8	10%	11%	89%	18%
36768840	Retina	12	8	10%	33%	67%	18%
36768875	Retina	9	5	6%	44%	56%	11%
36768907	Retina	9	6	8%	33%	67%	14%
36769563	Retina	7	6	8%	14%	86%	14%
36769641	Retina	12	11	14%	8%	92%	24%
37697547	Keypoint	51	21	27%	59%	41%	33%
37697714	Keypoint	51	25	32%	51%	49%	39%
37697946	Keypoint	48	25	32%	48%	52%	40%
37698009	Keypoint	53	24	31%	55%	45%	37%
37732318	Keypoint	48	25	32%	48%	52%	40%
37732355	Keypoint	53	24	31%	55%	45%	37%
37732415	Keypoint	43	23	29%	47%	53%	38%
37792158	Keypoint	41	22	28%	46%	54%	37%

Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	39	15	19%	62%	38%	26%
35857281	Faster	38	19	24%	50%	50%	33%
35857345	Faster	29	13	17%	55%	45%	24%
35857389	Faster	23	10	13%	57%	43%	20%
35857890	Faster	28	15	19%	46%	54%	28%
35857952	Faster	21	12	15%	43%	57%	24%
35858015	Faster	28	15	19%	46%	54%	28%
35858198	Faster	24	12	15%	50%	50%	24%
36761737	Faster	32	16	21%	50%	50%	29%
36761786	Faster	27	15	19%	44%	56%	29%
35858791	Mask	42	19	24%	55%	45%	32%
35858828	Mask	40	20	26%	50%	50%	34%
35858933	Mask	35	18	23%	49%	51%	32%
35859007	Mask	25	12	15%	52%	48%	23%
35859745	Mask	29	15	19%	48%	52%	28%
35861795	Mask	32	20	26%	38%	63%	36%
35861858	Mask	26	16	21%	38%	62%	31%
36494496	Mask	30	15	19%	50%	50%	28%
36761843	Mask	33	18	23%	45%	55%	32%
36762092	Mask	31	17	22%	45%	55%	31%
37129812	Mask	78	25	32%	68%	32%	32%

Tabelle B.8: Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	5	5	6%	0%	100%	12%
36768677	Retina	7	6	8%	14%	86%	14%
36768744	Retina	5	4	5%	20%	80%	10%
36768840	Retina	9	6	8%	33%	67%	14%
36768875	Retina	2	2	3%	0%	100%	5%
36768907	Retina	4	3	4%	25%	75%	7%
36769563	Retina	4	3	4%	25%	75%	7%
36769641	Retina	7	7	9%	0%	100%	16%
37697547	Keypoint	45	20	26%	56%	44%	33%
37697714	Keypoint	41	22	28%	46%	54%	37%
37697946	Keypoint	43	25	32%	42%	58%	41%
37698009	Keypoint	49	24	31%	51%	49%	38%
37732318	Keypoint	42	22	28%	48%	52%	37%
37732355	Keypoint	39	20	26%	49%	51%	34%
37732415	Keypoint	40	23	29%	43%	58%	39%
37792158	Keypoint	38	21	27%	45%	55%	36%

Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	32	13	17%	59%	41%	24%
35857281	Faster	35	18	23%	49%	51%	32%
35857345	Faster	22	11	14%	50%	50%	22%
35857389	Faster	17	10	13%	41%	59%	21%
35857890	Faster	23	14	18%	39%	61%	28%
35857952	Faster	19	11	14%	42%	58%	23%
35858015	Faster	26	15	19%	42%	58%	29%
35858198	Faster	22	12	15%	45%	55%	24%
36761737	Faster	28	15	19%	46%	54%	28%
36761786	Faster	24	15	19%	38%	63%	29%
35858791	Mask	37	17	22%	54%	46%	30%
35858828	Mask	33	17	22%	48%	52%	31%
35858933	Mask	28	15	19%	46%	54%	28%
35859007	Mask	21	12	15%	43%	57%	24%
35859745	Mask	26	13	17%	50%	50%	25%
35861795	Mask	29	18	23%	38%	62%	34%
35861858	Mask	22	13	17%	41%	59%	26%
36494496	Mask	25	15	19%	40%	60%	29%
36761843	Mask	29	18	23%	38%	62%	34%
36762092	Mask	30	17	22%	43%	57%	31%
37129812	Mask	65	24	31%	63%	37%	34%

Tabelle B.9: Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	3	3	4%	0%	100%	7%
36768677	Retina	4	4	5%	0%	100%	10%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	2	2	3%	0%	100%	5%
36768875	Retina	1	1	1%	0%	100%	3%
36768907	Retina	1	1	1%	0%	100%	3%
36769563	Retina	1	1	1%	0%	100%	3%
36769641	Retina	4	4	5%	0%	100%	10%
37697547	Keypoint	38	17	22%	55%	45%	29%
37697714	Keypoint	39	22	28%	44%	56%	38%
37697946	Keypoint	38	22	28%	42%	58%	38%
37698009	Keypoint	42	22	28%	48%	52%	37%
37732318	Keypoint	39	22	28%	44%	56%	38%
37732355	Keypoint	36	20	26%	44%	56%	35%
37732415	Keypoint	35	22	28%	37%	63%	39%
37792158	Keypoint	37	20	26%	46%	54%	35%

Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	23	10	13%	57%	43%	20%
35857281	Faster	22	11	14%	50%	50%	22%
35857345	Faster	16	8	10%	50%	50%	17%
35857389	Faster	12	9	12%	25%	75%	20%
35857890	Faster	18	12	15%	33%	67%	25%
35857952	Faster	16	10	13%	38%	63%	21%
35858015	Faster	21	13	17%	38%	62%	26%
35858198	Faster	22	12	15%	45%	55%	24%
36761737	Faster	25	15	19%	40%	60%	29%
36761786	Faster	22	15	19%	32%	68%	30%
35858791	Mask	23	10	13%	57%	43%	20%
35858828	Mask	25	12	15%	52%	48%	23%
35858933	Mask	22	14	18%	36%	64%	28%
35859007	Mask	17	10	13%	41%	59%	21%
35859745	Mask	20	12	15%	40%	60%	24%
35861795	Mask	27	17	22%	37%	63%	32%
35861858	Mask	20	12	15%	40%	60%	24%
36494496	Mask	24	15	19%	38%	63%	29%
36761843	Mask	27	18	23%	33%	67%	34%
36762092	Mask	26	17	22%	35%	65%	33%
37129812	Mask	53	22	28%	58%	42%	34%

Tabella B.10: Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	0	0	0%	0%	0%	0%
36768677	Retina	1	1	1%	0%	100%	3%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	0	0	0%	0%	0%	0%
36768875	Retina	0	0	0%	0%	0%	0%
36768907	Retina	0	0	0%	0%	0%	0%
36769563	Retina	0	0	0%	0%	0%	0%
36769641	Retina	2	2	3%	0%	100%	5%
37697547	Keypoint	33	15	19%	55%	45%	27%
37697714	Keypoint	32	19	24%	41%	59%	35%
37697946	Keypoint	30	18	23%	40%	60%	33%
37698009	Keypoint	35	20	26%	43%	57%	35%
37732318	Keypoint	33	20	26%	39%	61%	36%
37732355	Keypoint	30	17	22%	43%	57%	31%
37732415	Keypoint	30	21	27%	30%	70%	39%
37792158	Keypoint	33	19	24%	42%	58%	34%

Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	15	7	9%	53%	47%	15%
35857281	Faster	12	7	9%	42%	58%	16%
35857345	Faster	12	6	8%	50%	50%	13%
35857389	Faster	12	9	12%	25%	75%	20%
35857890	Faster	13	9	12%	31%	69%	20%
35857952	Faster	11	6	8%	45%	55%	13%
35858015	Faster	18	12	15%	33%	67%	25%
35858198	Faster	21	12	15%	43%	57%	24%
36761737	Faster	22	14	18%	36%	64%	28%
36761786	Faster	16	13	17%	19%	81%	28%
35858791	Mask	14	6	8%	57%	43%	13%
35858828	Mask	13	9	12%	31%	69%	20%
35858933	Mask	18	13	17%	28%	72%	27%
35859007	Mask	12	9	12%	25%	75%	20%
35859745	Mask	20	12	15%	40%	60%	24%
35861795	Mask	17	12	15%	29%	71%	25%
35861858	Mask	16	12	15%	25%	75%	26%
36494496	Mask	20	13	17%	35%	65%	27%
36761843	Mask	21	16	21%	24%	76%	32%
36762092	Mask	21	16	21%	24%	76%	32%
37129812	Mask	32	18	23%	44%	56%	33%

Tabella B.11: Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	0	0	0%	0%	0%	0%
36768677	Retina	0	0	0%	0%	0%	0%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	0	0	0%	0%	0%	0%
36768875	Retina	0	0	0%	0%	0%	0%
36768907	Retina	0	0	0%	0%	0%	0%
36769563	Retina	0	0	0%	0%	0%	0%
36769641	Retina	0	0	0%	0%	0%	0%
37697547	Keypoint	23	12	15%	48%	52%	24%
37697714	Keypoint	26	17	22%	35%	65%	33%
37697946	Keypoint	22	16	21%	27%	73%	32%
37698009	Keypoint	27	17	22%	37%	63%	32%
37732318	Keypoint	24	15	19%	38%	63%	29%
37732355	Keypoint	22	13	17%	41%	59%	26%
37732415	Keypoint	25	17	22%	32%	68%	33%
37792158	Keypoint	24	15	19%	38%	63%	29%

Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	117	19	34%	84%	16%	22%
35857281	Faster	92	21	38%	77%	23%	28%
35857345	Faster	129	22	39%	83%	17%	24%
35857389	Faster	127	20	36%	84%	16%	22%
35857890	Faster	122	21	38%	83%	17%	24%
35857952	Faster	79	16	29%	80%	20%	24%
35858015	Faster	104	19	34%	82%	18%	24%
35858198	Faster	85	15	27%	82%	18%	21%
36761737	Faster	139	27	48%	81%	19%	28%
36761786	Faster	83	20	36%	76%	24%	29%
35858791	Mask	102	20	36%	80%	20%	25%
35858828	Mask	105	21	38%	80%	20%	26%
35858933	Mask	149	24	43%	84%	16%	23%
35859007	Mask	108	19	34%	82%	18%	23%
35859745	Mask	80	19	34%	76%	24%	28%
35861795	Mask	123	25	45%	80%	20%	28%
35861858	Mask	79	20	36%	75%	25%	30%
36494496	Mask	119	22	39%	82%	18%	25%
36761843	Mask	138	25	45%	82%	18%	26%
36762092	Mask	86	24	43%	72%	28%	34%
37129812	Mask	223	28	50%	87%	13%	20%

Tabelle B.12: Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	208	26	46%	88%	13%	20%
36768677	Retina	190	19	34%	90%	10%	15%
36768744	Retina	233	25	45%	89%	11%	17%
36768840	Retina	200	20	36%	90%	10%	16%
36768875	Retina	185	19	34%	90%	10%	16%
36768907	Retina	167	16	29%	90%	10%	14%
36769563	Retina	223	19	34%	91%	9%	14%
36769641	Retina	157	17	30%	89%	11%	16%
37697547	Keypoint	195	31	55%	84%	16%	25%
37697714	Keypoint	200	30	54%	85%	15%	23%
37697946	Keypoint	204	31	55%	85%	15%	24%
37698009	Keypoint	190	29	52%	85%	15%	24%
37732318	Keypoint	140	26	46%	81%	19%	27%
37732355	Keypoint	170	25	45%	85%	15%	22%
37732415	Keypoint	154	33	59%	79%	21%	31%
37792158	Keypoint	185	31	55%	83%	17%	26%

Validierungsergebnisse bei 0% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	97	19	34%	80%	20%	25%
35857281	Faster	79	20	36%	75%	25%	30%
35857345	Faster	91	20	36%	78%	22%	27%
35857389	Faster	84	20	36%	76%	24%	29%
35857890	Faster	93	21	38%	77%	23%	28%
35857952	Faster	63	15	27%	76%	24%	25%
35858015	Faster	77	17	30%	78%	22%	26%
35858198	Faster	68	14	25%	79%	21%	23%
36761737	Faster	108	27	48%	75%	25%	33%
36761786	Faster	64	17	30%	73%	27%	28%
35858791	Mask	73	20	36%	73%	27%	31%
35858828	Mask	83	19	34%	77%	23%	27%
35858933	Mask	118	22	39%	81%	19%	25%
35859007	Mask	81	17	30%	79%	21%	25%
35859745	Mask	59	19	34%	68%	32%	33%
35861795	Mask	86	24	43%	72%	28%	34%
35861858	Mask	58	19	34%	67%	33%	33%
36494496	Mask	87	20	36%	77%	23%	28%
36761843	Mask	104	24	43%	77%	23%	30%
36762092	Mask	67	23	41%	66%	34%	37%
37129812	Mask	196	26	46%	87%	13%	21%

Tabelle B.13: Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	150	21	38%	86%	14%	20%
36768677	Retina	127	16	29%	87%	13%	17%
36768744	Retina	166	24	43%	86%	14%	22%
36768840	Retina	111	17	30%	85%	15%	20%
36768875	Retina	116	16	29%	86%	14%	19%
36768907	Retina	110	15	27%	86%	14%	18%
36769563	Retina	135	16	29%	88%	12%	17%
36769641	Retina	71	14	25%	80%	20%	22%
37697547	Keypoint	145	30	54%	79%	21%	30%
37697714	Keypoint	146	28	50%	81%	19%	28%
37697946	Keypoint	135	27	48%	80%	20%	28%
37698009	Keypoint	139	27	48%	81%	19%	28%
37732318	Keypoint	102	25	45%	75%	25%	32%
37732355	Keypoint	124	25	45%	80%	20%	28%
37732415	Keypoint	114	29	52%	75%	25%	34%
37792158	Keypoint	132	27	48%	80%	20%	29%

Validierungsergebnisse bei 10% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	73	17	30%	77%	23%	26%
35857281	Faster	63	18	32%	71%	29%	30%
35857345	Faster	61	17	30%	72%	28%	29%
35857389	Faster	51	18	32%	65%	35%	34%
35857890	Faster	63	19	34%	70%	30%	32%
35857952	Faster	47	14	25%	70%	30%	27%
35858015	Faster	55	17	30%	69%	31%	31%
35858198	Faster	44	14	25%	68%	32%	28%
36761737	Faster	77	23	41%	70%	30%	35%
36761786	Faster	45	15	27%	67%	33%	30%
35858791	Mask	64	17	30%	73%	27%	28%
35858828	Mask	72	18	32%	75%	25%	28%
35858933	Mask	81	20	36%	75%	25%	29%
35859007	Mask	55	16	29%	71%	29%	29%
35859745	Mask	45	18	32%	60%	40%	36%
35861795	Mask	64	22	39%	66%	34%	37%
35861858	Mask	44	18	32%	59%	41%	36%
36494496	Mask	57	17	30%	70%	30%	30%
36761843	Mask	72	24	43%	67%	33%	38%
36762092	Mask	54	21	38%	61%	39%	38%
37129812	Mask	160	26	46%	84%	16%	24%

Tabelle B.14: Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	59	16	29%	73%	27%	28%
36768677	Retina	47	15	27%	68%	32%	29%
36768744	Retina	53	15	27%	72%	28%	28%
36768840	Retina	46	12	21%	74%	26%	24%
36768875	Retina	39	11	20%	72%	28%	23%
36768907	Retina	46	12	21%	74%	26%	24%
36769563	Retina	51	11	20%	78%	22%	21%
36769641	Retina	33	14	25%	58%	42%	31%
37697547	Keypoint	100	25	45%	75%	25%	32%
37697714	Keypoint	102	24	43%	76%	24%	30%
37697946	Keypoint	92	26	46%	72%	28%	35%
37698009	Keypoint	94	25	45%	73%	27%	33%
37732318	Keypoint	71	23	41%	68%	32%	36%
37732355	Keypoint	86	22	39%	74%	26%	31%
37732415	Keypoint	79	27	48%	66%	34%	40%
37792158	Keypoint	87	25	45%	71%	29%	35%

Validierungsergebnisse bei 20% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	61	16	29%	74%	26%	27%
35857281	Faster	56	17	30%	70%	30%	30%
35857345	Faster	48	14	25%	71%	29%	27%
35857389	Faster	44	16	29%	64%	36%	32%
35857890	Faster	45	17	30%	62%	38%	34%
35857952	Faster	32	14	25%	56%	44%	32%
35858015	Faster	45	17	30%	62%	38%	34%
35858198	Faster	35	14	25%	60%	40%	31%
36761737	Faster	59	22	39%	63%	37%	38%
36761786	Faster	38	15	27%	61%	39%	32%
35858791	Mask	54	17	30%	69%	31%	31%
35858828	Mask	65	18	32%	72%	28%	30%
35858933	Mask	61	18	32%	70%	30%	31%
35859007	Mask	42	15	27%	64%	36%	31%
35859745	Mask	38	16	29%	58%	42%	34%
35861795	Mask	50	20	36%	60%	40%	38%
35861858	Mask	36	18	32%	50%	50%	39%
36494496	Mask	43	16	29%	63%	37%	32%
36761843	Mask	55	21	38%	62%	38%	38%
36762092	Mask	45	20	36%	56%	44%	40%
37129812	Mask	132	26	46%	80%	20%	28%

Tabelle B.15: Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	36	14	25%	61%	39%	30%
36768677	Retina	26	10	18%	62%	38%	24%
36768744	Retina	24	14	25%	42%	58%	35%
36768840	Retina	25	10	18%	60%	40%	25%
36768875	Retina	23	10	18%	57%	43%	25%
36768907	Retina	23	10	18%	57%	43%	25%
36769563	Retina	24	11	20%	54%	46%	28%
36769641	Retina	24	12	21%	50%	50%	30%
37697547	Keypoint	78	19	34%	76%	24%	28%
37697714	Keypoint	76	23	41%	70%	30%	35%
37697946	Keypoint	68	25	45%	63%	37%	40%
37698009	Keypoint	70	24	43%	66%	34%	38%
37732318	Keypoint	61	22	39%	64%	36%	38%
37732355	Keypoint	74	21	38%	72%	28%	32%
37732415	Keypoint	63	23	41%	63%	37%	39%
37792158	Keypoint	59	20	36%	66%	34%	35%

Validierungsergebnisse bei 30% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	51	14	25%	73%	27%	26%
35857281	Faster	48	17	30%	65%	35%	33%
35857345	Faster	42	14	25%	67%	33%	29%
35857389	Faster	37	15	27%	59%	41%	32%
35857890	Faster	37	15	27%	59%	41%	32%
35857952	Faster	26	13	23%	50%	50%	32%
35858015	Faster	38	16	29%	58%	42%	34%
35858198	Faster	32	14	25%	56%	44%	32%
36761737	Faster	46	19	34%	59%	41%	37%
36761786	Faster	34	15	27%	56%	44%	33%
35858791	Mask	50	17	30%	66%	34%	32%
35858828	Mask	54	16	29%	70%	30%	29%
35858933	Mask	49	18	32%	63%	37%	34%
35859007	Mask	35	14	25%	60%	40%	31%
35859745	Mask	33	15	27%	55%	45%	34%
35861795	Mask	40	19	34%	53%	48%	40%
35861858	Mask	30	17	30%	43%	57%	40%
36494496	Mask	38	15	27%	61%	39%	32%
36761843	Mask	46	21	38%	54%	46%	41%
36762092	Mask	37	19	34%	49%	51%	41%
37129812	Mask	105	24	43%	77%	23%	30%

Tabelle B.16: Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	17	10	18%	41%	59%	27%
36768677	Retina	12	7	13%	42%	58%	21%
36768744	Retina	17	12	21%	29%	71%	33%
36768840	Retina	16	9	16%	44%	56%	25%
36768875	Retina	13	8	14%	38%	62%	23%
36768907	Retina	15	9	16%	40%	60%	25%
36769563	Retina	12	8	14%	33%	67%	24%
36769641	Retina	17	12	21%	29%	71%	33%
37697547	Keypoint	65	19	34%	71%	29%	31%
37697714	Keypoint	56	22	39%	61%	39%	39%
37697946	Keypoint	61	24	43%	61%	39%	41%
37698009	Keypoint	59	23	41%	61%	39%	40%
37732318	Keypoint	53	21	38%	60%	40%	39%
37732355	Keypoint	57	20	36%	65%	35%	35%
37732415	Keypoint	48	22	39%	54%	46%	42%
37792158	Keypoint	45	19	34%	58%	42%	38%

Validierungsergebnisse bei 40% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	47	12	21%	74%	26%	23%
35857281	Faster	45	14	25%	69%	31%	28%
35857345	Faster	32	13	23%	59%	41%	30%
35857389	Faster	28	10	18%	64%	36%	24%
35857890	Faster	32	15	27%	53%	47%	34%
35857952	Faster	24	13	23%	46%	54%	33%
35858015	Faster	34	16	29%	53%	47%	36%
35858198	Faster	27	13	23%	52%	48%	31%
36761737	Faster	39	18	32%	54%	46%	38%
36761786	Faster	29	15	27%	48%	52%	35%
35858791	Mask	48	16	29%	67%	33%	31%
35858828	Mask	47	16	29%	66%	34%	31%
35858933	Mask	41	15	27%	63%	37%	31%
35859007	Mask	30	12	21%	60%	40%	28%
35859745	Mask	30	15	27%	50%	50%	35%
35861795	Mask	36	17	30%	53%	47%	37%
35861858	Mask	28	17	30%	39%	61%	40%
36494496	Mask	33	15	27%	55%	45%	34%
36761843	Mask	38	20	36%	47%	53%	43%
36762092	Mask	32	17	30%	47%	53%	39%
37129812	Mask	93	24	43%	74%	26%	32%

Tabelle B.17: Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	9	8	14%	11%	89%	25%
36768677	Retina	9	7	13%	22%	78%	22%
36768744	Retina	9	8	14%	11%	89%	25%
36768840	Retina	12	8	14%	33%	67%	24%
36768875	Retina	9	5	9%	44%	56%	15%
36768907	Retina	9	6	11%	33%	67%	18%
36769563	Retina	7	6	11%	14%	86%	19%
36769641	Retina	12	11	20%	8%	92%	32%
37697547	Keypoint	51	17	30%	67%	33%	32%
37697714	Keypoint	51	21	38%	59%	41%	39%
37697946	Keypoint	48	24	43%	50%	50%	46%
37698009	Keypoint	53	22	39%	58%	42%	40%
37732318	Keypoint	48	20	36%	58%	42%	38%
37732355	Keypoint	53	20	36%	62%	38%	37%
37732415	Keypoint	43	21	38%	51%	49%	42%
37792158	Keypoint	41	19	34%	54%	46%	39%

Validierungsergebnisse bei 50% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	39	11	20%	72%	28%	23%
35857281	Faster	38	14	25%	63%	37%	30%
35857345	Faster	29	13	23%	55%	45%	31%
35857389	Faster	23	10	18%	57%	43%	25%
35857890	Faster	28	15	27%	46%	54%	36%
35857952	Faster	21	12	21%	43%	57%	31%
35858015	Faster	28	15	27%	46%	54%	36%
35858198	Faster	24	12	21%	50%	50%	30%
36761737	Faster	32	16	29%	50%	50%	36%
36761786	Faster	27	15	27%	44%	56%	36%
35858791	Mask	42	15	27%	64%	36%	31%
35858828	Mask	40	16	29%	60%	40%	33%
35858933	Mask	35	14	25%	60%	40%	31%
35859007	Mask	25	12	21%	52%	48%	30%
35859745	Mask	29	15	27%	48%	52%	35%
35861795	Mask	32	17	30%	47%	53%	39%
35861858	Mask	26	16	29%	38%	62%	39%
36494496	Mask	30	15	27%	50%	50%	35%
36761843	Mask	33	18	32%	45%	55%	40%
36762092	Mask	31	17	30%	45%	55%	39%
37129812	Mask	78	23	41%	71%	29%	34%

Tabelle B.18: Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	5	5	9%	0%	100%	16%
36768677	Retina	7	6	11%	14%	86%	19%
36768744	Retina	5	4	7%	20%	80%	13%
36768840	Retina	9	6	11%	33%	67%	18%
36768875	Retina	2	2	4%	0%	100%	7%
36768907	Retina	4	3	5%	25%	75%	10%
36769563	Retina	4	3	5%	25%	75%	10%
36769641	Retina	7	7	13%	0%	100%	22%
37697547	Keypoint	45	16	29%	64%	36%	32%
37697714	Keypoint	41	19	34%	54%	46%	39%
37697946	Keypoint	43	24	43%	44%	56%	48%
37698009	Keypoint	49	22	39%	55%	45%	42%
37732318	Keypoint	42	19	34%	55%	45%	39%
37732355	Keypoint	39	17	30%	56%	44%	36%
37732415	Keypoint	40	21	38%	48%	53%	44%
37792158	Keypoint	38	19	34%	50%	50%	40%

Validierungsergebnisse bei 60% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	32	10	18%	69%	31%	23%
35857281	Faster	35	13	23%	63%	37%	29%
35857345	Faster	22	11	20%	50%	50%	28%
35857389	Faster	17	10	18%	41%	59%	27%
35857890	Faster	23	14	25%	39%	61%	35%
35857952	Faster	19	11	20%	42%	58%	29%
35858015	Faster	26	15	27%	42%	58%	37%
35858198	Faster	22	12	21%	45%	55%	31%
36761737	Faster	28	15	27%	46%	54%	36%
36761786	Faster	24	15	27%	38%	63%	38%
35858791	Mask	37	14	25%	62%	38%	30%
35858828	Mask	33	14	25%	58%	42%	31%
35858933	Mask	28	14	25%	50%	50%	33%
35859007	Mask	21	12	21%	43%	57%	31%
35859745	Mask	26	13	23%	50%	50%	32%
35861795	Mask	29	16	29%	45%	55%	38%
35861858	Mask	22	13	23%	41%	59%	33%
36494496	Mask	25	15	27%	40%	60%	37%
36761843	Mask	29	18	32%	38%	62%	42%
36762092	Mask	30	17	30%	43%	57%	40%
37129812	Mask	65	22	39%	66%	34%	36%

Tabelle B.19: Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	3	3	5%	0%	100%	10%
36768677	Retina	4	4	7%	0%	100%	13%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	2	2	4%	0%	100%	7%
36768875	Retina	1	1	2%	0%	100%	4%
36768907	Retina	1	1	2%	0%	100%	4%
36769563	Retina	1	1	2%	0%	100%	4%
36769641	Retina	4	4	7%	0%	100%	13%
37697547	Keypoint	38	14	25%	63%	37%	30%
37697714	Keypoint	39	19	34%	51%	49%	40%
37697946	Keypoint	38	21	38%	45%	55%	45%
37698009	Keypoint	42	20	36%	52%	48%	41%
37732318	Keypoint	39	19	34%	51%	49%	40%
37732355	Keypoint	36	17	30%	53%	47%	37%
37732415	Keypoint	35	20	36%	43%	57%	44%
37792158	Keypoint	37	18	32%	51%	49%	39%

Validierungsergebnisse bei 70% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	23	9	16%	61%	39%	23%
35857281	Faster	22	10	18%	55%	45%	26%
35857345	Faster	16	8	14%	50%	50%	22%
35857389	Faster	12	9	16%	25%	75%	26%
35857890	Faster	18	12	21%	33%	67%	32%
35857952	Faster	16	10	18%	38%	63%	28%
35858015	Faster	21	13	23%	38%	62%	34%
35858198	Faster	22	12	21%	45%	55%	31%
36761737	Faster	25	15	27%	40%	60%	37%
36761786	Faster	22	15	27%	32%	68%	38%
35858791	Mask	23	9	16%	61%	39%	23%
35858828	Mask	25	11	20%	56%	44%	27%
35858933	Mask	22	13	23%	41%	59%	33%
35859007	Mask	17	10	18%	41%	59%	27%
35859745	Mask	20	12	21%	40%	60%	32%
35861795	Mask	27	15	27%	44%	56%	36%
35861858	Mask	20	12	21%	40%	60%	32%
36494496	Mask	24	15	27%	38%	63%	38%
36761843	Mask	27	18	32%	33%	67%	43%
36762092	Mask	26	17	30%	35%	65%	41%
37129812	Mask	53	21	38%	60%	40%	39%

Tabelle B.20: Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	0	0	0%	0%	0%	0%
36768677	Retina	1	1	2%	0%	100%	4%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	0	0	0%	0%	0%	0%
36768875	Retina	0	0	0%	0%	0%	0%
36768907	Retina	0	0	0%	0%	0%	0%
36769563	Retina	0	0	0%	0%	0%	0%
36769641	Retina	2	2	4%	0%	100%	7%
37697547	Keypoint	33	12	21%	64%	36%	27%
37697714	Keypoint	32	17	30%	47%	53%	39%
37697946	Keypoint	30	18	32%	40%	60%	42%
37698009	Keypoint	35	19	34%	46%	54%	42%
37732318	Keypoint	33	18	32%	45%	55%	40%
37732355	Keypoint	30	16	29%	47%	53%	37%
37732415	Keypoint	30	20	36%	33%	67%	47%
37792158	Keypoint	33	17	30%	48%	52%	38%

Validierungsergebnisse bei 80% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
35857197	Faster	15	7	13%	53%	47%	20%
35857281	Faster	12	7	13%	42%	58%	21%
35857345	Faster	12	6	11%	50%	50%	18%
35857389	Faster	12	9	16%	25%	75%	26%
35857890	Faster	13	9	16%	31%	69%	26%
35857952	Faster	11	6	11%	45%	55%	18%
35858015	Faster	18	12	21%	33%	67%	32%
35858198	Faster	21	12	21%	43%	57%	31%
36761737	Faster	22	14	25%	36%	64%	36%
36761786	Faster	16	13	23%	19%	81%	36%
35858791	Mask	14	6	11%	57%	43%	17%
35858828	Mask	13	9	16%	31%	69%	26%
35858933	Mask	18	12	21%	33%	67%	32%
35859007	Mask	12	9	16%	25%	75%	26%
35859745	Mask	20	12	21%	40%	60%	32%
35861795	Mask	17	11	20%	35%	65%	30%
35861858	Mask	16	12	21%	25%	75%	33%
36494496	Mask	20	13	23%	35%	65%	34%
36761843	Mask	21	16	29%	24%	76%	42%
36762092	Mask	21	16	29%	24%	76%	42%
37129812	Mask	32	18	32%	44%	56%	41%

Tabelle B.21: Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenzahl

Netz-ID	Netztyp	Vorschläge	Richtig positiv	Recall	False Discovery Rate	Precision	F1-Score
36768636	Retina	0	0	0%	0%	0%	0%
36768677	Retina	0	0	0%	0%	0%	0%
36768744	Retina	0	0	0%	0%	0%	0%
36768840	Retina	0	0	0%	0%	0%	0%
36768875	Retina	0	0	0%	0%	0%	0%
36768907	Retina	0	0	0%	0%	0%	0%
36769563	Retina	0	0	0%	0%	0%	0%
36769641	Retina	0	0	0%	0%	0%	0%
37697547	Keypoint	23	10	18%	57%	43%	25%
37697714	Keypoint	26	16	29%	38%	62%	39%
37697946	Keypoint	22	16	29%	27%	73%	41%
37698009	Keypoint	27	17	30%	37%	63%	41%
37732318	Keypoint	24	15	27%	38%	63%	38%
37732355	Keypoint	22	13	23%	41%	59%	33%
37732415	Keypoint	25	16	29%	36%	64%	40%
37792158	Keypoint	24	14	25%	42%	58%	35%

Validierungsergebnisse bei 90% Schwellenwert für Klassenwahrscheinlichkeiten bei gekürzter Personenanzahl (Fortsetzung)

Typ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Faster	26%	26%	28%	28%	28%	27%	26%	26%	23%	20%
Retina	16%	18%	21%	22%	20%	17%	11%	5%	1%	0%
Mask	28%	30%	31%	32%	32%	31%	31%	30%	28%	25%
Keypoint	29%	33%	36%	36%	37%	37%	37%	36%	34%	30%

Tabelle B.22: F1-Scores nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11

Typ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Faster	78%	74%	67%	61%	57%	55%	50%	45%	41%	38%
Retina	89%	85%	71%	55%	37%	22%	15%	0%	0%	0%
Mask	76%	70%	65%	60%	55%	52%	49%	46%	42%	33%
Keypoint	79%	74%	66%	61%	55%	51%	47%	45%	42%	37%

Tabelle B.23: False Discovery Rate nach Netzarchitektur und Schwellenwert als Zusammenfassung der Tabellen B.2 bis B.11

Typ	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Faster	24%	27%	30%	32%	32%	31%	31%	31%	30%	26%
Retina	16%	19%	26%	28%	26%	22%	15%	7%	1%	0%
Mask	26%	30%	32%	34%	35%	35%	35%	35%	34%	32%
Keypoint	25%	29%	34%	36%	38%	39%	40%	39%	39%	36%

Tabelle B.24: F1-Score nach Netzarchitektur und Schwellenwert bei gekürzter Personenzahl als Zusammenfassung der Tabellen B.12 bis B.21

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	53%	49%	38%	31%	31%	27%	26%	22%	19%	15%
37697714	50%	47%	41%	36%	33%	32%	28%	28%	24%	22%
37697946	47%	40%	38%	35%	33%	32%	32%	28%	23%	21%
37698009	45%	41%	37%	35%	32%	31%	31%	28%	26%	22%
37732318	45%	42%	38%	35%	33%	32%	28%	28%	26%	19%
37732355	42%	41%	36%	35%	31%	31%	26%	26%	22%	17%
37732415	53%	45%	42%	36%	32%	29%	29%	28%	27%	22%
37792158	49%	42%	38%	31%	28%	28%	27%	26%	24%	19%

Tabelle B.25: Recall der Keypointnetze nach Schwellenwert als Ausschnitt aus den Tabellen B.2 bis B.11

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	55%	54%	45%	34%	34%	30%	29%	25%	21%	18%
37697714	54%	50%	43%	41%	39%	38%	34%	34%	30%	29%
37697946	55%	48%	46%	45%	43%	43%	43%	38%	32%	29%
37698009	52%	48%	45%	43%	41%	39%	39%	36%	34%	30%
37732318	46%	45%	41%	39%	38%	36%	34%	34%	32%	27%
37732355	45%	45%	39%	38%	36%	36%	30%	30%	29%	23%
37732415	59%	52%	48%	41%	39%	38%	38%	36%	36%	29%
37792158	55%	48%	45%	36%	34%	34%	34%	32%	30%	25%

Tabelle B.26: Recall der Keypointnetze nach Schwellenwert bei gekürzter Personenzahl als Ausschnitt aus den Tabellen B.12 bis B.21

Abb.	Breite	Höhe	Auflösung	Breite bei 100 dpi	Höhe bei 100 dpi	Fläche bei 100 dpi
3.1	3650 px	1500 px	300 dpi	93 cm	38 cm	3532 cm ²
3.3	1245 px	1200 px	72 dpi	32 cm	30 cm	964 cm ²
3.4	1603 px	1200 px	72 dpi	41 cm	30 cm	1241 cm ²
2.3	3177 px	1700 px	144 dpi	81 cm	43 cm	3484 cm ²
4.2	1600 px	742 px	96 dpi	41 cm	19 cm	766 cm ²
2.4	1227 px	1600 px	96 dpi	31 cm	41 cm	1267 cm ²

Tabelle B.27: Bildabmessungen der untersuchten Abbildungen

Abb.	Breite	Höhe	Auflösung	Breite bei 100 dpi	Höhe bei 100 dpi	Fläche bei 100 dpi
151577 / 1	709 px	278 px	96 dpi	18 cm	7 cm	126 cm ²
151577 / 2	709 px	271 px	96 dpi	18 cm	7 cm	126 cm ²
151577 / 3	709 px	268 px	96 dpi	18 cm	7 cm	126 cm ²
151577 / 4	709 px	266 px	96 dpi	18 cm	7 cm	126 cm ²
151577 / 5	709 px	256 px	96 dpi	18 cm	7 cm	126 cm ²
151577 / 6	709 px	256 px	96 dpi	18 cm	7 cm	126 cm ²
151578 / 1	709 px	208 px	96 dpi	18 cm	5 cm	90 cm ²
157046 / 1	709 px	237 px	96 dpi	18 cm	6 cm	108 cm ²
157047 / 1	709 px	296 px	96 dpi	18 cm	8 cm	144 cm ²
157048 / 1	709 px	257 px	96 dpi	18 cm	7 cm	126 cm ²
157682 / 1	709 px	333 px	144 dpi	18 cm	8 cm	144 cm ²
157999 / 1	709 px	243 px	144 dpi	18 cm	6 cm	108 cm ²
157999 / 2	709 px	225 px	144 dpi	18 cm	6 cm	108 cm ²
159240 / 1	709 px	210 px	144 dpi	18 cm	5 cm	90 cm ²
159240 / 2	703 px	295 px	144 dpi	18 cm	7 cm	126 cm ²
159364 / 1	709 px	479 px	144 dpi	18 cm	12 cm	216 cm ²
159869 / 1	669 px	199 px	144 dpi	17 cm	5 cm	85 cm ²
159869 / 2	709 px	502 px	144 dpi	18 cm	13 cm	234 cm ²
159869 / 3	709 px	509 px	144 dpi	18 cm	13 cm	234 cm ²
159869 / 4	709 px	535 px	144 dpi	18 cm	14 cm	252 cm ²
159869 / 5	709 px	563 px	144 dpi	18 cm	14 cm	252 cm ²
159869 / 6	709 px	571 px	144 dpi	18 cm	15 cm	270 cm ²
159869 / 7	709 px	484 px	144 dpi	18 cm	12 cm	216 cm ²
159869 / 8	709 px	531 px	144 dpi	18 cm	13 cm	234 cm ²
159869 / 9	663 px	537 px	144 dpi	17 cm	14 cm	238 cm ²

Tabelle B.28: Abmessungen der Abbildungen der Census-Datenbank

Abb.	Breite	Höhe	Auflösung	Breite bei 100 dpi	Höhe bei 100 dpi	Fläche bei 100 dpi
161355 / 1	709 px	361 px	144 dpi	18 cm	9 cm	162 cm ²
161774 / 1	709 px	252 px	144 dpi	18 cm	6 cm	108 cm ²
161774 / 2	709 px	212 px	144 dpi	18 cm	5 cm	90 cm ²
161878 / 1	709 px	225 px	144 dpi	18 cm	6 cm	108 cm ²
161891 / 1	703 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161891 / 2	703 px	534 px	144 dpi	18 cm	14 cm	252 cm ²
161892 / 1	703 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161892 / 2	703 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161892 / 3	703 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161892 / 4	703 px	290 px	144 dpi	18 cm	7 cm	126 cm ²
161892 / 5	703 px	644 px	144 dpi	18 cm	16 cm	288 cm ²
161893 / 1	703 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161893 / 2	709 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161893 / 3	709 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161893 / 4	709 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161893 / 5	709 px	473 px	144 dpi	18 cm	12 cm	216 cm ²
161893 / 6	709 px	405 px	144 dpi	18 cm	10 cm	180 cm ²
161893 / 7	582 px	795 px	144 dpi	15 cm	20 cm	300 cm ²
161893 / 8	515 px	625 px	144 dpi	13 cm	16 cm	208 cm ²
161894 / 1	709 px	265 px	144 dpi	18 cm	7 cm	126 cm ²
161894 / 2	709 px	436 px	144 dpi	18 cm	11 cm	198 cm ²
161895 / 1	709 px	480 px	144 dpi	18 cm	12 cm	216 cm ²
161895 / 2	649 px	715 px	144 dpi	16 cm	18 cm	288 cm ²
161895 / 3	709 px	791 px	144 dpi	18 cm	20 cm	360 cm ²
161895 / 4	709 px	341 px	144 dpi	18 cm	9 cm	162 cm ²
161908 / 1	709 px	209 px	144 dpi	18 cm	5 cm	90 cm ²
10077142 / 1	709 px	369 px	144 dpi	18 cm	9 cm	162 cm ²

Abmessungen der Abbildungen der Census-Datenbank (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	85%	76%	93%	37%	86%	55%	61%	58%	107%									
A.2	7%	25%	6%	25%	11%	12%	32%	3%	37%									
A.3	78%	86%	67%	80%	8%	21%	21%	29%	3%									
A.4	80%	79%	96%	67%	60%	39%	7%	46%	6%									
A.5	82%	72%	60%	68%	1%	30%	17%	32%	18%									
A.6	68%	73%	74%	24%	71%	0%	29%	0%	0%									
A.7	94%	86%	98%	20%	88%	65%	39%	59%	8%									
A.8	83%	76%	88%	22%	89%	41%	39%	28%	24%									
A.9	29%	30%	20%	8%	10%	0%	11%	0%	6%									
A.10	60%	56%	55%	50%	22%	12%	38%	0%	36%									
A.11	64%	58%	90%	23%	67%	25%	50%	0%	40%									
A.12	0%	0%	0%	0%	0%	0%	0%	0%	0%									
A.13	21%	35%	29%	26%	38%	0%	36%	0%	42%									
A.14	41%	57%	67%	40%	58%	30%	29%	39%	20%									
A.15	96%	101%	105%	49%	91%	56%	50%	17%	76%									
A.16	77%	70%	88%	5%	63%	0%	27%	0%	16%									
A.17	18%	20%	26%	0%	12%	22%	6%	0%	37%									
A.18	69%	67%	80%	28%	78%	48%	43%	51%	48%									
A.19	79%	86%	77%	38%	59%	45%	45%	37%	58%									
A.20	0%	0%	0%	0%	0%	0%	0%	0%	0%									
A.21	92%	97%	109%	25%	95%	59%	60%	54%	68%									
A.22	64%	62%	74%	8%	63%	15%	48%	26%	57%									
A.23	65%	52%	52%	0%	44%	0%	0%	0%	0%									
A.24	76%	54%	81%	2%	94%	38%	45%	19%	60%									

Tabelle B.30: Wahrscheinlichkeiten der Keypointerkennung für Netz 37697547

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	67%	89%	38%	13%	24%	3%	21%	0%
A.2	7%	57%	0%	19%	0%	46%	0%	0%
A.3	43%	10%	0%	0%	18%	21%	17%	21%
A.4	28%	28%	17%	10%	19%	7%	2%	21%
A.5	7%	21%	16%	0%	10%	0%	10%	0%
A.6	23%	0%	28%	5%	32%	1%	24%	9%
A.7	47%	30%	0%	0%	0%	0%	0%	3%
A.8	13%	5%	40%	64%	0%	33%	6%	20%
A.9	0%	0%	0%	0%	9%	0%	1%	0%
A.10	0%	46%	0%	0%	0%	0%	0%	0%
A.11	5%	72%	10%	22%	0%	10%	0%	4%
A.12	0%	0%	0%	0%	0%	5%	3%	7%
A.13	37%	48%	23%	33%	35%	35%	19%	4%
A.14	48%	42%	17%	28%	0%	24%	0%	0%
A.15	9%	71%	85%	81%	36%	65%	4%	7%
A.16	0%	34%	0%	0%	0%	18%	0%	12%
A.17	24%	60%	20%	15%	0%	13%	0%	0%
A.18	41%	58%	39%	23%	40%	29%	26%	13%
A.19	51%	43%	30%	33%	9%	16%	0%	0%
A.20	0%	0%	0%	0%	12%	0%	11%	18%
A.21	15%	103%	51%	46%	59%	33%	29%	30%
A.22	44%	55%	31%	38%	30%	26%	15%	8%
A.23	0%	0%	0%	0%	8%	20%	0%	0%
A.24	0%	97%	38%	38%	10%	11%	0%	5%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37697547 (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	103%	96%	117%	117%	31%	99%	62%	64%	57%	92%								
A.2	46%	40%	53%	53%	27%	31%	33%	47%	5%	53%								
A.3	80%	96%	75%	75%	79%	30%	1%	37%	1%	0%								
A.4	99%	108%	126%	126%	85%	95%	60%	50%	66%	5%								
A.5	84%	81%	86%	86%	78%	0%	28%	23%	7%	38%								
A.6	81%	87%	98%	98%	36%	87%	28%	27%	0%	2%								
A.7	85%	79%	101%	101%	12%	80%	49%	35%	46%	1%								
A.8	101%	100%	113%	113%	26%	77%	59%	47%	34%	22%								
A.9	62%	76%	68%	68%	44%	49%	21%	37%	0%	17%								
A.10	83%	93%	87%	87%	79%	22%	12%	40%	0%	48%								
A.11	52%	68%	83%	83%	20%	43%	25%	51%	0%	42%								
A.12	0%	8%	12%	12%	13%	27%	0%	0%	0%	0%								
A.13	0%	12%	0%	0%	0%	21%	0%	36%	0%	34%								
A.14	35%	52%	61%	61%	22%	51%	10%	10%	10%	39%								
A.15	96%	107%	119%	119%	44%	95%	59%	60%	22%	74%								
A.16	88%	73%	107%	107%	0%	79%	21%	50%	0%	40%								
A.17	48%	64%	72%	72%	21%	47%	27%	11%	8%	48%								
A.18	88%	91%	97%	97%	26%	98%	47%	57%	60%	67%								
A.19	90%	104%	99%	99%	49%	85%	57%	54%	47%	66%								
A.20	0%	9%	10%	10%	0%	0%	0%	0%	0%	0%								
A.21	122%	125%	133%	133%	32%	125%	66%	66%	64%	76%								
A.22	76%	65%	90%	90%	24%	81%	43%	57%	52%	69%								
A.23	60%	29%	42%	42%	2%	9%	0%	0%	0%	0%								
A.24	88%	55%	106%	106%	17%	100%	58%	55%	13%	61%								

Tabelle B.31: Wahrscheinlichkeiten der Keypointerkennung für Netz 37697714

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	69%	84%	34%	0%	8%	21%	0%	0%
A.2	20%	72%	0%	0%	0%	53%	0%	0%
A.3	38%	2%	0%	14%	3%	12%	17%	9%
A.4	41%	30%	36%	26%	8%	2%	0%	0%
A.5	1%	34%	1%	3%	0%	15%	0%	0%
A.6	6%	0%	34%	11%	6%	15%	5%	0%
A.7	26%	29%	0%	0%	0%	0%	0%	0%
A.8	7%	31%	40%	49%	0%	32%	41%	4%
A.9	0%	0%	0%	0%	20%	0%	7%	0%
A.10	0%	68%	4%	0%	0%	0%	0%	0%
A.11	0%	75%	29%	45%	6%	14%	4%	5%
A.12	0%	0%	0%	0%	0%	0%	0%	0%
A.13	16%	55%	0%	0%	15%	23%	9%	12%
A.14	19%	49%	0%	0%	0%	0%	3%	0%
A.15	29%	74%	71%	70%	61%	61%	40%	25%
A.16	0%	68%	0%	4%	0%	0%	0%	0%
A.17	2%	64%	40%	30%	22%	37%	26%	17%
A.18	45%	66%	53%	28%	26%	15%	11%	0%
A.19	47%	54%	72%	57%	59%	65%	15%	2%
A.20	0%	0%	17%	11%	18%	36%	16%	8%
A.21	27%	82%	71%	70%	55%	29%	0%	16%
A.22	50%	67%	50%	47%	52%	36%	17%	15%
A.23	0%	0%	0%	0%	8%	0%	0%	0%
A.24	0%	103%	54%	64%	28%	27%	0%	0%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37697714 (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	84%	88%	79%	79%	41%	95%	74%	52%	62%	73%								
A.2	41%	30%	39%	39%	15%	44%	27%	38%	13%	49%								
A.3	66%	98%	61%	61%	95%	11%	26%	40%	40%	6%								
A.4	82%	75%	82%	82%	65%	61%	45%	22%	58%	0%								
A.5	79%	65%	48%	48%	61%	11%	26%	10%	0%	0%								
A.6	81%	79%	83%	83%	39%	77%	16%	24%	0%	0%								
A.7	101%	61%	99%	99%	18%	90%	52%	46%	57%	9%								
A.8	92%	74%	106%	106%	22%	90%	36%	46%	26%	34%								
A.9	33%	28%	27%	27%	22%	51%	8%	21%	0%	19%								
A.10	58%	58%	53%	53%	66%	20%	0%	51%	0%	44%								
A.11	49%	39%	40%	40%	0%	14%	0%	17%	0%	49%								
A.12	17%	16%	34%	34%	36%	25%	0%	0%	0%	0%								
A.13	39%	48%	52%	52%	28%	19%	0%	41%	0%	56%								
A.14	44%	61%	62%	62%	26%	55%	15%	35%	13%	31%								
A.15	102%	104%	112%	112%	41%	82%	55%	46%	25%	79%								
A.16	103%	62%	106%	106%	13%	79%	33%	49%	3%	56%								
A.17	73%	76%	74%	74%	48%	67%	36%	23%	21%	22%								
A.18	75%	70%	74%	74%	16%	76%	43%	59%	55%	49%								
A.19	83%	87%	87%	87%	31%	67%	50%	47%	34%	64%								
A.20	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%								
A.21	123%	95%	117%	117%	18%	101%	70%	60%	61%	81%								
A.22	58%	49%	58%	58%	8%	51%	33%	57%	57%	82%								
A.23	63%	45%	48%	48%	5%	25%	0%	0%	0%	0%								
A.24	66%	37%	84%	84%	9%	90%	58%	42%	8%	64%								

Tabelle B.32: Wahrscheinlichkeiten der Keypointerkennung für Netz 37697946

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	66%	81%	33%	27%	22%	12%	4%	7%
A.2	23%	56%	0%	13%	0%	39%	7%	0%
A.3	44%	0%	0%	17%	13%	0%	4%	0%
A.4	53%	19%	43%	33%	7%	0%	0%	13%
A.5	12%	44%	1%	25%	0%	4%	0%	0%
A.6	21%	5%	33%	37%	33%	6%	4%	2%
A.7	33%	11%	37%	31%	25%	1%	0%	0%
A.8	3%	32%	41%	66%	0%	38%	2%	29%
A.9	0%	0%	0%	6%	0%	0%	0%	0%
A.10	9%	61%	4%	7%	0%	0%	1%	0%
A.11	0%	53%	0%	9%	0%	15%	18%	6%
A.12	0%	0%	0%	0%	0%	0%	0%	0%
A.13	9%	71%	0%	5%	15%	27%	5%	11%
A.14	30%	48%	21%	21%	10%	16%	0%	2%
A.15	41%	81%	68%	77%	39%	65%	21%	28%
A.16	7%	77%	13%	26%	0%	0%	10%	0%
A.17	9%	36%	20%	38%	7%	25%	20%	2%
A.18	31%	51%	44%	47%	21%	32%	24%	12%
A.19	44%	37%	53%	57%	42%	35%	22%	0%
A.20	0%	0%	0%	8%	0%	8%	4%	10%
A.21	4%	103%	67%	62%	54%	46%	4%	21%
A.22	63%	60%	43%	44%	38%	32%	24%	11%
A.23	0%	13%	0%	0%	0%	0%	0%	0%
A.24	0%	104%	49%	64%	58%	41%	18%	10%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37697946 (Fortsetzung)

Abbildung	Nase		Auge		Ohr		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	80%	67%	86%	31%	85%	61%	64%	74%	71%			
A.2	0%	4%	0%	11%	0%	21%	37%	12%	44%			
A.3	77%	98%	85%	68%	15%	51%	2%	26%	0%			
A.4	105%	107%	118%	93%	71%	39%	32%	57%	18%			
A.5	75%	67%	59%	59%	36%	25%	15%	0%	27%			
A.6	83%	89%	93%	47%	91%	2%	36%	0%	0%			
A.7	100%	83%	111%	2%	91%	64%	41%	83%	4%			
A.8	79%	73%	94%	18%	89%	42%	50%	34%	50%			
A.9	36%	38%	42%	0%	0%	0%	0%	15%	16%			
A.10	84%	92%	81%	80%	30%	21%	45%	31%	55%			
A.11	49%	30%	32%	5%	15%	0%	42%	0%	55%			
A.12	9%	23%	27%	9%	21%	0%	0%	0%	0%			
A.13	27%	50%	35%	21%	28%	0%	24%	0%	0%			
A.14	43%	62%	63%	10%	52%	20%	51%	21%	50%			
A.15	97%	106%	108%	37%	89%	52%	52%	29%	72%			
A.16	91%	88%	100%	0%	78%	27%	50%	0%	56%			
A.17	73%	71%	92%	36%	74%	36%	52%	0%	28%			
A.18	63%	48%	63%	4%	68%	52%	50%	75%	66%			
A.19	91%	94%	99%	34%	80%	47%	46%	49%	59%			
A.20	0%	0%	0%	0%	0%	0%	0%	0%	0%			
A.21	113%	91%	113%	16%	101%	59%	60%	59%	80%			
A.22	58%	35%	57%	9%	51%	27%	41%	48%	61%			
A.23	80%	46%	50%	0%	0%	0%	0%	0%	0%			
A.24	44%	29%	78%	15%	100%	56%	49%	15%	54%			

Tabelle B.33: Wahrscheinlichkeiten der Keypointerkennung für Netz 37698009

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	71%	66%	47%	22%	25%	0%	9%	7%
A.2	18%	57%	0%	2%	0%	47%	2%	1%
A.3	9%	28%	16%	0%	10%	0%	12%	0%
A.4	46%	41%	38%	46%	0%	5%	3%	18%
A.5	10%	24%	1%	10%	0%	0%	0%	0%
A.6	19%	0%	14%	23%	4%	16%	17%	18%
A.7	70%	10%	21%	0%	0%	0%	0%	0%
A.8	0%	11%	37%	61%	0%	46%	15%	24%
A.9	26%	9%	0%	0%	32%	33%	8%	16%
A.10	33%	54%	13%	15%	0%	3%	0%	0%
A.11	0%	58%	0%	16%	0%	8%	25%	15%
A.12	0%	0%	0%	0%	0%	0%	0%	0%
A.13	36%	13%	0%	0%	0%	0%	8%	5%
A.14	35%	55%	30%	33%	7%	5%	0%	0%
A.15	30%	77%	72%	75%	51%	76%	31%	28%
A.16	1%	84%	12%	40%	0%	0%	0%	0%
A.17	0%	51%	43%	35%	2%	25%	26%	29%
A.18	43%	61%	57%	46%	50%	21%	26%	6%
A.19	43%	53%	63%	65%	51%	45%	28%	11%
A.20	0%	0%	0%	0%	0%	0%	5%	16%
A.21	20%	99%	70%	72%	65%	34%	29%	25%
A.22	50%	17%	37%	26%	52%	27%	33%	18%
A.23	0%	0%	0%	0%	0%	0%	3%	0%
A.24	0%	82%	50%	54%	45%	40%	23%	14%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37698009 (Fortsetzung)

Abbildung	Nase		Auge		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	125%	105%	120%	120%	47%	101%	70%	72%	56%	103%				
A.2	112%	51%	91%	91%	5%	65%	30%	45%	10%	29%				
A.3	88%	96%	78%	78%	80%	1%	33%	38%	20%	25%				
A.4	95%	96%	104%	104%	88%	84%	58%	42%	62%	6%				
A.5	90%	92%	97%	97%	79%	40%	25%	26%	16%	32%				
A.6	86%	83%	86%	86%	51%	64%	20%	40%	0%	0%				
A.7	114%	91%	121%	121%	12%	96%	62%	54%	66%	27%				
A.8	101%	88%	109%	109%	36%	99%	38%	39%	29%	18%				
A.9	23%	37%	24%	24%	25%	48%	0%	3%	13%	19%				
A.10	94%	97%	96%	96%	88%	33%	30%	43%	0%	39%				
A.11	31%	22%	33%	33%	0%	24%	0%	40%	0%	45%				
A.12	0%	38%	8%	8%	0%	0%	0%	0%	0%	0%				
A.13	68%	70%	59%	59%	45%	33%	0%	10%	0%	0%				
A.14	61%	60%	64%	64%	37%	51%	19%	0%	18%	11%				
A.15	110%	130%	129%	129%	58%	107%	64%	58%	29%	84%				
A.16	85%	108%	136%	136%	15%	77%	29%	47%	0%	55%				
A.17	67%	78%	73%	73%	42%	63%	30%	45%	19%	47%				
A.18	86%	84%	91%	91%	29%	73%	52%	51%	64%	52%				
A.19	98%	97%	97%	97%	40%	94%	56%	67%	58%	74%				
A.20	7%	0%	0%	0%	10%	0%	0%	0%	0%	0%				
A.21	120%	101%	127%	127%	23%	100%	61%	56%	73%	71%				
A.22	79%	65%	70%	70%	18%	60%	25%	65%	37%	88%				
A.23	82%	67%	75%	75%	11%	38%	0%	0%	0%	0%				
A.24	78%	44%	102%	102%	28%	107%	64%	55%	32%	83%				

Tabelle B.34: Wahrscheinlichkeiten der Keypointerkennung für Netz 37732318

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	70%	88%	35%	33%	31%	0%	9%	0%
A.2	16%	83%	0%	17%	0%	46%	0%	0%
A.3	47%	0%	2%	8%	0%	9%	14%	4%
A.4	58%	43%	45%	21%	3%	0%	0%	18%
A.5	10%	43%	3%	4%	0%	0%	0%	0%
A.6	38%	0%	24%	3%	23%	13%	0%	0%
A.7	67%	7%	38%	19%	32%	5%	7%	6%
A.8	7%	6%	47%	52%	0%	40%	0%	10%
A.9	6%	15%	0%	0%	0%	0%	0%	0%
A.10	27%	63%	14%	21%	0%	0%	0%	0%
A.11	0%	77%	4%	10%	2%	6%	38%	14%
A.12	0%	0%	0%	0%	0%	0%	0%	1%
A.13	55%	14%	0%	0%	0%	0%	2%	0%
A.14	44%	35%	19%	8%	0%	0%	5%	0%
A.15	18%	95%	88%	97%	64%	73%	20%	7%
A.16	0%	81%	20%	44%	0%	0%	0%	0%
A.17	2%	67%	47%	50%	1%	33%	27%	14%
A.18	53%	40%	57%	54%	50%	32%	36%	5%
A.19	54%	71%	56%	55%	33%	34%	36%	0%
A.20	0%	0%	0%	0%	0%	2%	0%	11%
A.21	42%	85%	62%	79%	68%	55%	10%	22%
A.22	43%	36%	43%	57%	45%	30%	32%	27%
A.23	0%	30%	0%	0%	0%	0%	0%	0%
A.24	0%	92%	48%	40%	51%	25%	47%	13%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37732318 (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	91%	80%	96%	26%	94%	54%	62%	44%	87%									
A.2	72%	18%	35%	14%	35%	26%	44%	3%	46%									
A.3	83%	97%	69%	77%	6%	33%	31%	31%	0%									
A.4	104%	112%	122%	102%	99%	50%	24%	78%	0%									
A.5	95%	103%	70%	90%	16%	28%	32%	0%	20%									
A.6	79%	64%	87%	23%	82%	0%	23%	1%	6%									
A.7	104%	80%	115%	19%	97%	45%	41%	37%	0%									
A.8	91%	76%	97%	24%	90%	33%	41%	14%	46%									
A.9	36%	44%	24%	50%	39%	8%	12%	23%	7%									
A.10	90%	92%	66%	81%	24%	15%	43%	0%	52%									
A.11	49%	51%	80%	17%	70%	16%	40%	0%	59%									
A.12	0%	5%	5%	0%	0%	4%	0%	0%	0%									
A.13	62%	67%	67%	44%	45%	0%	50%	0%	64%									
A.14	65%	78%	81%	18%	75%	33%	48%	15%	49%									
A.15	119%	129%	134%	51%	105%	52%	58%	21%	85%									
A.16	100%	66%	124%	5%	91%	11%	41%	3%	62%									
A.17	83%	89%	100%	54%	86%	35%	21%	13%	35%									
A.18	97%	86%	88%	24%	81%	48%	67%	64%	75%									
A.19	103%	101%	99%	49%	95%	51%	52%	32%	73%									
A.20	15%	12%	14%	0%	0%	0%	0%	0%	0%									
A.21	124%	98%	128%	33%	123%	64%	72%	69%	82%									
A.22	91%	74%	87%	29%	71%	24%	60%	30%	83%									
A.23	61%	33%	44%	2%	25%	0%	0%	0%	0%									
A.24	74%	52%	99%	18%	82%	67%	61%	9%	73%									

Tabelle B.35: Wahrscheinlichkeiten der Keypointerkennung für Netz 37732355

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	56%	69%	47%	48%	15%	17%	13%	11%
A.2	23%	66%	0%	9%	0%	12%	0%	0%
A.3	43%	0%	6%	15%	6%	21%	17%	18%
A.4	68%	10%	40%	17%	1%	8%	0%	0%
A.5	39%	36%	13%	1%	0%	0%	0%	0%
A.6	31%	16%	26%	16%	10%	4%	14%	14%
A.7	32%	26%	25%	27%	6%	16%	5%	0%
A.8	11%	14%	31%	41%	0%	33%	9%	18%
A.9	7%	18%	0%	0%	0%	0%	0%	0%
A.10	0%	74%	6%	7%	6%	0%	0%	0%
A.11	0%	97%	28%	40%	0%	10%	29%	14%
A.12	0%	0%	17%	17%	3%	6%	7%	1%
A.13	32%	73%	21%	32%	21%	41%	23%	18%
A.14	52%	62%	35%	44%	0%	34%	0%	0%
A.15	14%	90%	72%	84%	54%	94%	7%	34%
A.16	0%	84%	9%	34%	0%	13%	0%	0%
A.17	0%	65%	39%	24%	40%	35%	26%	19%
A.18	57%	62%	69%	58%	47%	33%	40%	5%
A.19	38%	55%	58%	75%	53%	61%	30%	0%
A.20	0%	4%	10%	16%	12%	27%	0%	12%
A.21	11%	103%	80%	81%	77%	52%	22%	12%
A.22	48%	39%	28%	41%	33%	26%	27%	17%
A.23	0%	24%	0%	0%	0%	10%	0%	0%
A.24	0%	94%	47%	44%	37%	34%	0%	2%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37732355 (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	99%	92%	107%	107%	43%	90%	67%	68%	70%	105%								
A.2	13%	0%	10%	10%	21%	0%	38%	39%	26%	54%								
A.3	85%	98%	73%	73%	82%	13%	8%	35%	19%	0%								
A.4	102%	96%	95%	95%	90%	79%	46%	16%	46%	8%								
A.5	94%	98%	87%	87%	94%	19%	40%	31%	9%	55%								
A.6	82%	72%	81%	81%	34%	92%	0%	30%	0%	0%								
A.7	113%	94%	125%	125%	25%	101%	62%	58%	64%	20%								
A.8	99%	99%	116%	116%	56%	99%	65%	54%	30%	43%								
A.9	44%	55%	56%	56%	37%	23%	0%	19%	11%	12%								
A.10	101%	89%	89%	89%	77%	21%	31%	42%	0%	25%								
A.11	44%	38%	59%	59%	7%	30%	3%	40%	0%	65%								
A.12	9%	25%	27%	27%	44%	29%	2%	0%	0%	0%								
A.13	28%	27%	50%	50%	18%	37%	0%	44%	0%	68%								
A.14	57%	78%	96%	96%	37%	85%	34%	54%	27%	63%								
A.15	126%	140%	144%	144%	68%	117%	79%	59%	55%	102%								
A.16	107%	72%	110%	110%	6%	77%	28%	38%	0%	63%								
A.17	82%	82%	88%	88%	49%	84%	33%	8%	37%	63%								
A.18	96%	95%	108%	108%	29%	101%	58%	63%	76%	82%								
A.19	119%	123%	114%	114%	74%	98%	67%	64%	41%	72%								
A.20	0%	1%	0%	0%	9%	1%	0%	0%	0%	0%								
A.21	124%	120%	137%	137%	43%	120%	74%	73%	64%	85%								
A.22	50%	52%	60%	60%	21%	82%	35%	57%	42%	75%								
A.23	77%	62%	49%	49%	3%	42%	0%	0%	0%	0%								
A.24	82%	57%	99%	99%	25%	121%	76%	55%	17%	66%								

Tabelle B.36: Wahrscheinlichkeiten der Keypointerkennung für Netz 37732415

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	62%	79%	46%	33%	26%	13%	22%	4%
A.2	34%	71%	0%	2%	0%	16%	0%	0%
A.3	50%	15%	6%	20%	16%	7%	33%	25%
A.4	38%	31%	41%	26%	3%	0%	0%	0%
A.5	7%	40%	18%	15%	0%	0%	0%	0%
A.6	23%	0%	6%	0%	16%	10%	16%	16%
A.7	47%	0%	9%	7%	0%	0%	0%	0%
A.8	0%	0%	45%	56%	23%	51%	0%	32%
A.9	11%	17%	0%	0%	0%	0%	0%	0%
A.10	7%	67%	10%	0%	0%	0%	0%	0%
A.11	0%	77%	0%	0%	0%	0%	35%	0%
A.12	0%	0%	0%	0%	0%	0%	0%	0%
A.13	35%	73%	1%	14%	21%	36%	19%	17%
A.14	50%	84%	46%	51%	0%	9%	0%	0%
A.15	56%	103%	90%	97%	81%	106%	55%	52%
A.16	0%	91%	3%	12%	0%	3%	0%	16%
A.17	4%	67%	23%	15%	26%	29%	1%	0%
A.18	72%	75%	57%	43%	37%	26%	6%	3%
A.19	55%	83%	79%	67%	63%	59%	29%	13%
A.20	5%	0%	2%	0%	33%	32%	8%	21%
A.21	8%	115%	71%	69%	94%	66%	42%	17%
A.22	55%	59%	47%	55%	36%	33%	30%	23%
A.23	0%	0%	0%	0%	0%	0%	0%	0%
A.24	0%	111%	59%	59%	36%	33%	0%	0%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37732415 (Fortsetzung)

Abbildung	Nase		Auge		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts								
A.1	113%	109%	115%	115%	33%	106%	67%	61%	77%	106%	61%	77%	77%	106%				
A.2	55%	15%	51%	51%	20%	38%	37%	48%	19%	53%	48%	19%	19%	53%				
A.3	83%	100%	81%	81%	92%	15%	25%	50%	15%	18%	50%	15%	15%	18%				
A.4	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	44%	0%	0%				
A.5	106%	97%	91%	91%	88%	20%	35%	33%	26%	27%	33%	26%	26%	27%				
A.6	82%	74%	86%	86%	55%	73%	8%	8%	7%	0%	8%	7%	7%	0%				
A.7	103%	66%	103%	103%	8%	94%	51%	45%	52%	13%	45%	52%	52%	13%				
A.8	90%	73%	109%	109%	33%	100%	36%	42%	31%	26%	42%	31%	31%	26%				
A.9	44%	43%	46%	46%	34%	60%	2%	30%	0%	29%	30%	0%	0%	29%				
A.10	77%	91%	68%	68%	72%	39%	9%	41%	1%	47%	41%	1%	1%	47%				
A.11	42%	37%	61%	61%	8%	50%	20%	50%	9%	61%	50%	9%	9%	61%				
A.12	20%	18%	35%	35%	21%	13%	0%	0%	0%	0%	0%	0%	0%	0%				
A.13	57%	62%	70%	70%	52%	29%	0%	37%	0%	62%	37%	0%	0%	62%				
A.14	61%	52%	54%	54%	20%	48%	8%	15%	10%	40%	15%	10%	10%	40%				
A.15	113%	120%	124%	124%	60%	105%	69%	65%	39%	97%	65%	39%	39%	97%				
A.16	101%	99%	128%	128%	8%	85%	30%	53%	3%	67%	53%	3%	3%	67%				
A.17	79%	80%	79%	79%	42%	81%	38%	34%	12%	55%	34%	12%	12%	55%				
A.18	96%	92%	103%	103%	34%	85%	52%	41%	68%	54%	41%	68%	68%	54%				
A.19	93%	103%	106%	106%	48%	91%	51%	57%	32%	63%	57%	32%	32%	63%				
A.20	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%	1%				
A.21	131%	110%	127%	127%	31%	100%	71%	65%	71%	78%	65%	71%	71%	78%				
A.22	108%	73%	92%	92%	17%	83%	40%	58%	49%	86%	58%	49%	49%	86%				
A.23	72%	57%	56%	56%	0%	16%	0%	7%	0%	0%	7%	0%	0%	0%				
A.24	75%	37%	96%	96%	13%	116%	66%	60%	17%	78%	60%	17%	17%	78%				

Tabelle B.37: Wahrscheinlichkeiten der Keypointerkennung für Netz 37792158

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	74%	84%	43%	21%	24%	0%	0%	0%
A.2	17%	69%	0%	4%	0%	33%	0%	0%
A.3	28%	1%	9%	6%	4%	0%	0%	0%
A.4	52%	32%	29%	17%	0%	0%	0%	1%
A.5	16%	36%	0%	0%	1%	0%	0%	0%
A.6	24%	5%	9%	12%	28%	30%	18%	25%
A.7	42%	0%	30%	33%	31%	12%	9%	2%
A.8	10%	16%	52%	52%	5%	47%	0%	16%
A.9	10%	12%	0%	3%	0%	0%	0%	0%
A.10	0%	59%	0%	14%	0%	0%	0%	0%
A.11	0%	83%	30%	27%	23%	25%	23%	25%
A.12	0%	0%	0%	0%	0%	0%	0%	0%
A.13	0%	70%	8%	29%	33%	43%	17%	17%
A.14	29%	36%	7%	0%	0%	0%	0%	0%
A.15	53%	96%	85%	97%	80%	94%	32%	53%
A.16	9%	84%	35%	40%	0%	17%	0%	0%
A.17	8%	58%	34%	17%	19%	12%	0%	0%
A.18	57%	63%	51%	40%	36%	17%	34%	8%
A.19	44%	61%	55%	57%	43%	48%	18%	0%
A.20	0%	0%	0%	0%	2%	17%	0%	25%
A.21	35%	104%	73%	69%	73%	49%	31%	17%
A.22	51%	66%	40%	37%	28%	31%	27%	8%
A.23	0%	38%	0%	0%	0%	0%	0%	0%
A.24	0%	99%	62%	52%	46%	29%	24%	7%

Wahrscheinlichkeiten der Keypointerkennung für Netz 37792158 (Fortsetzung)

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	2	2	2	1	2	1	2	1	2	1	2	1	2	1	2	1
A.3	1	1	1	2	1	2	1	1	1	1	2	1	2	1	2	2
A.4	1	1	1	1	1	1	1	1	2	1	2	1	2	1	2	2
A.5	1	1	1	2	1	2	1	2	1	2	1	2	1	2	1	2
A.6	2	2	2	2	3	2	2	3	2	2	2	2	3	0	2	0
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2
A.8	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2
A.9	2	2	2	2	3	2	2	3	1	1	1	0	0	1	1	1
A.10	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1
A.11	1	1	2	1	1	1	1	1	1	1	0	0	0	1	1	1
A.12	3	3	3	0	3	0	3	0	3	3	0	0	0	0	0	0
A.13	2	2	1	2	3	2	1	2	3	1	1	3	1	1	1	1
A.14	1	1	2	1	1	1	2	1	1	1	2	1	2	1	1	1
A.15	1	1	2	1	1	1	2	1	1	1	2	1	2	1	1	1
A.16	1	1	2	1	2	1	2	1	3	1	1	0	0	1	1	1
A.17	1	1	3	1	2	1	3	1	2	2	1	0	0	1	1	1
A.18	1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1
A.19	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	1
A.20	3	3	3	0	3	0	3	0	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	2	1	1	1	1	1	1	1	1	1
A.22	1	2	2	1	1	1	2	1	1	1	1	1	1	1	1	1
A.23	1	2	0	1	3	1	0	1	3	3	3	0	0	3	3	3
A.24	1	2	0	1	1	1	0	1	1	1	1	2	2	1	1	1

Tabelle B.38: Ergebnisse der Keypointerkennung für Netz 37697547^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	2	2	3
A.2	2	1	3	2	3	2	0	0
A.3	3	2	3	3	1	2	1	2
A.4	1	2	1	1	2	2	2	2
A.5	1	2	1	3	1	0	1	0
A.6	1	0	2	2	1	1	1	2
A.7	1	2	3	3	3	0	3	2
A.8	2	2	2	2	0	2	2	2
A.9	0	3	0	3	0	3	2	0
A.10	3	1	0	3	0	0	0	0
A.11	2	1	1	1	3	1	3	2
A.12	0	0	0	0	0	2	2	2
A.13	2	1	1	1	1	1	1	1
A.14	2	1	1	1	3	1	0	3
A.15	1	1	1	1	1	1	1	1
A.16	0	1	3	3	0	2	0	2
A.17	0	1	2	2	0	1	0	0
A.18	1	1	1	1	1	1	1	1
A.19	2	1	1	1	1	1	3	3
A.20	3	3	3	3	2	3	2	1
A.21	1	1	1	1	1	1	1	1
A.22	1	1	1	1	1	2	1	2
A.23	0	3	0	0	2	2	0	0
A.24	0	1	1	1	1	2	3	2

Ergebnisse der Keypointerkennung für Netz 37697547 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.3	1	1	1	1	1	2	1	1	1	1	1	2	0	1	0	0
A.4	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2	2
A.5	1	1	1	1	1	0	1	1	1	1	1	2	1	1	1	1
A.6	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2
A.7	1	1	2	1	1	1	1	1	1	1	1	1	2	1	2	2
A.8	1	1	2	1	1	1	1	1	1	1	1	1	2	1	2	2
A.9	2	2	2	1	1	1	1	1	1	1	1	0	1	1	1	1
A.10	1	1	1	1	1	2	1	1	1	1	1	0	1	1	1	1
A.11	1	1	2	1	1	1	1	1	1	1	1	0	1	1	1	1
A.12	3	1	1	2	1	2	1	2	3	3	3	0	0	0	0	0
A.13	3	1	3	3	3	2	2	2	3	1	1	3	1	1	1	1
A.14	1	2	2	1	2	1	1	1	1	1	2	2	1	1	1	1
A.15	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1
A.17	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1	1
A.18	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	2	1	1	1	1	1	2	2	1	1	1	1
A.20	3	2	3	2	3	0	0	3	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.22	2	2	2	1	2	1	1	1	1	1	2	2	1	1	1	1
A.23	1	2	2	1	2	1	1	3	3	3	3	0	3	3	3	3
A.24	1	2	0	1	1	1	1	1	1	1	2	2	1	1	1	1

Tabelle B.39: Ergebnisse der Keypointerkennung für Netz 37697714^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	3	3	2	3	3
A.2	1	1	3	3	3	2	0	0
A.3	2	2	3	1	1	2	1	2
A.4	1	2	1	2	2	2	0	0
A.5	2	0	2	2	2	2	3	0
A.6	1	0	2	2	2	1	2	0
A.7	1	2	3	3	3	0	3	0
A.8	2	2	1	1	0	2	0	2
A.9	0	3	2	3	0	3	2	0
A.10	3	1	1	3	0	0	0	0
A.11	3	1	1	1	1	1	2	1
A.12	0	0	0	0	0	0	0	0
A.13	2	1	3	3	1	1	1	1
A.14	2	1	3	3	3	3	2	3
A.15	2	1	1	1	1	1	1	1
A.16	0	1	3	2	0	3	0	0
A.17	2	1	1	2	2	1	2	2
A.18	1	1	1	1	1	1	1	3
A.19	2	1	1	1	1	1	2	2
A.20	3	3	2	2	2	2	2	1
A.21	1	1	1	1	1	1	3	2
A.22	2	1	1	1	1	1	1	2
A.23	0	3	0	0	2	0	0	0
A.24	0	1	1	1	1	1	3	0

Ergebnisse der Keypointerkennung für Netz 37697714 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	1	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1
A.3	1	1	1	1	1	2	1	1	1	1	1	1	1	2	1	2
A.4	1	1	1	1	1	1	1	1	1	1	1	1	1	3	1	3
A.5	1	1	1	1	1	2	1	2	1	2	1	2	1	0	1	0
A.6	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	3
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1
A.8	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1
A.9	2	2	2	2	2	2	2	2	2	2	2	2	2	0	1	1
A.10	1	1	1	1	1	2	1	2	3	1	1	1	1	0	1	1
A.11	2	1	0	1	3	1	1	3	3	1	1	1	0	1	1	1
A.12	2	2	1	2	3	2	2	3	3	3	3	3	0	0	0	0
A.13	2	1	1	2	3	2	2	3	3	1	1	3	1	1	1	1
A.14	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1
A.15	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1
A.17	1	1	2	1	2	1	1	2	2	1	1	2	1	1	1	1
A.18	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1
A.20	3	3	3	0	3	0	3	0	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.22	1	1	2	2	1	1	1	1	1	1	1	2	1	1	1	1
A.23	1	1	2	1	2	1	1	1	3	3	3	0	3	3	3	3
A.24	1	1	2	1	2	1	1	1	1	1	1	2	1	2	1	1

Tabelle B.40: Ergebnisse der Keypointerkennung für Netz 37697946^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	2	2	2
A.2	1	1	3	2	3	2	2	0
A.3	1	0	3	1	2	0	2	0
A.4	1	2	1	1	2	0	0	0
A.5	2	2	2	2	3	2	3	0
A.6	1	2	1	1	1	1	2	2
A.7	1	2	1	1	1	2	3	0
A.8	2	2	1	1	0	2	2	1
A.9	0	3	0	1	0	2	0	0
A.10	1	1	2	1	0	0	0	0
A.11	3	1	3	1	3	1	2	1
A.12	0	0	0	0	0	0	0	0
A.13	2	1	3	1	1	1	1	1
A.14	2	1	1	1	2	1	0	1
A.15	1	1	1	1	1	1	1	1
A.16	2	1	1	1	0	3	2	0
A.17	2	1	1	1	2	1	2	2
A.18	1	1	1	1	1	2	1	2
A.19	2	1	1	1	1	1	1	3
A.20	3	3	3	2	3	1	2	1
A.21	1	1	1	1	1	1	1	1
A.22	2	1	1	1	1	2	2	2
A.23	0	1	0	0	0	0	0	0
A.24	0	1	1	1	1	1	1	2

Ergebnisse der Keypointerkennung für Netz 37697946 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	3	3	2	3	1	3	1	3	1	1	2	1	2	1	2	1
A.3	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	0
A.4	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2
A.5	1	1	1	2	1	2	1	2	1	2	1	2	0	1	1	1
A.6	2	2	2	2	2	2	2	2	2	2	2	2	3	0	2	0
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	2	1	2
A.8	1	1	2	1	1	1	1	1	1	1	1	1	1	2	1	2
A.9	2	2	0	3	3	3	3	3	3	3	3	3	2	1	2	1
A.10	1	1	1	2	1	2	1	2	1	1	1	1	1	1	1	1
A.11	2	1	2	1	3	1	2	1	3	1	0	1	0	1	1	1
A.12	1	1	1	2	3	2	1	2	3	3	3	0	0	0	0	0
A.13	1	1	1	2	3	2	2	2	3	1	3	3	3	3	3	3
A.14	1	1	2	1	1	1	1	1	1	1	2	2	2	1	1	1
A.15	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1
A.17	1	1	2	1	2	1	1	1	2	1	0	0	0	1	1	1
A.18	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	1	1	1	1	1	1	1	2	2	1	1	1
A.20	3	0	3	0	3	0	3	0	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.22	1	2	2	1	2	1	2	1	2	1	2	2	2	1	1	1
A.23	1	1	0	3	3	3	3	3	3	3	3	0	0	3	3	3
A.24	2	1	2	1	1	1	1	1	1	1	1	2	2	1	1	1

Tabelle B.41: Ergebnisse der Keypointerkennung für Netz 37698009^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	3	2	2
A.2	1	1	3	2	3	2	2	2
A.3	2	2	2	3	2	0	2	0
A.4	1	2	2	1	3	2	2	2
A.5	2	2	2	2	3	0	3	0
A.6	1	0	2	2	2	2	2	2
A.7	1	2	1	3	3	0	3	0
A.8	0	2	1	1	0	2	2	1
A.9	2	2	0	3	2	1	2	2
A.10	1	1	2	1	0	2	0	0
A.11	3	1	3	1	3	1	2	2
A.12	0	0	0	0	0	0	0	0
A.13	3	2	3	3	3	3	2	2
A.14	2	1	1	1	2	1	0	3
A.15	1	1	1	1	1	1	1	1
A.16	2	1	1	1	0	3	0	0
A.17	0	1	1	2	2	1	2	2
A.18	1	1	1	1	1	1	1	1
A.19	2	1	1	1	1	1	1	2
A.20	3	3	3	3	3	3	2	1
A.21	1	1	1	1	1	1	1	1
A.22	2	2	1	1	1	2	1	2
A.23	0	3	0	0	0	0	2	0
A.24	0	1	1	1	1	1	1	2

Ergebnisse der Keypointerkennung für Netz 37698009 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	1	1	2	1	1	1	1	1	1	1	1	1	2	1	1	1
A.3	1	1	1	1	1	2	1	1	1	1	1	1	2	2	2	2
A.4	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2
A.5	1	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1
A.6	2	2	2	2	2	2	2	2	2	2	2	2	3	0	0	0
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	2	2	2
A.8	1	1	2	1	1	1	1	1	1	1	1	1	1	2	2	2
A.9	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1
A.10	1	1	1	1	1	2	1	1	1	1	1	1	0	1	1	1
A.11	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1
A.12	3	2	3	0	3	0	3	0	3	3	3	0	0	2	2	2
A.13	1	1	1	1	1	2	1	2	3	1	1	3	3	3	3	3
A.14	2	1	2	1	1	1	1	1	1	3	3	2	2	1	1	1
A.15	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	2	1	1	1	1	1	1	1	1	0	0	1	1	1
A.17	1	1	1	1	1	1	1	1	2	1	1	2	2	1	1	1
A.18	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	1	1	1	1	1	1	1	2	2	1	1	1
A.20	2	3	2	0	2	0	0	0	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.22	1	1	2	2	1	2	2	2	1	1	1	2	2	1	1	1
A.23	1	1	2	1	1	1	1	1	3	3	3	0	0	3	3	3
A.24	1	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1

Tabelle B.42: Ergebnisse der Keypointerkennung für Netz 37732318^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	3	2	3
A.2	1	1	3	2	3	2	0	0
A.3	2	0	2	2	3	2	1	2
A.4	1	2	2	2	2	0	0	0
A.5	2	2	2	2	3	0	3	0
A.6	1	0	2	2	1	2	3	0
A.7	1	2	1	1	1	2	1	2
A.8	2	2	1	1	0	2	0	1
A.9	2	2	0	3	0	3	0	0
A.10	1	1	2	1	0	0	0	0
A.11	3	1	1	1	1	1	2	1
A.12	0	0	0	0	0	0	0	0
A.13	2	2	3	3	3	3	2	3
A.14	2	1	1	1	3	3	2	3
A.15	2	1	1	1	1	1	1	1
A.16	0	1	1	1	0	3	0	0
A.17	2	1	1	2	2	1	2	2
A.18	1	1	1	1	1	1	1	2
A.19	2	1	1	1	1	1	1	3
A.20	3	3	3	3	3	2	3	1
A.21	1	1	1	1	1	1	1	1
A.22	2	2	1	1	1	1	1	2
A.23	0	1	0	0	0	0	0	0
A.24	0	1	1	1	1	1	1	2

Ergebnisse der Keypointerkennung für Netz 37732318 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
A.2	1	1	2	1	1	2	1	1	1	1	2	1	1	1
A.3	1	2	1	2	1	2	1	1	1	1	2	0	2	0
A.4	1	1	1	1	1	1	1	1	1	1	1	3	1	3
A.5	1	1	1	2	1	2	1	1	1	1	0	2	1	2
A.6	2	2	2	2	3	2	3	2	2	2	1	2	2	2
A.7	1	1	2	1	1	1	1	1	1	1	1	0	1	0
A.8	1	1	2	1	1	1	1	1	1	1	2	2	2	2
A.9	2	2	2	2	2	2	2	1	1	1	2	1	1	1
A.10	1	1	1	1	1	2	1	1	1	1	0	1	1	1
A.11	1	1	2	1	1	1	1	1	1	1	0	1	1	1
A.12	3	2	3	0	2	0	2	3	3	3	0	2	2	2
A.13	2	2	1	2	3	2	2	3	1	1	3	1	1	1
A.14	1	1	2	1	2	1	1	2	1	1	2	1	1	1
A.15	1	1	2	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	2	1	1	1	1	1	1	1	2	1	1	1
A.17	1	1	1	1	1	1	1	1	1	1	2	1	1	1
A.18	1	1	2	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	1	1	1	1	1	1	2	1	1	1
A.20	2	2	3	0	3	0	3	3	3	3	3	3	3	3
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1
A.22	1	2	2	1	1	1	1	1	1	1	2	1	1	1
A.23	1	1	2	1	3	1	3	3	3	3	0	3	3	3
A.24	1	1	2	1	1	1	1	1	1	1	1	1	1	1

Tabelle B.43: Ergebnisse der Keypointerkennung für Netz 37732355^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	2	1	2
A.2	2	1	3	2	3	2	0	0
A.3	2	0	2	1	1	2	1	2
A.4	1	2	2	2	2	2	0	0
A.5	2	1	1	2	3	0	3	0
A.6	1	2	1	1	1	2	2	2
A.7	1	2	1	1	1	2	1	0
A.8	2	2	1	1	0	2	2	2
A.9	2	2	0	3	0	3	0	0
A.10	3	1	2	1	0	0	0	0
A.11	3	1	1	1	3	1	2	1
A.12	0	0	0	2	0	2	2	0
A.13	2	1	1	1	1	1	1	1
A.14	2	1	1	1	3	1	0	3
A.15	1	1	1	1	1	1	1	1
A.16	0	1	1	1	0	1	0	0
A.17	0	1	2	2	2	1	2	2
A.18	1	1	1	1	1	1	1	1
A.19	2	1	1	1	1	1	1	3
A.20	3	2	1	1	3	1	3	1
A.21	1	1	1	1	1	1	1	2
A.22	2	2	1	1	1	2	2	2
A.23	0	1	0	0	0	2	0	0
A.24	0	1	1	1	1	1	3	2

Ergebnisse der Keypointerkennung für Netz 37732355 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	1	0	2	1	1	3	1	1	1	1	2	1	2	1	1	1
A.3	1	1	1	1	1	2	1	1	1	1	2	1	2	0	1	0
A.4	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	2
A.5	1	1	1	1	1	2	1	1	1	1	2	1	2	1	1	1
A.6	2	2	2	2	2	2	3	2	2	2	3	2	3	0	2	0
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	2	1	2
A.8	1	1	2	1	1	1	1	1	1	1	2	1	2	2	1	2
A.9	2	2	2	2	2	2	3	1	1	1	2	1	2	1	1	1
A.10	1	1	1	1	1	2	1	1	1	1	1	1	0	1	1	1
A.11	1	1	2	1	1	1	1	1	1	1	1	1	0	1	1	1
A.12	1	1	1	1	1	2	1	1	1	3	1	0	0	0	0	0
A.13	2	2	2	2	2	2	3	2	2	1	3	1	3	1	1	1
A.14	1	1	2	1	2	1	1	1	1	1	2	1	2	1	1	1
A.15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	2	1	2	1	1	1	1	1	1	1	0	1	1	1
A.17	1	1	1	1	1	1	2	1	1	1	2	1	2	1	1	1
A.18	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1
A.20	3	2	2	0	2	2	3	2	2	3	3	3	3	3	3	3
A.21	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
A.22	1	1	2	1	2	1	1	1	1	1	2	1	2	1	1	1
A.23	1	1	2	1	2	1	3	1	1	3	3	3	0	3	3	3
A.24	1	1	2	1	2	1	1	1	1	1	1	2	2	1	1	1

Tabelle B.44: Ergebnisse der Keypointerkennung für Netz 37732415^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	2	1	2	2
A.2	1	1	3	2	3	2	0	0
A.3	2	2	2	1	1	2	1	2
A.4	1	2	1	1	2	0	0	0
A.5	2	1	2	2	3	0	3	0
A.6	2	0	2	3	2	2	2	2
A.7	1	0	1	1	3	0	3	0
A.8	0	0	1	1	2	2	0	2
A.9	2	2	0	3	0	3	0	0
A.10	1	1	2	3	0	0	0	0
A.11	3	1	3	3	3	3	2	3
A.12	0	0	0	0	0	0	0	0
A.13	2	1	1	1	1	1	1	1
A.14	2	1	1	1	3	1	0	3
A.15	1	1	1	1	1	1	1	1
A.16	0	1	2	2	0	2	0	2
A.17	2	1	2	2	2	1	2	0
A.18	1	1	1	1	1	1	2	2
A.19	2	1	1	1	1	1	1	2
A.20	2	3	2	3	2	2	2	1
A.21	1	1	1	1	1	1	1	1
A.22	2	1	1	1	1	2	2	2
A.23	0	3	0	0	0	0	0	0
A.24	0	1	1	1	1	1	3	0

Ergebnisse der Keypointerkennung für Netz 37732415 (Fortsetzung)^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Nase		Auge		Ohr		Ohr		Schulter		Schulter		Ellenbogen		Ellenbogen	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.2	1	1	2	2	1	2	1	1	1	1	2	1	1	1	2	1
A.3	1	1	1	2	1	2	1	1	1	1	2	1	2	2	1	2
A.4	3	3	2	3	3	3	3	3	3	3	1	3	1	3	3	3
A.5	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	1
A.6	2	2	2	2	2	2	2	2	2	2	1	2	1	0	1	0
A.7	1	1	2	1	1	1	1	1	1	1	1	1	1	2	1	2
A.8	1	1	2	1	1	1	1	1	1	1	2	1	2	2	2	2
A.9	1	2	2	1	1	1	1	1	1	1	1	1	0	1	1	1
A.10	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	1
A.11	1	1	2	1	1	1	1	1	1	1	2	1	2	1	1	1
A.12	2	2	1	2	3	2	2	2	3	3	3	0	0	0	0	0
A.13	2	2	1	2	3	2	2	2	3	1	1	3	1	1	1	1
A.14	1	2	2	1	1	2	1	1	1	1	2	2	1	1	1	1
A.15	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.16	1	1	2	1	1	1	1	1	1	1	2	2	1	1	1	1
A.17	1	1	1	1	2	1	1	1	1	1	2	2	1	1	1	1
A.18	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.19	1	1	2	1	1	1	1	1	1	1	2	2	1	1	1	1
A.20	3	0	3	0	3	0	0	0	3	3	3	3	2	2	2	2
A.21	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
A.22	1	1	2	1	1	1	1	1	1	1	2	2	1	1	1	1
A.23	1	1	0	1	3	1	1	1	3	1	0	0	3	3	3	3
A.24	1	2	2	1	1	1	1	1	1	1	2	2	1	1	1	1

Tabelle B.45: Ergebnisse der Keypointerkennung für Netz 37792158^b

^b Legende:

- 1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Abbildung	Hand		Hüfte		Knie		Fuß	
	Links	Rechts	Links	Rechts	Links	Rechts	Links	Rechts
A.1	1	1	1	1	1	3	3	3
A.2	1	1	3	2	3	2	0	0
A.3	2	2	2	2	1	0	3	0
A.4	1	2	1	1	3	0	0	2
A.5	2	2	3	3	1	0	3	0
A.6	1	2	2	2	2	2	2	2
A.7	1	0	1	1	1	2	1	2
A.8	2	2	1	1	2	2	0	2
A.9	2	2	0	2	0	3	0	0
A.10	3	1	0	1	0	0	0	0
A.11	3	1	1	1	1	1	2	1
A.12	0	0	0	0	0	0	0	0
A.13	3	1	1	1	1	1	1	1
A.14	2	1	1	3	3	3	0	3
A.15	1	1	1	1	1	1	1	1
A.16	2	1	1	1	0	1	0	0
A.17	2	1	1	1	2	1	0	0
A.18	1	1	1	2	1	2	1	1
A.19	2	1	1	1	1	1	1	3
A.20	3	3	3	3	2	1	3	1
A.21	1	1	1	1	1	1	1	1
A.22	2	1	1	1	1	2	1	2
A.23	0	1	0	0	0	0	0	0
A.24	0	1	1	1	1	2	1	2

Ergebnisse der Keypointerkennung für Netz 37792158 (Fortsetzung)^b

^b Legende:

1) Punkt korrekt erkannt 2) Punkt inkorrekt erkannt 3) Punkt inkorrekt nicht erkannt 4) Punkt korrekt nicht erkannt

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	80%	73%	61%	51%	40%	32%	24%	16%	12%	6%
37697714	80%	72%	64%	57%	51%	41%	32%	23%	17%	11%
37697946	86%	77%	69%	58%	48%	36%	26%	17%	11%	7%
37698009	79%	73%	66%	55%	48%	39%	26%	20%	13%	8%
37732318	82%	76%	70%	62%	52%	43%	35%	27%	20%	14%
37732355	88%	81%	70%	62%	52%	43%	36%	28%	22%	14%
37732415	85%	76%	69%	63%	55%	48%	39%	30%	22%	17%
37792158	84%	77%	70%	62%	52%	44%	33%	25%	19%	15%

Tabelle B.46: Recall der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	68%	75%	78%	87%	89%	78%	92%	94%	100%	100%
37697714	68%	74%	78%	84%	88%	78%	90%	92%	93%	97%
37697946	67%	73%	79%	81%	87%	79%	89%	89%	92%	95%
37698009	63%	69%	77%	79%	85%	77%	90%	90%	87%	89%
37732318	69%	74%	79%	83%	85%	79%	91%	97%	95%	100%
37732355	67%	72%	79%	83%	88%	79%	91%	96%	97%	100%
37732415	68%	72%	77%	83%	89%	77%	91%	95%	94%	96%
37792158	70%	75%	81%	84%	90%	81%	92%	94%	97%	100%

Tabelle B.47: Precision der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.

Netz-ID	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
37697547	73%	74%	68%	64%	55%	46%	38%	28%	21%	11%
37697714	74%	73%	70%	68%	64%	54%	48%	37%	28%	20%
37697946	75%	75%	74%	68%	62%	49%	40%	28%	19%	12%
37698009	70%	71%	71%	65%	61%	52%	41%	33%	23%	15%
37732318	75%	75%	74%	71%	64%	56%	51%	43%	34%	25%
37732355	76%	76%	74%	71%	66%	55%	51%	43%	35%	25%
37732415	76%	74%	73%	72%	68%	59%	54%	46%	36%	29%
37792158	76%	76%	75%	72%	66%	57%	48%	40%	32%	26%

Tabelle B.48: F1-Score der Keypointerkennung nach Netz und Schwellenwert als berechnete Ergebnisse der Tabellen B.30 bis B.45.