

Gestaltungsrichtlinien für montagegerechte Konstruktionen mit Hilfe von Deep Learning

Dissertation

zur Erlangung des akademischen Grades

**Doktoringenieurin
(Dr.-Ing.)**

von M. Sc. Johanna Gerlach

geb. am 28.07.1995 in Celle

genehmigt durch die Fakultät für Maschinenbau
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr.-Ing. Karl-Heinrich Grote
 Prof. Dr.-Ing. Manuel Löwer
 Prof. Dr.-Ing. habil. Frank Engelmann

Promotionskolloquium am 12.09.2023

Kurzreferat

Da nicht alle Montageprozesse wirtschaftlich und/ oder technisch automatisiert werden können, stellt die manuelle Montage in vielen Unternehmen noch immer einen erforderlichen Herstellungsschritt dar, welcher jedoch mit hohen Zeit- und Kostenfaktoren sowie weiteren Herausforderungen, wie einer wachsenden Variantenvielfalt verbunden ist. Um im internationalen Wettbewerb dennoch zu bestehen, sollte die manuelle Montage fortlaufend verbessert werden. Ein potenzieller Optimierungsansatz besteht durch die Integration von Künstlicher Intelligenz. Durch das Erfassen des Montageprozesses mittels Kamera und anschließender, auf Deep Learning basierender Objektdetektion ist es möglich, die Position, Lage und Art von Bauteilen zu bestimmen, um Informationen über Arbeitsschritte, Montagefehler oder den aktuellen Produktzustand abzuleiten.

So birgt der Einsatz von Deep Learning in der manuellen Montage große Potenziale, Montagezeiten zu reduzieren und Montagefehler zu vermeiden. Es stellt sich jedoch die Frage, inwieweit Produkte, welche nach dem aktuellen Stand der Technik konzipiert wurden, für den Einsatz von Deep Learning geeignet sind. Nur wenn sie durch bestehende Deep Learning Modelle eindeutig erkannt werden, ist ihre Anwendung in der manuellen Montage sinnvoll. Aktuelle Gestaltungsrichtlinien für montagegerechte Konstruktionen greifen diesen Aspekt bislang nicht auf.

Im Rahmen der vorliegenden Arbeit wird daher untersucht, welche konstruktiven Eigenschaften Baugruppen und Bauteile aufweisen sollten, um eine optimale Objekterkennung durch Deep Learning zu erzielen. Basierend auf der Funktionsweise der Deep Learning Technologie werden Hypothesen aufgestellt und mittels empirischer Untersuchungen verifiziert. Dabei werden Aspekte in Bezug auf die Gestaltung von Bauteilen, Baugruppen und Formelementen, sowie die Auswahl von Normteilen berücksichtigt. Durch die Anwendung auf reale Produktbeispiele wird darüber hinaus überprüft, ob die Ergebnisse der empirischen Untersuchung auf reale Montageumgebungen übertragbar sind. Schlussendlich werden Empfehlungen für eine montagegerechte Produktgestaltung auf Basis des Einsatzes von Deep Learning in Form von Gestaltungsrichtlinien abgeleitet. Ziel ist es, dem Konstrukteur der Zukunft ein Werkzeug an die Hand zu geben, mit welchem er den sich ändernden Rahmenbedingungen in der manuellen Montage begegnen kann.

Abstract

Since it is not possible to automate all assembly processes economically and/or technically, manual assembly is still a necessary manufacturing step in many companies. However, this is associated with high time and cost factors as well as further challenges such as a growing diversity of variants. In order to compete in the international market, manual assembly should be improved on a continuous basis. One potential optimization approach is the integration of artificial intelligence. By recording the assembly process using a camera and subsequent object detection based on Deep Learning, it is possible to determine the position, location and type of components in order to derive information about work steps, assembly errors or the current product condition.

Thus, the use of Deep Learning in manual assembly holds great potential for reducing assembly times and avoiding assembly errors. However, the question arises to what extent products that have been designed according to the current state of the art are suitable for the use of Deep Learning models. Only if they are clearly recognized by existing Deep Learning models, their application in manual assembly is reasonable. Current design guidelines for assembly-compatible designs do not address this aspect yet.

Therefore, the present work investigates the design properties that assemblies and components should have in order to achieve optimal object recognition through Deep Learning. Based on the principles of Deep Learning technology, hypotheses are formulated and verified by means of empirical investigations. Aspects related to the design of components, assemblies and form elements, as well as the selection of standard parts are considered. Furthermore, by using real product examples, it is verified whether the results of the empirical investigation can be transferred to real assembly environments. Finally, recommendations for an assembly-compatible product design based on the use of Deep Learning are derived in the form of design guidelines. The goal is to provide the future designer with a tool that enables him to cope with the changing conditions in manual assembly.

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Motivation zur Arbeit.....	1
1.2	Zielsetzung.....	3
1.3	Gliederung der Arbeit.....	4
2	Grundlagen.....	5
2.1	Montagegerechte Produktgestaltung.....	5
2.1.1	Der Produktentstehungsprozess.....	5
2.1.2	Montage in der industriellen Produktion.....	8
2.1.2.1	Montageoperationen.....	8
2.1.2.2	Automatisierungsgrad.....	10
2.1.3	Methoden zur montagegerechten Produktgestaltung.....	11
2.2	Deep Learning.....	13
2.2.1	Begriffliche Einordnung.....	13
2.2.2	Historische Einordnung.....	15
2.2.3	Grundlegende Funktionsweise.....	16
2.2.4	Mean Average Precision.....	20
3	Stand der Technik.....	24
3.1	Aktuelle Entwicklungen in der montagegerechten Produktgestaltung.....	24
3.2	Deep Learning in der Montage.....	24
4	Anforderungen an die Produktgestaltung.....	27
4.1	Einflussfaktoren.....	27
4.1.1	Bilder als Abfolge von Pixel-Werten.....	27
4.1.2	Interklassenvarianz.....	29
4.2	Hypothesen.....	30
4.2.1	Hypothesen zur Bauteilgestaltung.....	31
4.2.2	Hypothesen zur Baugruppengestaltung.....	34
4.2.3	Hypothesen zur Auswahl von Normteilen.....	35
4.2.4	Hypothesen zur Dimensionierung von Formelementen.....	37
5	Verifizierung der Hypothesen.....	40

5.1	Methodik.....	40
5.2	Identifikation der abhängigen und unabhängigen Variablen.....	40
5.3	Definition des Versuchsaufbaus.....	43
5.4	Auswahl des Objekterkennungsalgorithmus	45
5.5	Statistische Versuchsplanung.....	50
5.5.1	Festlegung der Faktorstufen und Faktorstufenkombinationen.....	50
5.5.2	Festlegung des Stichprobenumfangs.....	51
5.5.2.1	Umfang der Trainings- und Validierungsdatensätze	52
5.5.2.2	Anzahl der Iterationen	57
5.5.2.3	Stichprobenumfang – Bauteilgestaltung	61
5.5.2.4	Stichprobenumfang – Baugruppengestaltung	62
5.5.2.5	Stichprobenumfang – Auswahl von Normteilen	63
5.5.2.6	Stichprobenumfang – Dimensionierung von Formelementen.....	64
5.6	Versuchsdurchführung.....	65
5.6.1	Konzeption der Untersuchungsobjekte	65
5.6.1.1	Untersuchungsobjekte zur Bauteilgestaltung.....	65
5.6.1.2	Untersuchungsobjekte zur Baugruppengestaltung.....	68
5.6.1.3	Untersuchungsobjekte zur Auswahl von Normteilen	70
5.6.1.4	Untersuchungsobjekte zur Dimensionierung von Formelementen	72
5.6.2	Erstellung der Trainings- und Validierungsbilder	75
5.6.3	Annotation der Trainings- und Validierungsbilder.....	77
5.6.4	Implementierung des Objekterkennungsalgorithmus	78
5.6.5	Training- und Validierung der Modelle	80
5.7	Versuchsauswertung	83
5.7.1	Versuchsergebnisse zur Bauteilgestaltung	83
5.7.2	Versuchsergebnisse zur Baugruppengestaltung.....	86
5.7.3	Versuchsergebnisse zur Auswahl von Normteilen	90
5.7.4	Versuchsergebnisse zur Dimensionierung von Formelementen	101
5.7.5	Zusammenfassung der Versuchsergebnisse	103
6	Übertragbarkeit der Ergebnisse auf reale Produkte	104
6.1	Produktbeispiel I: Explosionsgeschützte Rohrleuchte	104

6.1.1	Produktbeschreibung.....	104
6.1.2	Montage mittels digitalem Assistenzsystem.....	106
6.1.3	Übertragbarkeit der Versuchsergebnisse.....	108
6.1.3.1	Bauteilgestaltung.....	109
6.1.3.2	Baugruppengestaltung	113
6.1.3.3	Auswahl von Normteilen.....	114
6.1.3.4	Dimensionierung von Formelementen	116
6.2	Produktbeispiel II: Würfel-Regal.....	119
6.2.1	Produktbeschreibung.....	119
6.2.2	Übertragbarkeit der Versuchsergebnisse.....	120
7	Ableitung konstruktiver Gestaltungsrichtlinien.....	124
8	Zusammenfassung	128
9	Literatur	131
A	Anhang	144
A.1	Erzeugnisgliederung der explosionsgeschützten Rohrleuchte	144
A.2	Programmcode – Video2image.....	145
A.3	Programmcode – Image2coco	146
A.4	Programmcode – Labelme2coco	149

Abbildungsverzeichnis

Abbildung 1-1: Die vier industriellen Revolutionen	2
Abbildung 1-2: Kostenfestlegung und Kostenverursachung im Unternehmen (nach [8]) ...	3
Abbildung 2-1: Technischer Produktlebenszyklus [9]	5
Abbildung 2-2: Hauptschritte des Entwicklungs- und Konstruktionsprozesses (nach [11])	6
Abbildung 2-3: Montageoperationen (in Anlehnung an [26]).....	9
Abbildung 2-4: Auswahlkriterien für Montagesysteme.....	11
Abbildung 2-5: Ansätze zur montagegerechten Produktgestaltung (nach [29])	12
Abbildung 2-6: Einordnung Deep Learning	14
Abbildung 2-7: Machine Learning versus Deep Learning	14
Abbildung 2-8: Beispiel für ein künstliches neuronales Netz	17
Abbildung 2-9: Feature-Hierarchie in CNNs	19
Abbildung 2-10: Beispiel einer Vorhersage durch einen Objekterkennungsalgorithmus ...	20
Abbildung 2-11: Intersection over Union	21
Abbildung 2-12: Precision-Recall-Kurve für Vorhersagen aus Abbildung 2-10	22
Abbildung 3-1: Anzahl der Suchergebnisse zum Suchbegriff "Deep Learning" in der Datenbank IEEE Xplore zum Stichtag 01.01.2023 im zeitlichen Verlauf	25
Abbildung 3-2: Anzahl der Suchergebnisse zu einzelnen Themen unter dem Überbegriff "Deep Learning" in der Datenbank IEEE Xplore zum Stichtag 01.01.2023.....	25
Abbildung 4-1: Darstellung eines SW-Bildes durch Pixelwerte [96].....	27
Abbildung 4-2: Gradientenbilder G_x und G_y nach Anwendung des Prewitt-Operators	28
Abbildung 4-3: SW-Bild eines Hauses a) Original b) Nach Prewitt-Operator (nach [97]) ..	29
Abbildung 4-4: Intra- und Interklassenvarianz (unter Verwendung von [98] - [101])	29
Abbildung 5-1: Über- und Unteranpassung	42
Abbildung 5-2: Versuchsaufbau bestehend aus Datengenerierung und -auswertung	44
Abbildung 5-3: Grundlegende Architektur von Two-Stage Detektoren (nach [123]).....	45
Abbildung 5-4: Grundlegende Architektur von One-Stage Detektoren (nach [123]).....	46
Abbildung 5-5: Vorhersagen eines Mask R-CNN (nach [121])	47
Abbildung 5-6: Architektur des R50-FPN-Netzwerks in Detectron2.....	49
Abbildung 5-7: Architektur eines Residual Networks (nach [133])	50

Abbildung 5-8: Erreichte mAP für das Modell „Form1-4“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder.....	54
Abbildung 5-9: Erreichte mAP für das Modell „Form1-4Weich“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder	54
Abbildung 5-10: Erreichte mAP für das Modell „Form3-3.3“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder.....	54
Abbildung 5-11: Erreichte mAP für das Modell „BG-Alu“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder.....	55
Abbildung 5-12: Erreichte mAP für das Modell „BG-Schwarz“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder	55
Abbildung 5-13: Erreichte mAP für das Modell „BG-Gemischt“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder	55
Abbildung 5-14: Vorversuchsergebnisse zur Iterationenanzahl – mAP	58
Abbildung 5-15: Vorversuchsergebnisse zur Iterationenanzahl – Trainingsfehler.....	59
Abbildung 5-16: Vorversuchsergebnisse zur Iterationenanzahl – Validierungsfehler.....	60
Abbildung 5-17: Aluminiumwürfel mit verschraubter Zylinderkopfschraube M6	72
Abbildung 5-18: Versuchsaufbau zur Aufnahme der Trainings- und Validierungsbilder....	76
Abbildung 5-19: Benutzeroberfläche – Labelme.....	78
Abbildung 5-20: Mittlere mAP für einfarbige und mehrfarbige Modelle – Rundumsicht	88
Abbildung 5-21: Mittlere mAP für einfarbige und mehrfarbige Modelle – Draufsicht	90
Abbildung 5-22: Mittlere mAP der Modelle SK+GS und ZK+LK in Abhängigkeit der Normteilgröße – Rundumsicht	93
Abbildung 5-23: Mittlere mAP der Modelle U, M, GS, SK, LK und ZK in Abhängigkeit der Normteilgröße – Rundumsicht	94
Abbildung 5-24: Mittlere mAP der Modelle SK+GS und ZK+LK in Abhängigkeit der Normteilgröße – Draufsicht	98
Abbildung 5-25: Mittlere mAP der Modelle U, M, GS, SK, LK und ZK in Abhängigkeit der Normteilgröße – Draufsicht	99
Abbildung 6-1: Explosionsgeschützte Rohrleuchte [155].....	104
Abbildung 6-2: Montagearbeitsplatz mit digitalem Assistenzsystem.....	107
Abbildung 6-3: Würfel-Regal in Weiß und Schwarz [160] [161]	119

Tabellenverzeichnis

Tabelle 2-1: Layer in CNNs.....	19
Tabelle 2-2: Definitionen: True Positive, False Positive und False Negative	21
Tabelle 2-3: Precision und Recall für Vorhersagen aus Abbildung 2-10	22
Tabelle 4-1: Null- und Alternativhypothesen zur Bauteilgestaltung.....	32
Tabelle 4-2: Null- und Alternativhypothese zur Baugruppengestaltung	34
Tabelle 4-3: Null- und Alternativhypothesen zur Auswahl von Normteilen.....	36
Tabelle 4-4: Null- und Alternativhypothesen zur Dimensionierung von Formelementen ...	38
Tabelle 5-1: Aufstellung der unabhängigen Variablen.....	43
Tabelle 5-2: Vortrainierte Faster R-CNN Modelle in Detectron2 (nach [131])	48
Tabelle 5-3: Versuchsplan a) Allgemein b) Bezüglich Hypothese BT1	51
Tabelle 5-4: Modelle zur Festlegung der Trainings- und Validierungsbilderanzahl	53
Tabelle 5-5: Anzahl der Trainings- und Validierungsbilder in der Hauptuntersuchung.....	57
Tabelle 5-6: Vorversuchsergebnisse – Bauteilgestaltung.....	61
Tabelle 5-7: Ableitung Stichprobenumfang – Bauteilgestaltung	62
Tabelle 5-8: Vorversuchsergebnisse – Baugruppengestaltung	62
Tabelle 5-9: Ableitung Stichprobenumfang – Baugruppengestaltung	63
Tabelle 5-10: Vorversuchsergebnisse – Auswahl von Normteilen.....	63
Tabelle 5-11: Ableitung Stichprobenumfang – Auswahl von Normteilen.....	64
Tabelle 5-12: Vorversuchsergebnisse – Dimensionierung von Formelementen	64
Tabelle 5-13: Ableitung Stichprobenumfang – Dimensionierung von Formelementen.....	65
Tabelle 5-14: Untersuchungsobjekte zur Bauteilgestaltung.....	66
Tabelle 5-15: Untersuchungsobjekte zur Baugruppengestaltung	69
Tabelle 5-16: Untersuchungsobjekte zur Auswahl von Normteilen.....	71
Tabelle 5-17: Untersuchungsobjekte zur Dimensionierung von Formelementen	73
Tabelle 5-18: Nutbreite, -tiefe und -länge der Untersuchungsobjekte (Hypothese FE3) ...	74
Tabelle 5-19: Notwendige Code-Anpassung für Windows-Implementierung	79
Tabelle 5-20: Versuchsergebnisse zur Bauteilgestaltung – Rundumsicht	84
Tabelle 5-21: Mann-Whitney-U-Test – Bauteilgestaltung – Rundumsicht.....	85
Tabelle 5-22: Versuchsergebnisse zur Bauteilgestaltung – Draufsicht	86

Tabelle 5-23: Versuchsergebnisse zur Baugruppengestaltung – Rundumsicht	87
Tabelle 5-24: Versuchsergebnisse zur Baugruppengestaltung – Draufsicht.....	89
Tabelle 5-25: Versuchsergebnisse zur Auswahl von Normteilen – Rundumsicht	91
Tabelle 5-26: Mann-Whitney-U-Test 1 – Auswahl von Normteilen – Rundumsicht.....	93
Tabelle 5-27: Mann-Whitney-U-Test 2 – Auswahl von Normteilen – Rundumsicht.....	95
Tabelle 5-28: Versuchsergebnisse zur Auswahl von Normteilen – Draufsicht	96
Tabelle 5-29: Mann-Whitney-U-Test 1 – Auswahl von Normteilen – Draufsicht.....	98
Tabelle 5-30: Mann-Whitney-U-Test 2 – Auswahl von Normteilen – Draufsicht.....	100
Tabelle 5-31: Versuchsergebnisse zu Formelementen – Rundumsicht.....	101
Tabelle 5-32: Einstichproben-Wilcoxon-Test – Formelemente – Rundumsicht.....	102
Tabelle 5-33: Versuchsergebnisse zu Formelementen – Draufsicht.....	102
Tabelle 5-34: Versuchsergebnisse – Zusammenfassung.....	103
Tabelle 6-1: Arbeitsplan zur Montage der explosionsgeschützten Rohrleuchte.....	105
Tabelle 6-2: Modelle der Rohrleuchte zur Bauteilgestaltung – Teil 1.....	110
Tabelle 6-3: Modelle der Rohrleuchte zur Bauteilgestaltung – Teil 2.....	111
Tabelle 6-4: Versuchsergebnisse zur Rohrleuchte – Bauteilgestaltung.....	112
Tabelle 6-5: mAP in Abhängigkeit des Hintergrundes	112
Tabelle 6-6: Modelle der Rohrleuchte zur Baugruppengestaltung.....	113
Tabelle 6-7: Versuchsergebnisse zur Rohrleuchte – Baugruppengestaltung	114
Tabelle 6-8: Modelle der Rohrleuchte zur Auswahl von Normteilen	115
Tabelle 6-9: Versuchsergebnisse zur Rohrleuchte – Normteile.....	116
Tabelle 6-10: Modelle der Rohrleuchte zur Dimensionierung von Formelementen.....	117
Tabelle 6-11: Versuchsergebnisse zur Rohrleuchte – Formelemente	119
Tabelle 6-12: Stückliste des Produktbeispiels Würfel-Regal.....	120
Tabelle 6-13: Modelle der Würfel-Regale.....	122
Tabelle 6-14: Versuchsergebnisse zu den Würfel-Regalen.....	123
Tabelle 7-1: Montagegerechte Gestaltungsrichtlinien zum Einsatz von Deep Learning..	125
Tabelle 7-2: Entfallende montagegerechte Gestaltungsrichtlinien.....	127

Abkürzungsverzeichnis

AL	Activation Layer
AP	Average Precision
AR	Augmented Reality
BG	Baugruppe
BN	Batch Normalization
BT	Bauteil
CL	Convolutional Layer
CNN	Convolutional Neural Network
COCO	Common Objects in Context
D	Durchmesser einer Bohrung
DFMA	Design for Manufacture and Assembly
DIN	Deutsches Institut für Normung
DL	Deep Learning
EF	Einfarbig
EN	Europäische Norm
F	Fase
FCL	Fully Connected Layer
FE	Formelement
FN	False Negative
FP	False Positive
FPN	Feature Pyramid Network

GS	Gewindestift
IEEE	Institute of Electrical and Electronics Engineers
IoU	Intersection over Union
ISO	International Organization for Standardization
KI	Künstliche Intelligenz
KNN	Künstliche neuronale Netze
LED	Light Emitting Diode
LK	Linsenkopf
LSTM	Long Short-Term Memory
M	Mutter
mAP	Mean Average Precision
MF	Mehrfarbig
ML	Machine Learning
MQTT	Message Queuing Telemetry Transport
N	Normteil
NMS	Non-Maximum Suppression
P	Precision
PL	Pooling Layer
R	Recall
R-CNN	Region-based Convolutional Neural Network
ReLU	Rectified Linear Unit
ResNet	Residual Network

ResNet-101	Residual Network mit 101 Schichten
ResNet-50	Residual Network mit 50 Schichten
ResNext-101	Next Generation Residual Networks mit 101 Schichten
Roi	Region of Interest
RPN	Region Proposal Network
SGD	Stochastical Gradient Descent
SPS	Speicherprogrammierbare Steuerung
SSD	Single Shot Multibox Detector
SW	Schwarz-Weiß
T	Tiefe einer Bohrung
TP	True Positive
U	Unterlegscheibe
VDI	Verein Deutscher Ingenieure
YOLO	You Only Look Once
ZK	Zylinderkopf

Symbolverzeichnis

Symbol Bedeutung

Lateinische Symbole

A	Pixel-Matrix des Originalbildes zur Anwendung des Prewitt-Operators
a_j	Funktionswert eines künstlichen Neuronen
b_j	Bias eines künstlichen Neuronen
G	Kantenstärke des Prewitt-Operators
G_x	Gradientenbild des Prewitt-Operators für horizontale Kanten
G_y	Gradientenbild des Prewitt-Operators für vertikale Kanten
g	Aktivierungsfunktion eines künstlichen Neuronen
H_0	Nullhypothese
H_1	Alternativhypothese
m	Anzahl Faktorstufenkombinationen
N	Anzahl Einzelversuche
n	Stichprobenumfang
P	Precision
p	Signifikanz
R	Recall
\bar{R}	Relative Spannweite
r	Effektstärke nach Cohen
s	Standardabweichung für eine Stichprobe

s^2	Varianz für eine Stichprobe
U	U-Statistik
w_{ij}	Gewichtungsfaktor eines künstlichen Neuronen
\bar{x}	Arithmetisches Mittel einer Stichprobe
x_a	Einzelmesswert
x_i	Eingabewert eines künstlichen Neuronen
x_{max}	Maximaler Messwert einer Stichprobe
x_{min}	Minimaler Messwert einer Stichprobe
Z	Z-Statistik

Griechische Symbole

α	Fehler 1. Art
β	Fehler 2. Art
$\Delta\mu$	Zu beweisender Effekt
σ	Standardabweichung für eine Zufallsvariable

1 Einleitung

1.1 Motivation zur Arbeit

Im Rahmen des Produktentstehungsprozesses stellt die manuelle Montage in Unternehmen oft einen erforderlichen Herstellungsschritt dar. Auch wenn sich seit der dritten industriellen Revolution (siehe Abbildung 1-1) ein Trend zur Automatisierung eingestellt hat, gibt es nach wie vor eine große Anzahl an Produkten, welche manuell montiert werden. Die Gründe dafür sind vielfältig. Beispielsweise erfordert die Automatisierung in den meisten Fällen einen sehr hohen Investitionsaufwand, welcher sich erst bei großen Absatzmengen amortisiert. Weiterhin gibt es Erzeugnisse, deren Montage sich aufgrund ihrer Komplexität nicht automatisieren lässt. Aus diesen Gründen kommt der manuellen Montage in Deutschland und Europa auch heute noch ein hoher Stellenwert zu.

Aus dem Prozess der manuellen Montage resultiert jedoch ein hoher Zeit- und Kostenaufwand für Unternehmen. Dieser Herausforderung stehen aufgrund der hohen Arbeitskosten insbesondere die Betriebe europäischer Industrieländer gegenüber. So waren 2018 die ersten elf Plätze der Länder mit den höchsten Arbeitskosten im verarbeitenden Gewerbe von europäischen Staaten besetzt. Mit 51,50 Euro je geleisteter Stunde weist die Schweiz im internationalen Vergleich die höchsten Arbeitskosten auf. Während Deutschland mit 41 Euro je geleisteter Stunde den fünften Rang belegt, liegen die Arbeitskosten in China bei lediglich 7 Euro je geleisteter Stunde. [1]

Eine weitere Herausforderung stellen die steigenden Anforderungen an die manuelle Montage durch eine wachsende Variantenvielfalt dar. Mit der dritten industriellen Revolution erfolgte eine vollständige Verschiebung der Marktsituation vom Verkäufer- zum Käufermarkt. Seither übersteigt das Angebot auf dem Markt die Nachfrage. Neben den veränderten Marktbedingungen ist seit der dritten industriellen Revolution außerdem ein gesellschaftlicher Trend zur Individualisierung zu verzeichnen (siehe Abbildung 1-1). Aus sozialwissenschaftlicher Sicht bedeutet Individualisierung, dass sich der Einzelne aus den Schicht- und Klassenmilieus herauslöst. Aus diesem Trend resultierte wiederum eine wachsende Heterogenität des Konsumverhaltens der Gesellschaft, sodass Unternehmen ihre Erzeugnisse zunehmend individualisieren mussten, um den Bedürfnissen ihrer Kunden gerecht zu werden und am Markt bestehen zu können. [2]

Es ist davon auszugehen, dass sich der Trend zur Individualisierung im Rahmen der vierten industriellen Revolution noch weiter verschärft. So wird in der Literatur – z. B. in [2] - [4] – aktuell diskutiert, inwieweit die zukünftige Produktion kundenindividueller Produkte in oder bis hin zur Losgröße 1 realisierbar und sinnvoll ist.

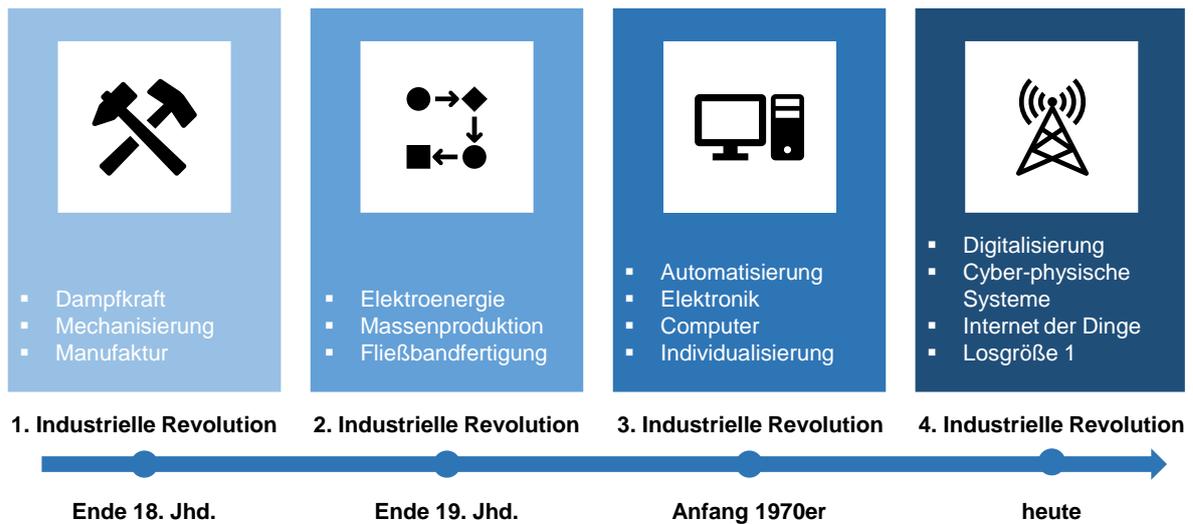


Abbildung 1-1: Die vier industriellen Revolutionen

Die aus der Individualisierung resultierende, wachsende Variantenvielfalt stellt die manuelle Montage jedoch vor große Herausforderungen. Mitarbeiter müssen immer flexibler und schneller für neue Produkte und Varianten angeleitet werden. Durch den gleichzeitig zunehmenden Fachkräftemangel in Deutschland, wird die Situation weiter verschärft. Der Notwendigkeit die Anlernzeit zu reduzieren, stehen u.a. hohe Fluktuationsraten [5] und sprachliche Barrieren [6] entgegen. Aufgrund des eingeschränkten Fachkräfteangebotes müssen viele Unternehmen außerdem auf weniger qualifizierte Mitarbeiter zurückgreifen.

Um diesen Herausforderungen zu begegnen und im internationalen Wettbewerb weiter bestehen zu können, werden in einer Vielzahl laufender Forschungsprojekte potenzielle Optimierungsansätze für die manuelle Montage untersucht. Umfassende neue Ansätze bietet die Integration von Objekterkennungsalgorithmen. Durch das Erfassen des Montageprozesses mittels Kamera und anschließender Objektdetektion ist es möglich, die Position, Lage und Art von Bauteilen zu bestimmen, sodass Informationen über Arbeitsschritte, Montagefehler oder den aktuellen Zustand des Produkts abgeleitet werden können. Deep Learning liegt vor, wenn die Objekte mithilfe von künstlichen neuronalen Netzen komplexer Struktur detektiert werden [7]. Durch einen vorgelagerten Trainingsprozess lassen sich Bilder aller Bauteile und typische Montagefehler in ein Objekterkennungsmodell integrieren, welches anschließend z. B. zur Überwachung oder Steuerung des Montageprozesses eingesetzt werden kann. Im Vergleich zu anderen Methoden zur Objekterkennung entfällt durch den Einsatz von Deep Learning das zeit-, wissens- und damit kostenintensive Feature-Engineering, siehe Kapitel 2.2.1.

Der Einsatz von Deep Learning in der manuellen Montage birgt große Potenziale, Montagezeiten zu reduzieren und Montagefehler zu vermeiden. Es stellt sich jedoch die Frage, inwieweit Produkte, welche nach dem aktuellen Stand der Technik konzipiert wurden, für den Einsatz von Deep Learning geeignet sind. Nur wenn Bauteile und Baugruppen durch bestehende Deep Learning Modelle eindeutig erkannt werden, ist ihre Anwendung in der manuellen Montage sinnvoll. Zum aktuellen Zeitpunkt bestehen zwar zahlreiche

Gestaltungsrichtlinien für eine montagegerechte Konstruktion, diese greifen jedoch nicht den Aspekt der Integration von Deep Learning auf. Auch aktuelle Untersuchungen im Bereich der Produktgestaltung betrachten diesen Kontext nicht.

Dass eine optimale Ausrichtung der Produktgestaltung auf die im Rahmen der Fertigung (inkl. Montage) eingesetzten Technologien für Unternehmen jedoch von sehr hoher Bedeutung ist, zeigt Abbildung 1-2. Während die Entwicklungs- und Konstruktionsabteilung lediglich 6 % der Selbstkosten verursacht, werden 70 % der Kosten durch die Arbeit dieser Abteilung festgelegt. Die Produktgestaltung bestimmt damit in erheblichem Maße die zukünftigen Selbstkosten eines Produktes und hat somit einen großen Einfluss auf den Erfolg eines Unternehmens.

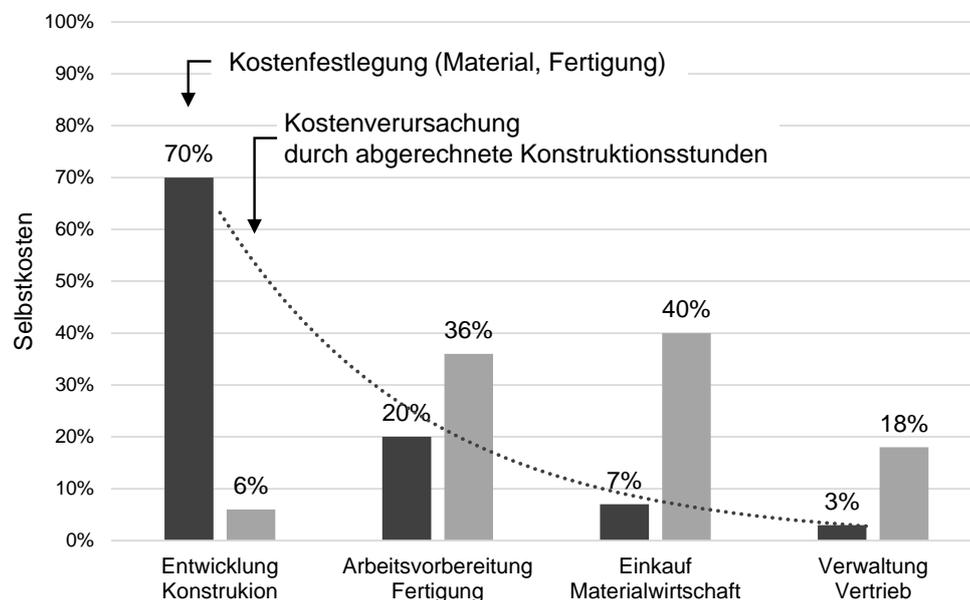


Abbildung 1-2: Kostenfestlegung und Kostenverursachung im Unternehmen (nach [8])

1.2 Zielsetzung

Um das Potential von Deep Learning in der manuellen Montage in vollem Umfang auszuschöpfen und die Selbstkosten bereits in der Konstruktionsabteilung zu optimieren, sollte die Produktgestaltung auf diesen Anwendungsfall ausgerichtet werden. Aufgrund der steigenden Ansprüche an die manuelle Montage ist davon auszugehen, dass zukünftig vermehrt Werkerassistenzsysteme zum Einsatz kommen werden, welche Deep Learning nutzen und die Thematik damit noch bedeutsamer wird.

Ziel der vorliegenden Arbeit ist es, die Lücke der aktuell bestehenden und etablierten Gestaltungsrichtlinien für montagegerechte Konstruktionen bezüglich des Einsatzes von Deep Learning in der manuellen Montage zu schließen. Durch umfangreiche empirische Untersuchungen soll ermittelt werden, inwieweit verschiedene Produkteigenschaften die

Detektierbarkeit durch Deep Learning beeinflussen. Als Ergebnis sollen Gestaltungsrichtlinien abgeleitet werden, um Produkte für den Einsatz von Deep Learning in der Montage effektiv und effizient gestalten zu können. Ziel ist es, dem Konstrukteur der Zukunft ein Werkzeug zur Verfügung zu stellen, welches die sich ändernden Rahmenbedingungen in der manuellen Montage berücksichtigt.

1.3 Gliederung der Arbeit

Um eine Ausgangsbasis zu schaffen, werden zunächst die wissenschaftlichen Grundlagen (Kapitel 2) sowie der Stand der Technik (Kapitel 3) hinsichtlich der montagegerechten Produktgestaltung und der Deep Learning Technologie analysiert und erläutert. Die Analyse der montagegerechten Produktgestaltung dient der Evaluation der aktuellen Gestaltungsrichtlinien. Sie bilden die Grundlage der vorliegenden Arbeit und sollen durch diese erweitert bzw. angepasst werden. Die Ausführung der Deep Learning Technologie stellt wiederum die Basis zur Ableitung von Optimierungs- bzw. Ergänzungsansätzen für aktuelle Gestaltungsrichtlinien dar.

Auf Grundlage dessen werden Anforderungen an die Gestaltung von Produkten, welche unter Zuhilfenahme von Deep Learning montiert werden sollen, definiert, siehe Kapitel 4. Dabei werden in einem ersten Schritt Einflussfaktoren erläutert, welche sich aus den Grundlagen der Deep Learning Technologie ergeben, um daraus in einem zweiten Schritt Hypothesen abzuleiten. Diese gliedern sich in Hypothesen zur Gestaltung von Bauteilen, Baugruppen und Formelementen, sowie zur Auswahl von Normteilen.

Zur Validierung der aufgestellten Thesen werden empirische Untersuchungen durchgeführt (Kapitel 5). Dazu werden zunächst die abhängigen und unabhängigen Variablen identifiziert und ein Versuchsaufbau definiert. Anschließend erfolgen die Auswahl des anzuwendenden Objekterkennungsalgorithmus sowie die statistische Versuchsplanung. Im Rahmen der darauf folgenden Versuchsdurchführung werden Untersuchungsobjekte konzipiert, Trainings- und Validierungsdatensätze erstellt, der ausgewählte Objekterkennungsalgorithmus implementiert und verschiedene, den Hypothesen entsprechende Modelle trainiert und validiert. Abschließend erfolgt eine umfangreiche statistische Auswertung der Versuchsergebnisse. Als Ergebnis liefert Kapitel 5 eine Übersicht zur Bestätigung oder Ablehnung aller betrachteten Hypothesen.

Da die empirische Untersuchung zur Reduktion von äußeren Störeinflüssen bewusst unter definierten Laborbedingungen durchgeführt wird, untersucht Kapitel 6 schließlich die Übertragbarkeit der Ergebnisse auf reale Produktbeispiele in realen Montageumgebungen. Dabei kommt neben einer explosionsgeschützten Leuchte aus dem industriellen Umfeld ein Regal aus dem Konsumentenbereich zum Einsatz.

Auf Grundlage der erzielten Ergebnisse werden in Kapitel 7 konstruktive Gestaltungsrichtlinien für montagegerechte Konstruktionen abgeleitet und in einer Übersicht zusammengestellt. Abschließend werden bestehende Gestaltungsempfehlungen hinsichtlich ihrer Relevanz für zukünftige, digitale Montageumgebungen analysiert.

2 Grundlagen

2.1 Montagegerechte Produktgestaltung

2.1.1 Der Produktentstehungsprozess

Der Produktentstehungsprozess ist ein wichtiger Bestandteil des technischen Produktlebenszyklus und lässt sich in die Phasen Produktplanung, Produktentwicklung und Realisierung/ Produktion einteilen [9], siehe Abbildung 2-1. Während die Produktplanung die allgemeine Gestaltung des Unternehmensangebots zum Ziel hat, ist die Produktentwicklung ein interdisziplinärer Unternehmensprozess zur Entwicklung eines marktfähigen Produkts, basierend auf einer Definition initialer Ziele und Anforderungen an das Produkt, die im Lauf des Prozesses kontinuierlich weiterentwickelt und iterativ angepasst werden. Die einzelnen Phasen sind zwar sequenziell aufeinander folgend dargestellt, sie werden jedoch meist iterativ und (teil-) parallelisiert durchlaufen.

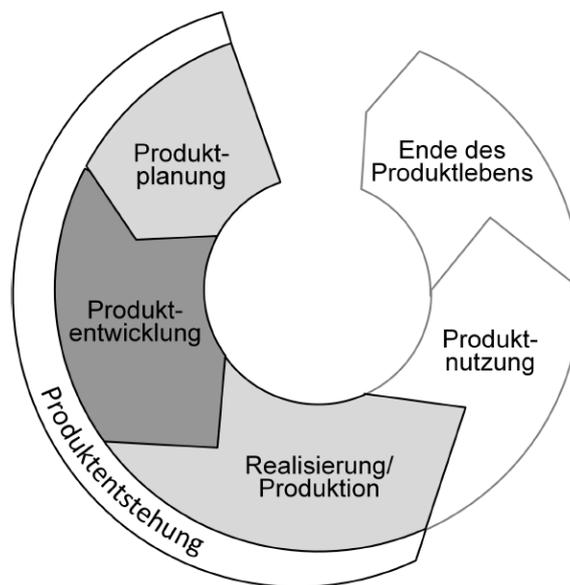


Abbildung 2-1: Technischer Produktlebenszyklus [9]

Um den umfangreichen und komplexen Produktentwicklungsprozess besser verständlich zu machen, wurde bereits seit Ende des neunzehnten Jahrhunderts Anstrengungen zur Systematisierung unternommen. Erste Überlegungen machte Reuleaux [10], indem er ein modellhaftes Verfahren zur kinematischen Synthese entwickelte. Einen anderen Ansatz lieferten Pahl und Beitz [11], indem sie alle Arbeitsschritte, welche im Rahmen der Produktplanung und Produktentwicklung durchgeführt werden, zum Entwicklungs- und Konstruktionsprozess zusammengefasst und diesen wiederum in vier Phasen eingeteilt haben, siehe Abbildung 2-2.

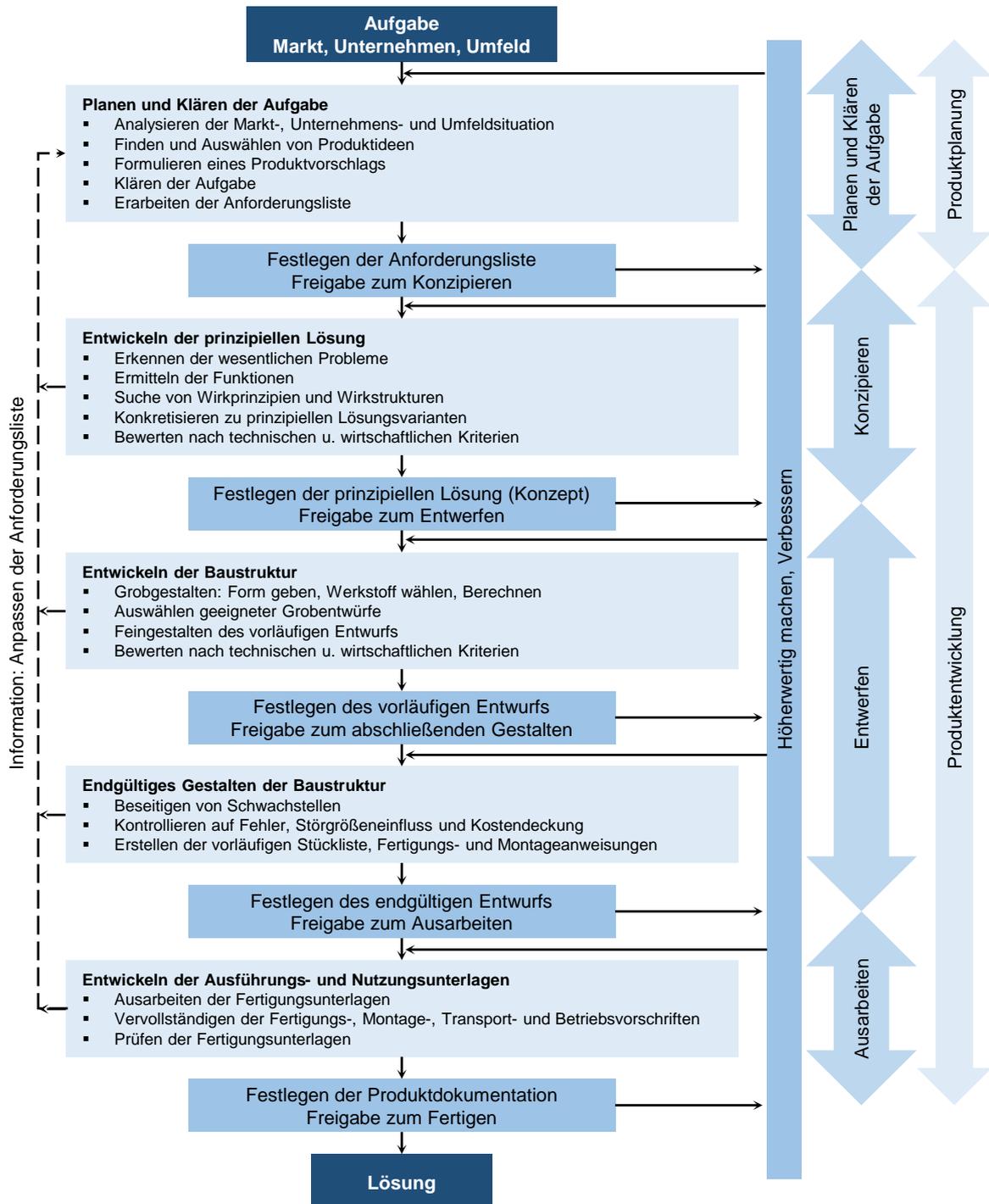


Abbildung 2-2: Hauptschritte des Entwicklungs- und Konstruktionsprozesses (nach [11])

Als erste Phasen definierten Pahl und Beitz das Planen und Definieren der Aufgabenstellung. Diese stellt die Basis der Entwicklungsaufgabe dar. Vor dem Beginn der eigentlichen Produktentwicklung (Phase 2-4) muss die Aufgabenstellung hinreichend geklärt sein. Hierzu sind Informationen über die Anforderungen, Bedingungen und ihre jeweilige Bedeutung zu beschaffen und in einer Anforderungsliste zu dokumentieren. Das Konzipieren stellt die zweite Phase des Entwicklungs- und Konstruktionsprozesses dar. Da im Rahmen der

nachfolgenden Phasen (Entwerfen und Ausarbeiten) grundlegende Mängel des Lösungsprinzips nur mit hohem Aufwand ausgeglichen werden können, sollte die Phase des Konzipierens konsequent durchlaufen werden. Somit soll sichergestellt werden, dass die bestmögliche Lösung erarbeitet wird. Als Ergebnis der zweiten Phase sollte ein Konzept sowie dessen Freigabe zum Entwerfen vorliegen.

Im Rahmen des Entwerfens (Phase 3) wird auf Grundlage der Wirkstruktur bzw. der prinzipiellen Lösung die Baustruktur nach technischen und wirtschaftlichen Gesichtspunkten eindeutig und vollständig erarbeitet. Als Ergebnis werden vorläufige Fertigungsunterlagen (Stückliste, Fertigungs- und Montageanweisungen) erstellt. Schlussendlich erfolgt das Ausarbeiten. Im Rahmen dieser letzten Phase wird die Baustruktur durch finale Angaben hinsichtlich der Form, Bemessung, Oberflächenbeschaffenheit aller Einzelteile sowie die Festlegung aller Werkstoffe finalisiert. Darüber hinaus erfolgt eine Überprüfung der Herstellmöglichkeiten und die Festlegung der endgültigen Kosten. Alle Fertigungs-, Montage-, Transport- und Betriebsvorschriften werden ausgearbeitet und abschließend geprüft.

Pahl und Beitz weisen darauf hin, dass die dargestellten Arbeitsschritte stark verallgemeinert sind und in der Praxis meist keine klare Trennung der Arbeitsschritte und Zwischenergebnisse erfolgt. Auch wenn die strikte Bearbeitung aller Arbeitsschritte in der Praxis meist nicht umsetzbar ist, dient die Einteilung des Produktentwicklungsprozesses als Orientierung und kann dadurch zu einer besseren Planung und einer reduzierten Fehleranfälligkeit führen.

Weitere Ansätze, den Produktentwicklungsprozess systematisch zu strukturieren, lieferten Wögerbauer [12], Kesselring [13], Hansen [14], Hubka [15], Koller [16], Roth [17] sowie Rodenacker [18]. Ziel aller Überlegungen war es, Konstrukteuren Hilfsmittel an die Hand zu geben, um Produkte effizienter zu entwickeln. Um einen deutschen Standard zu etablieren und die wissenschaftlichen Erkenntnisse zu harmonisieren, erschien 1986 die erste Auflage der VDI 2221 [19], welche aufgrund neuer Erkenntnisse zuletzt 2019 überarbeitet wurde. Neben der übergeordneten Richtlinie VDI 2221 [9] verweisen die Autoren auf zusätzliche Richtlinien wie die VDI 2222 [20] [21], VDI 2223 [22] und VDI 2206 [23], welche Sonderaspekte und Details der methodischen Produktentwicklung und Lösungsfindung spezifizieren.

So weist die VDI 2223 unter anderem auf die Nutzung von Gestaltungsrichtlinien (auch Konstruktionsregeln genannt) hin. Diese geben Konstrukteuren konkrete Hinweise auf potenzielle konstruktive Schwachstellen und schlagen geeignete Verbesserungen vor. Sie können als Checkliste zum Überprüfen von Entwürfen vor der Freigabe zum Ausarbeiten dienen. In der Fachliteratur sind sie meist als Lösungssammlungen oder in Konstruktionskatalogen dokumentiert. Die Gestaltungsrichtlinien können sich dabei auf eine Vielzahl von Aspekten beziehen. So gibt es bspw. Regeln zum beanspruchungsgerechten, kostengerechten, schweißgerechten oder fertigungsgerechten Konstruieren. Für die manuelle Montage relevant sind jedoch vor allem die Hinweise zur montagegerechten Produktgestaltung.

2.1.2 Montage in der industriellen Produktion

Gemäß Lotter und Wiendahl [24] setzen sich industriell hergestellte Produkte meist aus einer großen Anzahl von Einzelteilen zusammen, welche zu unterschiedlichen Zeitpunkten und mit verschiedenen Fertigungsverfahren hergestellt werden. Im Rahmen der Montage werden diese in einer bestimmten Zeit zu einem Produkt mit höherer Komplexität zusammengesetzt, welches eine definierte Funktion erfüllen soll. Während die Fertigung und Montage im Zeitalter der Manufakturen noch nicht getrennt voneinander betrachtet wurden, wird die Montage seit der zweiten industriellen Revolution als eigener Produktionsabschnitt verstanden. So führte bspw. das schnelle Wachstum der Automobilindustrie in den USA Anfang des 20. Jahrhunderts zu der Notwendigkeit, die Montage wirtschaftlicher zu gestalten, woraus sich die von Henry Ford 1908 erstmals eingeführte Fließmontage entwickelte. [24]

Nach der VDI-Richtlinie 2860 [25] werden unter dem Begriff der Montage alle Vorgänge zusammengefasst, die dem Zusammenbau von geometrisch bestimmten Körpern dienen. Dabei kann zwischen einer Vielzahl unterschiedlicher Vorgänge unterschieden werden. Die Montageplanung hat die Aufgabe, die erforderlichen Einzelfunktionen auszulegen und aufeinander abzustimmen. Da das Verständnis der einzelnen Vorgänge im Montageprozess die Grundlage für eine montagegerechte Produktgestaltung darstellt, werden die einzelnen Montageoperationen in Kapitel 2.1.2.1 näher erläutert. Darüber hinaus nimmt auch der Automatisierungsgrad Einfluss auf die montagegerechte Produktgestaltung, siehe Kapitel 2.1.2.2.

2.1.2.1 Montageoperationen

Die Montage eines Produktes besteht aus einer Vielzahl verschiedener Operationen, welche je nach Enderzeugnis, Stückzahl oder Automatisierungsgrad in unterschiedlicher Reihenfolge, Häufigkeit und Ausprägung auftreten können. Nach Bender und Gericke [26] können die Montageoperationen in Haupt- und Nebentätigkeiten, sowie wertschöpfende und nicht-wertschöpfende Tätigkeiten untergliedert werden, siehe Abbildung 2-3. Insgesamt wird zwischen den fünf Montageoperationen Fügen, Inbetriebnahme, Sonderoperationen, Handhaben und Hilfsprozesse unterschieden.

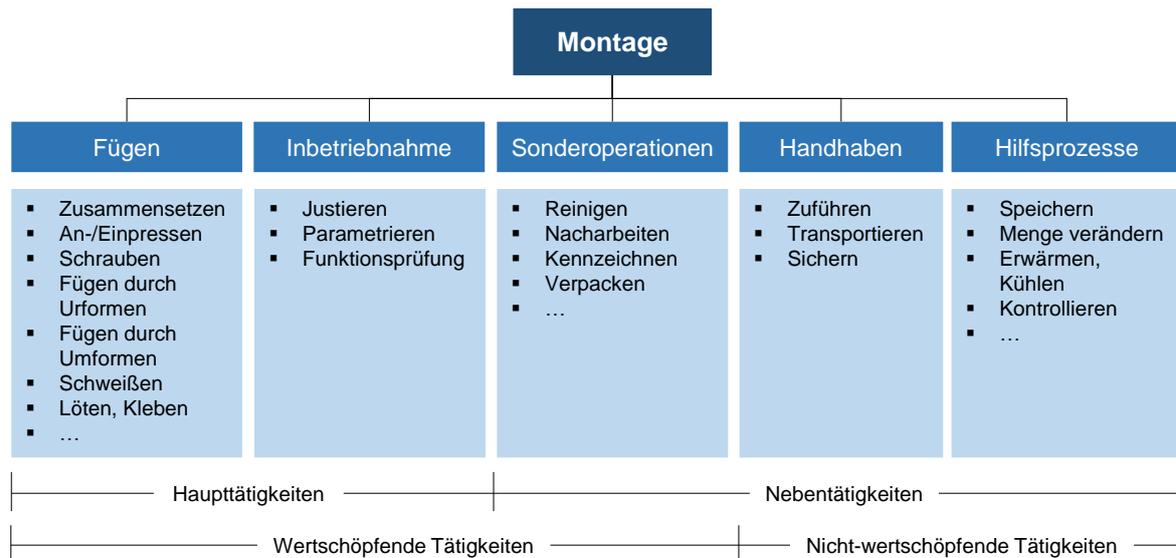


Abbildung 2-3: Montageoperationen (in Anlehnung an [26])

Gemäß DIN 8593-0 beschreibt die Montageoperation **Fügen** das „auf Dauer angelegte[] Verbinden oder sonstige[] Zusammenbringen von zwei oder mehreren Werkstücken geometrisch bestimmter fester Form oder von eben solchen Werkstoffen mit formlosem Stoff“ [27]. Dabei unterscheidet die Norm zwischen einer Vielzahl verschiedener Fügeverfahren, welche hinsichtlich Vorgehensweise und erzielter Effekte variieren. So kann der Zusammenhalt form-, kraft- oder stoffschlüssig erzeugt werden. Außerdem gibt es Fügeverfahren, welche sich Hilfselementen bedienen und Verfahren, welche ohne Zusatz auskommen. Das ausgewählte Fügeverfahren beeinflusst neben der Produktlebensdauer auch die Möglichkeiten zur einfachen Wartung und Reparatur, die Recyclingfähigkeit sowie die Wirtschaftlichkeit. Da die Ansprüche an eine montagegerechte Produktgestaltung stark von den eingesetzten Fügeverfahren abhängen, sollten diese im Produktentwicklungsprozess berücksichtigt werden.

Unter dem Begriff **Inbetriebnahme** werden alle Tätigkeiten zusammengefasst, welche das montierte Produkt in einen betriebsfähigen Zustand versetzen. Dabei wird zwischen einer Justierung, Parametrierung und Funktionsprüfung unterschieden. **Sonderoperationen** erzeugen zwar ebenfalls eine Wertsteigerung, können jedoch den Grundaufgaben des Fügens und der Inbetriebnahme nicht zugeordnet werden. Beispiele für Sonderoperationen sind das Reinigen, Verpacken oder Nacharbeiten des Endproduktes.

Die VDI-Richtlinie 2860 definiert das **Handhaben** als „das Schaffen, definierte Verändern oder vorübergehende Aufrechterhalten einer vorgegebenen räumlichen Anordnung von geometrisch bestimmten Körpern in einem Bezugskordinatensystem“ [25]. Es beinhaltet die Aufgabenbereiche Zuführen, Transportieren und Sichern. Während das Verbringen von Einzelteilen von der Materialbereitstellung des Montagesystems zum Montageprozess als Zuführen bezeichnet wird, beschreibt das Transportieren das Verbringen des Erzeugnisses zwischen einzelnen Montagestationen. Für bestimmte Montagevorgänge ist es außerdem sinnvoll, das Erzeugnis vor der Operation zu sichern, d. h. eine definierte räumliche Lage

vorrübergehend sicherzustellen. Um eine einfache Handhabung der Einzelteile und Baugruppen zu gewährleisten, sollte diese bereits im Rahmen der Produktgestaltung berücksichtigt werden. **Hilfsprozesse** sind zur Durchführung der Montageaufgabe zwar erforderlich, erzeugen dabei jedoch keine Wertsteigerung des Produktes. Aus diesem Grund sollten sie auf ein Minimum reduziert werden.

2.1.2.2 Automatisierungsgrad

Neben den zu berücksichtigenden Montageoperationen nimmt auch der angestrebte Automatisierungsgrad Einfluss auf die montagegerechte Produktgestaltung. Prinzipiell kann dabei zwischen manueller, teilautomatisierter und automatisierter Montage unterschieden werden. Die Auswahl des Automatisierungsgrades hängt von verschiedenen Faktoren ab. Neben der Produktgröße und -komplexität sind dabei die Stückzahlen, die Variantenvielfalt, der Schwierigkeitsgrad der Montageoperationen, die angestrebte Qualität und Prozesssicherheit sowie die Logistik und Materialbereitstellung zu berücksichtigen. [24] [28]

Die manuelle Montage eignet sich insbesondere für Produkte mit geringer Stückzahl. Aufgrund der hohen Flexibilität wird die manuelle Montage oft für Produkte mit hoher Variantenvielfalt oder Stückzahlschwankungen angewendet. Da die Montage nicht an einen Maschinentakt gekoppelt ist, besteht größerer Handlungs- und Entscheidungsspielraum für die Montagemitarbeiter. [28]

Die halbautomatische Montage wird auch als hybrides Montagesystem bezeichnet. Sie wird vorwiegend für Produkte mit mittlerer Variantenvielfalt und Stückzahlen eingesetzt, wenn eine manuelle Montage nicht mehr wirtschaftlich ist, sich eine vollständig automatisierte Montage jedoch noch nicht rentiert oder wenn sich einzelne Montageschritte technisch nicht automatisieren lassen. Meist werden die einzelnen Montageschritte auf mehrere Stationen aufgeteilt und manuelle sowie automatische Arbeitsanteile bspw. durch ein Montageband miteinander verkettet. [24] [28]

Im Rahmen der automatischen Montage erfolgt die Durchführung der Montageaufgabe mittels Maschinen. Hierbei können entweder Einzweckautomaten oder flexible Einheiten, welche mehrere Montageschritte bearbeiten, eingesetzt werden. Aufgrund des meist sehr hohen Investitionsaufwandes ist der Einsatz eines automatischen Montagesystems in der Regel erst bei sehr hohen Stückzahlen und einer geringer Variantenvielfalt wirtschaftlich. In Abbildung 2-4 ist der Einfluss der Investitionskosten, der angestrebten Flexibilität sowie der Stückzahl auf die Auswahl des Montagesystems dargestellt. [24]

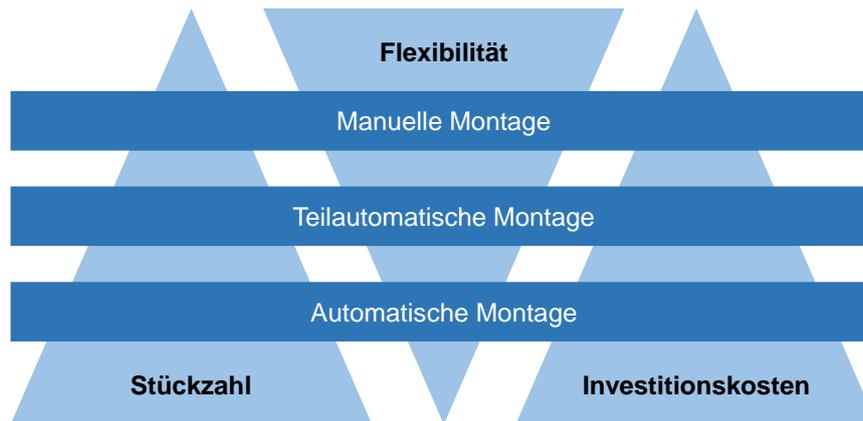


Abbildung 2-4: Auswahlkriterien für Montagesysteme

Der angestrebte Automatisierungsgrad sollte im Rahmen der Produktgestaltung berücksichtigt werden. Je nach Grad der Automatisierung müssen Produkte entsprechende Merkmale aufweisen, um eine optimale Montage zu gewährleisten. Während die manuelle Montage meist geringere Anforderungen an die Produktgestaltung stellt, steigen diese mit zunehmendem Automatisierungsgrad an [26]. So müssen Bauteile, welche automatisch geordnet und zugeführt werden sollen, entsprechend gestaltet werden und es sind Greifflächen für Zuführwerkzeuge zu berücksichtigen. Darüber hinaus sind die aus den eingesetzten Automaten resultierenden, eingeschränkten Fügerichtungen zu berücksichtigen. Wurden Produkte automatisierungsgerecht gestaltet, sind diese in der Regel auch manuell gut zu montieren [26].

2.1.3 Methoden zur montagegerechten Produktgestaltung

Um eine montagegerechte Produktgestaltung zu erzielen, sind in der Literatur viele verschiedene Strategien beschrieben. Gemäß Stoll [29] lassen sich diese Methoden in drei Kategorien unterteilen, siehe Abbildung 2-5. In der ersten Kategorie finden sich Methoden, welche die Montageproblematik in der Konstruktion von Produkten berücksichtigen. Durch Gestaltungsrichtlinien, Beispielsammlungen und Lösungskataloge soll das Bewusstsein für eine montagegerechte Konstruktion erhöht werden. Die Grundlagen dafür wurden bereits in den früher 1980er Jahren unter anderem durch die Werke von Pahl und Beitz [30], Andreasen [31], Hesse [32], Bäßler [33] und Gairola [34] geschaffen. So stellten Pahl und Beitz bereits 1986 die drei Hauptregeln „Vereinheitlichen“, „Vereinfachen“ und „Reduzieren“ auf, welche noch heute Grundlage einer montagegerechten Produktgestaltung sind [30].

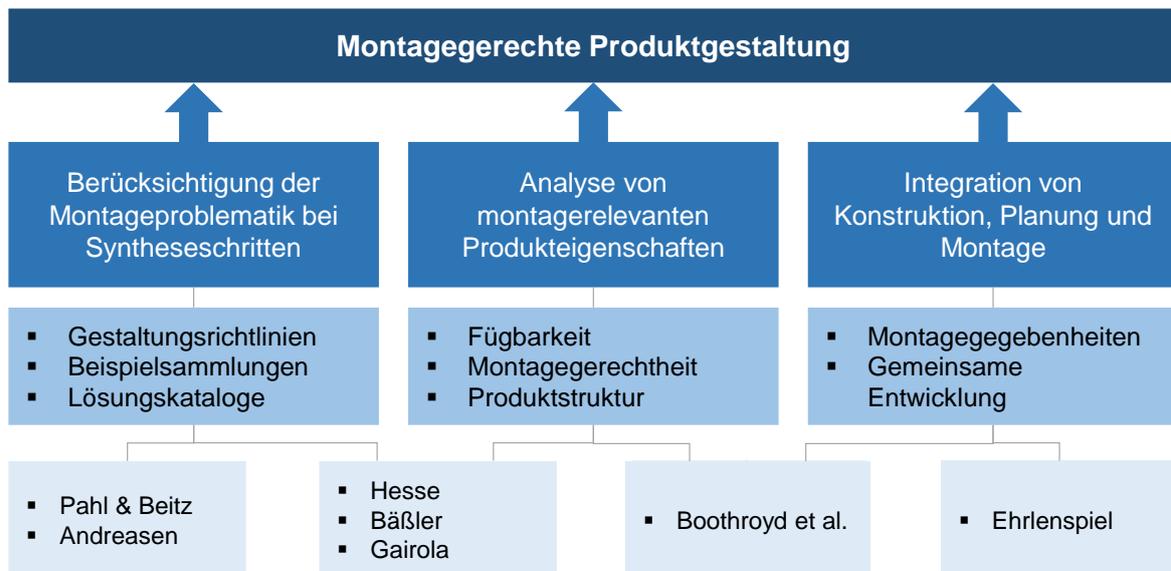


Abbildung 2-5: Ansätze zur montagegerechten Produktgestaltung (nach [29])

Der zweiten Kategorie lassen sich Methoden zuordnen, welche bestehende Konstruktionen hinsichtlich ihrer montagerelevanten Produkteigenschaften analysieren. Eine weit verbreitete Methode hierzu ist die „Design for Manufacture and Assembly (DFMA)“-Methode von Boothroyd, Dewhurst und Knight [35]. Diese wurde erstmals 1994 durch die Autoren beschrieben und seitdem vielfach weiterentwickelt. Ein wichtiger Bestandteil der Methode ist die Berechnung der Konstruktionseffizienz, welche unter anderem die Montagezeiten sowie die theoretisch maximal notwendige Anzahl von Einzelteilen berücksichtigt. Neben der DFMA-Methode gibt es weitere Ansätze, um die Montagegerechtheit eines Produktes zu analysieren. Den Werken von Hesse, Bäßler und Gairola können auch hierzu Hinweise entnommen werden.

Methoden, welche einen integrativen Ansatz verfolgen, werden gemäß Stoll in die dritte Kategorie eingeteilt. Durch organisatorische Maßnahmen sollen Zusammenarbeit und Kommunikation zwischen Konstruktion, Planung und Montage optimiert werden. Ehrlenspiel beschrieb diese Herangehensweise bereits 1995 im Rahmen seines Werkes „Integrierte Produktentwicklung“ [36], welches inzwischen in der sechsten Auflage erschienen ist [37] und als Standardwerk gilt. Heute ist das Vorgehen ebenfalls unter dem Begriff „Simultaneous Engineering“ bekannt. Um eine ideale, montagegerechte Produktgestaltung zu erzielen, sollten prinzipiell Methoden aller drei Kategorien miteinander kombiniert werden.

Im Rahmen der vorliegenden Arbeit sollen Gestaltungsrichtlinien für eine montagegerechte Konstruktion unter der Berücksichtigung des Einsatzes von einer auf Deep Learning basierenden Objekterkennung erarbeitet werden. Diese lassen sich klar der ersten Methodenkategorie zuordnen.

2.2 Deep Learning

2.2.1 Begriffliche Einordnung

Deep Learning (DL) beschreibt eine Teildisziplin des Machine Learning (ML), wobei sich beide Anwendungsfelder dem Fachgebiet der Künstlichen Intelligenz (KI) zuordnen lassen, siehe Abbildung 2-6. Für die Künstliche Intelligenz fehlt bis zum aktuellen Zeitpunkt eine weltweit anerkannte, allgemeingültige Definition. Es gibt jedoch eine Vielzahl von Erklärungsansätzen, welche die Künstliche Intelligenz in ähnlicher Weise beschreiben. So definierte der britische Computerpionier Alan M. Turing bereits 1950, dass ein technisches System als intelligent anzusehen ist, wenn es sich in seinen Antworten und dem Lösen von Problemen nicht von einem Menschen unterscheidet [38]. Auch aktuelle Einordnungen greifen diesen Aspekt auf, erweitern ihn jedoch um weitere Faktoren.

So verstehen Gethmann et al. Künstliche Intelligenz als die Fähigkeit von Computersystemen, auf sie zugeschnittene Aufgaben selbsttätig zu lösen, die aufgrund ihrer Komplexität bislang menschliche Fähigkeiten erforderten. Zu diesen Fähigkeiten zählen das Wahrnehmen (Sehen, Hören, Fühlen, etc.), Lernen, Wissen, Denken, Kommunizieren (Sprechen, Schreiben, etc.) und Handeln. KI-Systeme seien so ausgelegt, dass sie die beschriebenen Fähigkeiten nachahmen und somit komplexe Problemstellungen automatisch lösen können. [39]

Gemäß der High-Level Expert Group on Artificial Intelligence der Europäischen Kommission bezieht sich Künstliche Intelligenz auf Systeme, die intelligentes Verhalten zeigen, indem sie ihre Umgebung analysieren und Maßnahmen mit einem bestimmten Grad an Autonomie ergreifen, um spezifische Ziele zu erreichen. Dabei können die Systeme rein softwarebasiert sein und in der virtuellen Welt agieren (z. B. Sprachassistenten, Bildanalysesoftware, Suchmaschinen, Sprach- und Gesichtserkennungssysteme) oder in Hardware-Geräte eingebettet sein (z. B. fortschrittliche Roboter, autonome Autos, Drohnen oder Internet der Dinge-Anwendungen). [40]

Unter dem Begriff Machine Learning werden alle Methoden zusammengefasst, welche automatisch Muster in Daten erkennen und diese für die Vorhersage zukünftiger Daten oder andere Arten der Entscheidungsfindung nutzen können [41]. Deep Learning liegt vor, wenn zu diesem Zweck tiefe neuronale Netze eingesetzt werden [7].

Während das Lernen bzw. Auswählen von relevanten Merkmalen (Feature-Engineering) und die darauf aufbauende Klassifikation von Eingabedaten beim traditionellen Machine Learning in einzelnen Schritten erfolgt, werden diese beim Deep Learning in ein einziges Modell integriert, siehe Abbildung 2-7. Darüber hinaus zeichnet sich das Feature-Engineering beim traditionellen Machine Learning durch einen großen Anteil manueller Eingriffe aus. So ist im Rahmen des Feature-Engineering im Vorfeld manuell festzulegen, welche Merkmale für die verschiedenen Klassen relevant sind. Die Leistung des erstellten

Modells hängt damit nicht nur von der Optimierung der verwendeten Algorithmen ab, sondern wird auch stark von den zuvor definierten Merkmalen beeinflusst. Im Allgemeinen ist das Feature-Engineering sehr zeitaufwändig und erfordert großes Expertenwissen. [42]

Im Deep Learning werden hingegen tiefe neuronale Netze mit vielen verborgenen Schichten (siehe Kapitel 2.2.3) genutzt, um die relevanten Merkmale selbstständig aus den Eingangsdaten zu extrahieren. Sollen wie im Rahmen der vorliegenden Untersuchung visuelle Informationen verarbeitet werden, erfolgt die Extraktion von Merkmalen wie Kanten, Ecken, Konturen und Objektteilen Schicht für Schicht aus einer Vielzahl von Bildern. Diese Merkmalsrepräsentationen werden dann in die Klassifizierungsschicht eingegeben, um Klassifizierungs- und Regressionsaufgaben durchzuführen. Deep Learning zeichnet sich damit durch ein Minimum menschlicher Eingriffe aus, die Leistung des erstellten Modells wird jedoch stark von der Qualität der Eingangsdaten beeinflusst. [42]

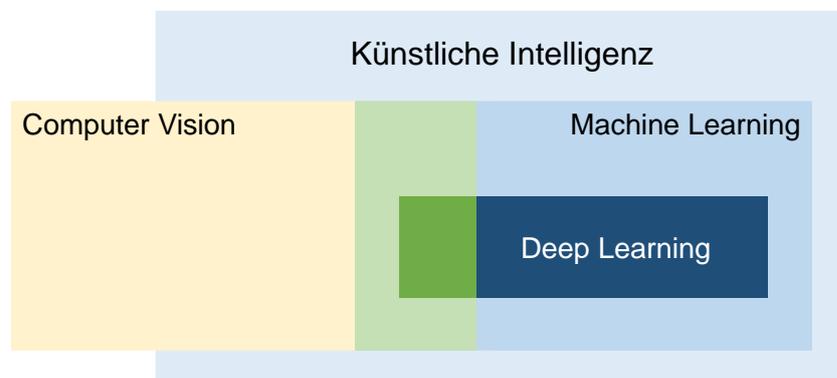


Abbildung 2-6: Einordnung Deep Learning

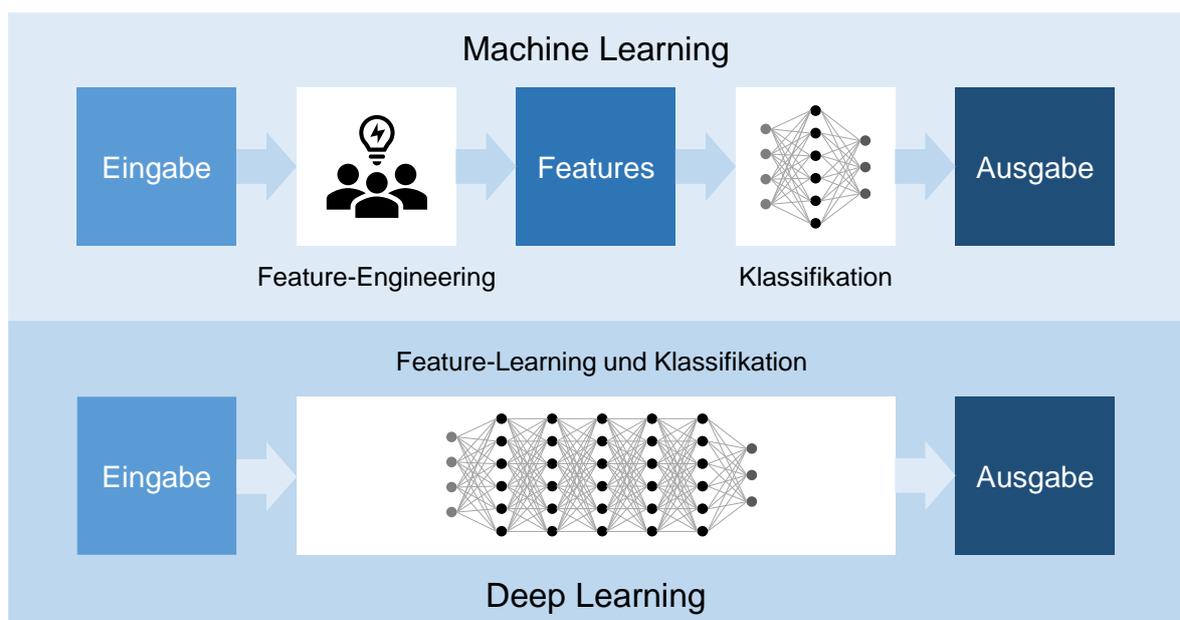


Abbildung 2-7: Machine Learning versus Deep Learning

Prinzipiell kann Deep Learning auf nahezu jede Art von Daten angewendet werden. Die bekanntesten Beispiele sind neben der Sprache (z. B. in Form von Tonaufnahmen oder Texten) die Anwendung auf Bild- und Videomaterial. Sollen wie im Rahmen der vorliegenden Arbeit visuelle Daten analysiert werden, ergibt sich eine Überschneidung von Deep Learning mit einer weiteren Teildisziplin der Künstlichen Intelligenz, der Computer Vision.

Unter dem Begriff der Computer Vision werden Methoden zusammengefasst, welche Computern die Fähigkeit verleihen, das komplexe menschliche Sehvermögen nachzubilden. Dieses umfasst neben der Klassifizierung von Objekten auch ihre Lokalisierung und Segmentierung. [7]

2.2.2 Historische Einordnung

Die Anfänge von Deep Learning gehen auf einfache lineare Modelle zurück, welche ihren Ursprung in der Neurowissenschaft haben. So machten der Neurologe Warren McCulloch und der Mathematiker Walter Pitts 1943 mit dem McCulloch-Pitts-Neuron einen frühen Versuch, die Gehirnfunktion mithilfe eines linearen Modells zu modellieren [43]. 1958 beschrieb der Psychologe Frank Rosenblatt in seinem Aufsatz „The perceptron: A probabilistic model for information storage and organization in the brain“ [44] erstmals das Perzeptron, welches auch heute noch die Grundlage von künstlichen neuronalen Netzen bildet. Diese und weitere in den 1940-1960er Jahren entwickelte einfache Lernalgorithmen [45] [46] haben die moderne Landschaft des maschinellen Lernens stark beeinflusst.

Nachdem eine 1969 von Marvin Minsky und Seymour Papert veröffentlichte mathematische Analyse die Grenzen des Perzeptron aufzeigte [47], nahm die Popularität der neuronalen Netze jedoch stark ab. Zwar gilt die Neurowissenschaft auch heute noch als wichtige Inspirationsquelle für Deep Learning Forscher, da die Funktionsweise des Gehirns jedoch auch heute noch nicht vollständig entschlüsselt ist, kommt ihr eine weit unbedeutendere Rolle zu. Die aktuelle Forschung zu Deep Learning lässt sich von vielen Bereichen inspirieren. Hierzu zählen insbesondere die Grundlagen der angewandten Mathematik (z. B. lineare Algebra, Wahrscheinlichkeitsrechnung, Informationstheorie und numerische Optimierung). [48]

Die zweite Forschungswelle auf dem Gebiet der neuronalen Netze begann in den 1980er Jahren. Im Kontext der Kognitionswissenschaft entstand der sogenannte Konnektionismus. Dieser basiert auf der zentralen Idee, dass eine große Anzahl einfacher Recheneinheiten intelligentes Verhalten erreichen kann, wenn sie miteinander vernetzt werden und führte zur Entdeckung mehrerer Schlüsselkonzepte, welche noch heute von zentraler Bedeutung sind [48]. Neben dem Konzept der verteilten Repräsentation (Distributed Representation) [49] wurden in dieser Zeit zum ersten Mal die Backpropagation [50] [51], sowie das LSTM-Netzwerk (Long Short-Term Memory) [52] angewendet. Mit der Programmierung eines Algorithmus für die Erkennung von handgeschriebenen Zahlen gelang es Yann Le Cun 1989 erstmals ein Deep Learning Modell mit Backpropagation anzuwenden, mit dessen Hilfe die Postleitzahlen des amerikanischen Postservices erkannt und verarbeitet werden

konnten [53]. Die zweite Welle der Forschung an neuronalen Netzen endete Mitte der 1990er Jahre, da die Forschung die unrealistischen Erwartungen ihrer Investoren nicht erfüllen konnte und daraufhin weitere Investitionen ausblieben [48]. Gleichzeitig wurden Fortschritte in anderen Bereichen des maschinellen Lernens erzielt, sodass sich der Fokus der Forschung verlagerte [48].

Die dritte Forschungswelle zu neuronalen Netzen begann 2006 mit einem Durchbruch von Geoffrey Hinton. Dieser nutzte die „Greedy Layer-Wise Pretraining“-Strategie um ein tiefes neuronales Netz zu trainieren und erzielte damit sehr gute Ergebnisse [54]. Auch andere Forschergruppen wiesen nach, dass die vorgestellte Strategie zu einem effizienten Trainieren verschiedener anderer tiefer neuronaler Netze angewendet werden kann und ihre Leistungsfähigkeit verbessert [55] [56]. Mit diesen Erkenntnissen begannen tiefe neuronale Netze konkurrierende KI-Systeme, welche auf anderen Technologien des Machine Learning beruhen, zu übertreffen. Aufgrund der großen Bedeutung der Tiefe von neuronalen Netzen etablierte sich zu dieser Zeit der Begriff Deep Learning [48].

Ein wichtiger Grund für die steigende Leistungsfähigkeit von Deep Learning in den letzten Jahren liegt in dem wachsenden Zugang zu sehr großen Datenmengen. Darüber hinaus haben sich die uns zur Verfügung stehenden Rechenressourcen sehr stark ausgeweitet, sodass deutlich größere neuronale Netze möglich sind. Diese sind in der Lage, bei komplexeren Aufgaben eine höhere Genauigkeit zu erreichen. Seit den 1980er Jahren hat sich die Erkennungs- und Vorhersagegenauigkeit daher stetig verbessert und Deep Learning fand in immer breiteren Anwendungsbereichen Verwendung. Während die ersten Deep Learning Modelle zur Erkennung einzelner Objekte in eng geschnittenen, extrem kleinen Bildern verwendet wurden, hat sich die Größe der Bilder sowie die Anzahl erkennbarer Objekte schrittweise erhöht. Inzwischen können moderne Netzwerke in der Regel mindestens 1.000 verschiedene Kategorien von Objekten auf hochauflösenden Fotos erkennen. [48]

Auch die Spracherkennung konnte durch den Einsatz von Deep Learning deutlich verbessert werden. Während die Fehlerquoten ab etwa 2000 stagnierten, führte die Einführung von Deep Learning zu einem deutlichen Rückgang. Teilweise konnten die Fehlerquoten um die Hälfte reduziert werden [57]–[59]. Weitere Anwendungsfelder liegen in der Erkennung von Fußgängern [60], der Segmentierung von Bildern [61] sowie der Klassifizierung von Verkehrsschildern [62]. Aufgrund ihrer Leistungsfähigkeit werden Deep Learning Modelle inzwischen außerdem vermehrt in der Wissenschaft [63], Medizin [64] [65] und Industrie [66] [67] eingesetzt.

2.2.3 Grundlegende Funktionsweise

Die Grundlage von Deep Learning stellen künstliche Neuronale Netze (KNN) dar. Diese sind den neuronalen Strukturen des menschlichen Gehirns nachempfunden und bestehen aus mehreren Recheneinheiten, die miteinander verknüpft sind. Diese künstlichen Neuronen erhalten jeweils eine gewichtete Summe aller mit ihnen verknüpften, vorangestellten Neuronen als Eingabewert und geben diesen ebenfalls gewichtet an alle nachfolgenden

Neuronen weiter, siehe Abbildung 2-8. Die künstlichen Neuronen werden auch als Knoten bezeichnet. Durch ihre Verknüpfung entstehen neben einer Eingabe und Ausgabeschicht (Input/ Output Layer) mehrere Zwischenschichten, welche als verborgene Schichten (Hidden Layer) bezeichnet werden. Dabei können beliebig viele Eingabeknoten, Hidden-Layer-Knoten und Ausgabeknoten miteinander kombiniert werden, um die Komplexität des neuronalen Netzes zu erhöhen.

Prinzipiell gilt, dass die Anzahl der Ausgabeknoten der Anzahl der zu detektierenden Klassen entspricht. Kommen sehr viele verborgene Schichten zum Einsatz, wird das Netzwerk als Deep Neural Network bezeichnet.

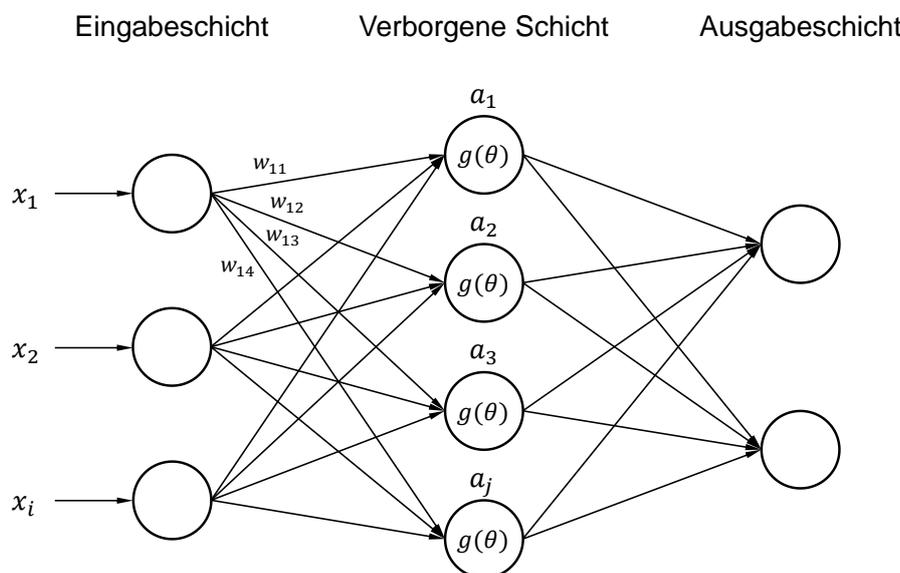


Abbildung 2-8: Beispiel für ein künstliches neuronales Netz

Die Funktionswerte a_j der Knoten in den Zwischenschichten ergeben sich aus den Eingabewerten x_i der vorgelagerten Knoten, welche jeweils mit dem zugehörigen Gewicht w_{ij} multipliziert, einem Bias b_j addiert und für alle verbundenen Eingabeknoten aufsummiert werden, siehe Gl. 1.

$$a_j = g\left(\sum_{i=1}^N w_{ij}x_i + b_j\right) \quad \text{Gl. 1}$$

Während das Gewicht bestimmt, inwieweit eine vorgelagerte Aktivierung verstärkt oder abgeschwächt wird, beeinflusst der Bias-Wert das Gewicht der einzelnen Neuronen im Gesamtkontext. Die korrekte Auswahl beider Größen hat maßgeblichen Einfluss auf die Leistungsfähigkeit eines neuronalen Netzes. Um die Werte zu bestimmen, gibt es eine Vielzahl verschiedener Methoden. Zu den weit verbreiteten Ansätzen zählt der stochastische Gradientenabstieg (Stochastical Gradient Descent (SGD)).

Ein Knoten gibt seinen Funktionswert nur dann an die nachfolgenden Neuronen weiter, wenn dieser einen definierten Schwellenwert übersteigt. Dieser Schwellenwert wird durch die Aktivierungsfunktion g bestimmt. In der Literatur wird dazu eine Vielzahl möglicher Funktionen beschrieben. Eine in neuronalen Netzen viel genutzte Aktivierungsfunktion ist die Rectified Linear Unit (ReLU).

Die optimalen Parameter werden im Zuge des Trainingsprozesses gelernt. Sollen visuelle Daten verarbeitet werden, wird dazu ein Trainingsdatensatz genutzt, welcher Bilder der zu identifizierenden Klassen enthält. Diese Daten werden in das neuronale Netz eingegeben, welches eine Vorhersage zu den abgebildeten Objektklassen trifft. Es wird ermittelt, wie stark die einzelnen Knoten im Netz aktiviert wurden. Durch den Abgleich dieser Vorhersage mit der tatsächlichen Objektklasse ergibt sich ein Fehler, welcher durch die Verlustfunktion berechnet wird. Anschließend misst der Algorithmus ausgehend von der Ausgangserschicht, wie groß der Einfluss eines jeden Knoten auf den Fehler war. Auf Grundlage dieser Berechnung werden die Parameter angepasst, bis der Fehler minimal ist. Je nach Anzahl und Komplexität der Objektklassen können dazu mehr oder weniger viele Iterations-schleifen notwendig sein. Dieser Prozess wird als Backpropagation bezeichnet [53].

Die bisherige Betrachtung beschränkt sich auf eindimensionale Knoten, welche einfache Zahlenwerte weitergeben. Im Rahmen der Bilderkennung werden dabei die Pixelwerte als Eingabe herangezogen, welche in der Regel für jedes Bild durch einen langen, eindimensionalen Vektor abgebildet werden. Daraus resultieren einige Schwierigkeiten. Zum einen ergeben die Bilder des gleichen Objektes an anderer Position auch andere Eingabevektoren, obwohl die Objekte gleich sind. Außerdem ist bei hochauflösenden Bildern eine sehr große Anzahl von Eingabeneuronen und damit eine Vielzahl von Gewichten und Biases notwendig. Dies wäre rechentechnisch nur sehr aufwändig handhabbar.

Um die beschriebenen Probleme zu lösen, werden gefaltete neuronale Netze (Convolutional Neural Network (CNN)) eingesetzt. Die Knoten eines CNNs können dreidimensionale Daten (Breite, Höhe und Farbtiefe) in Form von Vektoren halten [68]. Durch Faltungsoperationen werden die Bilddaten verdichtet, sodass sich die Anzahl der Parameter und der damit einhergehende Rechenaufwand reduziert. Um die Bilddaten zu verdichten, wird das gewichtete Mittel umliegender Pixel ermittelt und repräsentativ für die ursprüngliche Information angenommen. Um Objekte auf Bildern zu klassifizieren, nutzen CNNs sogenannte Features. Unter dem Begriff Feature werden Merkmale eines Objektes verstanden, durch welche dieses identifiziert werden kann. Während die ersten Schichten in einem CNN einfache Features wie Ecken oder Kanten erkennen, werden diese von Schicht zu Schicht komplexer, siehe Abbildung 2-9.

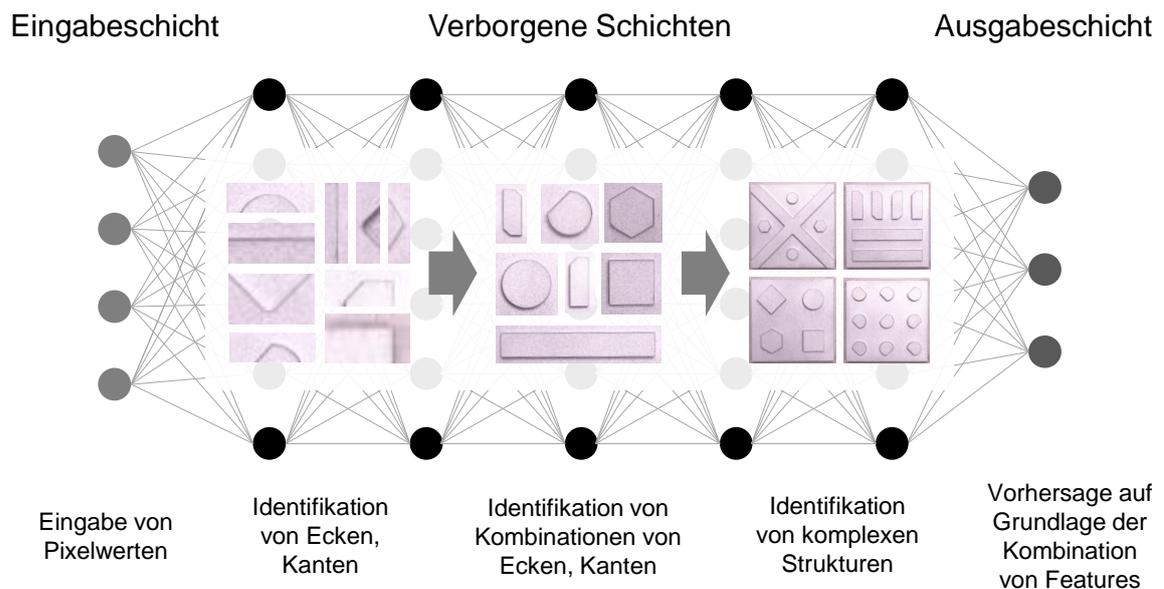


Abbildung 2-9: Feature-Hierarchie in CNNs

Darüber hinaus werden in CNNs verschiedene, spezialisierte Schichten eingesetzt. Diese sind unterschiedlich aufgebaut und dienen verschiedenen Funktionen, welche in Tabelle 2-1 erläutert werden. Durch eine unterschiedliche Kombination dieser Schichten können CNNs mit verschiedenen Eigenschaften gebildet werden.

Tabelle 2-1: Layer in CNNs

Typ	Beschreibung
Input Layer	Enthält die Pixelwerte des Eingabebildes.
Convolutional Layer (CL)	Verdichtet die Bildinformationen, indem verschiedene gewichtete Vektoren (=Filter) in definierter Schrittgröße über den Vektor aus Eingabewerten fahren und Skalarprodukte bilden und zusammenführen. Ergebnis ist ein Vektor, welcher neue Dimensionen in Breite, Höhe und Tiefe enthält, die sogenannte Feature Map.
Pooling Layer (PL)	Verkleinert die entstandenen Feature Maps weiter, um die Anzahl der Parameter und die damit notwendige Rechenleistung zu reduzieren. Im Vergleich zur CL enthält sie jedoch keine Gewichte.
Activation Layer (AL)	Enthalten Aktivierungsfunktionen, um dem Netzwerk einen nicht-linearen Anteil zu geben. Hierbei können verschiedene Aktivierungsfunktionen genutzt werden.
Fully Connected Layer (FCL)	Werden meist zum Ende eines CNN für die Klassifizierung genutzt und entsprechen einer CL mit einem Filter der Größe 1x1. Die Anzahl der Ausgangsneuronen im FCL entspricht der Anzahl der Klassen, welche detektiert werden sollen.

2.2.4 Mean Average Precision

Um die Erkennungsgenauigkeit von Objekterkennungsalgorithmen zu messen, gibt es eine Vielzahl verschiedener Möglichkeiten. In der Regel besteht das Ergebnis einer Objekterkennung aus einer Reihe von Bounding-Boxes sowie den zugehörigen Klassen und Konfidenzniveaus, siehe Abbildung 2-10. Das Ausgabe-Format variiert jedoch je nach eingesetztem Algorithmus. Auch die großen Wettbewerbe zur Objektdetektion nutzen verschiedene Metriken, um die Algorithmen der antretenden Teams zu bewerten [69] - [71].

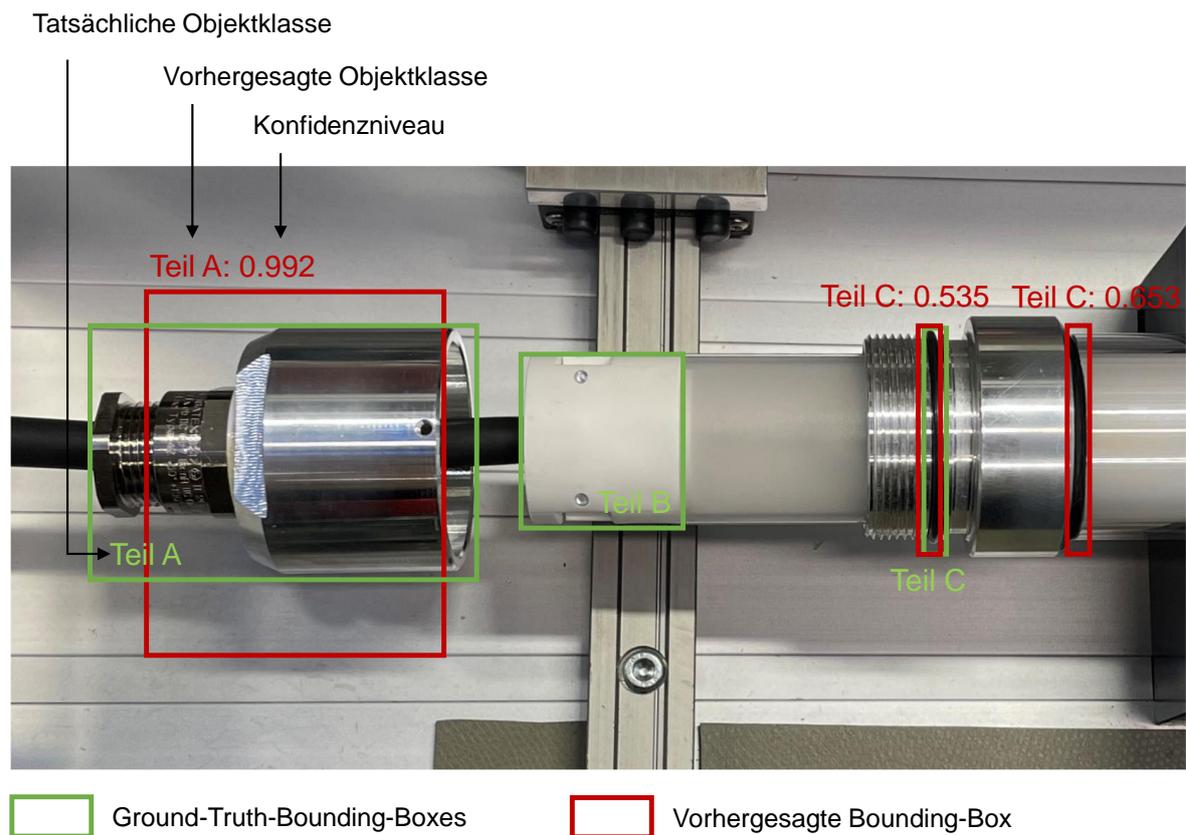


Abbildung 2-10: Beispiel einer Vorhersage durch einen Objekterkennungsalgorithmus

Die gängigste Metrik um die Genauigkeit von Objekterkennungsalgorithmen zu messen ist jedoch die Mean Average Precision (mAP), wobei diese ebenfalls auf verschiedenen Wegen berechnet werden kann [72]. Alle Berechnungsmethoden basieren dabei auf der gleichen Grundlage. Zunächst werden die vom Algorithmus bestimmten Bounding-Boxes mit den für den Validierungsdatensatz manuell definierten Ground-Truth-Bounding-Boxes verglichen und in die in Tabelle 2-2 dargestellten Kategorien eingeteilt.

Tabelle 2-2: Definitionen: True Positive, False Positive und False Negative

Kategorie	Erläuterung
True Positive	Korrekte Erkennung und Platzierung der Erkennung eines vorhandenen Objektes (Vorhergesagte Bounding-Box = Ground-Truth-Bounding-Box)
False Positive	Falsche Erkennung eines nicht vorhandenen Objektes oder falsch platzierte Erkennung eines vorhandenen Objektes
False Negative	Fehlende Erkennung eines vorhandenen Objektes (Nicht erkannte Ground-Truth-Bounding-Box)

Als Auswahlkriterium für eine korrekte Erkennung wird die Intersection over Union (IoU) verwendet. Diese misst die Überlappungsfläche zwischen der vorhergesagten Bounding-Box und der Ground-Truth-Bounding-Box (area of overlap) und setzt diese mit der gemeinsamen Fläche (area of union) ins Verhältnis, siehe Abbildung 2-11.

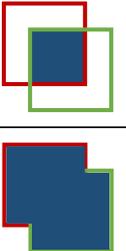
$$IoU = \frac{\text{area of overlap}}{\text{area of union}} = \frac{\text{Diagramm 1}}{\text{Diagramm 2}}$$


Abbildung 2-11: Intersection over Union

Durch den Vergleich der IoU mit einem zuvor definierten Schwellenwert t werden die vorhergesagten Bounding-Boxes als True Positive (TP) oder False Positive (FP) eingestuft. Wenn $IoU \geq t$ dann gilt die Vorhersage als TP, bei $IoU < t$ als FP. Wurde keine Bounding-Box vorhergesagt, obwohl eine Ground-Truth-Bounding-Box für den Validierungsdatensatz definiert wurde, zählt diese als False Negative (FN).

Auf Grundlage dieser Einteilung werden im Anschluss die Precision (P) und der Recall (R) gemäß Gl. 2 und Gl. 3 ermittelt. Während die Precision den Prozentsatz der richtigen positiven Vorhersagen darstellt und die Fähigkeit eines Modells beschreibt, nur relevante Objekte zu detektieren, beschreibt der Recall die Fähigkeit eines Modells, alle relevanten Objekte zu finden. Er ist der Prozentsatz der korrekten positiven Vorhersagen unter allen gegebenen Grundwahrheiten.

$$P = \frac{TP}{TP + FP} = \frac{TP}{\text{Alle Vorhersagen}} \quad \text{Gl. 2}$$

$$R = \frac{TP}{TP + FN} = \frac{TP}{\text{Alle Ground Truths}} \quad \text{Gl. 3}$$

Um die mAP zu erhalten, werden schlussendlich alle vorhergesagten Bounding-Boxes nach absteigenden Konfidenzwerten sortiert und die Anzahl der TP- und FP-Vorhersagen kumuliert. Unter Nutzung dieser kumulierten Werte werden dann in der entsprechenden Reihenfolge für jede Vorhersage Precision und Recall ermittelt und gegeneinander aufgetragen, sodass sich eine Precision-Recall-Kurve ergibt. Tabelle 2-3 und Abbildung 2-12 zeigen dies beispielhaft für die Vorhersagen aus Abbildung 2-10. Diese Betrachtung beschränkt sich auf die Vorhersagen aus einem Bild und dient nur als Erläuterungsbeispiel. In der Regel weisen Validierungsdatensätze eine deutlich größere Anzahl Bilder auf, sodass die Vorhersagen verschiedener Bilder geordnet nach ihrem Konfidenzniveau in der Tabelle zusammengefasst werden und Tabelle sowie Kurve deutlich umfangreicher sind.

Tabelle 2-3: Precision und Recall für Vorhersagen aus Abbildung 2-10

Vorhersage	Konfidenzniveau	TP	FP	TP (kumuliert)	FP (kumuliert)	FN	P	R
Teil A	0,992	1	0	1	0	2	1	$0, \bar{3}$
Teil C (1)	0,653	0	1	1	1	2	0,5	$0, \bar{3}$
Teil C (2)	0,535	1	0	2	1	1	$0, \bar{6}$	$0, \bar{6}$

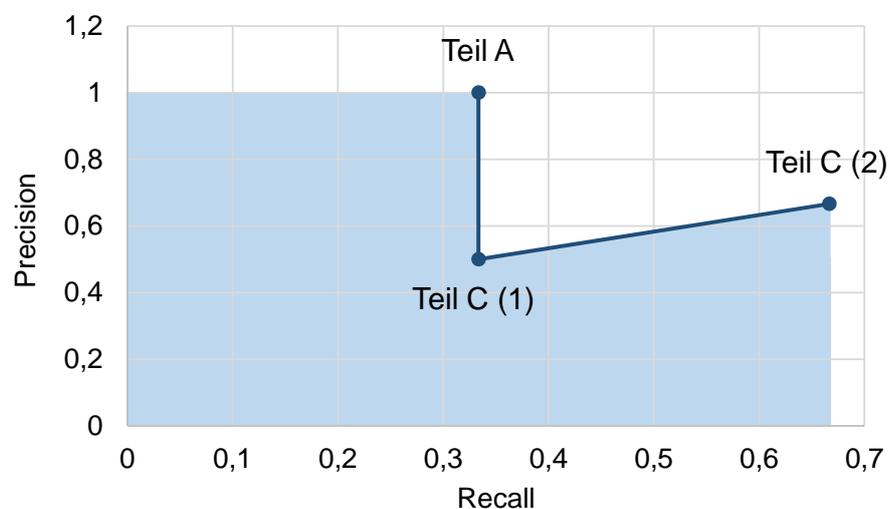


Abbildung 2-12: Precision-Recall-Kurve für Vorhersagen aus Abbildung 2-10

Um die Erkennungsgenauigkeiten von verschiedenen Algorithmen und Modellen anhand von einer einzelnen numerischen Zahl miteinander vergleichen zu können, wird im nächsten Schritt die Fläche unter der Precision-Recall-Kurve ermittelt. Diese stellt die Average Precision (AP) dar. Sie kann für alle Objektklassen erhoben werden und ergibt gemittelt die mAP.

Um die Fläche zu approximieren gibt es verschiedene Herangehensweisen, die je nach verwendetem Algorithmus und zwischen den Wettbewerben variieren und die mAP damit beeinflussen können. Gängige Methoden die Fläche unter der Kurve zu bestimmen sind die 11-Punkte-Interpolation sowie die Alle-Punkte-Interpolation. Darüber hinaus ist die resultierende mAP abhängig von dem für die IoU definierten Schwellenwert. Die mAP bei 50 % IoU (mAP@0.5) ist dabei in jedem Fall höher als oder gleich der mAP bei 95 % (mAP@0.95), da hier weniger oder maximal gleich viele Objekte als Positiv erkannt werden. Eine gängige Herangehensweise ist es, die mAP für alle Werte der IoU zwischen 50 % und 95 % in 5 % Schrittgröße zu ermitteln und den Durchschnitt zu bilden. Diese Metrik wird als mAP@[.50:.95] bezeichnet und – falls nicht anders gekennzeichnet – im Rahmen der vorliegenden Arbeit als mAP verwendet. [72]

3 Stand der Technik

3.1 Aktuelle Entwicklungen in der montagegerechten Produktgestaltung

Eine umfangreiche Literaturrecherche zu aktuellen Entwicklungen im Bereich der montagegerechten Produktgestaltung zeigt, dass über die in Kapitel 2 beschriebenen Grundlagen hinaus nur sehr begrenzt neue Ansätze zur montagegerechten Gestaltung von Produkten thematisiert werden.

So beziehen sich viele der identifizierten Paper auf die in Abschnitt 2.1.3 beschriebene DFMA-Methode von Boothroyd, Dewhurst und Knight [35]. Neben der Anwendung der Methode auf Fallbeispiele werden dabei verschiedene Optimierungspotenziale aufgezeigt. So wenden Salikan et al. [73] die DFMA-Methode auf ein Gartengerät an, während Butt et al. [74] aufzeigen, wie sich das Verfahren auf die Material- und Herstellkosten eines Förderbandsystems auswirkt. Eine Optimierung der Methode erfolgt durch Mora et al. [75]. Dabei entwickeln die Autoren ein auf der DFMA-Methode basierendes Verfahren für Produkte mit großen Abmessungen. Darüber hinaus besteht eine Vielzahl weiterer Veröffentlichungen zur Anwendung oder Optimierung der DFMA-Methode (z. B. [76] - [78]).

Ein weiterer aktueller Forschungsansatz besteht durch den Einsatz von digitalen Hilfsmitteln zur Optimierung der Montagegerechtheit. Bspw. erstellten Bouissiere et al. [79] ein mathematisches Modell, um kritische Fehler und Beurteilungen in Bezug auf den Produktzusammenbau zu erkennen. Als Anwendungsbeispiel werden dabei Flugzeugteile genutzt. Weitere Veröffentlichungen zum Einsatz von digitalen Hilfsmitteln für eine montagegerechte Produktgestaltung finden sich in [80] und [81]. Im Vergleich zur Anwendung der DFMA-Methode zeigt die Literaturrecherche eine deutlich geringere Anzahl Veröffentlichungen zu diesem Forschungsansatz.

Abschließend verdeutlicht die Recherche, dass die aktuellen Interessenschwerpunkte auf allgemeinen, nicht auf die Montagegerechtheit bezogenen Themen der Produktentwicklung liegen. Zwei große Trends stellen die agile Produktentwicklung sowie der Einsatz von digitalen Zwillingen dar. Alleine im deutschsprachigen Raum bestehen mannigfaltig Veröffentlichungen zur agilen Produktentwicklung (z. B. [82] - [86]) sowie zu digitalen Zwillingen (z. B. [87] - [89]).

3.2 Deep Learning in der Montage

Wie in Kapitel 2.2.2 beschrieben, werden Anwendungen, welche auf Deep Learning basieren, inzwischen in sehr vielen Bereichen genutzt oder ihre Einsatzpotenziale untersucht. Eine Literaturrecherche in der digitalen Bibliothek IEEE Xplore [90] zeigt, dass die Forschungsintensität im Bereich des Deep Learning seit 2016 sehr stark angestiegen ist, siehe Abbildung 3-1. Das Institute of Electrical and Electronics Engineers (IEEE) ist die weltweit größte technische Fachgesellschaft für die Gebiete Elektrik, Elektronik und Informatik sowie die verwandten Bereiche der Wissenschaft und Technologie [91] und veröffentlichte in den Jahren 2019 bis 2021 jährlich mehr als 5 Millionen Dokumente [92].

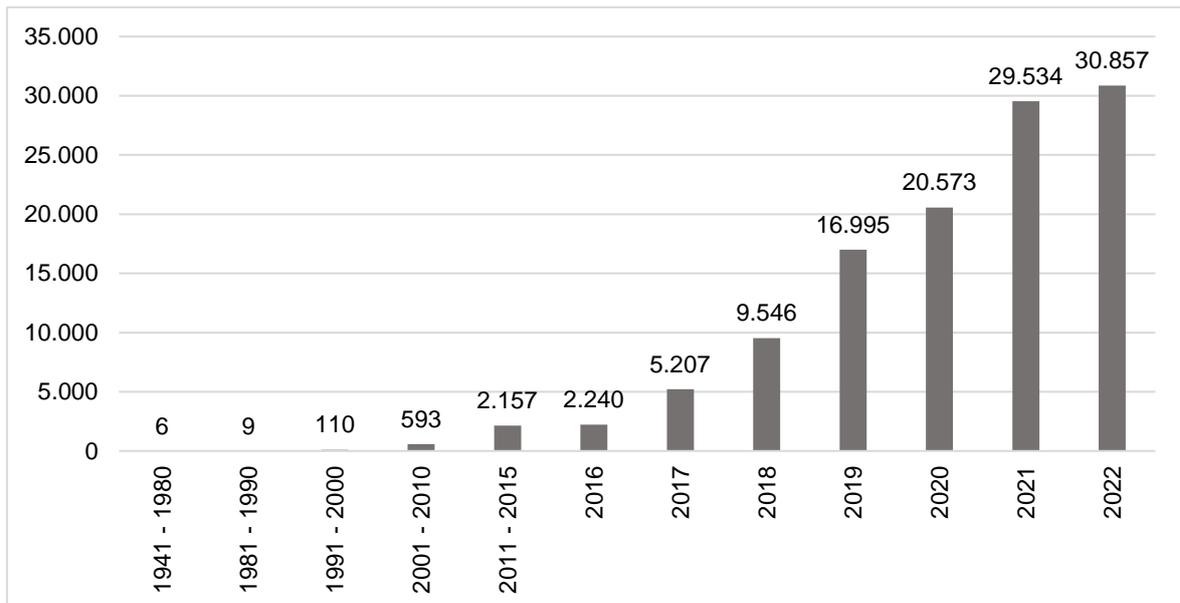


Abbildung 3-1: Anzahl der Suchergebnisse zum Suchbegriff "Deep Learning" in der Datenbank IEEE Xplore zum Stichtag 01.01.2023 im zeitlichen Verlauf

Die Erweiterung der Recherche auf typische Einsatzgebiete von Deep Learning zeigt außerdem, dass ein großer Forschungsschwerpunkt im Bereich der Medizin liegt, siehe Abbildung 3-2. Wenn auch die Forschungsintensität im Bereich der industriellen Anwendungen hoch ist, finden sich vergleichsweise wenig Veröffentlichungen zu Deep Learning in der Montage. Dennoch gibt es auch hier eine nicht zu vernachlässigende Anzahl von Forschungsansätzen.

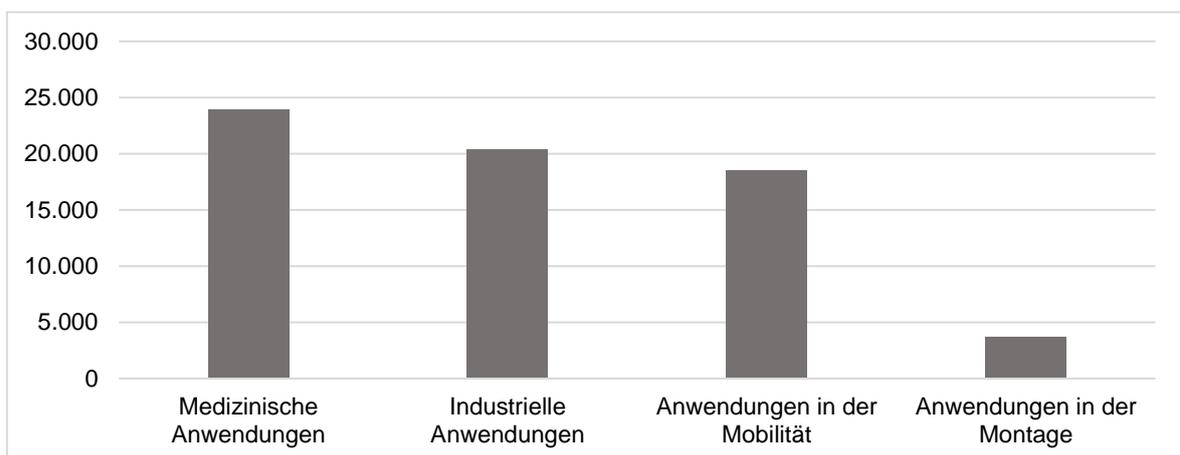


Abbildung 3-2: Anzahl der Suchergebnisse zu einzelnen Themen unter dem Überbegriff "Deep Learning" in der Datenbank IEEE Xplore zum Stichtag 01.01.2023 – Als Suchbegriffe wurden verwendet: Deep Learning AND (Medical OR Medicine OR Diseases OR Diagnostics OR Diagnosis), Deep Learning AND (Industrial OR Industry OR Manufacture OR Manufacturing OR Production OR Produce OR Fabrication OR Fabricate), Deep Learning AND (Mobility OR Cars OR Vehicle OR Automotive OR Driving OR Automobile OR Vessel), Deep Learning AND (Assembly OR Assembling OR Fitting OR Montage OR Installation OR Mounting)

Die Recherche nach konkreten Anwendungsbeispielen zeigt, dass Deep Learning bereits in diversen technischen Anwendungen für die manuelle Montage zum Einsatz kommt. So bietet das Unternehmen Kimoknow eine Augmented-Reality-Brille an, welche Bauteile durch den Einsatz von KI erkennt und die Mitarbeiter durch die Montage führt. Dabei werden Montagefehler automatisch detektiert und entsprechende Warnhinweise gegeben. Um den Implementierungsaufwand zu reduzieren, wird die KI auf Basis der CAD-Daten trainiert. [93]

Um Montagefehler zu reduzieren, forscht auch die Firma Bosch an einem Montageassistenzsystem mit integrierter Objekterkennung, welche auf dem Einsatz von Deep Learning basiert. Durch eine intelligente Kamera am Montagearbeitsplatz überprüft das System parallel zur Fertigung, ob der Montageprozess richtig abläuft und gibt entsprechende digitale Arbeitsanweisungen. Um den Trainingsprozess zu erleichtern, wird eine intelligente Bildauswertung eingesetzt, welche bereits während der Bildaufnahme des entsprechenden Bauteils die relevanten Bilder herausfiltert und diese automatisch annotiert. Aktuell wird das Montageassistenzsystem prototypisch in Zusammenarbeit mit der Mercedes-Benz Group AG am Beispiel einer Batteriemontage erprobt. [94]

Ein weiteres Beispiel für die Anwendung von Deep Learning zur Optimierung der Montage ist das Projekt NeuroCAD. Durch ein intelligentes CAD-Bauteile-Assessment auf Basis von KI soll die wirtschaftliche Montage einer Konstruktion ermittelt und relevante Bereiche markiert werden, um montagerelevante Eigenschaften bereits während der Konstruktionsphase zu bewerten [95]. Die vorgestellten Beispiele sind nur ein Auszug der im Rahmen der Recherche ermittelten Anwendungen. Insgesamt zeigt sich, dass es eine Vielzahl von Implementierungsansätzen gibt, diese jedoch größtenteils noch nicht über die Erprobungsphase hinaus gehen.

4 Anforderungen an die Produktgestaltung

4.1 Einflussfaktoren

4.1.1 Bilder als Abfolge von Pixel-Werten

Wie in Kapitel 2.2.3 beschrieben, nutzen Deep Learning Algorithmen zur Objekterkennung sogenannte Features, durch welche sich die Objekte einer Klasse auszeichnen. Zu diesen zählen z. B. Richtungsgradienten, Ecken, bestimmte geometrische Formen, Farben, etc. Für das Erlernen der Features werden Bilder bereitgestellt, für welche die Position und Klasse der abgebildeten Objekte bekannt sind. Auf Grundlage dieser Bilder werden im Rahmen des Trainingsprozesses objektklassenspezifische Features abgeleitet, welche für die anschließende Objektdetektion genutzt werden. Die Qualität eines Deep Learning Algorithmus wird demnach maßgeblich davon bestimmt, inwieweit sich die Features aus den Trainingsdaten ableiten lassen. Dies wird wiederum entscheidend durch die Art und Weise beeinflusst, wie Computer Bilder verarbeiten.

Für Computer bestehen Bilder aus einer Reihe von Pixel-Werten. Bei Schwarz-Weiß (SW)-Bildern hält jeder Pixel einen Wert (siehe Abbildung 4-1), während Farb-Bilder drei Werte (Rot-Grün-Blau) je Pixel aufweisen. Um Informationen aus den Pixel-Werten eines Bildes abzuleiten, werden unter anderem sogenannte Filter eingesetzt. Dabei gibt es eine Vielzahl verschiedener Filter, welche zu unterschiedlichen Zwecken eingesetzt werden. Prinzipiell nutzen Filter sogenannte Kernels, mit welchen ein Bild gefaltet wird. Dazu wird der Kernel Stück für Stück über das Originalbild geschoben und führt Berechnungen mit den benachbarten Pixeln durch.

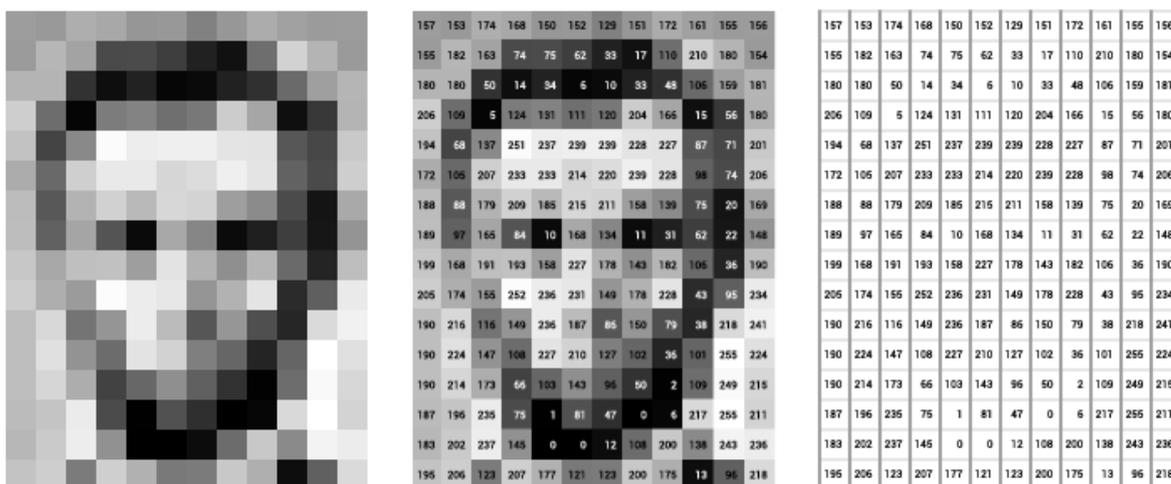


Abbildung 4-1: Darstellung eines SW-Bildes durch Pixelwerte [96]

Ein bekannter Filter, welcher für die Detektion von Kanten eingesetzt wird, ist bspw. der Prewitt-Operator. Dieser nutzt zwei 3x3-Kernels, um Pixelunterschiede in horizontaler und vertikaler Richtung zu identifizieren, siehe Gl. 4 und Gl. 5. Um G_x bzw. G_y zu berechnen, wird der entsprechende Kernel über das Eingabebild bewegt und die Pixelwerte werden

neu berechnet, siehe Abbildung 4-2. Um horizontale und vertikale Kanten kombiniert zu betrachten, können die beiden Gradientenbilder G_x und G_y anschließend durch Gl. 6 zur Kantenstärke G zusammengefasst werden. Abbildung 4-3 zeigt das Bild eines Gebäudes, auf welchem die horizontalen und vertikalen Kanten durch die Anwendung des Prewitt-Operators identifiziert wurden. [97]

$$G_x = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} * A \quad \text{Gl. 4}$$

$$G_y = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} * A \quad \text{Gl. 5}$$

$$G = \sqrt{G_x^2 + G_y^2} \quad \text{Gl. 6}$$

mit

G_x = Gradientenbild für horizontale Kanten

G_y = Gradientenbild für vertikale Kanten

A = Pixel-Matrix des Originalbildes

G = Kantenstärke

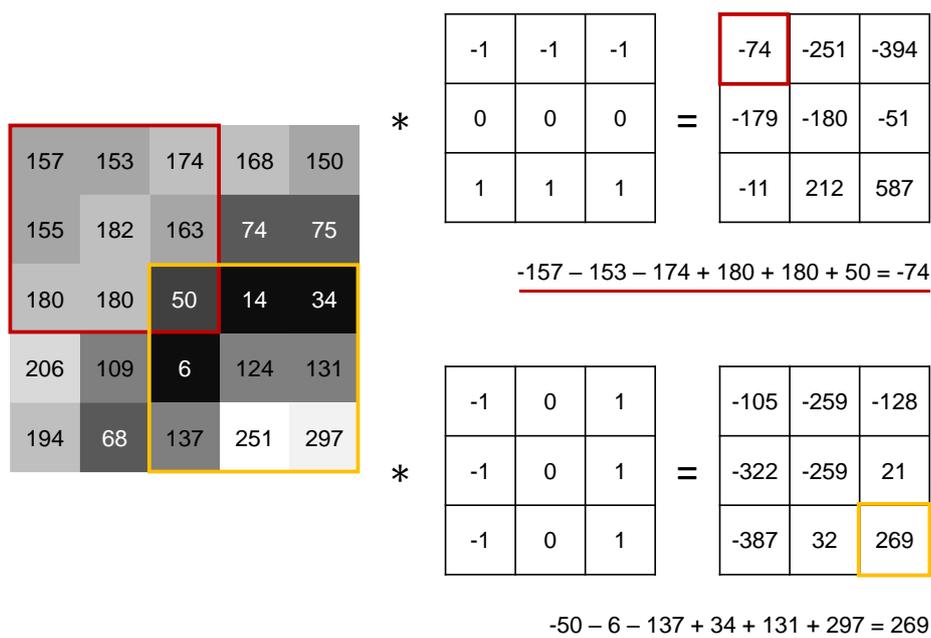


Abbildung 4-2: Gradientenbilder G_x und G_y nach Anwendung des Prewitt-Operators

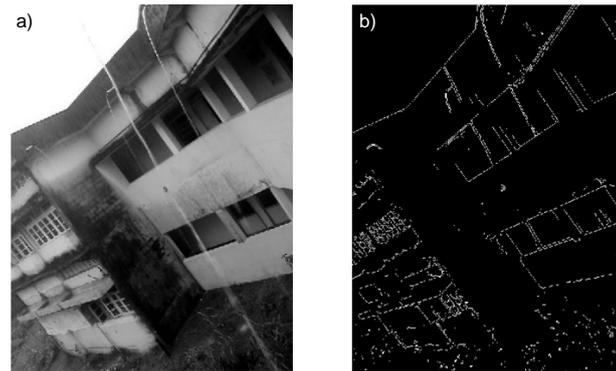


Abbildung 4-3: SW-Bild eines Hauses a) Original b) Nach Prewitt-Operator (nach [97])

Neben dem vorgestellten Filter gibt es eine Vielzahl weiterer Filter, Bilddeskriptoren und darauf aufbauende Methoden, welche zur Detektion verschiedener Bildmerkmale eingesetzt werden können. In der Regel basieren diese jedoch alle auf dem Vergleich benachbarter Pixel, wie im Beispiel dargestellt. Um eine hohe Erkennungsgenauigkeit zu erzielen, sollte bei der Produktgestaltung dementsprechend darauf geachtet werden, dass das optische Erscheinungsbild der Produkte zu Pixelwerten führt, in welchen Muster leicht erkannt werden. Hypothesen, welche diesen Aspekt berücksichtigen, werden in Kapitel 4.2 abgeleitet.

4.1.2 Interklassenvarianz

Weitere Einflussfaktoren auf die Performance von Deep Learning Algorithmen zur Objekterkennung bestehen durch die Intra- und Interklassenvarianz. Während die Intra-Klassenvarianz beschreibt, wie sehr sich die Objekte innerhalb einer Klasse unterscheiden, gibt die Interklassenvarianz an, in welchem Ausmaß die verschiedenen Klassen variieren, siehe Abbildung 4-4. Prinzipiell führt eine niedrige Intra-Klassenvarianz bei gleichzeitig hoher Interklassenvarianz dazu, dass Objekte gut detektiert werden können.

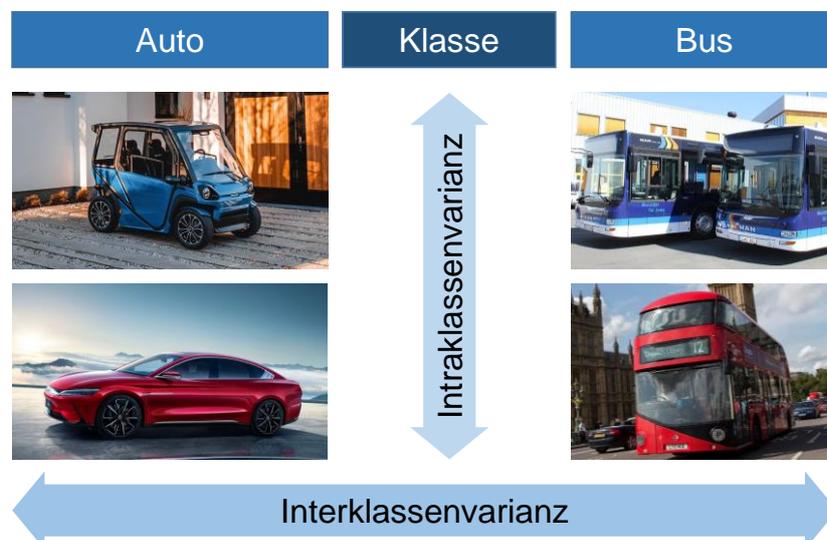


Abbildung 4-4: Intra- und Interklassenvarianz (unter Verwendung von [98] - [101])

Typische Datensätze, auf welche Deep Learning Algorithmen angewendet werden, enthalten Objekte mit einer hohen Intraklassenvarianz. So umfassen die weltweit bekannten Datensätze COCO [102] und ImageNet [103] unter anderem Objektklassen wie Menschen sowie verschiedene Tier- und Pflanzenarten. Diese unterscheiden sich stark von Individuum zu Individuum, woraus eine hohe Intraklassenvarianz resultiert. Im Vergleich dazu weisen technische Gebilde wie Bauteile, Normteile und Baugruppen eine sehr geringe Intraklassenvarianz auf. In der Regel zeichnen diese sich durch fest definierte Eigenschaften wie z. B. ihre Maße, Werkstoffe oder Oberflächengüte aus. Da im Rahmen der vorliegenden Arbeit ausschließlich technische Gebilde untersucht werden, kann die Intraklassenvarianz aus den weiteren Betrachtungen ausgenommen werden.

Eine hohe Interklassenvarianz ist hingegen in technischen Gebilden nicht zwangsläufig sichergestellt. Bauteile, Normteile und Baugruppen können sehr ähnliche Eigenschaften aufweisen. Weil daraus eine niedrige Interklassenvarianz resultiert, welche einen negativen Einfluss auf die Performance von Objekterkennungsalgorithmen hat, wird dieser Aspekt im Rahmen der Hypothesenableitung in Kapitel 4.2 berücksichtigt.

4.2 Hypothesen

Unter Beachtung der zuvor erläuterten Einflussfaktoren werden im vorliegenden Abschnitt Hypothesen bezüglich der optimalen Produktgestaltung aufgestellt, welche den Anforderungen an wissenschaftliche Hypothesen entsprechen. Zu diesen Anforderungen zählen neben der empirischen Überprüfbarkeit auch die Falsifizierbarkeit, der hinreichende Informationsgehalt, der logische Aufbau, die Präzision und Eindeutigkeit sowie die theoretische Fundierung [104].

Prinzipiell wird zwischen Unterschieds- und Zusammenhangshypothesen, ungerichteten und gerichteten Hypothesen sowie unspezifischen und spezifischen Hypothesen unterschieden. Während Unterschiedshypothesen einen Unterschied zwischen zwei oder mehreren Bedingungen formulieren, werden Zusammenhangshypothesen genutzt, um die Abhängigkeit zwischen zwei oder mehreren Variablen auszudrücken. Beide Formen lassen sich ungerichtet und gerichtet formulieren. Wenn es keine begründete Annahme über die Richtung des Unterschieds oder Zusammenhangs gibt, ist eine ungerichtete Hypothese anzugeben (*Ein Objekt wird mit einer anderen Genauigkeit detektiert, wenn ...*). Liegen jedoch Gründe für die Annahme einer Richtung vor, besteht die Möglichkeit eine gerichtete Hypothese zu formulieren (*Ein Objekt wird mit einer höheren Genauigkeit detektiert, wenn ...*). Eine spezifische Hypothese liegt vor, wenn ein exakter Unterschied zwischen zwei Variablen angegeben wird. Dies kann sowohl ungerichtet als auch gerichtet erfolgen (*Die mAP eines Modells unterscheidet sich um einen Wert von 3, wenn ... vs. Die mAP eines Modells liegt um einen Wert von 3 höher, wenn ...*). Kann kein exakter Unterschied angegeben werden, ist eine unspezifische Hypothese zu formulieren. Im Rahmen der vorliegenden Arbeit werden verschiedene Hypothesenarten genutzt, siehe Abschnitte 4.2.1 bis 4.2.4. [105]

Da die Gültigkeit einer Hypothese statistisch niemals bewiesen, sondern ausschließlich verworfen werden kann, ist jeweils eine Null- und Alternativhypothese aufzustellen. Als Nullhypothese wird dabei in der Regel das Gegenteil des zu beweisenden Effektes angenommen, während das zu zeigende Ergebnis als Alternative zur Null (=Alternativhypothese) angenommen wird. Lässt sich die Nullhypothese statistisch signifikant verwerfen, kann die Alternativhypothese angenommen werden. Die Nullhypothese wird demnach nur aufgestellt, um verworfen zu werden. Es ist jedoch zu beachten, dass eine nicht-verworfenen Nullhypothese nicht angenommen werden kann. [106]

Um Null- und Alternativhypothesen zur optimalen Produktgestaltung für den Einsatz von Deep Learning in der Montage aufzustellen, werden sowohl konventionelle als auch intuitiv betonte Methoden genutzt. Zunächst werden Hypothesen aus der Analyse der in Abschnitt 4.1 vorgestellten Einflussfaktoren abgeleitet. Um diese zu ergänzen, werden die Ergebnisse eines Brainstormings mit vier Teilnehmenden aus unterschiedlichen Fachgebieten herangezogen. Im Rahmen des Brainstormings erhielten die Teilnehmer eine kurze Einführung in die Funktionsweise von Objekterkennungsalgorithmen auf Basis von Deep Learning. Um die Mitwirkenden nicht zu beeinflussen, wurden die im Vorfeld aufgestellten Hypothesen nicht bekannt gegeben.

Für eine bessere Übersichtlichkeit werden die Hypothesen anschließend verschiedenen Teilaspekten der Produktgestaltung zugeordnet. Neben der Bauteilgestaltung (Abschnitt 4.2.1) werden die Baugruppungsgestaltung (Abschnitt 4.2.2), die Auswahl von Normteilen (Abschnitt 4.2.3) sowie die Dimensionierung von Formelementen (Abschnitt 4.2.4) im Folgenden getrennt betrachtet.

4.2.1 Hypothesen zur Bauteilgestaltung

Hypothesen, welche sich auf die allgemeine Gestaltung von Bauteilen beziehen, werden unter dem Teilaspekt der Bauteilgestaltung zusammengefasst. Dabei ist zwischen fünf gerichteten, unspezifischen Unterschiedshypothesen zu unterscheiden, siehe Tabelle 4-1.

Tabelle 4-1: Null- und Alternativhypothesen zur Bauteilgestaltung

Nr.	Nullhypothese	Alternativhypothese
BT1	<p>Bauteile, welche spiegelnde Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut detektiert, wie Bauteile, welche matte Oberflächen aufweisen.</p> $mAP_{BT-Spiegelnd} \geq mAP_{BT-Matt}$	<p>Bauteile, welche spiegelnde Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter detektiert, als Bauteile, welche matte Oberflächen aufweisen.</p> $mAP_{BT-Spiegelnd} < mAP_{BT-Matt}$
BT2	<p>Bauteile, welche dunkle Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut detektiert, wie Bauteile, welche helle Oberflächen aufweisen.</p> $mAP_{BT-Dunkel} \geq mAP_{BT-Hell}$	<p>Bauteile, welche dunkle Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter detektiert, als Bauteile, welche helle Oberflächen aufweisen.</p> $mAP_{BT-Dunkel} < mAP_{BT-Hell}$
BT3	<p>Bauteile, welche weiche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut detektiert, wie Bauteile, welche harte Konturen aufweisen.</p> $mAP_{BT-Weich} \geq mAP_{BT-Hart}$	<p>Bauteile, welche weiche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter detektiert, als Bauteile, welche harte Konturen aufweisen.</p> $mAP_{BT-Weich} < mAP_{BT-Hart}$
BT4	<p>Bauteile, welche ähnliche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut detektiert, wie Bauteile, welche sich hinsichtlich Ihrer Konturen nicht ähneln.</p> $mAP_{BT-Ähnlich} \geq mAP_{BT-NichtÄhnlich}$	<p>Bauteile, welche ähnliche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter detektiert, als Bauteile, welche sich hinsichtlich Ihrer Konturen nicht ähneln.</p> $mAP_{BT-Ähnlich} < mAP_{BT-NichtÄhnlich}$
BT5	<p>Bauteile, welche komplexe Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut detektiert, wie Bauteile, welche simple Konturen aufweisen.</p> $mAP_{BT-Komplex} \geq mAP_{BT-Simpel}$	<p>Bauteile, welche komplexe Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter detektiert, als Bauteile, welche simple Konturen aufweisen.</p> $mAP_{BT-Komplex} < mAP_{BT-Simpel}$

Wie in Abschnitt 4.1.1 erläutert, basieren Objekterkennungsalgorithmen in der Regel auf dem Vergleich benachbarter Pixel. Um eine hohe Erkennungsgenauigkeit zu erzielen, sollten Produkte so gestaltet werden, dass Muster über den Vergleich benachbarter Pixel abgeleitet werden können. Um Kanten, Ecken oder andere Muster zu erkennen, müssen sich die Werte benachbarter Pixel jedoch in ausreichendem Maß unterscheiden. Die Hypothesen BT1, BT2 und BT3 greifen diesen Aspekt auf.

So besteht die Vermutung, dass Bauteile, welche spiegelnde Oberflächen aufweisen, durch einen auf Deep Learning basierenden Objekterkennungsalgorithmus schlechter detektiert werden, als Bauteile mit matten Oberflächen, da durch die Spiegelung Ecken, Kanten und andere Muster optisch weniger ausgeprägt wahrnehmbar sein könnten. Gleiches könnte für Bauteile gelten, welche dunkle Oberflächen aufweisen. Da Licht durch dunkle Oberflächen stärker absorbiert wird als durch helle Oberflächen, könnten Konturen in diesem Fall weniger sichtbar sein, sodass Muster schlechter erkannt werden. Auch Bauteile, welche aufgrund von abgerundeten Kanten weiche Konturen ausweisen, könnten aus den zuvor genannten Gründen schlechter detektiert werden, als Bauteile mit harten Konturen.

Abschnitt 4.1.2 erläutert den Einfluss der Interklassenvarianz auf die Performance von Objekterkennungsalgorithmen. Prinzipiell erzielen Objekterkennungsmodelle höhere Genauigkeiten, wenn sich die Objekte unterschiedlicher Klassen klar voneinander unterscheiden lassen. Vor diesem Hintergrund wird die Hypothese formuliert, dass Bauteile, welche ähnliche Konturen aufweisen, besser durch einen Objekterkennungsalgorithmus detektiert werden, als Bauteile mit unterschiedlichen Konturen.

Die fünfte Hypothese zur Bauteilgestaltung basiert auf einer Kombination der beiden beschriebenen Einflussfaktoren. So besteht die Vermutung, dass Bauteile, welche komplexe Konturen aufweisen, durch einen Objekterkennungsalgorithmus schlechter detektiert werden, als Bauteile mit simplen Konturen. Neben den schwerer zu erkennenden Mustern in komplexen Konturen könnte die erschwerte optische Differenzierung zwischen Objekten komplexer Kontur zu einer schlechteren Objektdetektion führen.

Um die Hypothesen zu quantifizieren, wird die mAP genutzt. Wie in Abschnitt 2.2.4 erläutert, wird diese bei den meisten Deep Learning Algorithmen als Maß für die Erkennungsgenauigkeit eines Objekterkennungsalgorithmus herangezogen und soll daher auch in der vorliegenden Arbeit verwendet werden. Prinzipiell besteht durch alle Hypothesen zur Bauteilgestaltung die quantitative Frage, ob die mAP eines Modells mit Objekten einer bestimmten Eigenschaft kleiner als die mAP eines Modells mit Objekten entgegengesetzter Eigenschaften ist. Die Hypothese, dass Bauteile mit spiegelnden Oberflächen, durch einen Objekterkennungsalgorithmus schlechter detektiert werden, als Bauteile, welche matte Oberflächen aufweisen, würde sich bspw. in der in Gl. 7 dargestellten quantitativen Annahme widerspiegeln. Dabei ist jeder Alternativhypothese H_1 eine Nullhypothese H_0 gegenüberzustellen, welche das logische Gegenteil behauptet, siehe Gl. 8.

$$H_1: mAP_{BT-Spiegelnd} < mAP_{BT-Matt} \quad \text{Gl. 7}$$

$$H_0: mAP_{BT-Spiegelnd} \geq mAP_{BT-Matt} \quad \text{Gl. 8}$$

mit

H_1 = Alternativhypothese

H_0 = Nullhypothese

$mAP_{BT-Spiegelnd}$ = Mean Average Precision für ein Objekterkennungsmodell,
welches Objekte mit spiegelnden Oberflächen beinhaltet

$mAP_{BT-Matt}$ = Mean Average Precision für ein Modell,
welches Objekte mit matten Oberflächen beinhaltet

4.2.2 Hypothesen zur Baugruppengestaltung

Neben der Beschaffenheit einzelner Bauteile könnte auch die Kombination von Einzelteilen Einfluss auf die Erkennbarkeit durch einen auf Deep Learning basierenden Objekterkennungsalgorithmus nehmen. Aus diesem Grund wird auch der Aspekt der Baugruppengestaltung im Rahmen der Hypothesenbildung berücksichtigt. Hierzu wird eine gerichtete, unspezifische Unterschiedshypothese formuliert, siehe Tabelle 4-2.

Tabelle 4-2: Null- und Alternativhypothese zur Baugruppengestaltung

Nr.	Nullhypothese	Alternativhypothese
BG1	Die Einzelteile von Baugruppen, welche das gleiche optische Erscheinungsbild aufweisen, werden durch einen Objekterkennungsalgorithmus mindestens genauso gut unterschieden, wie die Einzelteile von Baugruppen, welche kein gleiches optisches Erscheinungsbild aufweisen. $mAP_{BG-Gleich} \geq mAP_{BG-NichtGleich}$	Die Einzelteile von Baugruppen, welche das gleiche optische Erscheinungsbild aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als die Einzelteile von Baugruppen, welche kein gleiches optisches Erscheinungsbild aufweisen. $mAP_{BG-Gleich} < mAP_{BG-NichtGleich}$

Die Grundlage für die Hypothese zur Baugruppengestaltung bildet der in Abschnitt 4.2.1 erläuterte Einflussfaktor der Mustererkennung aus dem Vergleich benachbarter Pixel. Bestehen die Einzelteile in Baugruppen aus dem gleichen Material bzw. Materialien, welche das gleiche optische Erscheinungsbild aufweisen, könnten sie aufgrund der geringen Unterschiede schlechter voneinander differenziert werden, als wenn sich ihre Materialien deutlich unterscheiden. Aus diesem Grund besteht die Vermutung, dass die Einzelteile von Baugruppen durch einen Objekterkennungsalgorithmus schlechter detektiert werden, wenn sie das gleiche optische Erscheinungsbild aufweisen, als wenn sich dieses zwischen den Einzelteilen unterscheidet. Unter dem Begriff des optischen Erscheinungsbildes werden dabei Aspekte wie der Werkstoff, die Farbe sowie die Oberflächengüte von Bauteilen zusammengefasst.

Um die Hypothese zur Baugruppengestaltung zu quantifizieren, wird analog zur Bauteilgestaltung (siehe Abschnitt 4.2.1) die mAP herangezogen. Der Alternativhypothese H_1 aus Gl. 9 wird die Nullhypothese H_0 aus Gl. 10 gegenübergestellt.

$$H_1: mAP_{BG-Gleich} < mAP_{BG-NichtGleich} \quad \text{Gl. 9}$$

$$H_0: mAP_{BG-Gleich} \geq mAP_{BG-NichtGleich} \quad \text{Gl. 10}$$

mit

$mAP_{BG-Gleich}$ = Mean Average Precision für ein Objekterkennungsmodell, welches Baugruppen aus Einzelteilen mit gleichem optischen Erscheinungsbild beinhaltet

$mAP_{BG-NichtGleich}$ = Mean Average Precision für ein Objekterkennungsmodell, welches Baugruppen aus Einzelteilen mit unterschiedlichen optischen Erscheinungsbildern beinhaltet

4.2.3 Hypothesen zur Auswahl von Normteilen

Auch Hypothesen, welche sich auf die Auswahl von Normteilen beziehen, werden unter einem eigenen Teilaspekt zusammengefasst. Zwar sind Normteile den Einzelteilen zuzuordnen, anders als die Hypothesen zur Bauteilgestaltung beziehen sich die Hypothesen hierbei jedoch nicht auf ihre konstruktive Gestaltung, sondern auf ihre Auswahl. Die Gestalt der Normteile ist vorgegeben und daher für die vorliegende Untersuchung nicht relevant. Insgesamt werden zwei ungerichtete, spezifische Zusammenhangshypothesen zur Auswahl von Normteilen aufgestellt, siehe Tabelle 4-3.

Tabelle 4-3: Null- und Alternativhypothesen zur Auswahl von Normteilen

Nr.	Nullhypothese	Alternativhypothese
N1	<p>Normteile vom gleichen Typ werden nicht verwechselt, wenn sich ihre Größe nur geringfügig unterscheidet.</p> <p>$mAP_{\text{Größe}} \geq 80 \%$</p>	<p>Normteile vom gleichen Typ werden verwechselt, wenn sich ihre Größe nur geringfügig unterscheidet.</p> <p>$mAP_{\text{Größe}} < 80 \%$</p>
N2	<p>Schrauben, welche einen ähnlichen Schraubenkopf aufweisen, werden nicht verwechselt.</p> <p>$mAP_{\text{Schraubenkopf}} \geq 80 \%$</p>	<p>Schrauben, welche einen ähnlichen Schraubenkopf aufweisen, werden verwechselt.</p> <p>$mAP_{\text{Schraubenkopf}} < 80 \%$</p>

Beide Hypothesen lassen sich aus dem in Abschnitt 4.2.1 erläuterten Einflussfaktor der Interklassenvarianz ableiten. Da eine geringe Interklassenvarianz in der Regel zu einer niedrigeren Performance von Objekterkennungsalgorithmen führt, ist anzunehmen, dass Normteile nicht zuverlässig detektiert werden, wenn sie eine zu große Ähnlichkeit aufweisen. Basierend auf dieser Annahme besteht die Vermutung, dass Normteile vom gleichen Typ verwechselt werden, wenn sich ihre Größe nur geringfügig unterscheidet. Da Schrauben als Normteile sehr große Anwendung finden, wird eine weitere, auf Schrauben beschränkte Hypothese aufgestellt. So wird vermutet, dass Schrauben, welche einen ähnlichen Kopf aufweisen, durch Objekterkennungsalgorithmen verwechselt werden.

Um die beiden Hypothesen zur Auswahl von Normteilen zu quantifizieren, wird analog zur Bauteil- und Baugruppengestaltung (siehe Abschnitte 4.2.1 und 4.2.2) die mAP herangezogen. Im Vergleich zur Bauteil- und Baugruppengestaltung handelt es sich bei den zwei Hypothesen jedoch nicht um Unterschieds-, sondern um Zusammenhangshypothesen. Aus diesem Grund werden keine mAP-Werte miteinander verglichen, sondern die erzielte mAP einem definierten Grenzwert gegenübergestellt.

Als Grenzwert ist dabei ein Wert auszuwählen, welcher eine zuverlässige Detektion von Objekten gewährleistet. Aufgrund der hohen Komplexität und Individualität der Anwendungen von Objekterkennungsalgorithmen wird in der Literatur jedoch kein konkreter Mindestwert für die mAP definiert. Je nach Anwendungsfall kann eine höhere oder niedrigere mAP für eine zuverlässige Objektdetektion ausreichen. Um dennoch einen Grenzwert für die mAP auszuwählen, wurden verschiedene Anwendungsfälle betrachtet.

So nutzten Iqbal et al. verschiedene, auf Deep Learning basierende Objekterkennungsalgorithmen, um Treibgut zu detektieren, welches hydraulische Strukturen und Durchlässe blockiert. Dabei erzielten sie eine maximale mAP von 84 %. Die Anwendbarkeit des Modells wurde durch zwei Fallbeispiele nachgewiesen. [107]

Saleem et al. trainierten verschiedene Objekterkennungsalgorithmen, um kranke Pflanzen zu detektieren. Mit einer maximal erreichten mAP von 73,07 % konnten dabei alle betrachteten Klassen gesunder und kranker Blätter identifiziert werden. [108]

Insgesamt wurden sieben verschiedene Forschungsprojekte herangezogen, welche eine mAP zwischen 37,0 % und 92,8 % erreichen [107] - [113]. Dabei weisen sechs von sieben Projekten eine mAP von mindestens 72,9 % auf. Als arithmetisches Mittel ergibt sich eine mAP von 75,4 %, während der Median 81,0 % beträgt. Im Rahmen der vorliegenden Arbeit wird eine mAP größer 80 % als ausreichend hoch angenommen.

Unter Berücksichtigung des definierten Grenzwertes ergeben sich beispielhaft für die erste Hypothese zur Auswahl von Normteilen die in Gl. 11 dargestellte Alternativhypothese H_1 sowie die entgegengesetzte Nullhypothese H_0 , siehe Gl. 12.

$$H_1: mAP_{Größe} < 80 \% \quad \text{Gl. 11}$$

$$H_0: mAP_{Größe} \geq 80 \% \quad \text{Gl. 12}$$

mit

$mAP_{Größe}$ = Mean Average Precision für ein Objekterkennungsmodell, welches Normteile des gleichen Typs mit unterschiedlicher Größe beinhaltet

4.2.4 Hypothesen zur Dimensionierung von Formelementen

In der letzten Unterkategorie werden Hypothesen zusammengefasst, welche sich auf die Dimensionierung von Formelementen beziehen. Zwar zählen Formelemente wie Bohrungen, Nuten, Fasen oder Radien ebenfalls zur Bauteilgestalt, anders als die Hypothesen zur Bauteilgestaltung, welche sich auf die allgemeine Gestaltung von Bauteilen beziehen, adressieren die Hypothesen jedoch spezifische Formelemente. Für eine bessere Übersichtlichkeit sind sie einer eigenen Unterkategorie zugeordnet. Insgesamt werden fünf ungerichtete, spezifische Zusammenhangshypothesen zur Dimensionierung von Formelementen formuliert, siehe Tabelle 4-4.

Tabelle 4-4: Null- und Alternativhypothesen zur Dimensionierung von Formelementen

Nr.	Nullhypothese	Alternativhypothese
FE1	Bauteile werden nicht verwechselt, wenn die Maße ihrer Fasen sich nur geringfügig unterscheiden. $mAP_{\text{Fase}} \geq 80 \%$	Bauteile werden verwechselt, wenn die Maße ihrer Fasen sich nur geringfügig unterscheiden. $mAP_{\text{Fase}} < 80 \%$
FE2	Bauteile werden nicht verwechselt, wenn die Maße ihrer Radien sich nur geringfügig unterscheiden. $mAP_{\text{Radius}} \geq 80 \%$	Bauteile werden verwechselt, wenn die Maße ihrer Radien sich nur geringfügig unterscheiden. $mAP_{\text{Radius}} < 80 \%$
FE3	Bauteile werden nicht verwechselt, wenn die Maße ihrer Nuten sich nur geringfügig unterscheiden. $mAP_{\text{Nut}} \geq 80 \%$	Bauteile werden verwechselt, wenn die Maße ihrer Nuten sich nur geringfügig unterscheiden. $mAP_{\text{Nut}} < 80 \%$
FE4	Bauteile werden nicht verwechselt, wenn die Tiefe ihrer Bohrungen sich nur geringfügig unterscheiden. $mAP_{\text{Bohrungstiefe}} \geq 80 \%$	Bauteile werden verwechselt, wenn die Tiefe ihrer Bohrungen sich nur geringfügig unterscheiden. $mAP_{\text{Bohrungstiefe}} < 80 \%$
FE5	Bauteile werden nicht verwechselt, wenn der Durchmesser ihrer Bohrungen sich nur geringfügig unterscheiden. $mAP_{\text{Bohrungsdurchmesser}} \geq 80 \%$	Bauteile werden verwechselt, wenn der Durchmesser ihrer Bohrungen sich nur geringfügig unterscheiden. $mAP_{\text{Bohrungsdurchmesser}} < 80 \%$

Wie die Hypothesen zur Auswahl von Normteilen werden die fünf Hypothesen zur Dimensionierung von Formelementen aus dem in Abschnitt 4.2.1 erläuterten Einflussfaktor der Interklassenvarianz abgeleitet. Auch für Formelemente ist anzunehmen, dass sie nicht zuverlässig detektiert werden, wenn sie eine zu große Ähnlichkeit aufweisen, da eine geringe Interklassenvarianz in der Regel zu einer niedrigeren Performance von Objekterkennungsalgorithmen führt. Basierend auf dieser Annahme wird vermutet, dass Bauteile verwechselt werden, wenn sich die in ihnen enthaltenen Formelemente nur geringfügig unterscheiden. Dabei werden neben Fasen auch Radien, Nuten und Bohrungen betrachtet. Je zu untersuchendem Formelement wird mindestens eine Hypothese formuliert. Da Bohrungen bei gleichem Durchmesser unterschiedliche Tiefen aufweisen können, werden beide Bohrungsmerkmale durch jeweils eine Hypothese berücksichtigt.

Um die fünf Hypothesen zu quantifizieren, wird analog zu den Hypothesen zur Auswahl von Normteilen die mAP herangezogen und mit dem in Abschnitt 4.2.3 definierten Grenzwert verglichen. Beispielhaft für die erste Hypothese zur Dimensionierung von Formelementen wird die in Gl. 13 dargestellte Alternativhypothese H_1 der in Gl. 14 aufgeführten Nullhypothese H_0 gegenübergestellt.

$$H_1: mAP_{Fase} < 80 \% \quad \text{Gl. 13}$$

$$H_0: mAP_{Fase} \geq 80 \% \quad \text{Gl. 14}$$

mit

mAP_{Fase} = Mean Average Precision für ein Objekterkennungsmodell,
welches Bauteile mit Fasen unterschiedlicher Größe beinhaltet

5 Verifizierung der Hypothesen

5.1 Methodik

Um die Gültigkeit der in Kapitel 4 formulierten Hypothesen zu untersuchen, werden im Rahmen der vorliegenden Arbeit Experimente durchgeführt und statistisch ausgewertet. Im ersten Schritt werden alle zu berücksichtigenden abhängigen und unabhängigen Variablen identifiziert und die zu variierenden Faktoren ausgewählt. Diese bilden die Grundlage für die anschließende Festlegung des Versuchsaufbaus.

Einen wesentlichen Bestandteil stellt die Auswahl des anzuwendenden, auf Deep Learning basierenden, Objekterkennungsalgorithmus dar. Um die Anzahl der durchzuführenden Versuche zu bestimmen, erfolgt anschließend die statistische Versuchsplanung, an welche sich die Versuchsdurchführung sowie die Auswertung der Versuchsergebnisse anschließt. Im Rahmen der Versuchsdurchführungen werden unter anderem den Hypothesen entsprechende Versuchsobjekte konzipiert. Insgesamt werden die folgenden sechs Arbeitsschritte definiert und in den nachfolgenden Abschnitten 5.2 bis 5.7 im Detail erläutert:

- Identifikation der abhängigen und unabhängigen Variablen
- Definition des Versuchsaufbaus
- Auswahl des anzuwendenden Objekterkennungsalgorithmus
- Statistische Versuchsplanung
- Versuchsdurchführung
- Versuchsauswertung

5.2 Identifikation der abhängigen und unabhängigen Variablen

Die Grundlage für die Durchführung von Versuchen zur Überprüfung der formulierten Hypothesen bildet die Festlegung der abhängigen und unabhängigen Variablen. Dabei wird definiert, welche Einflussgrößen im Rahmen der Versuche zu variieren und welche Zielgrößen zu erfassen sind. Die Zielgrößen beschreiben das Ergebnis eines Versuchs und hängen von den Einflussgrößen ab. Sie werden daher als abhängige Variablen bezeichnet.

Im Rahmen der vorliegenden Arbeit wird untersucht, inwieweit verschiedene konstruktive Produkteigenschaften die Genauigkeit und Zuverlässigkeit von auf Deep Learning basierenden Objekterkennungsalgorithmen beeinflussen. Als abhängige Variable ist eine Größe zu wählen, welche die Performance eines Objekterkennungsalgorithmus beschreibt. Um eine statistische Auswertung zu ermöglichen, sollte diese Größe quantitativ erfassbar sein. Wie in Abschnitt 2.2.4 erläutert, wird durch die mAP angegeben, mit welcher Genauigkeit Objekte innerhalb eines Modells korrekt klassifiziert und an der richtigen Position detektiert werden. Die mAP beschreibt diese Genauigkeit durch einen einzelnen Zahlenwert. Sie ist damit statistisch auszuwerten und wird als abhängige Variable definiert.

Als Einflussgrößen werden alle Größen identifiziert, welche Einfluss auf die mAP nehmen. Dabei wird zwischen Steuer- und Störgrößen unterschieden. Während Steuergrößen im Rahmen der Versuchsdurchführung bewusst eingestellt werden können, lassen sich Störgrößen nicht beeinflussen. Der Versuchsaufbau sollte so gestaltet werden, dass der Einfluss von Störgrößen minimal ist. Da Einflussgrößen nicht von anderen Variablen abhängen, werden sie auch als unabhängige Variablen bezeichnet. Im Gegensatz zu abhängigen Variablen müssen sie für eine statistische Auswertung nicht zwingend quantitativ erfassbar sein. Im Rahmen der Versuchsplanung ist festzulegen, welche Einflussgrößen variiert und welche konstant gehalten werden sollen.

Als erste Einflussgröße werden die konstruktiven Produkteigenschaften definiert. Die in Kapitel 4 formulierten Hypothesen beschreiben den Einfluss dieser Eigenschaften auf die Performance von Objekterkennungsalgorithmen. Neben der Farbe von Produkten beziehen sie sich z. B. auf die Form und das Reflexionsvermögen der Objektoberfläche. Um die Hypothesen zu untersuchen, sind die konstruktiven Produkteigenschaften im Rahmen der Versuche zu variieren.

Weiteren Einfluss auf die Zielgröße nimmt der eingesetzte Objekterkennungsalgorithmus. Je nach Aufbau eines Algorithmus unterscheidet sich die Qualität der Objektdetektion, sodass sich unterschiedliche Werte für die mAP ergeben, siehe Abschnitt 5.4. Der Vergleich verschiedener Algorithmen ist durch eine Vielzahl internationaler Publikationen hinreichend untersucht [114] - [117] und soll nicht Teil der vorliegenden Arbeit sein. Da der Einsatz verschiedener Objekterkennungsalgorithmen für die Untersuchung der Hypothesen keine Relevanz hat, wird diese Einflussgröße konstant gehalten. Die Auswahl des anzuwendenden Objekterkennungsalgorithmus wird in Abschnitt 5.4 erläutert.

Eine weitere Einflussgröße stellt die Anzahl der Trainings- und Validierungsbilder dar. Wie in Abschnitt 2.2.3 erläutert, werden auf Deep Learning basierende Objekterkennungsalgorithmen mit einem Trainingsdatensatz bestehend aus Bildern, für welche die Klasse und Position der zu detektierenden Objekte bekannt ist, trainiert. Anschließend wird das trainierte Objekterkennungsmodell auf die Bilder des Validierungsdatensatzes angewendet und die mAP über den Vergleich der Vorhersagen mit den tatsächlichen Klassen und Positionen ermittelt. In der Regel führen größere Trainingsdatensätze dabei zu genaueren Modellen, da somit mehr Informationen über die einzelnen Klassen in die Modelle einfließen. Auch die Modellvalidierung wird mit zunehmender Größe des Validierungsdatensatzes genauer. Schlussendlich nimmt die Anzahl der verwendeten Trainings- und Validierungsbilder Einfluss auf die Zielgröße mAP. Da ihre Variation für die Untersuchung der Hypothesen keine Relevanz hat, wird sie im Sinne der Vergleichbarkeit einzelner Modelle konstant gehalten. Die Festlegung der konkreten Anzahl der Trainings- und Validierungsbilder wird in Abschnitt 5.5.2.1 beschrieben.

Neben der Bilderanzahl, nehmen auch die Perspektiven, aus welcher die Trainings- und Validierungsbilder aufgenommen werden, Einfluss auf die Erkennungsgenauigkeit eines trainierten Modells. Prinzipiell hängt die Qualität der Objekterkennung von der Qualität der

Trainingsbilder ab. Sollen Objekte von allen Seiten detektiert werden, sollte der Trainingsdatensatz entsprechende Aufnahmen enthalten. Da die Perspektiven, aus welcher die Trainings- und Validierungsbilder aufgenommen werden, für die Untersuchung der Hypothesen nicht von Interesse sind, werden sie im Rahmen der Versuche konstant gehalten.

Die Qualität der Trainings- und Validierungsbilder wird weiterhin durch die Auflösung der Bilder sowie die Beleuchtungsintensität während der Aufnahme beeinflusst. Je geringer die Auflösung der Bilder ist, desto weniger genau lassen sich Muster aus dem Vergleich benachbarter Pixel ableiten. Eine ungünstige Beleuchtungsintensität könnte dazu führen, dass Konturen nicht erkannt werden. Beide Größen nehmen damit Einfluss auf die Qualität der Objekterkennung und werden im Rahmen der Versuche konstant gehalten.

Eine weitere Einflussgröße besteht durch die Umgebung, in welcher die Trainings- und Validierungsbilder aufgezeichnet werden. So birgt eine den zu detektierenden Objekten zu ähnliche Produktumgebung die Gefahr, dass im Validierungsprozess Elemente der Umgebung fälschlicherweise als Objekte detektiert werden. Ob der Bildhintergrund einen negativen Einfluss auf die Objekterkennung hat, hängt zwar vom Einzelfall ab. Um die Vergleichbarkeit zwischen verschiedenen Objekterkennungsmodellen zu gewährleisten, wird der Hintergrund im Rahmen der Versuche dennoch neutral und konstant gehalten.

Als letzte Einflussgröße wird die Anzahl der Iterationen herangezogen, mit welcher ein Objekterkennungsmodell trainiert wird. Wie in Abschnitt 2.2.3 erläutert, werden Objekterkennungsalgorithmen über mehrere Iterationsschleifen trainiert, um die Parameter des neuronalen Netzes über die Berechnung einer Verlustfunktion Schritt für Schritt zu optimieren. In der Regel wird dabei die Anzahl der zu trainierenden Iterationsschleifen durch den Anwender vorgegeben. Dabei gilt es, die richtige Anzahl Iterationen auszuwählen, sodass eine Über- und Unteranpassung (Over- und Underfitting) des Modells vermieden wird, siehe Abbildung 5-1.

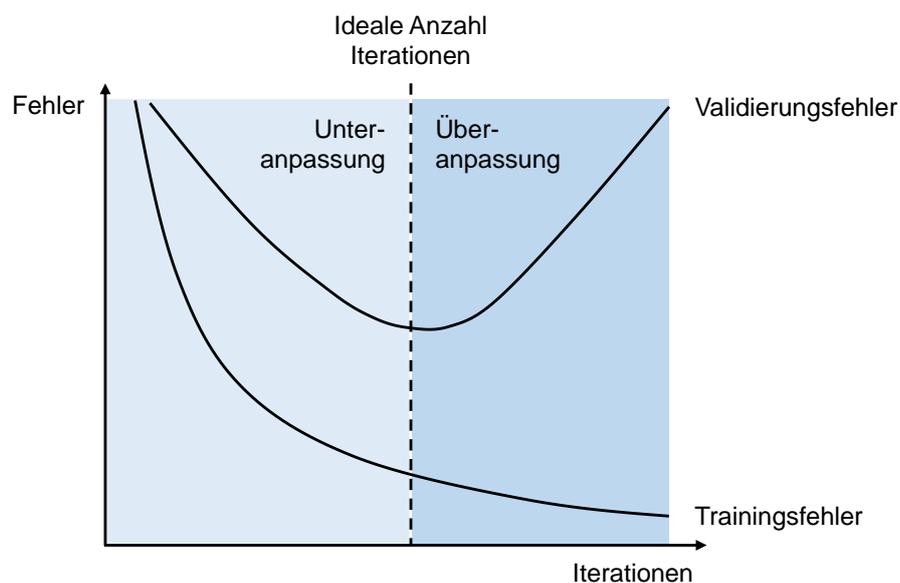


Abbildung 5-1: Über- und Unteranpassung

Erfolgt das Training über eine zu große Anzahl von Iterationen, passen sich Modelle zu sehr an den Trainingsdatensatz an und lassen sich schlecht auf neue, ungesehene Daten anwenden. In diesem Fall wurde das Modell überangepasst, sodass die Bilder des Validierungsdatensatzes nicht mehr zuverlässig detektiert werden können. Eine Unteranpassung tritt hingegen auf, wenn Modelle über eine zu geringe Anzahl Iterationen trainiert werden, sodass die Parameter des neuronalen Netzes noch nicht ausreichend optimiert wurden. Da die Anzahl der Iterationen die Objekterkennungsmodelle und damit die Zielgröße mAP maßgeblich beeinflusst, wird sie im Sinne der Vergleichbarkeit einzelner Modelle konstant gehalten. Die Festlegung der konkreten Iterationsanzahl wird in Abschnitt 5.5.2.2 beschrieben. Tabelle 5-1 zeigt eine Übersicht aller Einflussgrößen.

Tabelle 5-1: Aufstellung der unabhängigen Variablen

Nr.	Einflussgröße	Art	Wert
1	Konstruktive Produkteigenschaften	Steuergröße	Variiert
2	Verwendetes Objekterkennungsmodell	Steuergröße	Konstant
3	Anzahl Trainings- und Validierungsbilder	Steuergröße	Konstant
4	Anzahl Iterationen	Steuergröße	Konstant
5	Perspektive der Trainings- und Validierungsbilder	Steuergröße	Konstant
6	Bildhintergrund	Steuergröße	Konstant
7	Bildqualität	Steuergröße	Konstant
8	Beleuchtungsintensität	Störgröße	Konstant

5.3 Definition des Versuchsaufbaus

Der im Folgenden beschriebene Versuchsaufbau berücksichtigt die in Abschnitt 5.2 identifizierten abhängigen und unabhängigen Variablen und ermöglicht eine Variation der konstruktiven Produkteigenschaften, während die übrigen Einflussgrößen konstant bleiben. Prinzipiell kommen dabei Versuchsobjekte zum Einsatz, welche den in Kapitel 4 aufgestellten Hypothesen entsprechende konstruktive Eigenschaften aufweisen und durch einen Objekterkennungsalgorithmus zu detektieren sind. Je Objekteigenschaft wird ein Modell trainiert, sodass die resultierenden mAPs der Modelle miteinander verglichen werden können. Die Konzipierung der Versuchsobjekte wird in Abschnitt 5.6.1 erläutert.

Grundsätzlich bestehen die Versuche aus zwei Teilen. In einem ersten Schritt erfolgt die Datengenerierung, indem Bilder von den Untersuchungsobjekten erstellt werden. Diese werden in einem zweiten Schritt (Datenauswertung) verwendet, um Objekterkennungsmodelle zu trainieren, siehe Abbildung 5-2.

Während die Erstellung der Trainings- und Validierungsdatensätze eine physische Aktion voraussetzt, findet die Datenauswertung ausschließlich digital statt.

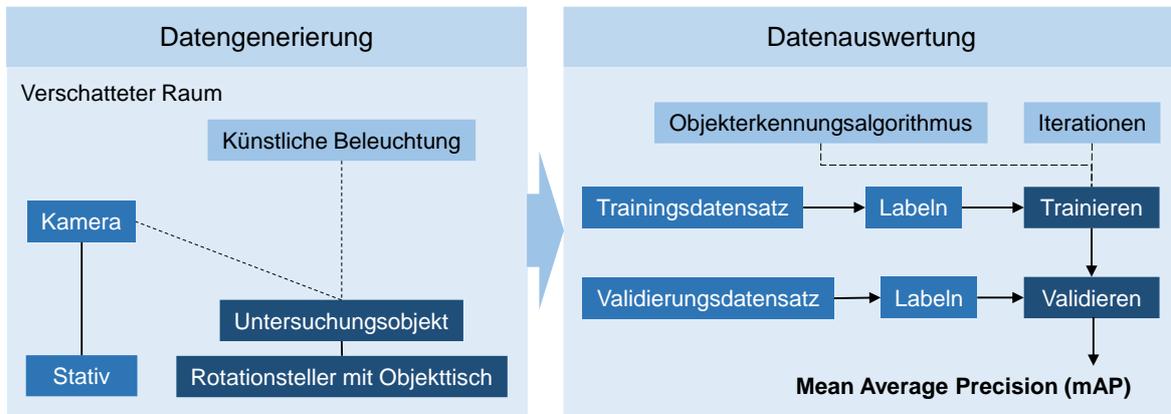


Abbildung 5-2: Versuchsaufbau bestehend aus Datengenerierung und -auswertung

Im Rahmen der Datengenerierung werden Bilder der verschiedenen Versuchsobjekte aufgezeichnet. Um alle Objekte aus der gleichen Perspektive aufzunehmen, wird die Kamera mit Hilfe eines festen Stativs auf diese ausgerichtet. Durch das Stativ ist es außerdem möglich, verschiedene Aufnahmewinkel einzustellen. In der Praxis ist sowohl eine vollständige Rundumsicht – z. B. durch den Einsatz von Smart Glasses – als auch die Beschränkung auf die Draufsicht – z. B. durch eine fest installierte Kamera oberhalb eines Montagearbeitsplatzes – denkbar. Aus diesem Grund wird im Rahmen der vorliegenden Untersuchung neben einer Versuchsreihe mit Bildern aus verschiedenen Perspektiven auch eine Versuchsreihe durchgeführt, welche ausschließlich Bilder aus der Draufsicht beinhaltet.

Damit die Bilder im Rahmen der Datenauswertung digital verarbeitet werden können, wird eine Kamera ausgewählt, welche ein entsprechendes Datenformat ausgibt, siehe Abschnitt 5.6.2. Weiterhin wird durch die Einstellung der Kamera sichergestellt, dass alle Bilder eine konstante Auflösung aufweisen.

Um die Beleuchtungsintensität konstant zu halten, werden die Bilder in einem ausschließlich mit künstlichem Licht beleuchteten Raum aufgezeichnet und die Lichtintensität kontinuierlich überwacht. Damit die Objekte aus allen Richtungen aufgezeichnet werden, ohne die Kamera zu bewegen, werden die Versuchsobjekte auf einem rotierenden Objektisch platziert. Um den Bildhintergrund konstant zu halten, weist der Objektisch eine neurale, gleichbleibende Oberfläche auf. Die im Zuge der Datengenerierung erzeugten Bilder werden anschließend zur Datenauswertung verwendet.

Für diesen Zweck werden sie zunächst in einen Trainings- und einen Validierungsdatensatz aufgeteilt. Dabei wird wie in der Praxis üblich eine Aufteilung von 80 % Trainingsdaten zu 20 % Validierungsdaten gewählt [118] [119]. Anschließend werden beide Datensätze manuell durch Informationen zu den abgebildeten Objektklassen und -positionen ergänzt.

Dieser Prozess wird als Labeln oder Annotation bezeichnet und ist notwendig, da im Trainings- und Validierungsprozess die vorhergesagten Bounding-Boxes mit den tatsächlichen Ground-Truth-Bounding-Boxes abgeglichen werden, siehe Abschnitt 2.2.4.

Sind alle Bilder gelabelt, erfolgt der Trainingsprozess. Dabei wird ein Objekterkennungsalgorithmus mit Hilfe des Trainingsdatensatzes über eine definierte Anzahl Iterationen trainiert. Die Auswahl des anzuwendenden Algorithmus wird im Abschnitt 5.4 erläutert. Die Festlegung der Iterationsanzahl erfolgt in Abschnitt 5.5.2.2. Im Anschluss an den Trainingsprozess wird das entstandene Objekterkennungsmodell auf den Validierungsdatensatz angewendet. Ergebnis des Validierens ist schließlich die mAP, welche als Zielgröße dient. Grundsätzlich besteht jedoch auch die Möglichkeit, das Modell bereits während des Trainingsprozesses in regelmäßigen Iterationsschritten auf den Validierungsdatensatz anzuwenden. Die konkrete Umsetzung des beschriebenen Versuchsaufbaus wird im Rahmen der Versuchsdurchführung in Kapitel 5.6 erläutert.

5.4 Auswahl des Objekterkennungsalgorithmus

Wie in Kapitel 2.2.3 beschrieben, gibt es eine Vielzahl Kombinationsmöglichkeiten der speziellen Schichten eines künstlichen neuronalen Netzes und damit eine große Auswahl verschiedener Objekterkennungsalgorithmen. Prinzipiell können diese in One-Stage und Two-Stage Detektoren eingeteilt werden. Two-Stage Detektoren wie Faster Region-based Convolutional Neural Networks (R-CNN) [120] oder Mask R-CNN [121] verwenden ein Region Proposal Network (RPN), um in der ersten Stufe relevante Bildregionen (Region of Interest (RoI)) zu identifizieren und diese zur Objektklassifizierung und Bounding-Box-Regression weiterzuleiten, siehe Abbildung 5-3. Two-Stage-Detektoren erreichen damit die höchsten Genauigkeitsraten, sind aber in der Regel langsamer. [122]

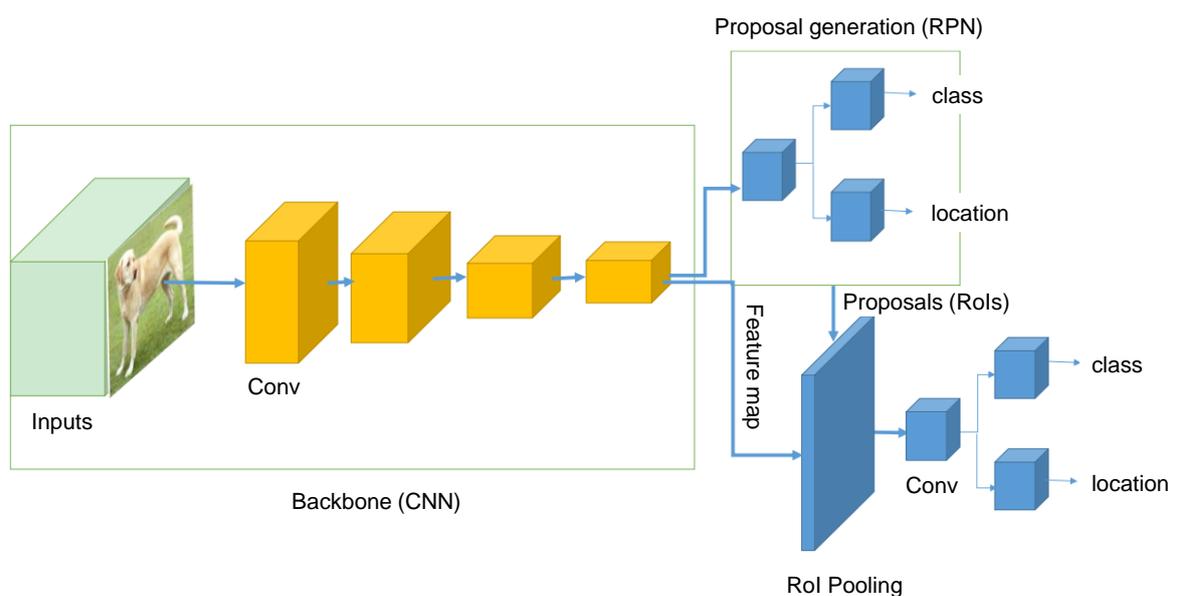


Abbildung 5-3: Grundlegende Architektur von Two-Stage Detektoren (nach [123])

Im Gegensatz dazu behandeln One-Stage Detektoren wie YOLO (You Only Look Once) [124] und SSD (Single Shot Multibox Detector) [125] die Objekterkennung als ein einfaches Regressionsproblem, indem sie ein Eingabebild nehmen und die Klassenwahrscheinlichkeiten und Bounding-Box-Koordinaten lernen, siehe Abbildung 5-4. One-Stage-Detektoren erreichen damit zwar geringere Genauigkeitsraten, sind aber deutlich schneller als zweistufige Objektdetektoren. [122]

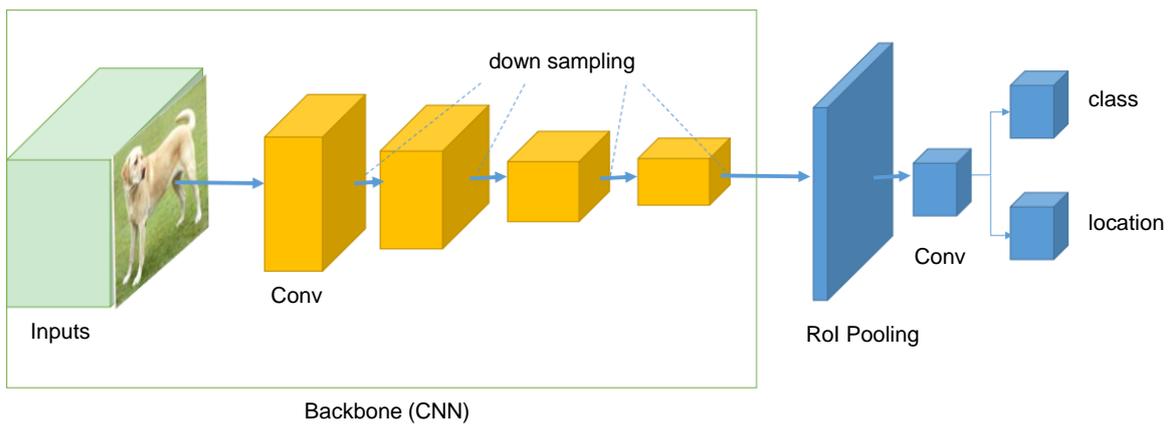


Abbildung 5-4: Grundlegende Architektur von One-Stage Detektoren (nach [123])

Im Rahmen des Forschungsprojektes kommt es insbesondere auf eine hohe Genauigkeit der Vorhersagen an, um auch kleine Unterschiede in der Erkennung von Objekten mit verschiedenen Eigenschaften festzustellen. Das Modell soll nicht in einer Echtzeit-Anwendung zum Einsatz kommen, sodass keine hohe Erkennungsgeschwindigkeit notwendig ist. Aus diesem Grund wird ein Two-Stage Detektor für die vorliegende Untersuchung ausgewählt. Die gängigsten Two-Stage Detektoren sind Faster R-CNN und Mask R-CNN. Während Faster R-CNN lediglich Objektklassen und Bounding-Boxes ausgeben, liefern Mask R-CNN zusätzlich eine Objektmaske, siehe Abbildung 5-5. Mask R-CNN werden zur Instanz-Segmentierung eingesetzt. Da sie jedoch aufgrund der genaueren Objektlokalisierung weitere Rechenressourcen benötigen und die Ausgabe von einfachen Bounding-Boxes für die Anwendung in einem Montageassistenzsystem in der Regel ausreicht, wird in der vorliegenden Untersuchung ein Faster R-CNN genutzt.

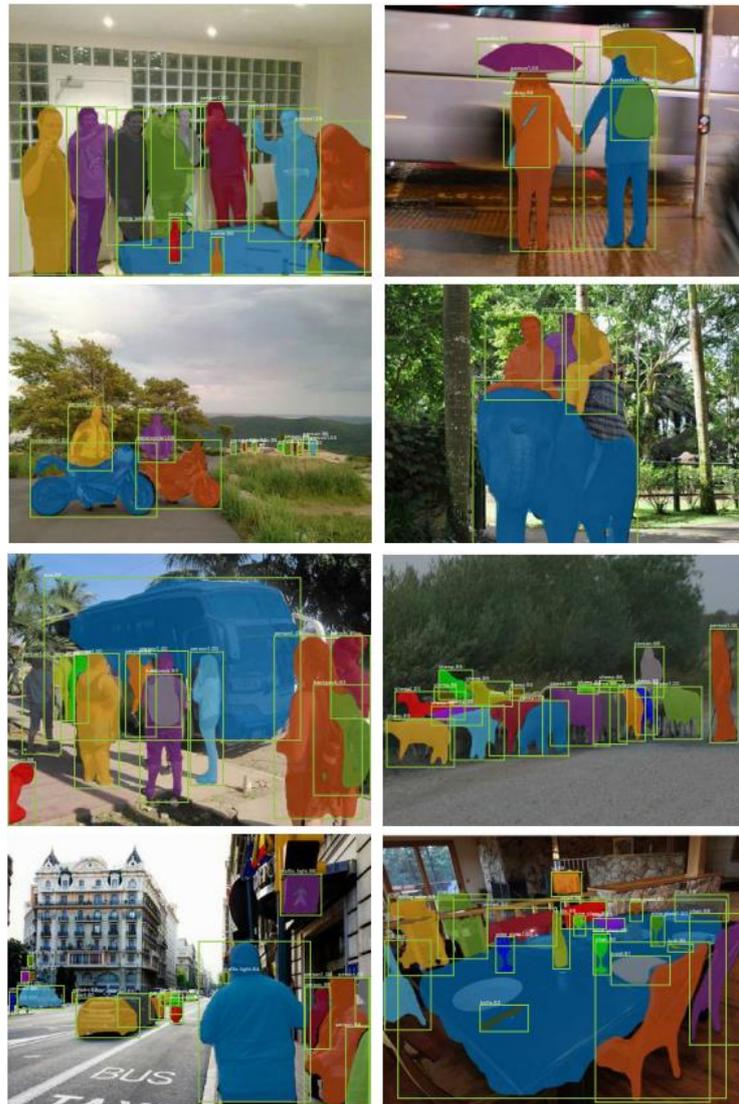


Abbildung 5-5: Vorhersagen eines Mask R-CNN (nach [121])

Da die Erstellung eines neuronalen Netzes sehr zeit- und arbeitsaufwändig ist, wurden mit der Zeit Bibliotheken aufgebaut, welche die wesentlichen Komponenten wie Schichten, Initialisierungsstrategien, Verlustfunktionen, Optimierer, usw. in einer einfach zu bedienenden Schnittstelle zusammenführen. Zwei weit verbreitete, als Open Source frei zur Verfügung stehende Programmiersbibliotheken sind TensorFlow [126] und PyTorch [127]. Ein Vergleich beider Bibliotheken zeigt, dass TensorFlow zwar mehr Programmieroptionen bietet, insgesamt jedoch deutlich komplexer als PyTorch ist [128]. Sie wird daher insbesondere in kommerziellen Anwendungen genutzt. Die Programmiersbibliothek PyTorch zeichnet sich hingegen durch eine höhere Benutzerfreundlichkeit sowie kürzere Trainingszeiten aus [128]. Aus diesen Gründen wird im Forschungsumfeld meist PyTorch eingesetzt und auch für die vorliegende Arbeit genutzt.

Während PyTorch als übergeordnete Programmbibliothek für verschiedene Anwendungen im Bereich des Deep Learning genutzt werden kann, fokussiert sich die Bibliothek Detectron2 auf den Bereich der Computer Vision [129]. Basierend auf PyTorch bietet Detectron2 eine Vielzahl von etablierten Objekterkennungsalgorithmen. Darüber hinaus stellt die Bibliothek vortrainierte Modelle zur Verfügung. Diese wurden im Vorfeld mit dem COCO-Datensatz, welcher ca. 333.000 Bilder, 1,5 Millionen Objektinstanzen und 80 Objektklassen beinhaltet [102], trainiert und können anschließend um weitere Objektklassen erweitert werden. Da die vortrainierten Modelle bereits eine Vielzahl unterschiedlicher Features gelernt haben, können deutlich kleinere Datensätze zur Erweiterung des Modells genutzt werden. Darüber hinaus verkürzt sich die Trainingszeit um ein Vielfaches. [130]

Tabelle 5-2 zeigt die in der Programmbibliothek Detectron2 enthaltenen, auf dem COCO-Datensatz über 37 Epochen vortrainierten Faster R-CNN Modelle. Diese greifen auf verschiedene Backbone-Kombinationen zurück. Dabei steht R50 für die Verwendung eines Residual Network mit 50 Schichten (ResNet-50), R101 für die Verwendung eines ResNet-101 und X101 für die Verwendung eines Next Generation Residual Networks mit ebenfalls 101 Schichten (ResNext-101). Die unterschiedlichen Residual Networks werden wiederum mit verschiedenen Backbones kombiniert (C4, DC5 oder FPN). Da das Modell mit dem Namen R50-FPN die beste Kombination aus hoher AP und geringer Trainingszeit bietet, wird dieses für die vorliegende Untersuchung ausgewählt.

Tabelle 5-2: Vortrainierte Faster R-CNN Modelle in Detectron2 (nach [131])

Name	Trainingszeit [Sekunde/Iteration]	Inferenzzeit [Sekunde/Bild]	Speicher für Training [GB]	AP
R50-C4	0.543	0.104	4.8	38.4
R50-DC5	0.378	0.070	5.0	39.0
R50-FPN	0.209	0.038	3.0	40.2
R101-C4	0.619	0.139	5.9	41.1
R101-DC5	0.452	0.086	6.1	40.6
R101-FPN	0.286	0.051	4.1	42.0
X101-FPN	0.638	0.098	6.7	43.0

Das zur Anwendung kommende Faster R-CNN besteht aus einem ResNet-50 kombiniert mit einem Feature Pyramid Network (FPN) und einem Region Proposal Network (RPN), siehe Abbildung 5-6. ResNets wurden 2015 durch He et. al [132] entdeckt und haben zu einem maßgeblichen Qualitätssprung in der Objekterkennung geführt. Sie sind seitdem zu einem Standard für tiefe CNNs geworden. Im Gegensatz zu klassischen CNNs nutzen ResNets Übersprungsverbindungen (skip connections), um Signale direkt an tiefere Netzwerkschichten weiterzuleiten, ohne von allen Schichten verarbeitet zu werden, siehe

Abbildung 5-7. Dabei entstehen viele kleine Subnetzwerke (Residual Units). Nach jeder Faltung (Convolution) wird außerdem eine Batch Normalization (BN) und die ReLU Aktivierungsfunktion durchgeführt. Durch das Überspringen von Schichten reduziert sich die Komplexität des Netzwerkes. An Stellen, welche eine hohe Komplexität erfordern, kann das Netzwerk aber dennoch auf die notwendigen Schichten zurückgreifen. Das ResNet-50 besitzt fünf Convolutional Layer Blöcke, die sich aus jeweils ca. zehn Convolutional Layern zusammensetzen. Darüber hinaus haben sich ResNets mit 34, 101 und 152 Netzwerkschichten etabliert, wobei das ResNet-50 für Objekterkennungsmethoden die beste Erkennungsleistung bei gleichzeitig akzeptabler Rechendauer erreicht [131].

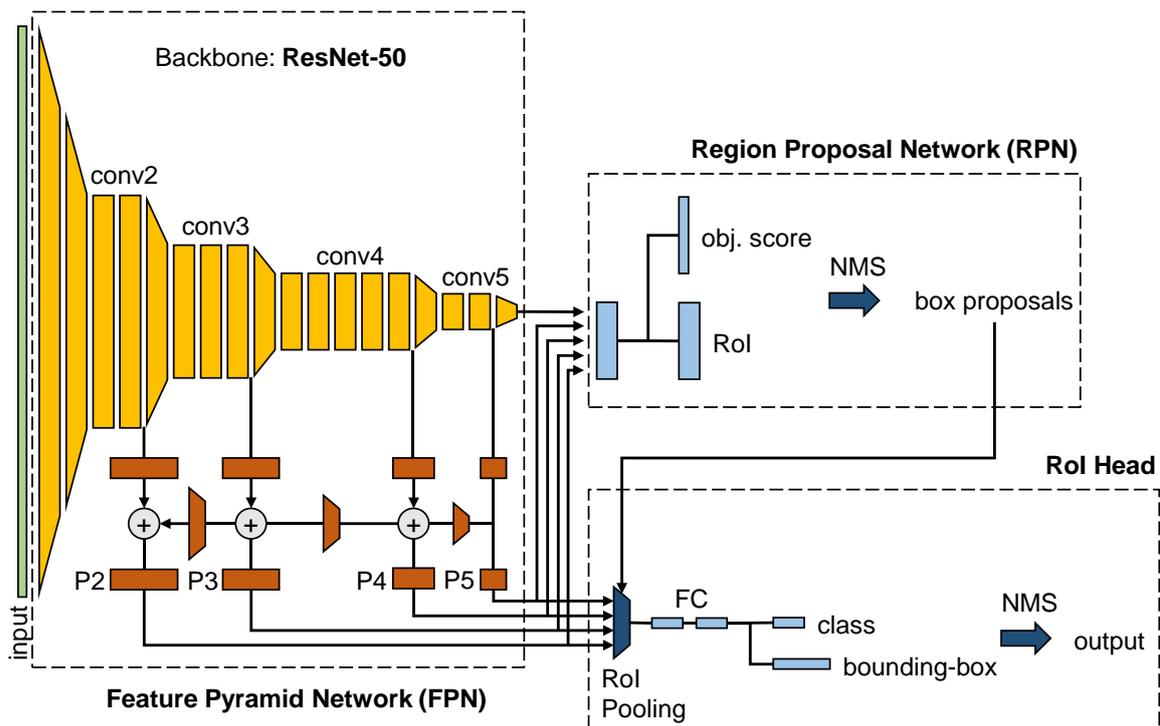


Abbildung 5-6: Architektur des R50-FPN-Netzwerks in Detectron2

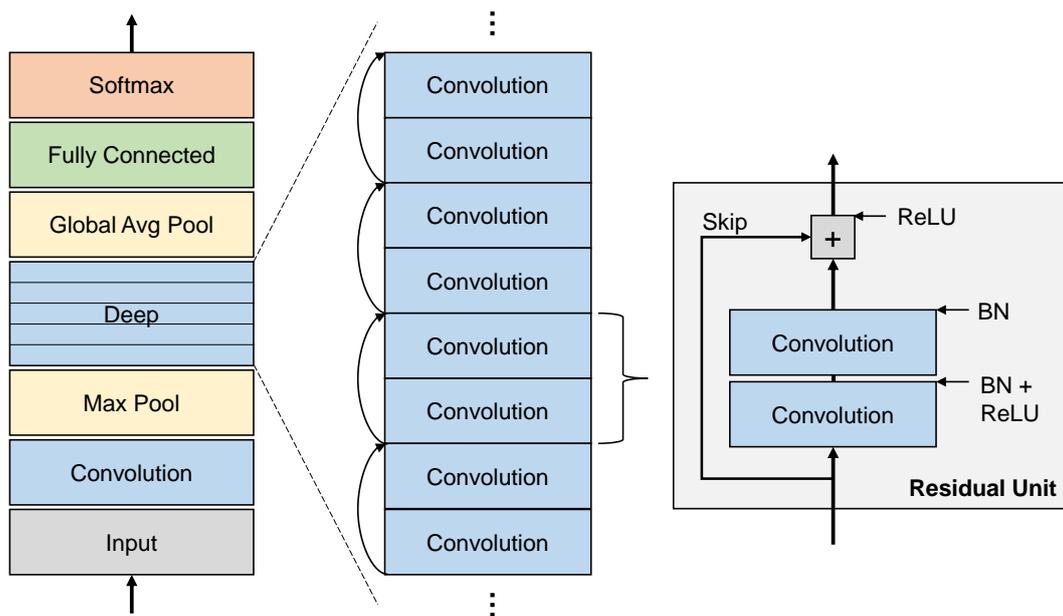


Abbildung 5-7: Architektur eines Residual Networks (nach [133])

Durch das eingesetzte FPN werden die Features nicht nur aus der letzten Faltungsschicht (conv5), sondern aus allen Faltungsschichten (conv2 – conv5) entnommen und an das RPN weitergeleitet, sodass Objekte verschiedener Größen deutlich besser detektiert werden. Dazu nutzt das FPN die Feature Maps der Convolutional Layer nicht direkt, sondern addiert zu jeder Feature Map noch die Feature Map der folgenden Convolutional Layer, sodass sich die Pyramid Layer (P2 – P5) ergeben, siehe Abbildung 5-6. Im Vergleich zu einem Faster R-CNN ohne FPN ließ sich die AP beim Validieren um fünf Punkte erhöhen [134].

Die zum Einsatz kommende Netzwerkarchitektur verwendet außerdem Non-Maximum Suppression (NMS). Durch das RoI-Pooling werden in der Regel mehrere korrekt klassifizierte Bounding Boxes mit leichtem Versatz um das zu detektierende Objekt ausgegeben. Um diese zu reduzieren, wird im Rahmen der NMS die IoU aller sich überlagernder Bounding Boxes erfasst und die Vorhersage mit der größten Überlagerung als einzige Box ausgewählt [135].

5.5 Statistische Versuchsplanung

5.5.1 Festlegung der Faktorstufen und Faktorstufenkombinationen

Um statistisch gesicherte Ergebnisse zu erzielen, erfolgt eine umfangreiche statistische Versuchsplanung. Im ersten Schritt wird die Anzahl der Faktorstufen sowie ihre Kombination festgelegt. Da im Rahmen der Festlegung der abhängigen und unabhängigen Variablen definiert wurde, dass ausschließlich die konstruktiven Produkteigenschaften zu variieren sind (siehe Abschnitt 5.2), werden diese als einziger Faktor berücksichtigt. Es handelt sich um einen qualitativen Faktor, dessen Stufen sich nicht durch Zahlenwerte auf einer Messskala beschreiben lassen. Vielmehr bilden die Faktorstufen verschiedene

qualitative Eigenschaften ab. Da das Ziel der vorliegenden Arbeit darin besteht, eine Vielzahl unterschiedlicher konstruktiver Produkteigenschaften zu untersuchen und eine erste Tendenz bezüglich ihres Einflusses auf die Objekterkennung festzustellen, werden je Produkteigenschaft nur zwei Faktorstufen betrachtet.

Im nächsten Schritt werden die Faktorstufenkombinationen festgelegt. Da lediglich ein Faktor in zwei Stufen zu untersuchen ist, resultiert daraus ein einfaktorieller, zweifachgestufter Versuchsplan, siehe Tabelle 5-3 a). Je Hypothese ist ein eigener Versuchsplan aufzustellen. Tabelle 5-3 b) zeigt diesen für die Hypothese BT1. Die weiteren Versuchspläne folgen demselben Prinzip. Im Sinne der Übersichtlichkeit wird auf die Darstellung aller Versuchspläne verzichtet.

Tabelle 5-3: Versuchsplan a) Allgemein b) Bezüglich Hypothese BT1

a)		b)	
Nr.	Konstruktive Produkteigenschaft	Nr.	Konstruktive Produkteigenschaft
1	Eigenschaft 1	1	Spiegelnde Bauteiloberfläche
2	Eigenschaft 2	2	Matte Bauteiloberfläche

5.5.2 Festlegung des Stichprobenumfangs

Im zweiten Schritt der statistischen Versuchsplanung erfolgt die Definition des Stichprobenumfangs. Um die Ableitung statistisch gesicherter Erkenntnisse aus den Versuchen zu gewährleisten, ist eine ausreichend große Anzahl Einzelversuche durchzuführen. Die Anzahl der durchzuführenden Versuche hängt dabei von der angestrebten Genauigkeit, mit welcher der Effekt eines Faktors bestimmt werden soll und der Streuung der Versuchsergebnisse, aufgrund zufälliger Unterschiede ab.

Gemäß Kleppmann lässt sich der Stichprobenumfang für Versuche mit zweistufigen Faktoren durch Gl. 15 und Gl. 16 schätzen. Der Faktor 60 gilt dabei für ein Signifikanzniveau von $\alpha = 1\%$ und einen Fehler 2. Art von $\beta = 10\%$. [136]

$$N \approx 60 \cdot \left(\frac{\sigma}{\Delta\mu} \right)^2 \quad \text{Gl. 15}$$

$$n = \frac{N}{m} \quad \text{Gl. 16}$$

mit

σ = Standardabweichung

$\Delta\mu$ = zu beweisender Effekt

N = Anzahl Einzelversuche

m = Anzahl Faktorstufenkombinationen

n = Stichprobenumfang

Gl. 15 zeigt, dass der Stichprobenumfang von der Höhe der Standardabweichung abhängt. Liegen bereits Erfahrungswerte zu geplanten Versuchen vor, kann die Standardabweichung in der Regel gut geschätzt werden. Im Fall der vorliegenden Untersuchung bestehen jedoch keinerlei Vorkenntnisse. Aus diesem Grund werden Vorversuche durchgeführt, aus welchen sich die Standardabweichung ermitteln lässt. Der Aufbau der Vorversuche entspricht dabei den Hauptversuchen, siehe Abschnitt 5.3. Um den Aufwand zu minimieren, werden im Rahmen der Vorversuche nur ausgewählte Hypothesen und damit eine begrenzte Anzahl an repräsentativen Untersuchungsobjekten berücksichtigt.

Wie in Abschnitt 5.2 beschrieben, sind außerdem Vorversuche notwendig, um die Anzahl der zu verwendenden Trainings- und Validierungsbilder sowie Iterationen festzulegen. Da die Vorversuche zur Festlegung der Standardabweichung unter den gleichen Parametern wie die Hauptversuche durchzuführen sind, werden zunächst Bilderanzahl (siehe Abschnitt 5.5.2.1) sowie die Anzahl an notwendigen Iterationen (siehe Abschnitt 5.5.2.2) untersucht und festgelegt. Im Anschluss werden diese Parameter genutzt, um die Standardabweichungen in Bezug auf die Bauteilgestaltung, die Baugruppengestaltung, die Auswahl von Normteilen und die Dimensionierung von Formelementen zu ermitteln und daraus den notwendigen Stichprobenumfang abzuleiten, siehe Abschnitt 5.5.2.3 bis 5.5.2.6.

5.5.2.1 Umfang der Trainings- und Validierungsdatensätze

Im Rahmen von umfangreichen Vorversuchen wird die Anzahl der in der Hauptuntersuchung zu verwendenden Trainings- und Validierungsbilder definiert. Im Sinne einer guten Übertragbarkeit der Ergebnisse kommen dabei die Untersuchungsobjekte der Hauptuntersuchung zum Einsatz. Diese werden in Abschnitt 5.6.1 detailliert beschrieben. Um den Aufwand der Voruntersuchung auf einem realisierbaren Niveau zu halten, werden nicht alle Objekte genutzt. Stattdessen werden einzelnen Objekte ausgewählt, welche die Übrigen repräsentieren.

Um den Einfluss der Bilderanzahl auf die Erkennungsgenauigkeit des Objekterkennungsalgorithmus zu ermitteln, werden Modelle trainiert, welche unterschiedliche Objektklassen detektieren sollen. Insgesamt werden sechs repräsentative Modelle ausgewählt, siehe Tabelle 5-4. Die in Spalte a) enthaltenen Modelle verkörpern dabei Objekte, welche einzeln auf den Trainings- und Validierungsbildern abgebildet sind, siehe Abschnitte 5.6.1.1 und 5.6.1.4. Im Gegensatz dazu stehen die in Spalte b) aufgeführten Modelle für Objekte, welche in einer Baugruppe zusammenhängen, sodass mehrere Objekte auf einem Bild zu detektieren sind, siehe Abschnitt 5.6.1.2.

Tabelle 5-4: Modelle zur Festlegung der Trainings- und Validierungsbilderanzahl

	a)			b)		
Modell	Form1-4	Form1-4Weich	Form3-3.3	BG-Alu	BG-Schwarz	BG-Gemischt
Objekt 1	Form1	Form1Weich	Form3	PlatteAlu	PlatteSchwarz	PlatteAlu
Objekt 2	Form2	Form2Weich	Form3.1	Form1Alu	Form1Schwarz	Form1Schwarz
Objekt 3	Form3	Form3Weich	Form3.2	Form2Alu	Form2Schwarz	Form2Rot
Objekt 4	Form4	Form4Weich	Form3.3	Form3Alu	Form3Schwarz	Form3Schwarz
Objekt 5				Form4Alu	Form4Schwarz	Form4Rot

Für den Trainings- und Validierungsprozess werden schließlich unterschiedliche Bilderanzahlen genutzt. Die Anzahl der Bilder ergibt sich dabei aus der Anzahl der Perspektiven, aus welchen die Objekte aufgenommen wurden. Prinzipiell werden je Betrachtungswinkel 60 Bilder aufgezeichnet. Die Grundlage hierfür besteht durch den in Abschnitt 5.3 vorgestellten Versuchsaufbau. In diesem wird festgelegt, dass alle Objekte auf einem Drehteller positioniert und in Rotation versetzt werden, um in regelmäßigen Abständen Bilder aufzuzeichnen. Durch diese Vorgehensweise lassen sich reproduzierbare Aufnahmen der Untersuchungsobjekte aus allen Richtungen erstellen. Durch die Rotationsgeschwindigkeit des Drehtellers von einer Umdrehung je Minute und einer Bildrate von einem Bild je Sekunde ergeben sich 60 Bilder je Perspektive. Neben der Draufsicht, welche in allen Modellen enthalten ist, werden eine bis neun weitere Perspektiven für die Bilderstellung genutzt, sodass sich insgesamt 120 bis 600 Bilder je Objekt und Modell ergeben. Diese werden für den Trainings- und Validierungsprozess in 80 % Trainings- und 20 % Validierungsbilder unterteilt. Abbildung 5-8 bis Abbildung 5-13 zeigen die in Abhängigkeit der genutzten Trainings- und Validierungsbilderanzahl erzielten Erkennungsgenauigkeiten.

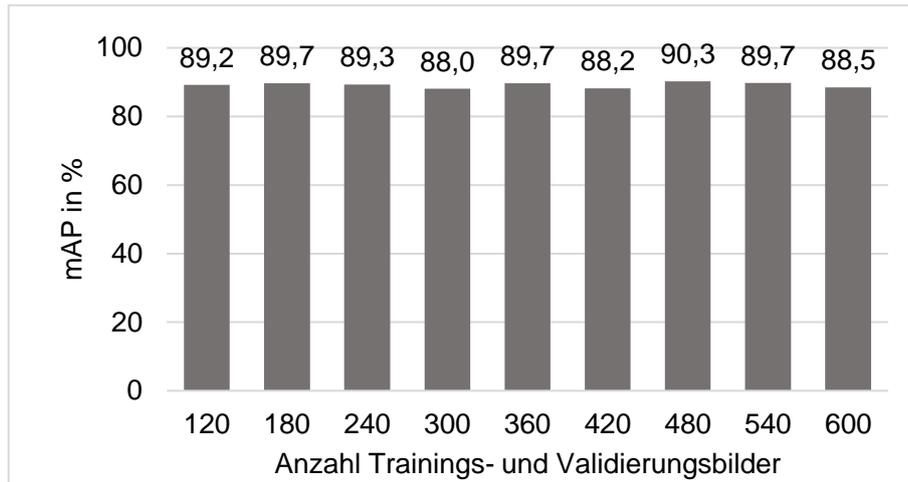


Abbildung 5-8: Erreichte mAP für das Modell „Form1-4“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

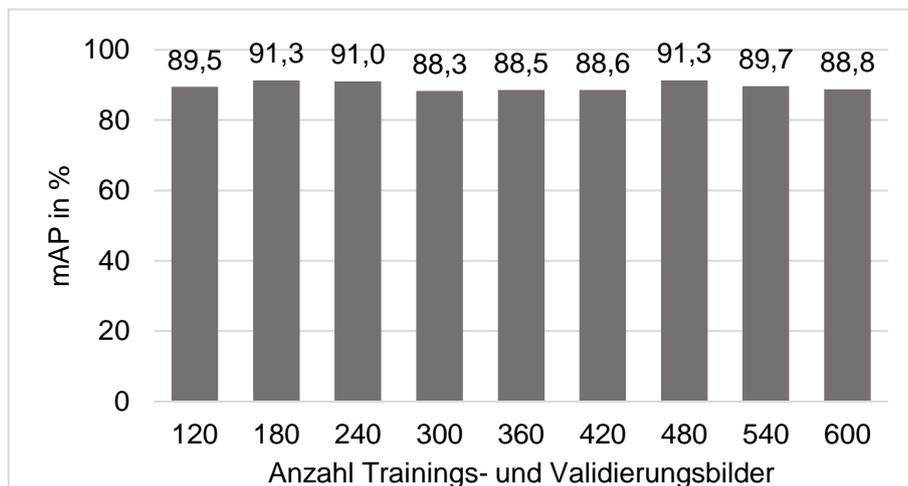


Abbildung 5-9: Erreichte mAP für das Modell „Form1-4Weich“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

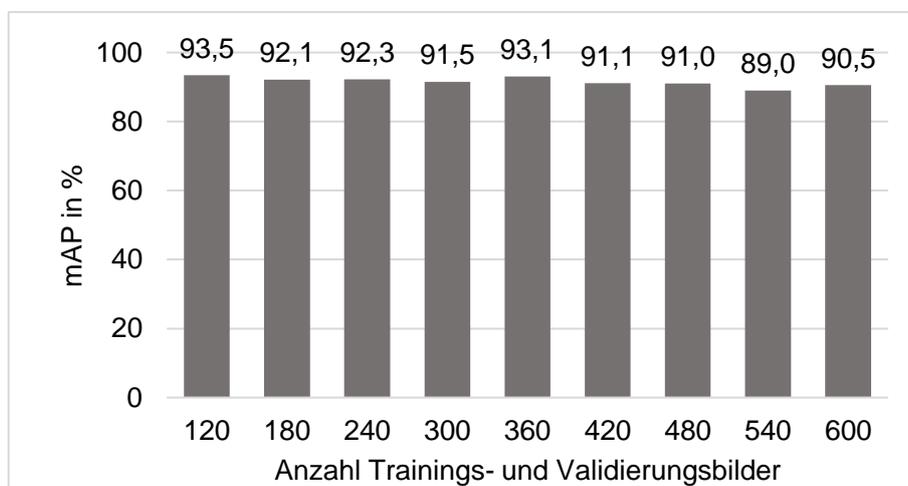


Abbildung 5-10: Erreichte mAP für das Modell „Form3-3.3“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

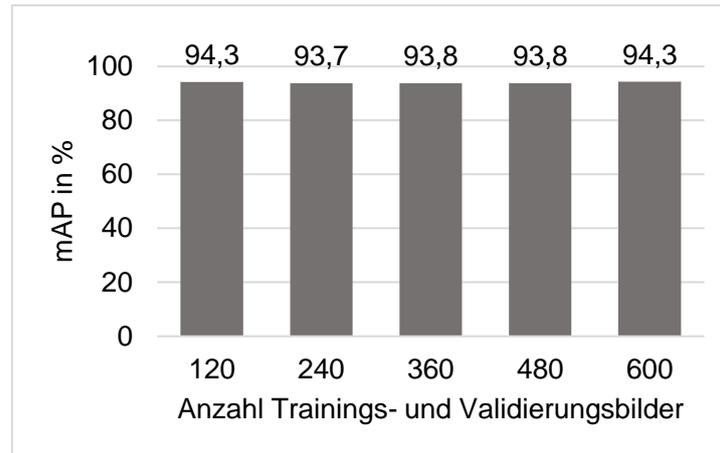


Abbildung 5-11: Erreichte mAP für das Modell „BG-Alu“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

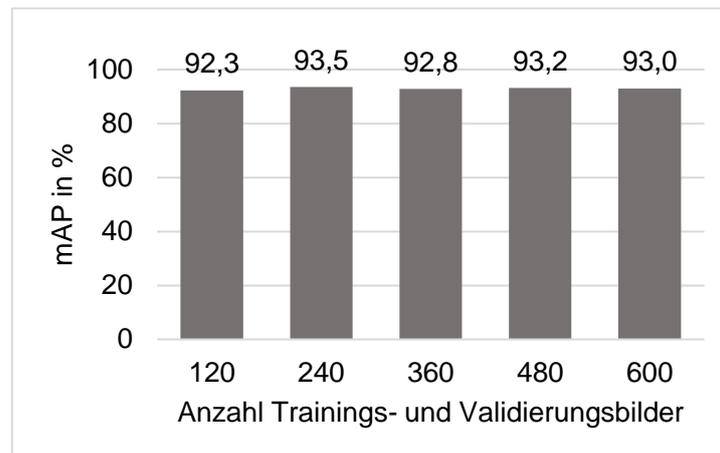


Abbildung 5-12: Erreichte mAP für das Modell „BG-Schwarz“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

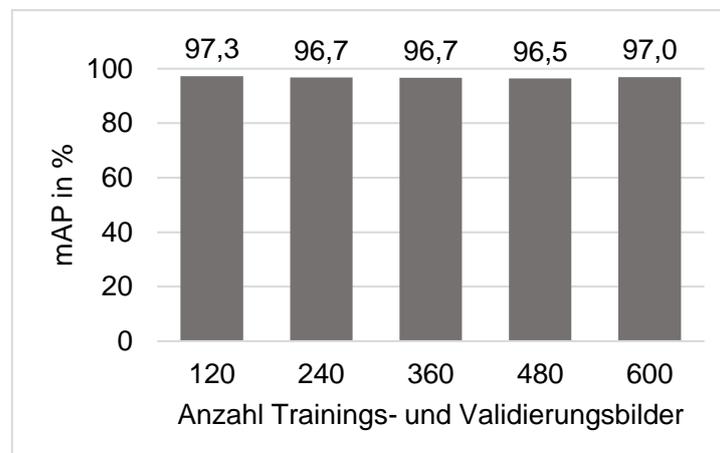


Abbildung 5-13: Erreichte mAP für das Modell „BG-Gemischt“ in Abhängigkeit der Anzahl der verwendeten Trainings- und Validierungsbilder

Die Ergebnisse der Voruntersuchung zeigen, dass die Anzahl der Trainings- und Validierungsbilder in der vorliegenden Untersuchung einen sehr geringen Einfluss auf die mAP hat. Sowohl bei 120 als auch 600 Bildern je Objekt ergeben sich für alle Modelle sehr hohe Erkennungsgenauigkeiten. Dennoch nimmt die Aussagekraft eines Modells mit steigender Bilderanzahl zu, da für den Validierungsprozess mehr Bilder zur Verfügung stehen. Die Berechnung der mAP bezieht sich dann auf eine größere Datenbasis, sodass der Einfluss einzelner Vorhersagen abnimmt. Da die Vorversuche zeigen, dass zehn Perspektiven im Rahmen der Versuchsdurchführung gut umzusetzen sind, werden die Bilder im Rahmen der Hauptuntersuchung aus zehn Perspektiven aufgezeichnet.

Weiterhin verdeutlicht die Voruntersuchung, dass Trainings- und Validierungsbilder eine sehr große Ähnlichkeit aufweisen. Dies ist unter anderem der einheitlichen Objektumgebung und -positionierung geschuldet, durch welche Störeinflüsse minimiert werden. Eine große Ähnlichkeit von Trainings- und Validierungsbildern kann jedoch dazu führen, dass ein Modell im Validierungsprozess zwar eine hohe mAP erzielt, in der konkreten Anwendung jedoch eine schlechte Performance aufweist.

Um die Ähnlichkeit der Trainings- und Validierungsbilder zu reduzieren, ohne den Versuchsaufbau zu verändern, werden im Rahmen der Hauptuntersuchung Trainings- und Validierungsbilder aus unterschiedlichen Perspektiven erfasst. Dabei werden jeweils 20 % der Bilder aus den Validierungsbilderperspektiven und 80 % der Bilder aus den Trainingsbilderperspektiven genutzt, sodass sich insgesamt 300 Bilder je Objekt und Modell ergeben, siehe Tabelle 5-5. Zwar halbiert sich damit die Anzahl der verwendeten Bilder im Vergleich zu den Vorversuchen, sodass die Aussagekraft etwas sinkt. Da die Vorversuche jedoch zeigen, dass die absolute Bilderanzahl die Erkennungsgenauigkeit nicht maßgeblich beeinflusst, wird dies im Sinne einer geringeren Ähnlichkeit zwischen Trainings- und Validierungsbildern in Kauf genommen.

Tabelle 5-5: Anzahl der Trainings- und Validierungsbilder in der Hauptuntersuchung

Perspektive	Anzahl Bilder	80% Training	20% Validierung
1	60	48	-
2	60	-	12
3	60	48	-
4	60	-	12
5	60	48	-
6	60	-	12
7	60	48	-
8	60	-	12
9	60	48	-
10	60	-	12
Summe	600	240	60
Summe Trainings- und Validierungsbilder		300	

5.5.2.2 Anzahl der Iterationen

Im nächsten Schritt erfolgt die Durchführung von Vorversuchen zur Bestimmung der notwendigen Iterationenanzahl. Dazu werden die im vorherigen Abschnitt ausgewählten, repräsentativen Modelle (siehe Tabelle 5-4) jeweils über 100.000 Iterationen trainiert und in Abständen von 1.000 Iterationen validiert. Neben der mAP (siehe Abbildung 5-14) wird der Trainings- (siehe Abbildung 5-15) sowie der Validierungsverlust (siehe Abbildung 5-16) erfasst, um eine Iterationsanzahl festzulegen, welche eine Über- sowie Unteranpassung der Modelle verhindert. Im Sinne einer guten Übertragbarkeit der Ergebnisse auf die Hauptuntersuchung, werden die Bilder für den Trainings- und Validierungsprozess dabei nach den im vorherigen Abschnitt festgelegten Bedingungen (siehe Tabelle 5-5) aufgezeichnet.

In Abbildung 5-14 wird deutlich, dass sich die mAP der Modelle Form1-4, Form1-4Weich, BG-Alu, BG-Schwarz und BG-Gemischt nach spätestens 25.000 Iterationen einem für jedes Modell spezifischem Maximalwert annähert. Lediglich für das Modell Form3-3.3 ergibt sich eine deutliche Abweichung hinsichtlich des Kurvenverlaufs. So erreicht dieses Modell in den frühen Iterationen deutlich geringere mAP-Werte und verzeichnet einen flacheren Anstieg. Die Annäherung an einen Maximalwert wird erst ab 50.000 Iterationen erreicht. Weiterhin zeigen sich deutlich größere Sprünge der mAP im Vergleich zu den anderen Modellen.

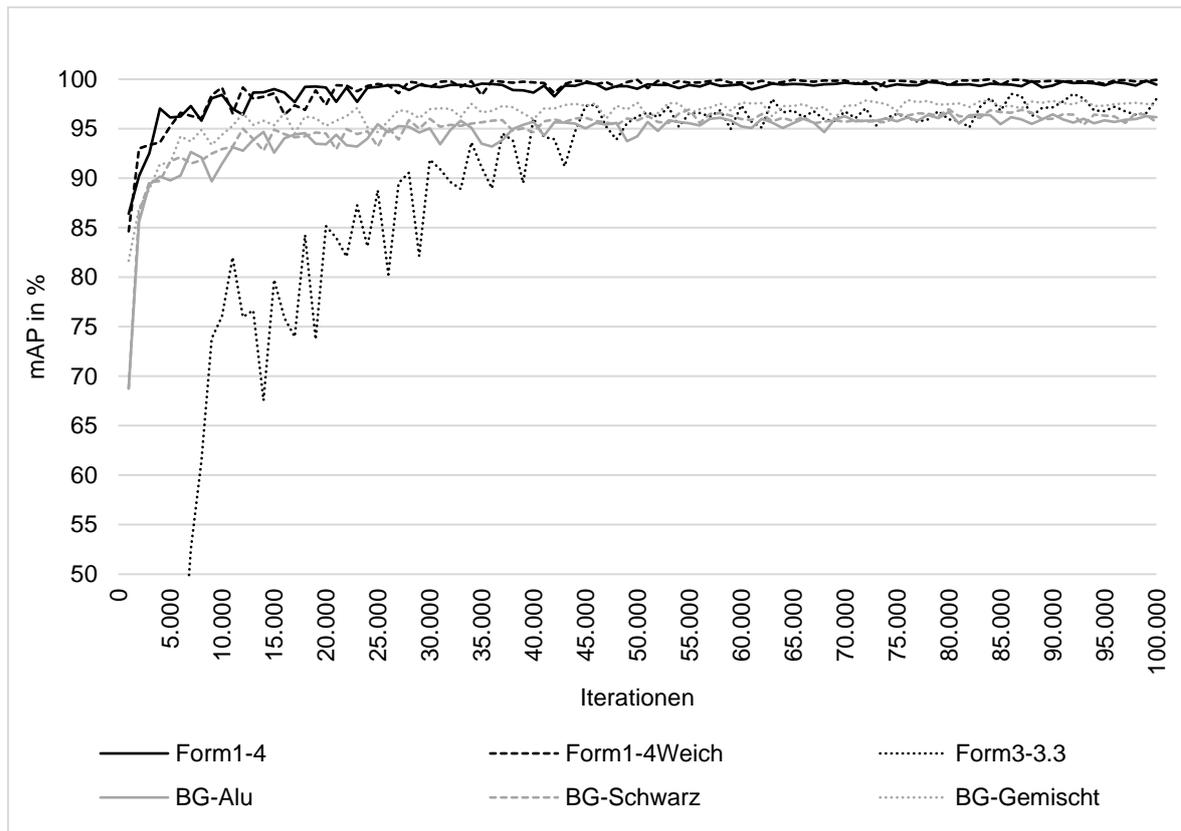


Abbildung 5-14: Vorversuchsergebnisse zur Iterationenanzahl – mAP

Abbildung 5-15 zeigt, dass der Trainingsfehler aller betrachteten Modelle in den ersten 30.000 Iterationen stark abnimmt und sich anschließend langsam einem Minimum annähert. Dabei weisen die Modelle, welche Baugruppen enthalten, insgesamt höhere Werte als die bauteilbeinhaltenden Modelle auf. Außerdem wird deutlich, dass der Trainingsfehler des Modells Form3-3.3 im Vergleich zu den Modellen Form1-4 und Form1-4Weich deutlich langsamer fällt und sich erst nach ca. 30.000 Iterationen dem gleichen Minimalwert annähert. Obwohl sich der Trainingsfehler aller Modelle nach 30.000 Iterationen nur noch geringfügig reduziert, zeigt Abbildung 5-15, dass auch nach 100.000 Iterationen kein Graph seinen Minimalwert erreicht.

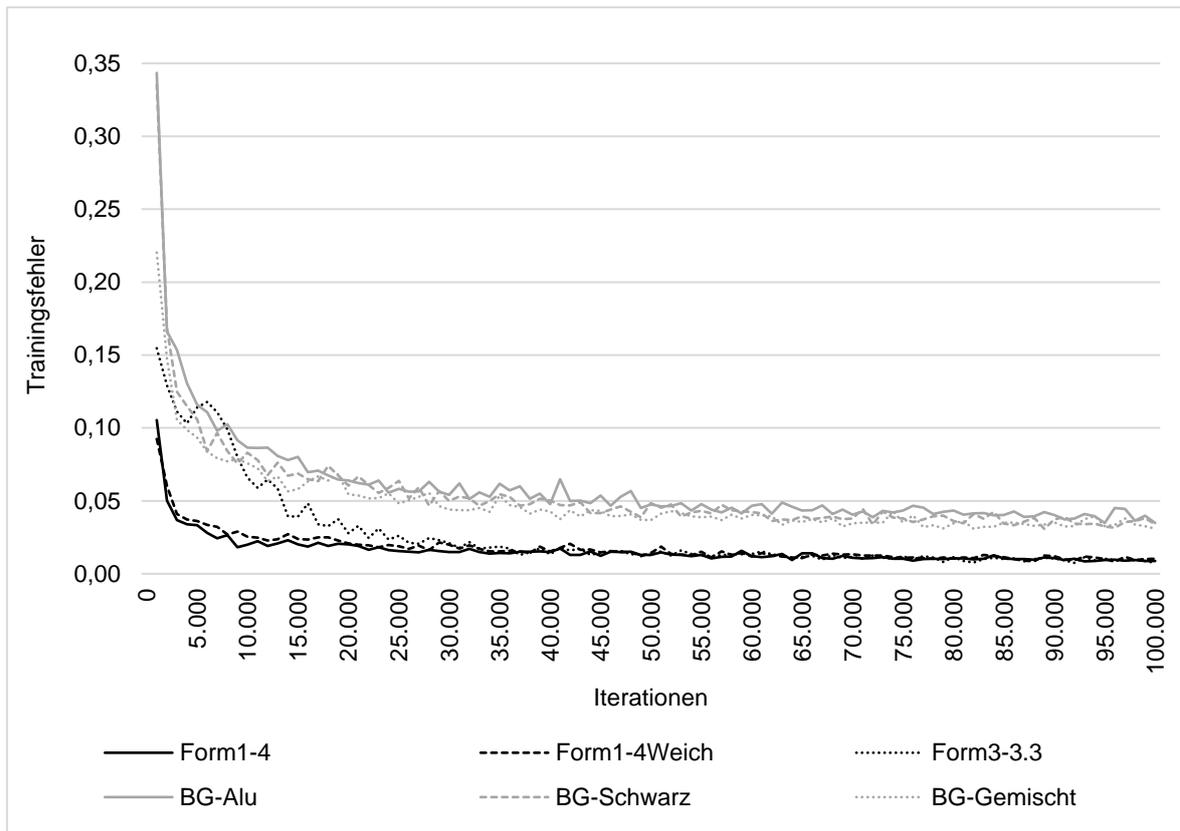


Abbildung 5-15: Vorversuchsergebnisse zur Iterationenanzahl – Trainingsfehler

Auch der Validierungsfehler weist zu Beginn des Trainings- und Validierungsprozesses den größten Abfall auf, siehe Abbildung 5-16. Anders als der Trainingsfehler unterscheidet sich der Kurvenverlauf der einzelnen Modelle dabei jedoch stärker voneinander. So sinkt der Validierungsfehler der Modelle Form1-4 und Form1-4Weich bis zu einer Iterationenanzahl von 40.000 deutlich und nähert sich anschließend langsam einem Minimalwert. Der Validierungsfehler des Modells Form3-3.3 nimmt hingegen bis ca. 60.000 Iterationen deutlich ab und nähert sich anschließend einem im Vergleich zu den anderen bauteilenthaltenden Modellen höheren Minimalwert. Eine noch spätere Annäherung an ein Minimum weist der Validierungsfehler der Modelle BG-Alu, BG-Schwarz und BG-Gemischt auf. Dieser nimmt bis ca. 75.000 Iterationen deutlich ab.

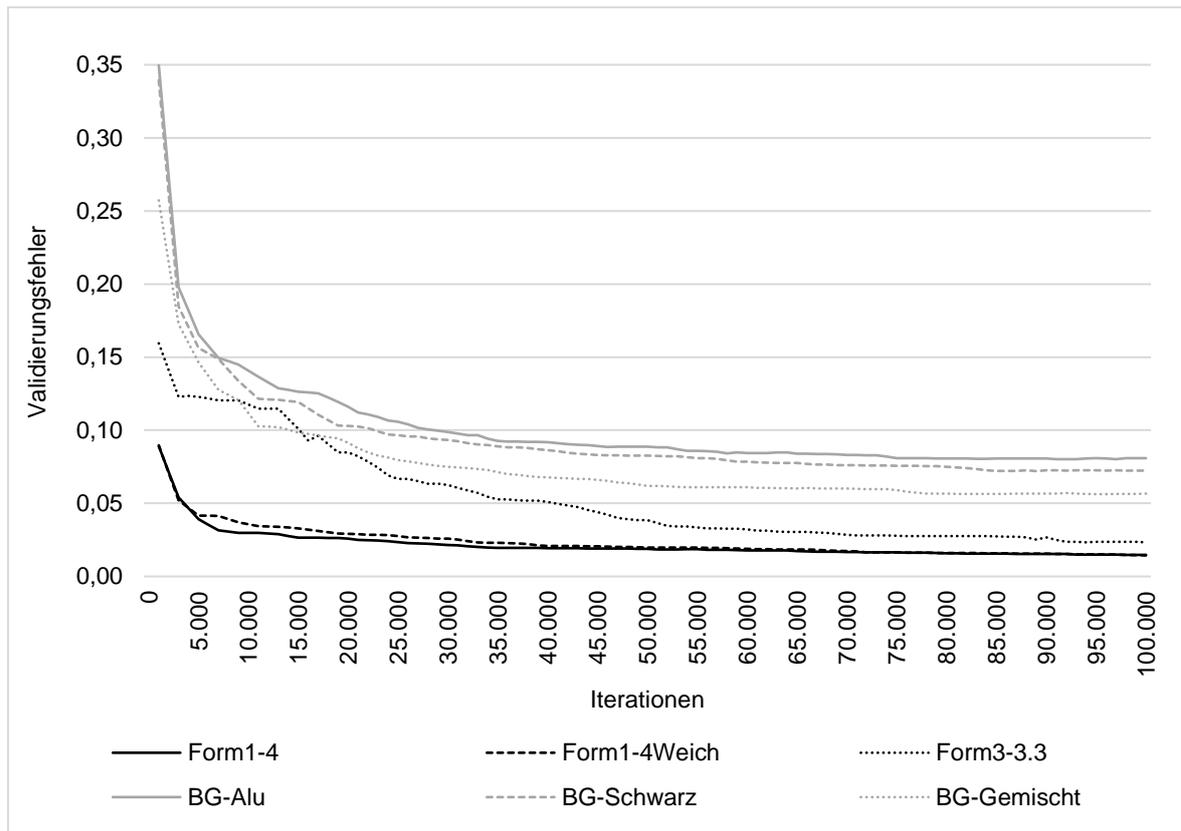


Abbildung 5-16: Vorversuchsergebnisse zur Iterationenanzahl – Validierungsfehler

Insgesamt wird deutlich, dass der Validierungsverlust über die betrachteten Iterationen kontinuierlich sinkt und sich kein Minimum wie in Abbildung 5-1 abzeichnet, welches in der Regel die Schwelle zwischen Unter- und Überanpassung aufzeigt, siehe Abschnitt 5.2. Dennoch demonstriert die Annäherung der mAP-Werte an einen für das jeweilige Modell spezifischen Maximalwert, dass keine Unteranpassung vorliegt.

Weiterhin zeigt sich, dass die Trainings- und Validierungsergebnisse des Modells Form3-3.3 deutlich von den Ergebnissen der anderen Modelle abweichen. Im Gegensatz zu diesen nähert sich das Modell erst nach 50.000 Iterationen einer maximalen mAP an. Dennoch erreicht es nach 40.000 Iterationen eine mAP von 96,02 %, welche nur um 2,5 %P vom absoluten Maximum (98,5 %) abweicht. Da die Anzahl der Iterationen den notwendigen Rechenaufwand maßgeblich beeinflusst und die übrigen Modelle sich bereits nach spätestens 20.000 Iterationen einem Maximum annähern, wird für die Hauptuntersuchung eine Anzahl von 40.000 Iterationen definiert. Damit lassen sich auf dem verwendeten System je zu trainierendem Modell mindestens zwei Stunden Rechenzeit einsparen. Zusätzlich erfolgt im Rahmen der Hauptuntersuchung während jedes Trainingsprozesses alle 1.000 Iterationen eine Validierung, sodass der Verlauf der mAP erfasst wird. Sollte sich zeigen, dass die Ergebnisse nach 40.000 Iterationen noch eine maßgebliche Steigung aufweisen, wird die Hauptuntersuchung hinsichtlich der umzusetzenden Iterationen nachträglich angepasst.

5.5.2.3 Stichprobenumfang – Bauteilgestaltung

Auf Grundlage der festgelegten Anzahl an Trainings- und Validierungsbildern sowie Iterationen, werden Vorversuche zur Bestimmung der Standardabweichung durchgeführt. Da sowohl die Hypothesen als auch die daraus abgeleitete Hauptuntersuchung in die vier Kategorien Bauteilgestaltung, Baugruppengestaltung, Auswahl von Normteilen und Dimensionierung von Formelementen untergliedert sind, wird diese Unterteilung auch für die Vorversuche angewendet.

Als erstes erfolgt die Auswahl dreier repräsentativer Modelle zur Bauteilgestaltung und die Durchführung von jeweils fünf Versuchen inkl. Bilderstellung, Annotation sowie Trainings- und Validierungsprozess, um die Standardabweichung zu ermitteln, siehe Tabelle 5-6. Dabei werden neben dem Modell Form1-4, welches als Referenzmodell von besonderer Bedeutung ist, Modelle ausgewählt, welche den Aspekt der Oberflächenbeschaffenheit (Form1-4Schwarz) sowie der unterschiedlichen Konturen (Form3-3.3) abbilden. Die Modelle und die darin enthaltenen Objekte werden in Abschnitt 5.6.1.1 detailliert beschrieben.

Tabelle 5-6: Vorversuchsergebnisse – Bauteilgestaltung

Modell	mAP [%]					\bar{x} [%]	s [%]
	V1	V2	V3	V4	V5		
Form1-4	94,79	93,68	92,96	92,60	94,31	93,67	0,91
Form1-4Schwarz	99,11	99,72	99,53	99,02	98,90	99,26	0,35
Form3-3.3	82,07	80,03	85,16	80,60	77,14	81,00	2,94

Unter Nutzung der ermittelten Standardabweichung wird anschließend der erforderliche Stichprobenumfang abgeleitet, siehe Tabelle 5-7. Dabei wird Gl. 15 aus Abschnitt 5.5.1 angewendet und als zu beweisender Effekt eine Differenz von 10 %P angenommen. Die Anzahl der Faktorstufenkombinationen kann Abschnitt 5.5.1 entnommen werden und beträgt für alle durchzuführenden Versuche einen Wert von zwei.

Tabelle 5-7 zeigt, dass sich für das Modell Form3-3.3 mit 2,59 notwendigen Versuchen der größte Stichprobenumfang ergibt. Auch wenn die beiden anderen Modelle deutlich geringere Werte aufweisen, wird für die Hauptuntersuchung zur Bauteilgestaltung ein Stichprobenumfang von drei Versuchen je Modell festgelegt, um die Vergleichbarkeit der Versuchsergebnisse sicherzustellen und die Übersichtlichkeit zu erhöhen.

Tabelle 5-7: Ableitung Stichprobenumfang – Bauteilgestaltung

Modell	Form1-4	Form1-4Schwarz	Form3-3.3
Zu beweisender Effekt	10 %P	10 %P	10 %P
Einzelversuche	0,50	0,07	5,18
Faktorstufenkombinationen	2	2	2
Stichprobenumfang	0,25	0,04	2,59

5.5.2.4 Stichprobenumfang – Baugruppengestaltung

Auch für die Ermittlung des Stichprobenumfangs für die Versuche zur Baugruppengestaltung werden drei repräsentative Modelle ausgewählt und jeweils fünf Versuche durchgeführt. Tabelle 5-8 zeigt die Versuchsergebnisse sowie die daraus resultierenden Standardabweichungen. Durch die Modellauswahl werden Baugruppen mit Bauteilen gleicher (Einfarbig) und unterschiedlicher (Mehrfarbig) sowie heller (Alu) und dunkler (Schwarz) Materialfarbe abgebildet. Die in den einzelnen Modellen enthaltenen Baugruppen und Bauteile werden in Abschnitt 5.6.1.2 detailliert beschrieben.

Tabelle 5-8: Vorversuchsergebnisse – Baugruppengestaltung

Modell	mAP [%]					\bar{x} [%]	s [%]
	V1	V2	V3	V4	V5		
Einfarbig_Alou	92,71	96,39	97,80	95,75	95,78	95,69	1,86
Einfarbig_Schwarz	94,34	95,24	96,20	94,05	94,98	94,96	0,84
Mehrfarbig	95,68	97,46	97,66	96,24	96,41	96,69	0,84

Tabelle 5-9 zeigt den notwendigen Stichprobenumfang je Modell, welcher sich aus der ermittelten Standardabweichung ergibt. Als zu beweisender Effekt wird eine Differenz von 10 %P angenommen. Es wird deutlich, dass sich der höchste notwendige Stichprobenumfang mit einem Wert von 1,04 für das Modell Einfarbig_Alou ergibt. Auch wenn die Ergebnisse darauf hindeuten, dass zwei Stichproben ausreichen, werden im Rahmen der Hauptuntersuchung drei Versuche je Modell durchgeführt. Durch die Erhöhung des Stichprobenumfangs wird die Versuchsdurchführung vereinheitlicht, sodass sich die Übersichtlichkeit erhöht und das Fehlerrisiko reduziert.

Tabelle 5-9: Ableitung Stichprobenumfang – Baugruppengestaltung

Modell	Einfarbig_Alu	Einfarbig_Schwarz	Mehrfarbig
Zu beweisender Effekt	10 %P	10 %P	10 %P
Einzelversuche	2,08	0,42	0,43
Faktorstufenkombinationen	2	2	2
Stichprobenumfang	1,04	0,21	0,21

5.5.2.5 Stichprobenumfang – Auswahl von Normteilen

Um den Stichprobenumfang der Hauptuntersuchung zur Auswahl von Normteilen zu definieren, werden ebenfalls drei repräsentative Modelle ausgewählt, jeweils fünf Versuche durchgeführt und die Standardabweichung ermittelt, siehe Tabelle 5-10. Die in der Tabelle aufgeführten Modelle kommen zum Einsatz, da die enthaltenen Schrauben alle Nennweiten abdecken. Darüber hinaus sind sowohl Schrauben enthalten, welche vollständig im Bauteil versenkt werden, als auch Schrauben, deren Köpfe aus dem Bauteil herausragen. Weiterhin decken die Modelle sowohl die Hypothese der geringen Größendifferenz (N1) als auch die Hypothese der ähnlichen Schraubenköpfe (N2) ab. Die in den einzelnen Modellen enthaltenen Objekte werden in Abschnitt 5.6.1.3 detailliert beschrieben.

Tabelle 5-10: Vorversuchsergebnisse – Auswahl von Normteilen

Modell*	mAP [%]					\bar{x} [%]	s [%]
	V1	V2	V3	V4	V5		
LK_M6 + ZK_M6	95,94	97,79	98,64	98,76	98,36	97,90	1,16
U_M8 + U_M10	96,43	94,07	92,97	96,73	90,65	94,17	2,52
GS_M3 + GS_M4	73,46	75,88	80,23	78,13	80,44	77,63	2,97

*LK = Linsenkopf; ZK = Zylinderkopf; GS = Gewindestift; U = Unterlegscheibe

Unter Nutzung der berechneten Standardabweichung erfolgt die Ermittlung der erforderlichen Stichprobenumfänge je Modell, siehe Tabelle 5-11. Als zu beweisender Effekt wird dabei wie zuvor eine Differenz von 10 %P angenommen. Die Tabelle zeigt, dass sich mit einem Wert von 2,65 der höchste notwendige Stichprobenumfang für das Modell GS_M3 + GS_M4 ergibt. Obwohl die beiden anderen Modelle deutlich geringere Werte aufweisen, wird für die Hauptuntersuchung zur Auswahl von Normteilen ein Stichprobenumfang von drei Versuchen je Modell festgelegt, um die Vergleichbarkeit der Versuchsergebnisse sicherzustellen und die Übersichtlichkeit zu erhöhen. Darüber hinaus entspricht

dies dem Stichprobenumfang der Hauptuntersuchungen zur Bauteil- und Baugruppengestaltung. Hieraus resultiert eine weitere, positive Beeinflussung der Übersichtlichkeit sowie des Fehlerrisikos, da die Versuchsdurchführung einheitlicher gestaltet ist.

Tabelle 5-11: Ableitung Stichprobenumfang – Auswahl von Normteilen

Modell*	LK_M6 + ZK_M6	U_M8 + U_M10	GS_M3 + GS_M4
Zu beweisender Effekt	10 %P	10 %P	10 %P
Einzelversuche	0,81	3,82	5,31
Faktorstufenkombinationen	2,00	2,00	2,00
Stichprobenumfang	0,40	1,91	2,65
*LK = Linsenkopf; ZK = Zylinderkopf; GS = Gewindestift; U = Unterlegscheibe			

5.5.2.6 Stichprobenumfang – Dimensionierung von Formelementen

Abschließend werden drei repräsentative Modelle zur Dimensionierung von Formelementen ausgewählt und jeweils fünf Versuche durchgeführt, um die Standardabweichung zu ermitteln, siehe Tabelle 5-12. Die in der Tabelle aufgeführten Modelle kommen zum Einsatz, da die enthaltenen Objekte sowohl den Einfluss der Nutgeometrie (N4 + N6), als auch der Kantenform (F1 + F1,5) sowie der Bohrungsgeometrie (T3 + T4,5) beinhalten. Darüber hinaus sind sowohl kleine Größensprünge – Fase mit 1x45° gegenüber einer Fase mit 1,5x45° – als auch ausgeprägtere Größensprünge – Nut mit einer Breite von 4 mm gegenüber einer Nut mit einer Breite von 6 mm – berücksichtigt. Die konkreten, in den Modellen enthaltenen Objekte werden in Abschnitt 5.6.1.3 detailliert beschrieben.

Tabelle 5-12: Vorversuchsergebnisse – Dimensionierung von Formelementen

Modell*	mAP [%]					\bar{x} [%]	s [%]
	V1	V2	V3	V4	V5		
N4 + N6	96,66	100	100	100	100	99,33	1,49
F1 + F1,5	100	100	100	100	100	100	0
T3 + T4,5	99,95	100	100	100	99,79	99,95	0,09
*F = Fase; N = Nut; T = Bohrungstiefe							

Tabelle 5-13 zeigt den notwendigen Stichprobenumfang, welcher sich für die drei betrachteten Modelle aus der ermittelten Standardabweichung ergibt. Als zu beweisender Effekt wird wie zuvor eine Differenz von 10 %P angenommen. Es wird deutlich, dass das Modell F1 + F1,5 mit einem Wert von 0,67 den höchsten erforderlichen Stichprobenumfang aufweist. Dennoch werden im Rahmen der Hauptuntersuchung je Modell drei Versuche durchgeführt, um die Versuchsdurchführung zu vereinheitlichen. Indem die Hauptuntersuchungen zu allen Kategorien den gleichen Stichprobenumfang aufweisen, wird die Übersichtlichkeit erhöht und das Fehlerrisiko reduziert.

Tabelle 5-13: Ableitung Stichprobenumfang – Dimensionierung von Formelementen

Modell*	N4 + N6	F1 + F1,5	T3 + T4,5
Zu beweisender Effekt	10 %P	10 %P	10 %P
Einzelversuche	0,01	1,34	0
Faktorstufenkombinationen	2	2	2
Stichprobenumfang	0,003	0,67	0
*F = Fase; N = Nut; T = Bohrungstiefe			

5.6 Versuchsdurchführung

5.6.1 Konzeption der Untersuchungsobjekte

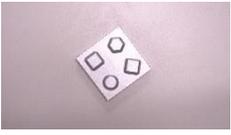
Um die in Abschnitt 4.2 aufgestellten Hypothesen zu untersuchen, werden im Rahmen der Versuchsdurchführung zunächst Untersuchungsobjekte konzipiert bzw. ausgewählt und beschafft. Da die verschiedenen Hypothesenkategorien durch unterschiedliche Objektarten bzw. Objektkombinationen zu untersuchen sind, werden diese auch im Zuge der Versuchsdurchführung einzeln betrachtet. Neben Objekten, welche zur Überprüfung der Hypothesen zur Bauteilgestaltung geeignet sind (siehe Abschnitt 5.6.1.1), werden Objekte zur Baugruppengestaltung (siehe Abschnitt 5.6.1.2), zur Auswahl von Normteilen (siehe Abschnitt 5.6.1.3) und zur Dimensionierung von Formelementen (siehe Abschnitt 5.6.1.4) ausgewählt.

5.6.1.1 Untersuchungsobjekte zur Bauteilgestaltung

Zur Überprüfung der Hypothesen bezüglich der Bauteilgestaltung werden insgesamt 36 Untersuchungsobjekte konzipiert und beschafft, siehe Tabelle 5-14. Wie in Abschnitt 4.2.1 beschrieben, beziehen sich die Hypothesen auf verschiedene Produkteigenschaften, wie spiegelnde und dunkle Oberflächen sowie komplexe, weiche und ähnliche Konturen. Prinzipiell wird je Hypothese mindestens ein Modell trainiert und validiert, welches Objekte mit den der These entsprechenden Eigenschaften enthält. Als Kompromiss zwischen einer

großen Objektvielfalt und einem vertretbaren Aufwand zur Datengenerierung wird eine Anzahl von vier Objekten je Modell definiert. Da es sich bei den Hypothesen zur Bauteilgestaltung um Unterschiedshypothesen handelt, welche den Vergleich zweier Modelle mit Objekten entgegengesetzter Eigenschaften erfordern (siehe Abschnitt 4.2.1), werden zusätzlich vier Referenzobjekte konzipiert.

Tabelle 5-14: Untersuchungsobjekte zur Bauteilgestaltung

Hypothese	Objekt 1	Objekt 2	Objekt 3	Objekt 4
Referenz	Form1 	Form2 	Form3 	Form4 
BT1 – Spiegelnd	Form1Spiegelnd 	Form2Spiegelnd 	Form3Spiegelnd 	Form4Spiegelnd 
BT2 – Dunkel	Form1Schwarz 	Form2Schwarz 	Form3Schwarz 	Form4Schwarz 
BT3 – Weich	Form1Weich 	Form2Weich 	Form3Weich 	Form4Weich 
BT4 – Ähnlich	Form1 	Form1.1 	Form1.2 	Form1.3 
BT4 – Ähnlich	Form2 	Form2.1 	Form2.2 	Form2.3 
BT4 – Ähnlich	Form3 	Form3.1 	Form3.2 	Form3.3 
<i>Fortsetzung auf der nächsten Seite</i>				

Fortsetzung Tabelle 5-14

Hypothese	Objekt 1	Objekt 2	Objekt 3	Objekt 4
BT4 – Ähnlich	Form4 	Form4.1 	Form4.2 	Form4.3 
BT5 – Komplex	Form5 	Form6 	Form7 	Form8 

Die in Tabelle 5-14 enthaltenen Objekte bilden die in Bezug auf die Bauteilgestaltung zu überprüfenden Hypothesen konstruktiv ab. Grundlage für alle Objekte sind quadratische Aluminiumblöcke, welche verschiedene geometrische Formen auf ihrer Oberfläche aufweisen. Da angenommen wird, dass eine hohe Konturenkomplexität zu einer reduzierten mAP führt, werden sehr einfache geometrische Formen genutzt. Darüber hinaus werden vier sehr unterschiedliche geometrische Formen ausgewählt, um den Einfluss ähnlicher Konturen auszuschließen. Prinzipiell stellt die Gestaltung der geometrischen Formen sicher, dass ausschließlich der Einfluss der jeweils zu untersuchenden Produkteigenschaften erfasst wird. Um den Aufwand zur Datengenerierung zu reduzieren, lassen sich die Referenzobjekte als Referenz für alle Hypothesen heranziehen. Die Untersuchungsobjekte der Hypothesenmodelle unterscheiden sich jeweils nur in der durch die Hypothese adressierten Eigenschaft.

So variieren die Objekte zur Hypothese BT1 lediglich in der Reflexionsfähigkeit ihrer Oberflächen. Während die Referenzformen (Form1 – Form4) durch nach dem Fräsprozess anschließendes Sandstrahlen eine hohe Rauheit und damit matte Oberflächen aufweisen, werden die Hypothesenobjekte (Form1Spiegelnd – Form4Spiegelnd) nach dem Fräsen in mehreren Schichten mit einem Chromlack überzogen, sodass ihre Oberflächen eine hohe Reflexionsfähigkeit innehaben. Durch das angewendete Airbrush-Verfahren wird eine homogene Schichtdicke gewährleistet. Die geometrischen Formen sind in beiden Modellen identisch.

Auch die Objekte zur Hypothese BT2 weisen die gleichen Formen wie die Referenzobjekte auf. Dazu werden mit den Referenzobjekten identische Formen gefertigt und nach dem Sandstrahlen mittels schwarzer Farbe lackiert. Im Vergleich zu den Referenzobjekten besitzen Form1Schwarz, Form2Schwarz, Form3Schwarz und Form4Schwarz damit eine deutlich dunklere Oberfläche. Um die Hypothese BT3 zu untersuchen, kommen Untersuchungsobjekte mit weichen Konturen zum Einsatz. Dazu werden die Referenzobjekte als Grundlage genutzt und die Kanten der geometrischen Formen der Aluminiumblockoberseite im Fräsprozess mit einem Radius von 2 mm abgerundet. Damit sich die

Hypothesen- und Referenzobjekte ausschließlich hinsichtlich ihrer Kanten unterscheiden, werden Form1Weich, Form2Weich, Form3Weich und Form4Weich anschließend auf die gleiche Weise sandgestrahlt wie die Formen1-4.

Für die Hypothese BT4 werden Objekte genutzt, welche sehr ähnliche geometrische Formen aufweisen. Um die Vergleichbarkeit mit dem Referenzmodell sicherzustellen, werden die Formen1-4 als Grundlage herangezogen. Durch minimale Veränderungen wird zusätzlich zur Form1 die Form1.1, die Form1.2 und die Form1.3 erstellt. Da nicht auszuschließen ist, dass Form1 im Vergleich zu den Formen2-4 besonders gut oder schlecht detektiert wird, werden auch diese durch kleine Veränderungen um jeweils drei weitere Formen ergänzt, siehe Tabelle 5-14. Während sich die Geometrien der Hypothesen- und Referenzobjekte unterscheiden, stellen die gleichen Prozessparameter beim Fräsen und Sandstrahlen sicher, dass die Objektflächen identische Eigenschaften aufweisen. Dadurch wird gewährleistet, dass die Untersuchung ausschließlich den Einfluss der Objektgeometrie erfasst.

Abschließend werden Objekte konzipiert, welche zur Untersuchung der Hypothese BT5 notwendig sind. Da im Rahmen der Hypothese zu untersuchen ist, welchen Einfluss die Komplexität der Konturen auf die Erkennungsgenauigkeit hat, weisen Form5, Form6, Form7 und Form8 deutlich komplexere geometrische Konturen auf, als die Formen1-4. Auch für die Objekte zu Hypothese BT5 wird durch die gleichen Prozessparameter beim Fräsen und Sandstrahlen sichergestellt, dass die Oberflächen der Hypothesenobjekte die gleichen Eigenschaften aufweisen wie die Referenzobjekte. Alle Objekte sind aus Aluminium EN AW-2007 gefertigt. Die vollständigen technischen Zeichnungen können [137] entnommen werden.

5.6.1.2 Untersuchungsobjekte zur Baugruppengestaltung

Um die Hypothese zur Baugruppengestaltung zu untersuchen, werden die im vorherigen Abschnitt beschriebenen Formen als Grundlage genutzt. Da im Rahmen der Hypothese zu untersuchen ist, ob Einzelteile von Baugruppen, welche das gleiche optische Erscheinungsbild aufweisen, durch einen Objekterkennungsalgorithmus mit geringerer Genauigkeit detektiert werden, als Einzelteile von Baugruppen, welche kein gleiches optisches Erscheinungsbild aufweisen, kommen verschiedene Baugruppen mit den entsprechenden Eigenschaften zum Einsatz. Prinzipiell werden jeweils zwei Aluminiumblöcke mit unterschiedlichen Geometrien durch eine Grundplatte zu einer Baugruppe zusammengefasst, siehe Tabelle 5-15. Da es sich bei der Hypothese zur Baugruppengestaltung um eine Unterschiedshypothese handelt, welche den Vergleich zweier Modelle mit Objekten entgegengesetzter Eigenschaften erfordert (siehe Abschnitt 4.2.2), wird jeder Baugruppe (Objekte Modell 1) eine Referenzbaugruppe (Objekte Modell 2) gegenübergestellt. Im Vergleich zu den Untersuchungsobjekten zur Bauteilgestaltung lässt sich hierbei kein einheitliches Referenzmodell für alle Hypothesen nutzen. Da nicht auszuschließen ist, dass die Farbe der Grundplatte die Höhe der mAP beeinflusst, weil sie sich z. B. besser vom Hintergrund abhebt, beinhalten die Baugruppen der Referenzmodelle jeweils die gleiche Grundplatte.

Tabelle 5-15: Untersuchungsobjekte zur Baugruppengestaltung

Hypothese	Objekte Modell 1		Objekte Modell 2	
BG1 – Optische Erscheinung	Platte, Form1, Form2 	Platte, Form3, Form4 	Platte, Form1Schwarz, Form2Rot 	Platte, Form3Blau, Form4Grün 
BG1 – Optische Erscheinung	PlatteSchwarz, Form1Schwarz, Form2Schwarz 	PlatteSchwarz, Form3Schwarz, Form4Schwarz 	PlatteSchwarz, Form1, Form2Rot 	PlatteSchwarz, Form3Blau, Form4Grün 
BG1 – Optische Erscheinung	PlatteWeiß, Form1Weiß, Form2Weiß 	PlatteWeiß, Form3Weiß, Form4Weiß 	PlatteWeiß, Form1Schwarz, Form2Rot 	PlatteWeiß, Form3Blau, Form4Grün 
BG1 – Optische Erscheinung	PlatteRot, Form1Rot, Form2Rot 	PlatteRot, Form3Rot, Form4Rot 	PlatteRot, Form1Schwarz, Form2 	PlatteRot, Form3Blau, Form4Grün 
BG1 – Optische Erscheinung	PlatteBlau, Form1Blau, Form2Blau 	PlatteBlau, Form3Blau, Form4Blau 	PlatteBlau, Form1Schwarz, Form2Rot 	PlatteBlau, Form3, Form4Grün 
BG1 – Optische Erscheinung	PlatteGrün, Form1Grün, Form2Grün 	PlatteGrün, Form3Grün, Form4Grün 	PlatteGrün, Form1Schwarz, Form2Rot 	PlatteGrün, Form3Blau, Form4 

Prinzipiell werden je Modell zwei Baugruppen konzipiert, welche sich in der Kombination der Grundformen unterscheiden. Die erste Baugruppe besteht jeweils aus einer Grundplatte, welche die Grundformen 1 und 2 aufnimmt. Für die zweite Baugruppe finden stattdessen die Grundformen 3 und 4 Verwendung. Während alle Bauteile für das erste Modell das gleiche optische Erscheinungsbild aufweisen, kommen im Referenzmodell Bauteile mit unterschiedlichen Erscheinungsbildern zum Einsatz.

Obwohl Abschnitt 4.2.2 nur eine Hypothese zur Baugruppengestaltung formuliert, zeigt Tabelle 5-15 insgesamt sechs verschiedene Modelle mit jeweils zugehörigem Referenzmodell. Grund dafür ist der nicht auszuschließende Einfluss der Bauteilfarbe. So kann im Vorfeld nicht ausgeschlossen werden, dass die Bauteilfarbe die Erkennungsgenauigkeit beeinflusst (siehe Hypothese BT2 in Abschnitt 4.2.1). Aus diesem Grund werden Baugruppen in verschiedenen Farben untersucht. Neben einer durch Sandstrahlen erzeugten matten Aluminiumoberfläche werden die Baugruppen in den Farben Schwarz, Weiß, Rot, Grün und Blau getestet. Dazu werden die Bauteile nach dem Fräsen und Sandstrahlen durch das zuvor beschriebene Airbrush-Verfahren mit verschiedenfarbigen Lacken überzogen.

Die technischen Zeichnungen der Aluminiumblöcke entsprechen den Formen 1-4. Zusätzlich zu diesen, können die technische Zeichnung der Grundplatte sowie die Zusammenbauzeichnungen der Baugruppen [137] entnommen werden.

5.6.1.3 Untersuchungsobjekte zur Auswahl von Normteilen

Da sich die Hypothesen zur Auswahl von Normteilen auf die Eigenschaften verschiedener Normteile beziehen, kommen statt selbst konzipierter Untersuchungsobjekte, kommerziell erhältliche Teile zum Einsatz. Prinzipiell werden dabei jeweils zwei Normteile ausgewählt, welche durch einen Objekterkennungsalgorithmus zu unterscheiden sind. Im Vergleich zu den Hypothesen zur Bauteil- und Baugruppengestaltung, handelt es sich bei den Hypothesen zur Auswahl von Normteilen um Zusammenhangshypothesen, welche überprüft werden, indem die Versuchsergebnisse mit einem konkreten Grenzwert verglichen werden (siehe Abschnitt 4.2.3). Aus diesem Grund entfällt die Auswahl von Referenzobjekten. Tabelle 5-16 zeigt alle für die Überprüfung der Hypothesen ausgewählten Normteile.

Tabelle 5-16: Untersuchungsobjekte zur Auswahl von Normteilen

Hypothese	Objekt 1	Objekt 2
N1	Linsenkopfschraube M3, M5 bzw. M8	Linsenkopfschraube M4, M6 bzw. M10
N1	Senkkopfschraube M3, M5 bzw. M8	Senkkopfschraube M4, M6 bzw. M10
N1	Zylinderkopfschraube M3, M5 bzw. M8	Zylinderkopfschraube M4, M6 bzw. M10
N1	Gewindestift M3, M5 bzw. M8	Gewindestift M4, M6 bzw. M10
N1	Mutter M3, M5 bzw. M8	Mutter M4, M6 bzw. M10
N1	Unterlegscheibe M3, M5 bzw. M8	Unterlegscheibe M4, M6 bzw. M10
N2	Zylinderkopfschraube M3, M6 bzw. M10	Linsenkopfschraube M3, M6 bzw. M10
N2	Senkkopfschraube M3, M6 bzw. M10	Gewindestift M3, M6 bzw. M10

Prinzipiell besteht eine sehr große Vielfalt an möglichen Normteilen. Da die Berücksichtigung aller verfügbaren Normteile den Rahmen der vorliegenden Arbeit jedoch deutlich übersteigen würde, begrenzt sich die Auswahl der zu betrachtenden Normteile auf Schrauben, Muttern sowie die zugehörigen Unterlegscheiben. Diese werden branchenübergreifend in einer sehr großen Anzahl unterschiedlicher Produkte eingesetzt.

Weiterhin beschränkt sich die Auswahl der Schrauben auf die Variation ihrer Kopfform. Theoretisch zeichnen sich Schrauben neben dieser auch durch die Antriebsart (Innensechsrund, Innensechskant, etc.) aus. Die Ausweitung der Untersuchung auf beide Teilaspekte würde den Rahmen der vorliegenden Arbeit jedoch überschreiten. Aus diesem Grund berücksichtigt die Auswahl der Normteile ausschließlich den Einfluss verschiedener Kopfformen.

Wie in Abschnitt 4.2.3 erläutert, ist für die Überprüfung der Hypothese N1 zu untersuchen, ob Normteile vom gleichen Typ verwechselt werden, wenn sie in verschiedenen Größen auftreten. Als Untersuchungsobjekte kommen daher verschiedene, in der Praxis gängige Normteile in unterschiedlichen Größen zum Einsatz. Neben Linsenkopfschrauben nach ISO 7380-1 finden Senkkopfschrauben nach ISO 10642, Zylinderkopfschrauben nach ISO 4762, Gewindestifte nach ISO 4026, Muttern nach ISO 4032 und Unterlegscheiben nach ISO 7089 Verwendung. Um den Einfluss der absoluten Normteilgröße zu berücksichtigen, werden jeweils die Nennweiten M3 und M4, M5 und M6 sowie M8 und M10 kombiniert. So wird bspw. eine Linsenkopfschraube der Größe M3 mit einer Linsenkopfschraube der Größe M4 zusammengebracht und in einem Objekterkennungsmodell erfasst. Insgesamt werden je Versuch 18 Modelle trainiert, um die Hypothese N1 zu untersuchen.

Hypothese N2 berücksichtigt die Fragestellung, ob Schrauben mit einem ähnlichen Schraubenkopf verwechselt werden. Aus diesem Grund kommen Schraubenköpfe zum Einsatz, welche eine große Ähnlichkeit aufweisen. Da Zylinderkopfschrauben und Linsenkopfschrauben insbesondere in der Draufsicht ähnliche geometrische Merkmale aufweisen,

werden diese als erstes Objektpaar ausgewählt. Konkret finden Zylinderkopfschrauben nach ISO 4762 sowie Linsenkopfschrauben nach ISO 7380-1 Verwendung. Weiterhin werden Senkkopfschrauben und Gewindestifte kombiniert, da beide in der Regel während der Montage im Bauteil versenkt werden und ihr Kopf anschließend nur noch zum Teil sichtbar ist. Bei den ausgewählten Schrauben handelt es sich um Senkkopfschrauben nach ISO 10642 sowie um Gewindestifte nach ISO 4026.

Da nicht auszuschließen ist, dass die Nennweite der betrachteten Schrauben die mAP beeinflusst, werden die Schraubenpaare jeweils in den Größen M3, M6 und M10 betrachtet. So wird bspw. eine Linsenkopfschraube der Größe M3 mit einer Zylinderkopfschraube der gleichen Größe in einem Objekterkennungsmodell erfasst. Insgesamt werden demnach je Versuch sechs Modelle trainiert, um die Hypothese N2 zu untersuchen.

Da Schrauben in Baugruppen in der Regel im verschraubten Zustand auftreten, werden zusätzlich zu den ausgewählten Normteilen Objekte konstruiert, welche die Schrauben, Muttern und Unterlegscheiben aufnehmen. Dabei handelt es sich um einfache Aluminium-Würfel, welche Gewindebohrungen unterschiedlicher Nennweite aufweisen. Abbildung 5-17 zeigt beispielhaft eine Zylinderkopfschraube der Nennweite M6 im eingeschraubten Zustand. Die technischen Zeichnungen der Aluminiumwürfel können [137] entnommen werden.

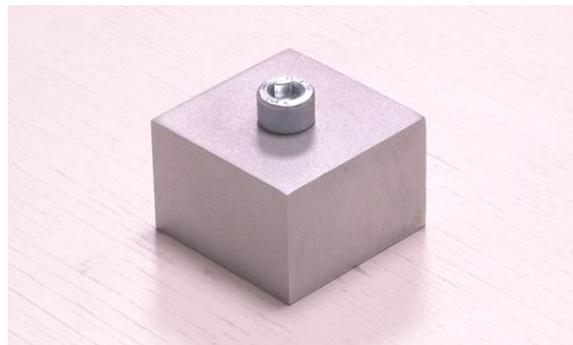


Abbildung 5-17: Aluminiumwürfel mit verschraubter Zylinderkopfschraube M6

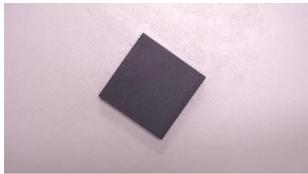
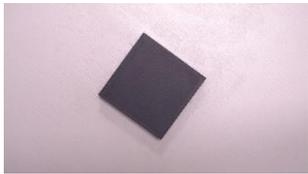
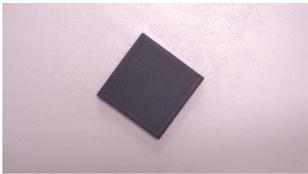
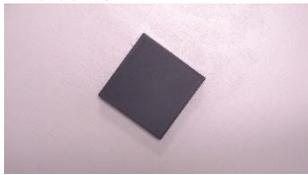
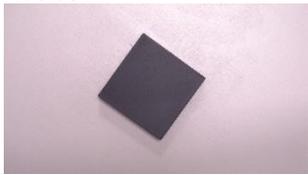
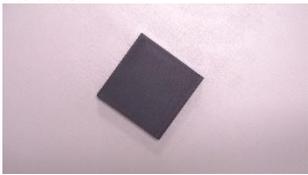
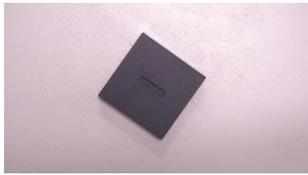
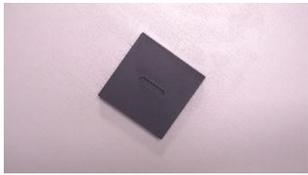
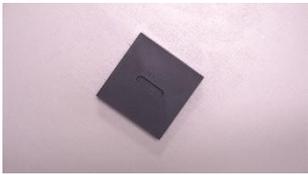
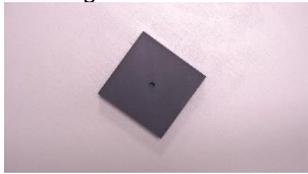
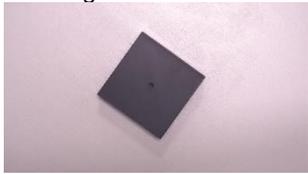
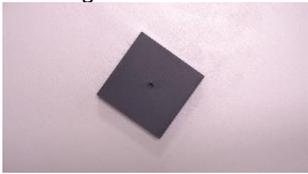
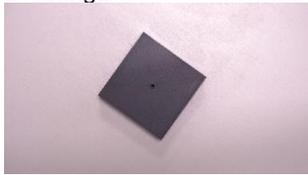
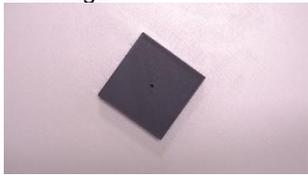
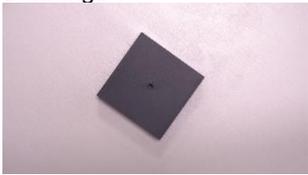
5.6.1.4 Untersuchungsobjekte zur Dimensionierung von Formelementen

Zur Überprüfung der Hypothesen zur Dimensionierung von Formelementen kommen insgesamt 15 Untersuchungsobjekte zum Einsatz, siehe Tabelle 5-17. Wie in Abschnitt 4.2.4 beschrieben, beziehen sich die Hypothesen dabei auf die Dimensionierung verschiedener Formelemente wie Fasen, Radien, Nuten und Bohrungen. Da die Geometrien von Formelementen in der Regel durch Normen eindeutig definiert sind, unterscheiden sich die einzelnen Untersuchungsobjekte ausschließlich hinsichtlich ihrer Dimensionierung.

Prinzipiell werden je Hypothese zwei Modelle trainiert und validiert, welche jeweils zwei Objekte mit unterschiedlich dimensionierten Formelementen beinhalten, siehe Tabelle 5-17. Als Grundform werden dabei quadratische Blöcke mit einer Seitenlänge von 60 mm

und einer Höhe von 15 mm genutzt, welche durch verschiedene, den Thesen entsprechende Formelemente wie Radien, Fasen, etc. ergänzt werden. Um zu untersuchen, welchen Einfluss die Größendifferenz der Formelemente zweier Objekte hat, ist ihr Größensprung im ersten Modell jeweils kleiner als im zweiten Modell. Da es sich bei den Hypothesen zur Dimensionierung von Formelementen um Zusammenhangshypothesen handelt, deren Überprüfung durch einen Vergleich der Untersuchungsergebnisse mit einem definierten Grenzwert erfolgt (siehe Abschnitt 4.2.4), werden keine Referenzobjekte benötigt.

Tabelle 5-17: Untersuchungsobjekte zur Dimensionierung von Formelementen

Hypothese	Objekte Modell 1		Objekte Modell 2
FE1 – Fase	Fase1.5x45 	Fase1x45 	Fase2x45 
FE2 – Radius	Radius1.5 	Radius1 	Radius2 
FE3 – Nut	Nut5 	Nut4 	Nut6 
FE4 – Bohrung Tiefe	BohrungT4.5 	BohrungT3 	BohrungT6 
FE5 – Bohrung Durchmesser	BohrungD3.5 	BohrungD3 	BohrungD4 

Die erste Hypothese zur Dimensionierung von Formelementen (FE1) bezieht sich auf die Dimensionierung von Fasen. Zur Überprüfung der These wird untersucht, ob Bauteile verwechselt werden, wenn die Maße ihrer Fasen sich nur geringfügig unterscheiden. Aus diesem Grund kommen drei Untersuchungsobjekte zum Einsatz, welche Fasen unterschiedlicher Größe aufweisen.

Die Grundlage für alle Objekte stellt wie oben beschrieben ein Quader dar, dessen Kanten jeweils mit einer Fase unterschiedlicher Ausprägung versehen ist. Während der Fasenwinkel bei allen drei Objekten 45° beträgt, unterscheiden sich die Fasen hinsichtlich ihrer Fassenbreite. Neben einer Breite von 1 mm findet eine Breite von 1,5 mm und 2 mm Verwendung. Um verschiedene Größensprünge zu untersuchen, wird neben einem Modell, welches Quader mit den Fassenbreiten 1 mm und 1,5 mm enthält, ein Modell mit den Fassenbreiten 1 mm und 2 mm erstellt.

Für die Überprüfung der Hypothese FE2 kommen Objekte zum Einsatz, dessen Kanten durch unterschiedliche Radien abgerundet sind. Dazu werden die Kanten des Grundquaders jeweils mit Radien von 1 mm, 1,5 mm bzw. 2 mm versehen. Wie bei den Objekten zur Hypothese FE1 enthalten die Modelle jeweils Untersuchungsobjekte mit den Radien 1 mm und 1,5 mm bzw. 1 mm und 2 mm, um verschiedene Größensprünge zu untersuchen.

Die dritte Hypothese zur Dimensionierung von Formelementen (FE3) bezieht sich auf die Dimensionierung von Nuten. Es wird untersucht, ob Bauteile verwechselt werden, wenn die Maße ihrer Nuten sich nur geringfügig unterscheiden. Aus diesem Grund finden drei Untersuchungsobjekte Verwendung, welche Nuten unterschiedlicher Größe aufweisen. Da bei der Dimensionierung von Nuten im Regelfall auf bestehende Normen zurückgegriffen wird, orientieren sich die Maße der angewendeten Nuten an der Norm DIN 6885-1. Während sich Nuttiefen und -breiten durch die Norm ergeben, ist ihre Länge in einem gewissen Rahmen selbst festzulegen. Tabelle 5-18 zeigt die Maße der Nuten der drei Untersuchungsobjekte zur Hypothese FE3. Aus den oben genannten Gründen werden im ersten Modell die Untersuchungsobjekte mit den Nutbreiten 4 mm und 5 mm detektiert, während das zweite Modell die Nutbreiten 4 mm und 6 mm beinhaltet.

Tabelle 5-18: Nutbreite, -tiefe und -länge der Untersuchungsobjekte (Hypothese FE3)

	Nut4	Nut5	Nut6
Breite	4 mm	5 mm	6 mm
Tiefe	2,5 mm	3 mm	3,5 mm
Länge	18 mm	18 mm	18 mm

Durch die Hypothesen F4 und F5 wird schließlich geprüft, ob Bauteile verwechselt werden, wenn sich die Dimensionierung ihrer Bohrungen nur geringfügig unterscheidet. Dabei wird der Einfluss der Bohrungstiefe (F4) und des Bohrungsdurchmessers (F5) untersucht. Um die beiden Aspekte zu überprüfen, kommen Untersuchungsobjekte zum Einsatz, welche Bohrungen unterschiedlicher Tiefe und Durchmesser aufweisen. Grundlage für alle Objekte ist der oben beschriebene Quader, in welchen mittig eine Bohrung eingebracht wird. Die Untersuchungsobjekte zur Hypothese FE4 weisen einen Bohrungsdurchmesser von 4 mm auf. Die Bohrungstiefe variiert zwischen 3 mm, 4 mm und 6 mm. Für die Untersuchungsobjekte zur Hypothese FE5 wird eine konstante Bohrungstiefe von 3 mm definiert. Der Bohrungsdurchmesser variiert zwischen den drei Objekten und beträgt 3 mm, 3,5 mm bzw. 4 mm.

Auch in Bezug auf Bohrungstiefe und -durchmesser soll der Einfluss verschiedener Größensprünge betrachtet werden. Aus diesem Grund werden die Objekte mit der geringsten Tiefe bzw. dem kleinsten Durchmesser jeweils mit den zweitgrößten und den größten Tiefen bzw. Durchmessern in einem Modell zusammengefasst.

Die vorgestellten Untersuchungsobjektpaare berücksichtigen zunächst nur zwei verhältnismäßig kleine Größendifferenzen zwischen den einzelnen Formelementen. Um den Versuchsaufwand zu reduzieren, wird im Rahmen der vorliegenden Arbeit zunächst festgestellt, ob Objekterkennungsalgorithmen in der Lage sind, Bauteile mit diesen sehr geringen Größendifferenzen zuverlässig zu detektieren. Ein solches Ergebnis wäre auf größere Differenzen übertragbar. Sollten die Versuchsergebnisse zeigen, dass die betrachteten Objekte verwechselt werden, könnten weitere Untersuchungsobjekte konzipiert werden. Bei diesen würden die Größendifferenzen sukzessive erhöht werden, um festzustellen, ab welcher Differenz eine zuverlässige Objektdetektion erreicht wird.

Im Gegensatz zu den Objekten für die Hypothesen zur Bauteil- und Baugruppengestaltung, beziehen sich die Hypothesen zur Dimensionierung von Formelementen ausschließlich auf geometrische Formen. Die Oberflächenbeschaffenheit sowie -farbe ist nicht relevant. Aufgrund der zur Verfügung stehenden Fertigungsmöglichkeiten werden die Objekte daher nicht gefräst, sondern mittels Schmelzschichtungsverfahren auf einem 3D-Drucker des Typs Ultimaker S5 [138] hergestellt. Die technischen Zeichnungen können [137] entnommen werden.

5.6.2 Erstellung der Trainings- und Validierungsbilder

Wie im Rahmen der Voruntersuchung festgelegt, werden im Rahmen der Hauptuntersuchung von allen zu betrachtenden Untersuchungsobjekten jeweils 240 Trainings- und 60 Validierungsbilder aus insgesamt zehn unterschiedlichen Perspektiven aufgezeichnet, siehe Abschnitt 5.5.2.1. Dazu wird ein schwingungsgedämpfter elektrischer Drehteller mit einem Durchmesser von 25 cm verwendet und die Auflagefläche durch eine mattweißlackierte Holzplatte (Durchmesser = 60 cm) vergrößert, siehe Abbildung 5-18. Weiterhin weist der Drehteller eine konstante Rotationsgeschwindigkeit von einer Umdrehung je Minute auf.

Die Aufnahme der Bilder erfolgt durch eine Kamera des Typs Sony Alpha 6300 mit einem SELP1650 Objektiv. Um den Bilderstellungsaufwand zu reduzieren, werden anstatt einzelner Bilder kompakte Videos aufgezeichnet. Um die Untersuchungsobjekte aus allen Richtungen aufzuzeichnen, werden basierend auf der Rotationsgeschwindigkeit des Drehtellers Videos mit einer Länge von 60 Sekunden erstellt. Durch das im Anhang auf Seite 145 aufgeführte Programm [139] erfolgt anschließend eine automatische Extraktion von einem Bild je Sekunde.

Wie in Abschnitt 5.5.2.1 erläutert, werden die Untersuchungsobjekte aus insgesamt zehn verschiedenen Perspektiven aufgezeichnet. Ihre Einstellung erfolgt durch ein Rollei C5i Dreibeinstativ mit integrierten Wasserwaagen. Damit die einzelnen Perspektiven der aufgezzeichneten Bilder für alle zu detektierenden Objekte übereinstimmen, wird das Stativ erst bewegt, nachdem alle relevanten Videos erstellt sind. Auch die Kameraeinstellungen (Belichtung, Zoom, Fokus, Auflösung, etc.) werden während der gesamten Aufzeichnung nicht verändert.



Abbildung 5-18: Versuchsaufbau zur Aufnahme der Trainings- und Validierungsbilder

Um die Lichtstärke konstant zu halten und natürliche Schwankungen des Sonnenlichtes auszuschließen, werden die Bilder in einem abgedunkelten Raum aufgezeichnet. Durch ein kalibriertes Luxmeter vom Typ LX-10 der Firma Voltcraft mit einem Messbereich von 0 bis 199900 lx wird die Lichtstärke zusätzlich überwacht und händisch aufgezeichnet. Dabei wird für jede Videoaufnahme die maximal auftretende Lichtstärke erfasst. Die Auswertung der aufgezeichneten Lichtstärke zeigt eine maximale Schwankung von 40 lux. Bei einem Mittelwert aller aufgezeichneten Lichtstärken von 714,7 lux ergibt sich eine Standardabweichung von 8,5 lux. Eine vollständige Übersicht der ermittelten Lichtstärken kann [140] entnommen werden.

Da die Modelle zur Bauteilgestaltung jeweils vier unterschiedliche Untersuchungsobjekte enthalten, werden für diese insgesamt 960 Trainings- und 240 Validierungsbilder je Modell aufgezeichnet. Im Gegensatz dazu beinhalten die Modelle zur Baugruppengestaltung jeweils fünf verschiedene Objekte. Diese werden jedoch in nur zwei unterschiedlichen Baugruppen zusammengefasst, sodass sich trotz der fünf zu detektierenden Einzelteile nur 480 Trainings- und 120 Validierungsbilder je Modell ergeben. Im Gegensatz zu den Bauteilen enthalten sie jedoch drei verschiedene Objektklassen je Bild. Die Modelle zur Auswahl von Normteilen und zur Dimensionierung von Formelementen enthalten jeweils zwei verschiedene, einzeln aufzuzeichnende Untersuchungsobjekte. Aus diesem Grund werden ebenfalls 480 Trainings- und 120 Validierungsobjekte je Modell aufgezeichnet. Im Vergleich zu den Bildern zur Baugruppengestaltung ist jedoch nur eine Objektklasse je Bild enthalten.

5.6.3 Annotation der Trainings- und Validierungsbilder

Für den Trainings- und Validierungsprozess müssen alle Trainings- und Validierungsbilder mit Informationen zur Klasse und Position der abgebildeten Objekte versehen werden. Die Bilder erhalten je abgebildetem Objekt eine Bounding-Box, welche die Objektposition im Bild wiedergibt und mit der entsprechenden Objektklasse versehen wird.

Um den Annotationsaufwand zu reduzieren, werden die Bilder für die Untersuchungen zur Bauteilgestaltung und Dimensionierung von Formelementen mithilfe eines Programms gelabelt, welches die Bounding-Boxes auf Grundlage der Objektkonturen selbstständig erstellt und die Objektklasse aus dem Dateinamen ableitet. Das verwendete Programm [141] ist dem Anhang auf Seite 146 zu entnehmen. Als Ergebnis gibt es eine JSON-Datei aus, welche die Klassen- und Positionsinformationen aller ausgewählten Bilder zusammenfasst. Diese wird im Folgenden als `coco.json` bezeichnet.

Auf die Bilder zur Baugruppengestaltung und Auswahl von Normteilen ist das beschriebene Programm nicht anwendbar, da es je Bild nur eine Bounding-Box erstellen kann bzw. eindeutige Konturen voraussetzt. Die Bilder zur Baugruppengestaltung beinhalten hingegen mehrere zu detektierende Objekte je Bild. Die Untersuchungsobjekte zur Auswahl von Normteilen sind, wie in Abschnitt 5.6.1.3 beschrieben, in Aluminiumwürfel eingeschraubt, sodass das Programm keine eindeutige Kontur erkennen kann. Aus diesem Grund kann für die Annotation dieser Bilder nicht auf das zuvor beschriebene Programm zurückgegriffen werden. Stattdessen werden die Bilder manuell gelabelt.

Für die manuelle Annotation kommt das Open Source Programm Labelme zum Einsatz. Der Download des Programms erfolgt gemeinsam mit der Implementierung des Objekterkennungsalgorithmus und wird im folgenden Abschnitt 5.6.4 erläutert. Abbildung 5-19 zeigt die Benutzeroberfläche des Programms. Im Zentrum steht das jeweils zu annotierende Bild, auf welchem händisch die zu detektierenden Objekte mit Rechtecken markiert und die entsprechenden Objektklassen mit diesen verknüpft werden. Für eine bessere Übersichtlichkeit ist dabei jeder Objektklasse eine Farbe zugeordnet.

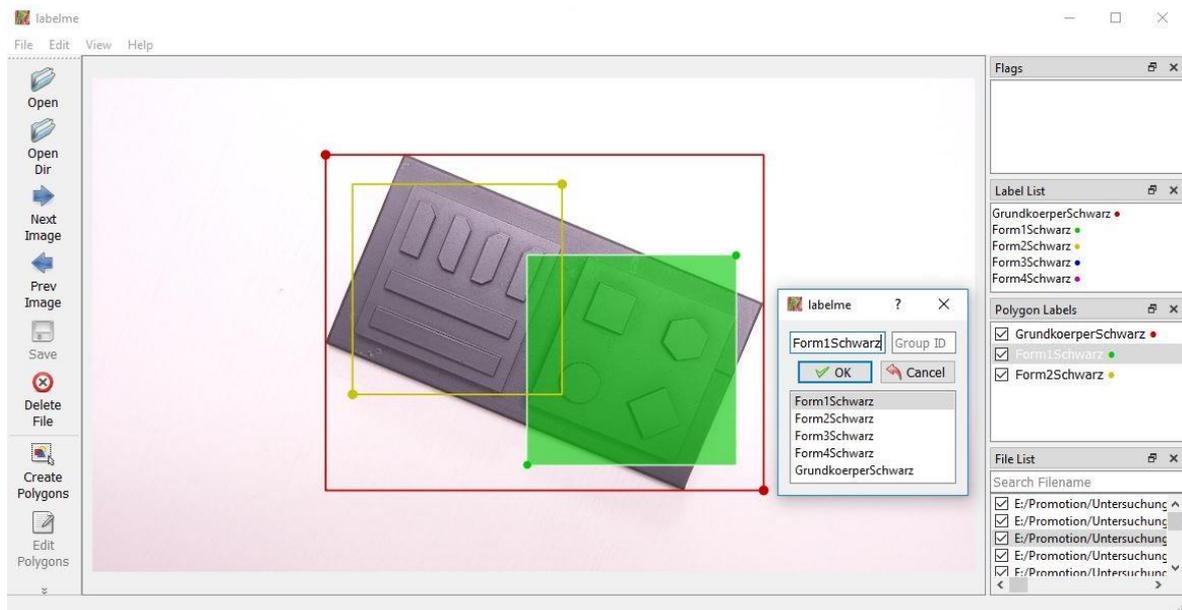


Abbildung 5-19: Benutzeroberfläche – Labelme

Anders als das zuvor beschriebene, automatische Annotationsprogramm speichert Labelme zu jedem Bild eine eigene JSON-Datei ab, welche die Klassen- und Positionsinformationen der abgebildeten Objekte enthält. Für den Trainings- und Validierungsprozess wird jedoch sowohl für den Trainings- als auch den Validierungsdatensatz eine JSON-Datei benötigt, welche die Klassen- und Positionsinformationen aller im Datensatz enthaltenen Bilder zusammenfasst. Aus diesem Grund findet das im Anhang auf Seite 149 aufgeführte Programm [142] Verwendung, welches mehrere JSON-Dateien in einer einzelnen Datei zusammenfasst und eine coco.json ausgibt.

5.6.4 Implementierung des Objekterkennungsalgorithmus

Um das in Kapitel 5.4 beschriebene Faster R-CNN anzuwenden, kommt die Programm-bibliothek Detectron2 in der Version 0.1.3 zum Einsatz. Zwar ist eine Implementierung von Detectron2 in Windows nicht vorgesehen, über verschiedene Workarounds kann dies jedoch ermöglicht werden [143]. Die Installation auf einer Windowsumgebung wird beispielhaft an einem System mit einer Nvidia Quadro P2000 Grafikkarte und Windows 10 Enterprise 2016 LTSC beschrieben.

Da für den Einsatz von Detectron2 verschiedene Python Libraries benötigt werden, ist es sinnvoll mit einem Paketverwaltungsprogramm zu arbeiten, welches gleichzeitig virtuelle Python Umgebungen zur Verfügung stellt. Im Rahmen der vorliegenden Arbeit wird zu diesem Zweck Anaconda [144] genutzt und in einer Python 3.6 Umgebung gearbeitet. Damit das Programm für den Trainingsprozess auf die Grafikkarte zurückgreifen kann, wird anschließend der von Nvidia bereitgestellte CUDA-Treiber [145] in der Version 10.1 installiert. Darüber hinaus wird das Programm Microsoft Visual Studio 2019 Community [146] heruntergeladen, um das Modul „Desktopentwicklung mit C++“ zu installieren. Dieses enthält die notwendigen MS Visual C++ build tools für die Nutzung von Detectron2. Darüber

hinaus wird die Open Source Computer Vision Bibliothek OpenCV [147] in der Version 4.4.0 mit CUDA-Unterstützung heruntergeladen. Anschließend werden PyTorch in der Version 1.4.0 und Torchvision in der Version 0.5.0 in der aktivierten Anaconda-Umgebung durch folgende Befehle installiert:

```
conda install pytorch==1.4.0 torchvision==0.5.0 -c pytorch
```

Danach erfolgt die Installation der FAIR Modulbibliothek FVCore durch folgenden Befehl:

```
pip install git+https://github.com/facebookresearch/fvcore
```

Außerdem werden die Module Cython, welches die Ausführung von C++ Code ermöglicht und PyCocoTools, welches die COCO API enthält, durch folgende Befehle installiert:

```
pip install cython
```

```
pip install
git+https://github.com/philferriere/cocoapi.git#subdirectory=PythonAPI
```

Um die installierten Module an die Windows-Umgebung anzupassen, müssen anschließend die in Tabelle 5-19 aufgeführten Dateien umgeschrieben werden.

Tabelle 5-19: Notwendige Code-Anpassung für Windows-Implementierung

Dateipfad	Programmzeile	Original-Code	Neuer Code
{Pfad zur Anaconda-Umgebung}\Lib\site-packages\torch\include\torch\csrc\jit\argument_spec.h	190	static constexpr size_t DEPTH_LIMIT = 128;	static const size_t DEPTH_LIMIT = 128;
{Pfad zur Anaconda-Umgebung}\Lib\site-packages\torch\include\pybind11\cast.h	1449	explicit operator type&() { return *(this->value); }	explicit operator type&() { return *((type*)this->value); }

Anschließend wird das detectron2 GitHub Repository in der aktivierten Anaconda-Umgebung über folgenden Befehl heruntergeladen:

```
git clone https://github.com/facebookresearch/detectron2.git
```

Außerdem wird das Visual Studio C++ Build Tool über den nachfolgenden Befehl aktiviert:

```
C:\Program Files (x86)\Microsoft Visual  
Studio\2019\Enterprise\VC\Auxiliary\Build\vcvars64.bat
```

Schlussendlich erfolgt die Installation von Detectron2 durch folgenden Befehl:

```
cd detectron2
```

Neben Detectron2 wird außerdem ein Hilfsprogramm installiert, welches zur Erstellung der Trainingsdaten notwendig ist. Zu diesem Zweck kommt das Open Source Tool Labelme [148] zum Einsatz, mit welchem die Position und Klasse von Objekten beschrieben und im COCO-Format ausgegeben wird. Die Installation des Tools erfolgt in der gleichen aktivierten Anaconda-Umgebung über folgenden Befehl:

```
pip install labelme
```

Das Programm kann anschließend über folgenden Befehl aufgerufen werden:

```
conda activate labelme
```

5.6.5 Training- und Validierung der Modelle

Unter Verwendung der vorbereiteten virtuellen Anaconda-Umgebung wird zum Trainieren und Validieren der Objekterkennungsmodelle folgendes Programm in Visual Studio Code [149] ausgeführt:

```
1  from detectron2.data.datasets import register_coco_instances
2  from torch.utils.cpp_extension import CUDA_HOME
3  from detectron2 import model_zoo
4  from detectron2.data import DatasetCatalog, MetadataCatalog
5  from detectron2.config import get_cfg
6  from detectron2.evaluation import inference_context
7  from detectron2.data import build_detection_train_loader
8  import os
9  from MyTrainer import MyTrainer
10
11 object_train = "Beispielmodell_train"
12 object_val = "Beispielmodell_val"
13
14 model_to_use = "COCO-Detection/faster_rcnn_R_50_FPN_3x.yaml"
15
16 image_path = "Dateipfad_Trainingsdatensatz"
17 coco_file = "Dateipfad_Trainingsdatensatz/coco.json"
18 register_coco_instances(object_train, {}, coco_file, image_path)
19
20 val_image_path = "Dateipfad_Validierungsdatensatz"
21 val_coco_file = "Dateipfad_Validierungsdatensatz/coco.json"
22 register_coco_instances(object_val, {}, val_coco_file, val_image_path)
23
24 dataset_dicts = DatasetCatalog.get(object_train)
25 dataset_metadata = MetadataCatalog.get(object_train)
```

```
26
27 cfg = get_cfg()
28
29 def run():
30     cfg.merge_from_file(model_zoo.get_config_file(model_to_use))
31     cfg.DATASETS.TRAIN = (object_train,)
32     cfg.DATALOADER.NUM_WORKERS = 2
33     cfg.MODEL.DEVICE = "cuda"
34     cfg.MODEL.WEIGHTS = model_zoo.get_checkpoint_url(model_to_use)
35     cfg.SOLVER.IMS_PER_BATCH = 1
36     cfg.SOLVER.BASE_LR = 0.00025
37     cfg.SOLVER.MAX_ITER = 40000
38     cfg.MODEL.ROI_HEADS.NUM_CLASSES = len(dataset_metadata.thing_classes)
39     cfg.OUTPUT_DIR = image_path
40     cfg.DATASETS.TEST = (object_val,)
41     cfg.TEST.EVAL_PERIOD = 1000
42     os.makedirs(cfg.OUTPUT_DIR, exist_ok=True)
43     trainer = MyTrainer(cfg)
44     trainer.resume_or_load(resume=False)
45     trainer.train()
46
47 if __name__ == '__main__':
48     run()
```

In den Programmzeilen 1 – 9 werden zunächst alle notwendigen Bibliotheken importiert. Anschließend wird in den Zeilen 11 und 12 der Name des jeweils zu trainierenden Modells angegeben. Als Nächstes (Programmzeile 14) erfolgt die Angabe des zu verwendenden Algorithmus. Wie in Abschnitt 5.4 erläutert, können mittels Detectron2 verschiedene vortrainierte Modelle angewendet werden. Im Rahmen der Untersuchung wird das Modell R50-FPN genutzt.

In den Programmzeilen 16 – 22 werden dann die Trainings- und Validierungsdatensätze registriert. Dazu ist jeweils der Dateipfad des Trainings- und Validierungsordners sowie der zugehörigen coco.json-Datei anzugeben. Um falsche Zuordnungen zu vermeiden, befinden sich die Trainings- bzw. Validierungsbilder sowie die zugehörigen JSON-Dateien dabei stets im selben Ordner. Die Programmzeilen 16 – 17 und 20 – 21 werden für jedes zu trainierende Modell angepasst, während die Zeilen 18 und 22 unverändert bleiben.

Durch die Zeilen 24 und 25 werden die Informationen aus dem zuvor definierten Trainingsdatensatz in den Datensatz- und Metadatenkatalog geladen. Dazu zählt insbesondere die Anzahl und Art der zu detektierenden Objektklassen. Anschließend wird in Zeile 27 die Standardkonfigurationsdatei [150] für Detectron2 geladen. Diese enthält alle für den Trainings- und Validierungsprozess notwendigen Informationen.

In den Programmzeilen 29 – 48 erfolgt schließlich der eigentliche Trainings- und Validierungsprozess. Dazu wird die für alle in Detectron2 enthaltenen Modelle geltende Standardkonfigurationsdatei durch weitere Informationen ergänzt bzw. spezifiziert. Durch

Zeile 30 wird diese zunächst mit einer speziell für das Modell R50-FPN geltenden Konfigurationsdatei [151] zusammengeführt. Anschließend wird der Name des Trainingsdatensatzes in die Konfigurationsdatei geschrieben, sodass beide miteinander verknüpft sind (Zeile 31).

Durch Zeile 32 wird definiert, dass die Trainings- und Validierungsdaten mit zwei CPU-Kernen gleichzeitig geladen werden. Je rechnerstärker das eingesetzte System ist, desto mehr Kerne können an dieser Stelle für den Trainings- und Validierungsprozess verwendet werden. Die Programmzeile 33 bestimmt das Gerät, auf welchem das neuronale Netz trainiert wird. Durch den Befehl *cuda* wird die Grafikkarte ausgewählt.

Um die Gewichte und Biase des vortrainierten Modells zu laden, wird anschließend die Programmzeile 34 ausgeführt. Hierbei werden die zu Beginn eines jeden Trainingsprozesses zufällig initialisierten Gewichte und Biase durch die des zuvor ausgewählten, vortrainierten Modells ersetzt.

Die Programmzeilen 35 – 37 definieren dann verschiedene Trainingsparameter wie die Anzahl der betrachteten Bilder je Iteration, die Basislernrate sowie die Anzahl der zu trainierenden Iterationen. Durch Zeile 38 wird anschließend die Anzahl der Objektklassen aus dem Metadatenkatalog geladen und durch Zeile 39 festgelegt, dass die Trainings- und Validierungsergebnisse im Trainingsordner zu speichern sind.

Wie in Zeile 31 für den Trainingsdatensatz wird in Zeile 40 der Name des Validierungsdatensatzes in die Konfigurationsdatei geschrieben. Anschließend wird durch Zeile 41 definiert, dass alle 1.000 Iterationen ein Validierungsprozess durchgeführt wird. Programmzeile 42 gewährleistet, dass im Fall eines nicht existierenden Ausgabeordners ein neuer Ordner erstellt wird.

Durch Programmzeile 43 wird schließlich das Trainingsprogramm geladen und die finale Konfigurationsdatei an dieses übergeben. Im Vergleich zum von Detectron2 vorgegebenen Standardtrainer wird hierbei ein angepasstes Trainingsprogramm (MyTrainer) durchgeführt. Dieses unterscheidet sich vom Standardtrainer jedoch lediglich dadurch, dass alle 1.000 Iterationen ein Validierungsprozess durchgeführt wird. Durch Zeile 44 wird schließlich ein neuronales Netz erstellt und diesem die vortrainierten Gewichte und Biase zugeordnet, sodass durch Zeile 45 der Trainingsprozess gestartet werden kann. Die Programmzeilen 47 – 48 nehmen keinen Einfluss auf den Trainings- und Validierungsprozess. Sie gewährleisten, dass das Programm nur im Mainthread ausgeführt wird.

Das Ergebnis eines Trainings- und Validierungsprozesses ist neben einem Modell im PTH-Format eine Tabelle im JSON-Format, welche alle Prozessparameter wie die AP je Objektklasse, die zusammenfassende mAP, den Trainingsverlust sowie den Validierungsverlust enthält. Für die Auswertung wird diese Datei in das Excel-Format konvertiert.

5.7 Versuchsauswertung

5.7.1 Versuchsergebnisse zur Bauteilgestaltung

Als Ergebnis der Hauptuntersuchung wird für jeden Versuch die maximal erreichte mAP erfasst. Da die Objekterkennungsmodelle alle 1.000 Iterationen validiert werden, weisen die Versuche insgesamt jeweils 40 mAP-Werte auf. Da die letzte, nach 40.000 Iterationen errechnete mAP aufgrund leichter Schwankungen nicht zwangsläufig den höchstmöglichen Wert beschreibt, wird die maximal erzielte mAP herangezogen. Dies gewährleistet die Vergleichbarkeit einzelner Objekterkennungsmodelle.

Wie in Abschnitt 5.5.1 definiert, werden je betrachtetem Modell insgesamt drei Versuche durchgeführt. Darüber hinaus erfolgt zusätzlich zu einer Versuchsreihe mit Trainings- und Validierungsbilder aus der Rundumsicht eine Versuchsreihe, in welcher die Untersuchungsobjekte ausschließlich aus der Draufsicht dargestellt sind, siehe Abschnitt 5.3. Auf Grundlage der maximalen mAP je Versuche, wird für jedes Modell das arithmetische Mittel \bar{x} nach Gl. 17, die Standardabweichung s nach Gl. 18 sowie die Varianz s^2 nach Gl. 19 ermittelt. Da die statistische Aussagekraft der Standardabweichung und Varianz aufgrund der geringen Stichprobenanzahl begrenzt ist, wird als zusätzliches Maß für die Streuung der Messwerte die relative Spannweite \bar{R} nach Gl. 20 berechnet.

$$\bar{x} = \frac{1}{n} \sum_{a=1}^n x_a \quad \text{Gl. 17}$$

$$s = \sqrt{\frac{1}{n-1} \sum_{a=1}^n (x_a - \bar{x})^2} \quad \text{Gl. 18}$$

$$s^2 = \frac{1}{n-1} \sum_{a=1}^n (x_a - \bar{x})^2 \quad \text{Gl. 19}$$

$$\bar{R} = \frac{x_{max} - x_{min}}{\bar{x}} \cdot 100\% \quad \text{Gl. 20}$$

Tabelle 5-20 zeigt die im Rahmen der Versuche zur Bauteilgestaltung aus der Rundumsicht maximal erzielte mAP je Versuch. Zunächst wird deutlich, dass die erfassten Versuchsergebnisse eine insgesamt sehr geringe Variabilität aufweisen. So beträgt die maximale relative Streubreite 5,33 % (siehe Modell 2-2.3) und sechs von neun Versuchsreihen weisen eine relative Streubreite kleiner als 2 % auf.

In Bezug auf die Qualität der Objekterkennung zeigt sich, dass alle trainierten Objekterkennungsmodelle sehr hohe Erkennungsgenauigkeiten erzielen. Die höchste mittlere mAP wird dabei mit einem Wert von 99,82 % durch das Modell 1-4Schwarz erreicht. Die niedrigste mittlere mAP weist das Modell 3-3.3 auf. Mit einem Wert von 96,80 % liegt die Erkennungsgenauigkeit jedoch auch für dieses Modell deutlich über dem in Abschnitt 4.2.3 definierten Grenzwert von 80 %.

Tabelle 5-20: Versuchsergebnisse zur Bauteilgestaltung – Rundumsicht

These	Modell	V1*	V2*	V3*	\bar{x}^*	s	s^2	\bar{R}^*
Referenz	1-4	99,72	98,73	97,11	98,52	1,32	1,73	2,65
BT1	1-4Spiegelnd	99,92	99,48	99,18	99,53	0,37	0,14	0,75
BT2	1-4Schwarz	99,89	99,95	99,63	99,82	0,17	0,03	0,32
BT3	1-4Weich	99,71	99,07	99,40	99,39	0,32	0,10	0,64
BT4	1-1.3	99,64	99,55	99,13	99,44	0,27	0,07	0,51
	2-2.3	99,49	98,49	94,30	97,42	2,75	7,58	5,33
	3-3.3	97,92	96,85	95,64	96,80	1,14	1,30	2,35
	4-4.3	99,86	97,99	98,07	98,64	1,06	1,12	1,90
BT5	5-8	99,88	98,76	98,45	99,03	0,75	0,57	1,45
* in %								

Um nachzuweisen, dass die erzielten Erkennungsgenauigkeiten den Schwellenwert von 80 % statistisch signifikant übersteigen, wird mithilfe der Statistik-Software SPSS Statistics des Unternehmens IBM [152] ein Einstichproben-Wilcoxon-Test auf die Versuchsergebnisse angewendet. Dieser dient als nicht-parametrische Alternative zum Einstichproben t-Test und kann im Gegensatz zu diesem auf nicht normalverteilte Messwerte angewendet werden [153]. Eine Normalverteilung kann nicht angenommen werden, da die Versuchsergebnisse durch die maximal erzielbare mAP von 100% versuchsbedingt in ihrer Höhe begrenzt sind.

Der Einstichproben-Wilcoxon-Test prüft, ob sich der Median einer vorliegenden Verteilung von einem bestimmten Wert unterscheidet. Verglichen mit dem Referenzmedian von 80 % weisen die Versuchsergebnisse zur Bauteilgestaltung aus der Rundumsicht ($n = 27$) einen signifikant höheren Median von 99,18 % auf. ($Z = 4,541$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,874$ und entspricht einem starken Effekt.

Um zu überprüfen, ob die Versuchsergebnisse der Referenzversuchsreihe statistisch signifikant von den Ergebnissen der den Hypothesen entsprechenden Versuchsreihen abweichen, wird der Mann-Whitney-U-Test angewendet. Dieser wird als nicht-parametrische

Alternative zum ungepaarten t-Test angewendet und eignet sich zur Analyse nicht normalverteilter Versuchsergebnisse [153]. Der Test prüft, ob die zentralen Tendenzen zweier unabhängiger Stichproben verschieden sind. Auch für den Mann-Whitney-U-Test findet die Software SPSS Statistics Verwendung.

Tabelle 5-21 zeigt die Ergebnisse der Mann-Whitney-U-Tests für die Hypothesen zur Bauteilgestaltung. Um die verschiedenen Hypothesen einzeln zu testen, werden die Versuchsergebnisse des Referenzmodells (Stichprobe 1) den mAP-Werten der Hypothesenmodelle (Stichprobe 2) gegenübergestellt. Neben der U -Statistik zeigt die Tabelle die errechnete Z -Statistik sowie die zweiseitige, exakte Signifikanz p . Ausgehend von einem angestrebten Signifikanzniveau von $\alpha = 0,05$ kann damit für keine der betrachteten Hypothesen ein signifikanter Unterschied zwischen den Stichproben festgestellt werden, da für alle gilt, dass $p > \alpha$. Die in Abschnitt 4.2.1 aufgestellten Nullhypothesen zur Bauteilgestaltung können damit nicht abgelehnt werden.

Tabelle 5-21: Mann-Whitney-U-Test – Bauteilgestaltung – Rundumsicht

These	Stichprobe 1	Stichprobe 2	U	Z	p
BT1	1-4	1-4Spiegelnd	2	-1,091	0,400
BT2	1-4	1-4Schwarz	1	-1,528	0,200
BT3	1-4	1-4Weich	3	-0,655	0,700
BT4	1-4	1-1.3; 2-2.3; 3-3.3; 4-4.3	15	-0,433	0,734
BT5	1-4	5-8	3	-0,655	0,700

Werden für den Trainings- und Validierungsprozess ausschließlich Bilder aus der Draufsicht genutzt, ergeben sich im Vergleich zur Rundumsicht noch höhere Erkennungsgenauigkeiten, siehe Tabelle 5-22. Im Rahmen der Hauptuntersuchung zur Bauteilgestaltung erzielen in diesem Fall 25 von 27 trainierten Objekterkennungsmodellen die maximal erreichbare mAP von 100 %. Lediglich die Modelle 1-4Spiegelnd und 4-4.3 erreichen in je einem Versuch eine niedrigere mAP. Mit einem Wert von 99,74 % bzw. 99,76 % ist diese jedoch ebenfalls sehr hoch.

Tabelle 5-22: Versuchsergebnisse zur Bauteilgestaltung – Draufsicht

These	Modell	V1*	V2*	V3*	\bar{x}^*	s	s^2	\bar{R}^*
Referenz	1-4	100	100	100	100	0	0	0
BT1	1-4SpiegelInd	99,74	100	100	99,91	0,15	0,02	0,26
BT2	1-4Schwarz	100	100	100	100	0	0	0
BT3	1-4Weich	100	100	100	100	0	0	0
BT4	1-1.3	100	100	100	100	0	0	0
	2-2.3	100	100	100	100	0	0	0
	3-3.3	100	100	100	100	0	0	0
	4-4.3	99,76	100	100	99,92	0,14	0,02	0,24
BT5	5-8	100	100	100	100	0	0	0
* in %								

Auch für die Untersuchung zur Bauteilgestaltung aus der Draufsicht zeigt der Einstichproben-Wilcoxon-Test, dass die Versuchsergebnisse ($n = 27$) im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (100 %) aufweisen. ($Z = 4,972$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,957$ und entspricht ebenfalls einem starken Effekt.

Die Anwendung des Mann-Whitney-U-Tests ist für die Ergebnisse aus der Draufsicht nicht sinnvoll, da nur zwei von 27 Messwerten von der maximal erreichbaren mAP von 100 abweichen. Der Mann-Whitney-U-Test setzt die Bildung von Rangfolgen voraus, welche bei identischen Werten nicht möglich ist. In Bezug auf die Hypothesen zeigen jedoch auch die absoluten Werte aus Tabelle 5-22, dass keine erheblichen Unterschiede zwischen den Referenz- und Hypothesenmodellen bestehen. Damit sind die Nullhypothesen zur Bauteilgestaltung auch für den Fall nicht abzulehnen, dass die Trainings- und Validierungsbilder aus der Draufsicht aufgezeichnet werden.

5.7.2 Versuchsergebnisse zur Baugruppengestaltung

Die Versuchsergebnisse zur Baugruppengestaltung zeigen sehr hohe Erkennungsgenauigkeiten für alle untersuchten Modelle. Für die Versuche, welche Trainings- und Validierungsbilder aus der Rundumsicht nutzen, enthält Tabelle 5-23 die maximale mAP je Versuch sowie das arithmetische Mittel, die Standardabweichung, die Varianz und die relative Streubreite je Modell.

Es wird deutlich, dass die Versuchsergebnisse eine sehr geringe Variabilität aufweisen. So beträgt die maximale relative Streubreite 2,22 %, während alle weiteren Versuchsreihen eine relative Streubreite kleiner als 2 % aufweisen. Darüber hinaus zeigt sich, dass alle trainierten Objekterkennungsmodelle eine mAP von mindestens 95 % erzielen.

Tabelle 5-23: Versuchsergebnisse zur Baugruppengestaltung – Rundumsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
BG1	EF_Alu	96,02	95,50	96,13	95,88	0,34	0,11	0,66
	EF_Schwarz	96,42	96,72	96,67	96,60	0,16	0,02	0,30
	EF_Weiss	95,43	94,52	95,54	95,16	0,56	0,31	1,07
	EF_Rot	95,51	95,99	95,59	95,70	0,26	0,07	0,50
	EF_Gruen	94,57	96,70	95,85	95,70	1,07	1,15	2,22
	EF_Blau	97,02	97,37	96,20	96,86	0,60	0,36	1,21
	MF_Alu	97,71	98,10	97,94	97,92	0,20	0,04	0,40
	MF_Schwarz	97,05	97,83	97,62	97,50	0,40	0,16	0,80
	MF_Weiss	97,52	97,73	97,58	97,61	0,11	0,01	0,21
	MF_Rot	97,53	97,88	97,46	97,63	0,23	0,05	0,43
	MF_Gruen	96,59	96,74	96,64	96,65	0,08	0,01	0,15
	MF_Blau	98,04	97,61	97,56	97,74	0,27	0,07	0,50

* EF = Einfarbig; MF = Mehrfarbig; ** in %

Ein Einstichproben-Wilcoxon-Test für alle Versuchsergebnisse ($n = 36$) zeigt, dass die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (96,73 %) aufweisen. ($Z = 5,232$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,872$ und entspricht einem starken Effekt.

Weiterhin zeigen die Versuchsergebnisse, dass die einfarbigen Modelle im Durchschnitt niedrigere Erkennungsgenauigkeiten aufweisen, als die Modelle, welche Baugruppen aus verschiedenfarbigen Bauteile beinhalten. So ergibt sich für die Versuchsergebnisse aller einfarbigen Modelle ein arithmetisches Mittel von 95,99 %, während dieses für alle mehrfarbigen Modelle 97,51 % beträgt, siehe Abbildung 5-20.

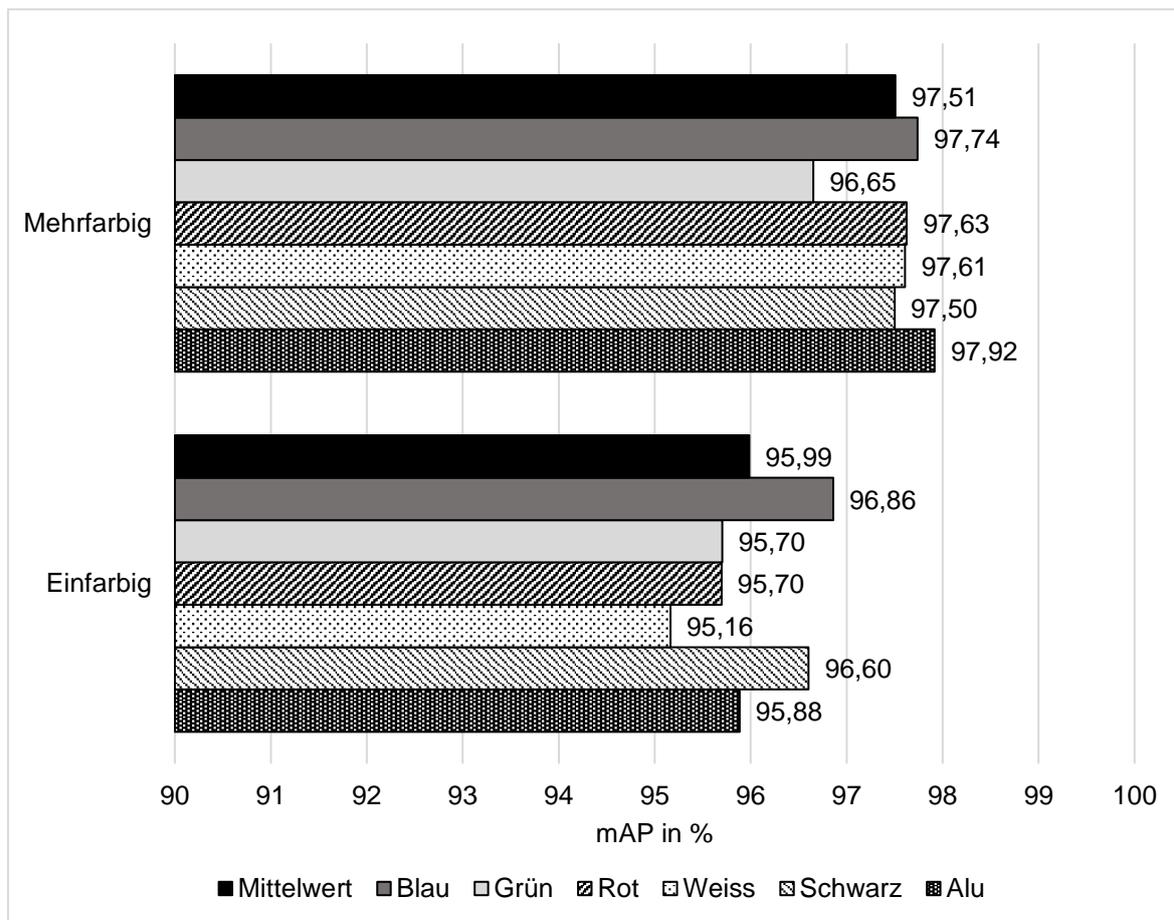


Abbildung 5-20: Mittlere mAP für einfarbige und mehrfarbige Modelle – Rundumsicht

Ein Mann-Whitney-U-Test, welcher die Versuchsergebnisse aller einfarbigen Modelle ($n = 18$) den Erkennungsgenauigkeiten aller mehrfarbigen Modelle ($n = 18$) gegenüberstellt, bestätigt diese Beobachtung. Der Test zeigt einen signifikanten Unterschied mit einem starken Effekt nach Cohen. ($U = 13$; $Z = -4,714$; $p < 0,001$; $r = 0,786$) Die in Abschnitt 4.2.2 aufgestellte Nullhypothese ($mAP_{BG-Gleich} \geq mAP_{BG-NichtGleich}$) kann damit abgelehnt und die Alternativhypothese ($mAP_{BG-Gleich} < mAP_{BG-NichtGleich}$) angenommen werden.

Kommen für den Trainings- und Validierungsprozess ausschließlich Bilder aus der Draufsicht zum Einsatz, ergeben sich im Vergleich zur Rundumsicht noch höhere Erkennungsgenauigkeiten, siehe Tabelle 5-24. So erzielen 16 von 36 trainierten Objekterkennungsmodellen die maximal erreichbare mAP von 100 % und das arithmetische Mittel aller Versuchsergebnisse beträgt 99,11 %. Im Vergleich dazu erzielen die Ergebnisse aus der Rundumsicht einen Mittelwert von 96,75 %.

Tabelle 5-24: Versuchsergebnisse zur Baugruppengestaltung – Draufsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
BG1	EF_Alou	97,92	98,61	99,81	98,78	0,96	0,91	1,91
	EF_Schwarz	100	99,70	100	99,90	0,17	0,03	0,30
	EF_Weiss	93,33	93,78	98,89	95,33	3,09	9,55	5,84
	EF_Rot	97,88	99,77	100	99,22	1,16	1,35	2,13
	EF_Gruen	98,92	99,58	98,86	99,12	0,40	0,16	0,72
	EF_Blau	96,40	100	100	98,80	2,08	4,32	3,64
	MF_Alou	97,66	100	100	99,22	1,35	1,83	2,36
	MF_Schwarz	100	100	100	100	0	0	0,00
	MF_Weiss	100	99,83	100	99,94	0,10	0,01	0,17
	MF_Rot	99,75	99,74	100	99,83	0,15	0,02	0,26
	MF_Gruen	99,70	99,52	100	99,74	0,24	0,06	0,48
	MF_Blau	98,39	100	100	99,46	0,93	0,86	1,62

* EF = Einfarbig; MF = Mehrfarbig; ** in %

Auch für die Untersuchung zur Baugruppengestaltung aus der Draufsicht zeigt der Einstichproben-Wilcoxon-Test, dass die Versuchsergebnisse (n = 36) im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (99,79 %) aufweisen. ($Z = 5,287$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,881$ und entspricht einem starken Effekt.

Wie die Versuchsergebnisse aus der Rundumsicht, zeigen auch die Ergebnisse aus der Draufsicht, dass die einfarbigen Modelle im Durchschnitt niedrigere Erkennungsgenauigkeiten aufweisen, als die Modelle welche Baugruppen aus verschiedenfarbigen Bauteile beinhalten. So ergibt sich für die Versuchsergebnisse aller einfarbigen Modelle ein arithmetisches Mittel von 98,52 %, während dieses für alle mehrfarbigen Modelle 99,70 % beträgt, siehe Abbildung 5-21.

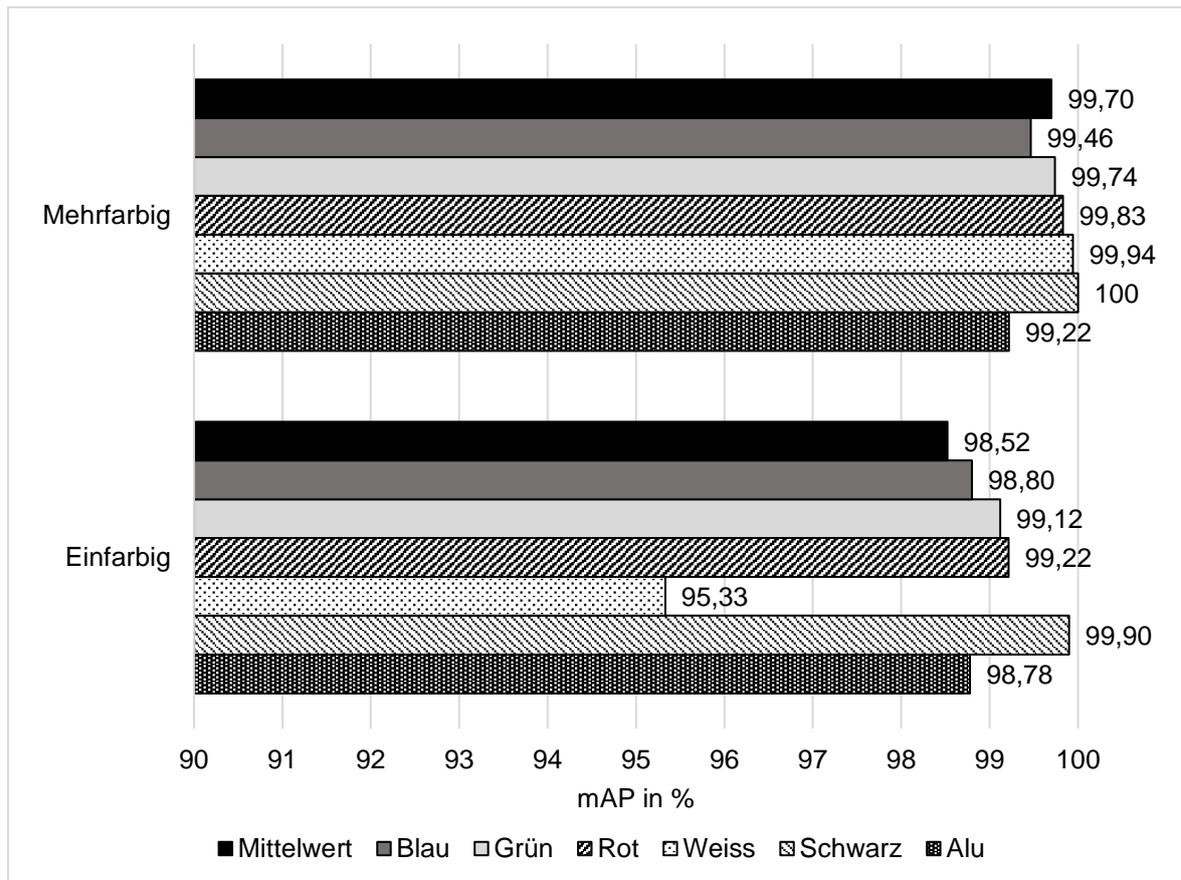


Abbildung 5-21: Mittlere mAP für einfarbige und mehrfarbige Modelle – Draufsicht

Ein Mann-Whitney-U-Test, welcher die Erkennungsgenauigkeiten der einfarbigen Modelle ($n = 18$) den mAP-Werten aller mehrfarbigen Modelle ($n = 18$) gegenüberstellt, bestätigt diese Beobachtung. Der Test zeigt einen signifikanten Unterschied, welcher im Vergleich zu den Ergebnissen aus der Draufsicht jedoch nur eine mittlere Effektstärke nach Cohen [154] aufweist. ($U = 91$; $Z = -2,352$; $p = 0,018$; $r = 0,392$)

Die in Abschnitt 4.2.2 aufgestellte Nullhypothese ($mAP_{BG-Gleich} \geq mAP_{BG-NichtGleich}$) kann damit auch für den Fall abgelehnt und die Alternativhypothese ($mAP_{BG-Gleich} < mAP_{BG-NichtGleich}$) angenommen werden, dass die Trainings- und Validierungsbilder ausschließlich aus der Draufsicht aufgezeichnet werden.

5.7.3 Versuchsergebnisse zur Auswahl von Normteilen

Tabelle 5-25 zeigt die im Rahmen der Versuche zur Auswahl von Normteilen aus der Rundumsicht maximal erzielte mAP je Versuch. Zunächst wird deutlich, dass die erfassten Versuchsergebnisse eine höhere Variabilität aufweisen als die Messwerte zur Bauteil- und Baugruppengestaltung. So beträgt die maximale relative Streubreite 13,95 % und neun von 24 Versuchsreihen weisen eine relative Streubreite größer als 5 % auf. Darüber hinaus lässt sich beobachten, dass die relativen Streubreiten für Versuchsreihen mit versenkten Normteilen (Gewindestift und Senkkopfschraube) höhere Werte annehmen als für Versuchsreihen, welche nicht-versenkte Normteile beinhalten.

In Bezug auf die erzielten Erkennungsgenauigkeiten zeigt sich, dass auch die Versuche zur Auswahl von Normteilen aus der Rundumsicht insgesamt hohe mAP-Werte erzielen. Zwar weisen sie im Vergleich zu den Versuchsergebnissen zur Bauteil- und Baugruppengestaltung teilweise deutlich niedrigere Werte auf, jedoch erreichen alle betrachteten Objekterkennungsmodelle eine maximale mAP von mindestens 80 % und liegen damit über dem in Abschnitt 4.2.3 definierten Schwellenwert. Die niedrigste mittlere mAP ergibt sich mit einem Wert von 81,64 % für das Modell GS_M3 + GS_M4, während das Modell LK_M5 + LK_M6 mit einer mittleren mAP von 99,55 % die höchste Erkennungsgenauigkeit aufweist.

Tabelle 5-25: Versuchsergebnisse zur Auswahl von Normteilen – Rundumsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
N1	ZK_M3 + ZK_M4	98,04	97,38	99,00	98,14	0,81	0,66	1,65
	ZK_M5 + ZK_M6	98,88	97,90	100	98,92	1,05	1,11	2,13
	ZK_M8 + ZK_M10	99,44	97,52	99,70	98,89	1,19	1,41	2,20
	LK_M3 + LK_M4	97,52	97,31	98,12	97,65	0,42	0,18	0,83
	LK_M5 + LK_M6	99,75	99,15	99,75	99,55	0,34	0,12	0,60
	LK_M8 + LK_M10	99,65	99,34	99,57	99,52	0,16	0,03	0,31
	SK_M3 + SK_M4	92,21	88,80	94,03	91,68	2,65	7,05	5,71
	SK_M5 + SK_M6	94,80	84,70	95,35	91,62	5,99	35,92	11,62
	SK_M8 + SK_M10	96,42	88,19	95,59	93,40	4,53	20,50	8,81
	GS_M3 + GS_M4	82,18	79,32	83,41	81,64	2,10	4,41	5,01
	GS_M5 + GS_M6	87,23	82,46	82,79	84,16	2,67	7,11	5,67
	GS_M8 + GS_M10	91,09	79,12	87,08	85,77	6,09	37,10	13,95
	M_M3 + M_M4	96,32	95,96	98,01	96,76	1,09	1,20	2,12
	M_M5 + M_M6	98,10	98,22	98,97	98,43	0,47	0,22	0,89
	M_M8 + M_M10	99,55	97,72	99,58	98,95	1,06	1,13	1,88
	U_M3 + U_M4	93,67	92,44	94,62	93,58	1,09	1,19	2,33
	U_M5 + U_M6	95,68	96,79	98,04	96,84	1,18	1,40	2,44
	U_M8 + U_M10	95,36	96,91	98,51	96,93	1,58	2,49	3,26

Fortsetzung auf der nächsten Seite

Fortsetzung Tabelle 5-25

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
N2	ZK_M3 + LK_M3	94,72	96,34	97,49	96,18	1,39	1,94	2,88
	SK_M3 + GS_M3	86,90	81,86	87,92	85,56	3,24	10,52	7,08
	ZK_M6 + LK_M6	98,57	97,79	99,72	98,69	0,97	0,95	1,96
	SK_M6 + GS_M6	90,88	84,70	90,07	88,55	3,36	11,26	6,97
	ZK_M10 + LK_M10	100	98,20	100	99,40	1,04	1,08	1,81
	SK_M10 + GS_M10	91,26	86,38	90,70	89,45	2,67	7,14	5,46

* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe; ** in %

Um zu überprüfen, ob die Erkennungsgenauigkeiten statistisch signifikant über dem definierten Grenzwert von 80 % liegen, kommt der Einstichproben-Wilcoxon-Test zur Anwendung. Dabei wird zwischen den in Abschnitt 4.2.3 aufgestellten Hypothesen N1 ($mAP_{\text{Größe}} \geq 80\%$) und N2 ($mAP_{\text{Schraubenkopf}} \geq 80\%$) differenziert. Für die Überprüfung der ersten Hypothese zur Normteilauswahl werden alle Versuchsreihen einbezogen, welche Normteile vom gleichen Typ in unterschiedlichen Größen beinhalten ($n = 54$). Der Test zeigt, dass die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (97,11 %) aufweisen. ($Z = 6,367$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,866$ und entspricht einem starken Effekt.

Zur Überprüfung der Hypothese N2 werden alle Versuchsreihen einbezogen, welche Schrauben der gleichen Größe mit einem ähnlichen Kopf beinhalten ($n = 18$). Auch für diese Stichprobe zeigt der Test, dass die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (92,99 %) aufweisen. ($Z = 3,724$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,878$ und entspricht einem starken Effekt. Die Nullhypothesen N1 und N2 können damit nicht abgelehnt werden.

Neben den Aussagen zu den im Vorfeld aufgestellten Hypothesen, lassen sich weitere Beobachtungen aus den Versuchsergebnissen ableiten. So wird bspw. deutlich, dass Modelle mit Normteilen kleiner Nennweiten im Durchschnitt niedrigere Erkennungsgenauigkeiten aufweisen, als Modelle mit größeren Nennweiten. So ergibt sich für die Versuchsergebnisse aller Modelle, welche Normteile der Nennweite M3 aufweisen, ein arithmetisches Mittel von 90,87 %, während dieses für Modelle mit Normteilen der Nennweite M6 bzw. M10 einen Wert von 93,62 % bzw. 94,42 % beträgt, siehe Abbildung 5-22.

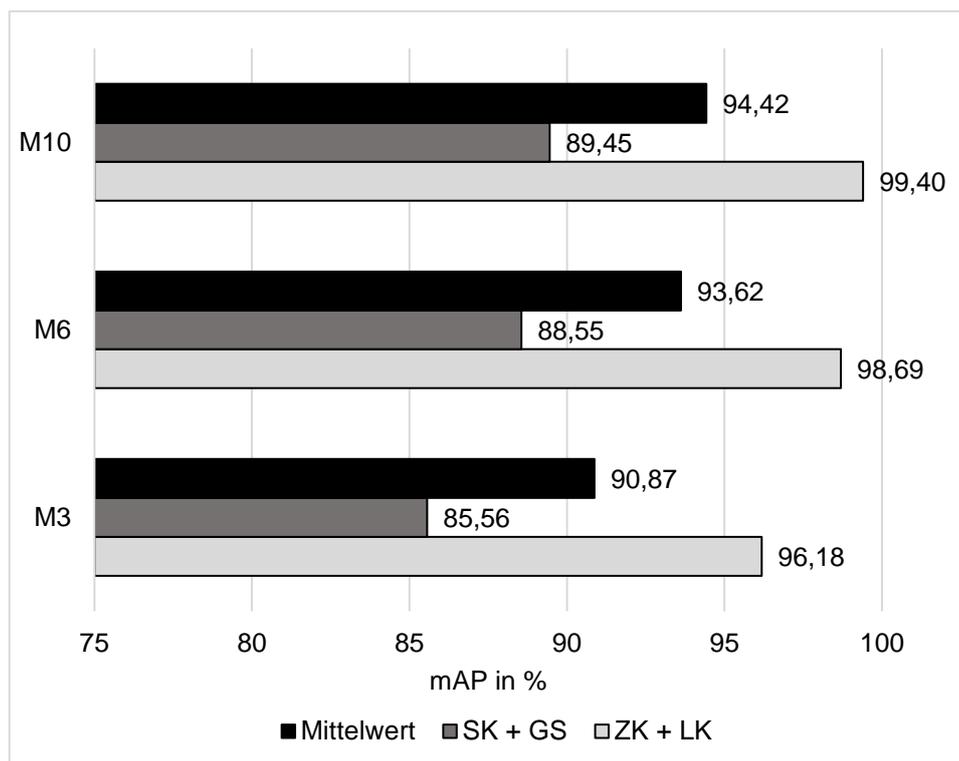


Abbildung 5-22: Mittlere mAP der Modelle SK+GS und ZK+LK in Abhängigkeit der Normteilgröße – Rundumsicht

Um zu überprüfen, ob diese Beobachtung statistisch signifikant ist, wird der Mann-Whitney-U-Test für verschiedene Normteilgrößenkombinationen angewendet. Neben den Versuchsergebnisse der Größen M3 und M6, werden die erzielten Erkennungsgenauigkeiten der Größen M3 und M10 sowie M6 und M10 gegenübergestellt. Tabelle 5-26 zeigt die Ergebnisse der Mann-Whitney-U-Tests für die verschiedenen Größenkombinationen. Ausgehend von einem angestrebten Signifikanzniveau von $\alpha = 0,05$ kann für keine der betrachteten Kombinationen ein signifikanter Unterschied zwischen den Stichproben festgestellt werden, da für alle gilt, dass $p > \alpha$.

Tabelle 5-26: Mann-Whitney-U-Test 1 – Auswahl von Normteilen – Rundumsicht

Stichprobe 1* (n=6)	Stichprobe 2* (n=6)	<i>U</i>	<i>Z</i>	<i>p</i>
SK_M3 + GS_M3; ZK_M3 + LK_M3	SK_M6 + GS_M6; ZK_M6 + LK_M6	11	-1,121	0,310
SK_M3 + GS_M3; ZK_M3 + LK_M3	SK_M10 + GS_M10; ZK_M10 + LK_M10	11	-1,123	0,290
SK_M6 + GS_M6; ZK_M6 + LK_M6	SK_M10 + GS_M10; ZK_M10 + LK_M10	14	-0,642	0,554

* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe

Der Vergleich der Modelle mit Normteilen der gleichen Art zeigt einen ähnlichen Effekt. So weisen die Modelle mit Normteilen der Nennweiten M3 und M4 ein arithmetisches Mittel von 93,24 % auf, während dieses für die Modelle mit Normteilen der Nennweiten M5 und M6 bzw. M8 und M10 einen Mittelwert von 94,92 % bzw. 95,58 % ergibt, siehe Abbildung 5-23.

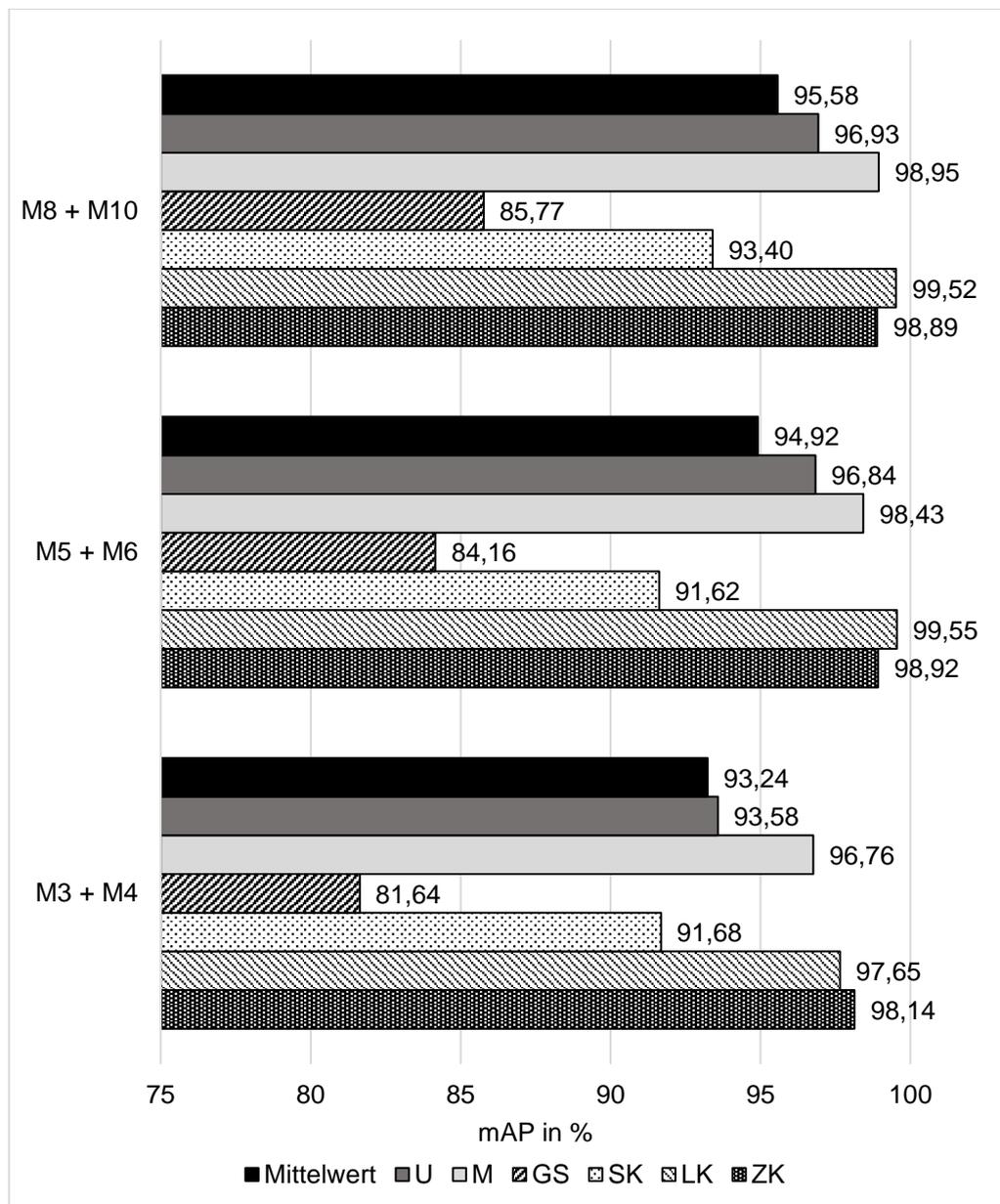


Abbildung 5-23: Mittlere mAP der Modelle U, M, GS, SK, LK und ZK in Abhängigkeit der Normteilgröße – Rundumsicht

Für verschiedene Größenkombinationen durchgeführte Mann-Whitney-U-Tests zeigen jedoch auch für diese Modelle, dass der beobachtete Unterschied nicht signifikant ist. Die Ergebnisse der Tests können Tabelle 5-27 entnommen werden. Es wird deutlich, dass alle erzielten p -Werte oberhalb des angestrebten Signifikanzniveaus von 0,05 liegen.

Tabelle 5-27: Mann-Whitney-U-Test 2 – Auswahl von Normteilen – Rundumsicht

Stichprobe 1* (n=18)	Stichprobe 2* (n=18)	<i>U</i>	<i>Z</i>	<i>p</i>
ZK_M3 + ZK_M4; LK_M3 + LK_M4; SK_M3 + SK_M4; GS_M3 + GS_M4; M_M3 + M_M4; U_M3 + U_M4	ZK_M5 + ZK_M6; LK_M5 + LK_M6; SK_M5 + SK_M6; GS_M5 + GS_M6; M_M5 + M_M6; U_M5 + U_M6	107	-1,740	0,085
ZK_M3 + ZK_M4; LK_M3 + LK_M4; SK_M3 + SK_M4; GS_M3 + GS_M4; M_M3 + M_M4; U_M3 + U_M4	ZK_M8 + ZK_M10; LK_M8 + LK_M10; SK_M8 + SK_M10; GS_M8 + GS_M10; M_M8 + M_M10; U_M8 + U_M10	103	-1,867	0,064
ZK_M5 + ZK_M6; LK_M5 + LK_M6; SK_M5 + SK_M6; GS_M5 + GS_M6; M_M5 + M_M6; U_M5 + U_M6	ZK_M8 + ZK_M10; LK_M8 + LK_M10; SK_M8 + SK_M10; GS_M8 + GS_M10; M_M8 + M_M10; U_M8 + U_M10	153	-0,285	0,791
* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe				

Neben den Unterschieden der Erkennungsgenauigkeiten in Abhängigkeit von der Normteilgröße deuten Abbildung 5-22 und Abbildung 5-23 auf einen weiteren Effekt hin. Es zeigt sich ein deutlicher Unterschied zwischen Modellen mit versenkten und nicht versenkten Normteilen. So erzielen die Modelle, welche Gewindestifte und Senkkopfschrauben beinhalten, im Durchschnitt deutlich geringere Erkennungsgenauigkeiten als die übrigen Modelle. Während die Versuchsergebnisse der Modelle mit Gewindestiften und Senkkopfschrauben ein arithmetisches Mittel von 87,98 % ergeben, beträgt dieses für die übrigen Modelle 97,90 %.

Ein Mann-Whitney-U-Test, welcher die Erkennungsgenauigkeiten der Modelle mit versenkten Bauteilen ($n = 27$) den mAP-Werten aller Modelle mit nicht versenkten Normteilen ($n = 45$) gegenüberstellt, bestätigt diese Beobachtung. Der Test zeigt einen signifikanten Unterschied, welcher eine große Effektstärke nach Cohen aufweist. ($U = 24$; $Z = -6,787$; $p < 0,001$; $r = 0,800$)

Werden für den Trainings- und Validierungsprozess ausschließlich Bilder aus der Draufsicht genutzt, ergeben sich im Vergleich zur Rundumsicht noch höhere Erkennungsgenauigkeiten, siehe Tabelle 5-28. Im Rahmen der Hauptuntersuchung zur Auswahl von Normteilen erzielen in diesem Fall 47 von 72 trainierten Objekterkennungsmodellen die maximal erreichbare mAP von 100 %. Die niedrigste mittlere Erkennungsgenauigkeit erzielt auch hier das Modell GS_M3 + GS_M4. Mit einer mittleren mAP von 94,28 % weist das Modell mit Bildern aus der Draufsicht jedoch eine deutlich höhere Erkennungsgenauigkeit auf als das gleiche Modell mit Bildern aus der Rundumsicht (84,16 %).

Tabelle 5-28: Versuchsergebnisse zur Auswahl von Normteilen – Draufsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
N1	ZK_M3 + ZK_M4	99,45	96,65	98,29	98,13	1,41	1,98	2,85
	ZK_M5 + ZK_M6	100	100	100	100	0	0	0
	ZK_M8 + ZK_M10	100	100	100	100	0	0	0
	LK_M3 + LK_M4	99,45	99,32	100	99,59	0,36	0,13	0,69
	LK_M5 + LK_M6	100	100	100	100	0	0	0
	LK_M8 + LK_M10	100	100	100	100	0	0	0
	SK_M3 + SK_M4	99,09	97,85	99,39	98,78	0,81	0,66	1,55
	SK_M5 + SK_M6	100	100	100	100	0	0	0
	SK_M8 + SK_M10	100	100	100	100	0	0	0
	GS_M3 + GS_M4	93,99	95,50	93,37	94,28	1,10	1,20	2,26
	GS_M5 + GS_M6	99,45	97,22	100	98,89	1,47	2,17	2,81
	GS_M8 + GS_M10	100	100	100	100	0	0	0
	M_M3 + M_M4	98,80	99,35	98,84	98,99	0,31	0,09	0,55
	M_M5 + M_M6	99,45	100	100	99,82	0,32	0,10	0,55
	M_M8 + M_M10	100	100	100	100	0	0	0
	U_M3 + U_M4	100	100	100	100	0	0	0
U_M5 + U_M6	100	100	100	100	0	0	0	
U_M8 + U_M10	100	100	100	100	0	0	0	
N2	ZK_M3 + LK_M3	100	99,49	98,84	99,44	0,58	0,34	1,17
	SK_M3 + GS_M3	94,40	95,22	96,88	95,50	1,26	1,59	2,59
	ZK_M6 + LK_M6	99,49	100	100	99,83	0,30	0,09	0,51
	SK_M6 + GS_M6	99,52	100	98,76	99,43	0,62	0,39	1,25
	ZK_M10 + LK_M10	100	100	100	100	0	0	0
	SK_M10 + GS_M10	100	100	100	100	0	0	0

* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe; ** in %

Um zu überprüfen, ob die Erkennungsgenauigkeiten auch statistisch signifikant über dem Grenzwert von 80 % liegen, wenn die Aufnahme der Trainings- und Validierungsbilder ausschließlich aus der Draufsicht erfolgt, kommt der Einstichproben-Wilcoxon-Test zur Anwendung. Für die Überprüfung der ersten Hypothese zur Normteilauswahl (N1) werden wie bei der Auswertung der Ergebnisse aus der Rundumsicht alle Versuchsreihen einbezogen, welche Normteile vom gleichen Typ in unterschiedlichen Größen beinhalten ($n = 54$). Der Test zeigt auch für die Draufsicht, dass die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (100 %) aufweisen. ($Z = 6,659$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,906$ und entspricht einem starken Effekt.

Zur Überprüfung der Hypothese N2 werden alle Versuchsreihen einbezogen, welche Schrauben der gleichen Größe mit einem ähnlichen Kopf beinhalteten ($n = 18$). Auch für diese Stichprobe zeigt der Einstichproben-Wilcoxon-Test, dass die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % einen signifikant höheren Median (100 %) aufweisen. ($Z = 3,799$; $p < 0,001$) Die Effektstärke nach Cohen [154] liegt bei $r = 0,895$ und entspricht ebenfalls einem starken Effekt. Die Nullhypothesen N1 und N2 sind damit auch für den Fall, dass Trainings- und Validierungsbilder aus der Draufsicht aufgezeichnet werden, nicht abzulehnen.

Wie die Versuchsergebnisse zur Auswahl von Normteilen aus der Rundumsicht zeigen auch die Ergebnisse aus der Draufsicht weitere, über die Hypothesen hinausgehende Effekte. So wird deutlich, dass Normteile mit größeren Nennweiten und Normteile, welche nicht in einem Bauteil versenkt werden, höhere Erkennungsgenauigkeiten aufweisen, als Normteile mit kleineren Nennweiten bzw. Normteile, welche vollständig versenkt werden. Beispielsweise ergibt sich für die Modelle mit Normteilen der Nennweite M3 ein arithmetisches Mittel von 97,47 %, während dieses für Modelle mit Normteilen der Nennweite M6 bzw. M10 einen Wert von 99,63 % bzw. 100 % beträgt, siehe Abbildung 5-24.

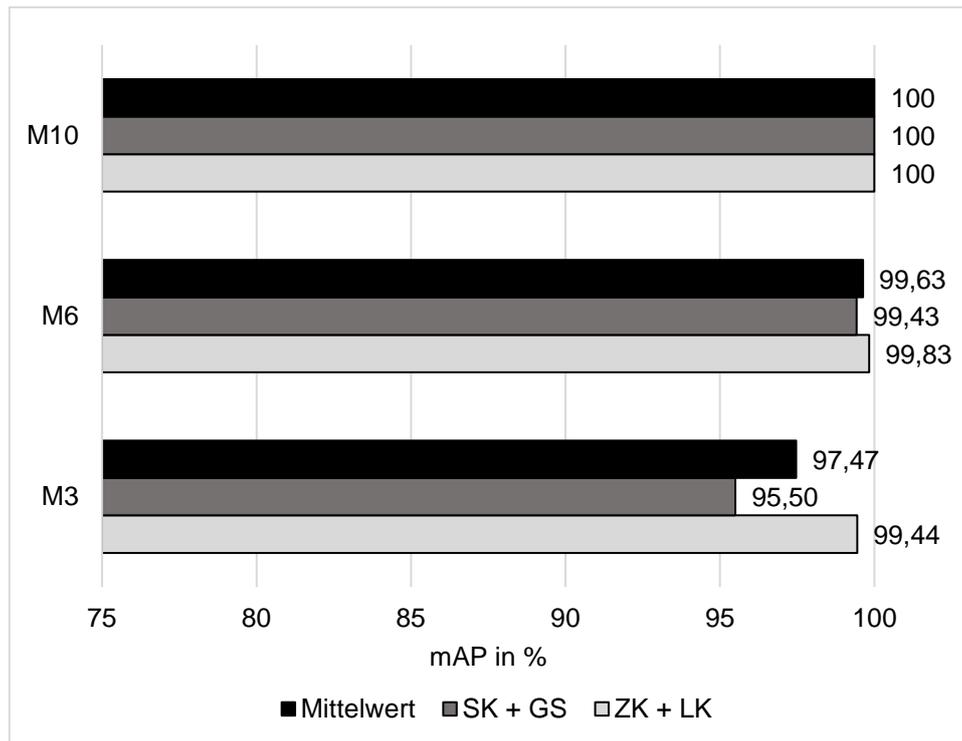


Abbildung 5-24: Mittlere mAP der Modelle SK+GS und ZK+LK in Abhängigkeit der Normteilgröße – Draufsicht

Um die statistische Signifikanz dieser Beobachtung auch für die Untersuchung aus der Draufsicht zu überprüfen, werden Mann-Whitney-U-Tests für die Normteilgrößenkombinationen M3 und M6, M3 und M10 sowie M6 und M10 durchgeführt. Tabelle 5-29 zeigt die Ergebnisse. Ausgehend von einem angestrebten Signifikanzniveau von $\alpha = 0,05$ kann für die Kombination M3 und M10 ein signifikanter Unterschied zwischen den Stichproben festgestellt werden. Für die anderen beiden Kombinationen gilt $p > \alpha$. Obwohl die Versuchsergebnisse aus der Rundumsicht für die Größerkombination M3 und M10 eine höhere Differenz der mittleren mAP aufweisen, lässt sich für diese keine statistische Signifikanz belegen. Diese Beobachtung ist damit zu erklären, dass die Messwerte aus der Rundumsicht deutlich stärker variieren als die Versuchsergebnisse aus der Draufsicht.

Tabelle 5-29: Mann-Whitney-U-Test 1 – Auswahl von Normteilen – Draufsicht

Stichprobe 1* (n=6)	Stichprobe 2* (n=6)	<i>U</i>	<i>Z</i>	<i>p</i>
SK_M3 + GS_M3; ZK_M3 + LK_M3	SK_M6 + GS_M6; ZK_M6 + LK_M6	7	-1,796	0,082
SK_M3 + GS_M3; ZK_M3 + LK_M3	SK_M10 + GS_M10; ZK_M10 + LK_M10	3	-2,678	0,015
SK_M6 + GS_M6; ZK_M6 + LK_M6	SK_M10 + GS_M10; ZK_M10 + LK_M10	9	-1,892	0,182

* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe

Der Vergleich der Modelle mit Normteilen der gleichen Art zeigt auch im Rahmen der Untersuchung aus der Draufsicht einen ähnlichen Effekt. So erzielen die Modelle mit Normteilen der Nennweiten M3 und M4 ein arithmetisches Mittel von 98,30 %, während dieses für die Modelle mit Normteilen der Nennweiten M5 und M6 bzw. M8 und M10 einen Mittelwert von 99,78 % bzw. 100 % ergibt, siehe Abbildung 5-25.

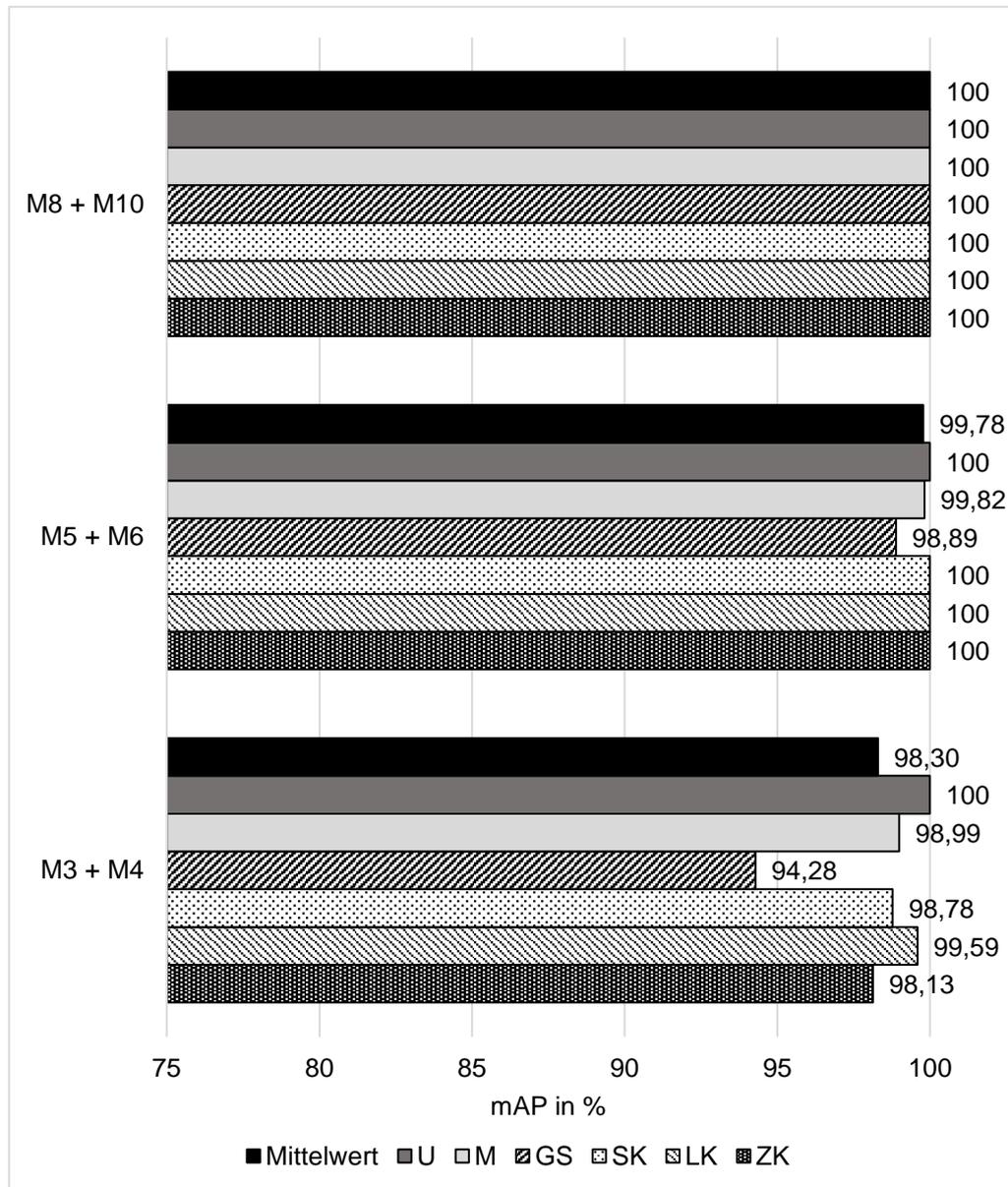


Abbildung 5-25: Mittlere mAP der Modelle U, M, GS, SK, LK und ZK in Abhängigkeit der Normteilgröße – Draufsicht

Um die statistische Signifikanz zu überprüfen, kommen auch in Bezug auf diese Beobachtungen Mann-Whitney-U-Tests zur Anwendung. Dabei werden neben den Versuchsergebnissen der Größen M3+M4 und M5+M6 auch die erzielten Erkennungsgenauigkeiten der Größen M3+M4 und M8+M10 sowie M5+M6 und M8+M10 gegenübergestellt. Die Ergebnisse der Tests können Tabelle 5-30 entnommen werden. Es wird deutlich, dass

sowohl für die Größenkombination M3+M4 und M5+M6 ($r = 0,617$) als auch für die Kombination M3+M4 und M8+M10 ($r = 0,756$) ein signifikanter Unterschied mit hoher Effektstärke nach Cohen [154] zwischen den betrachteten Stichproben festzustellen ist. Für die Größenkombination M5+M6 und M8+M10 konnte hingegen kein signifikanter Unterschied belegt werden, da gilt $p > \alpha$.

Auch für die betrachteten Größenkombinationen weisen die Versuchsergebnisse aus der Rundumsicht zwar höhere Differenzen der mittleren mAP auf. Trotzdem kann für diese im Gegensatz zu den Messwerten aus der Draufsicht keine statistische Signifikanz nachgewiesen werden, da sie deutlich höhere Varianzen aufweisen.

Tabelle 5-30: Mann-Whitney-U-Test 2 – Auswahl von Normteilen – Draufsicht

Stichprobe 1* (n=18)	Stichprobe 2* (n=18)	<i>U</i>	<i>Z</i>	<i>p</i>
ZK_M3 + ZK_M4; LK_M3 + LK_M4; SK_M3 + SK_M4; GS_M3 + GS_M4; M_M3 + M_M4; U_M3 + U_M4	ZK_M5 + ZK_M6; LK_M5 + LK_M6; SK_M5 + SK_M6; GS_M5 + GS_M6; M_M5 + M_M6; U_M5 + U_M6	54	-3,702	<0,001
ZK_M3 + ZK_M4; LK_M3 + LK_M4; SK_M3 + SK_M4; GS_M3 + GS_M4; M_M3 + M_M4; U_M3 + U_M4	ZK_M8 + ZK_M10; LK_M8 + LK_M10; SK_M8 + SK_M10; GS_M8 + GS_M10; M_M8 + M_M10; U_M8 + U_M10	36	-4,537	<0,001
ZK_M5 + ZK_M6; LK_M5 + LK_M6; SK_M5 + SK_M6; GS_M5 + GS_M6; M_M5 + M_M6; U_M5 + U_M6	ZK_M8 + ZK_M10; LK_M8 + LK_M10; SK_M8 + SK_M10; GS_M8 + GS_M10; M_M8 + M_M10; U_M8 + U_M10	135	-1,782	0,229
* ZK = Zylinderkopf; LK = Linsenkopf; SK = Senkkopf; GS = Gewindestift; M = Mutter; U = Unterlegscheibe				

Zwar zeigt sich auch für die Untersuchungen aus der Draufsicht ein Unterschied zwischen Modellen mit versenkten und nicht-versenkten Normteilen. Im Vergleich zu den Versuchsergebnissen aus der Rundumsicht ist dieser jedoch deutlich weniger ausgeprägt. So erzielen die Modelle, welche Gewindestifte und Senkkopfschrauben beinhalten, ein arithmetisches Mittel von 98,54 %, während dieses für die übrigen Modelle 99,72 % ergibt.

Ein Mann-Whitney-U-Test, welcher die Erkennungsgenauigkeiten der Modelle mit versenkten Bauteilen ($n = 27$) den mAP-Werten aller Modelle mit nicht versenkten Normteilen ($n = 45$) gegenüberstellt, bestätigt diese Beobachtung. Der Test zeigt einen signifikanten Unterschied, welcher jedoch im Vergleich zu demselben Test für die Ergebnisse aus der Rundumsicht nur eine geringe Effektstärke mit Tendenz zu einem mittleren Effekt nach Cohen [154] aufweist. ($U = 436,5$; $Z = -2,341$; $p = 0,019$; $r = 0,276$)

5.7.4 Versuchsergebnisse zur Dimensionierung von Formelementen

Schlussendlich zeigt Tabelle 5-31 die im Rahmen der Versuche zur Dimensionierung von Formelementen aus der Rundumsicht maximal erzielte mAP je Versuch. Es wird deutlich, dass alle untersuchten Modelle sehr hohe Erkennungsgenauigkeiten bei einer geringen Streuung der Werte erzielen. So erreichen 20 von 30 Versuchen die maximale mAP von 100 %. Die niedrigste mittlere mAP weist das Modell D3_D3,5 auf, wobei dieses mit einer mAP von 99,7 % dennoch einen sehr hohen Wert erzielt.

Tabelle 5-31: Versuchsergebnisse zu Formelementen – Rundumsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s^2	\bar{R} **
FE1	F1x45_F1,5x45	100	100	100	100	0	0	0
	F1x45_F2x45	100	100	100	100	0	0	0
FE2	R1_R15	99,98	100	99,69	99,89	0,17	0,03	0,31
	R1_R2	100	100	100	100	0	0	0
FE3	N4_N5	99,92	100	99,82	99,91	0,09	0,01	0,18
	N4_N6	100	100	100	100	0	0	0
FE4	T3_T4,5	99,75	100	100	99,92	0,14	0,02	0,25
	T3_T6	100	99,88	99,74	99,87	0,13	0,02	0,26
FE5	D3_D3,5	99,95	99,76	99,39	99,70	0,28	0,08	0,56
	D3_D4	100	100	100	100	0	0	0

* R = Radius; N = Nut; F = Fase; T = Tiefe; D = Durchmesser; ** in %

Um zu überprüfen, ob die Erkennungsgenauigkeiten statistisch signifikant über dem in Abschnitt 4.2.4 definierten Grenzwert von 80 % liegen, wird je Hypothese ein Einstichproben-Wilcoxon-Test angewendet. Tabelle 5-32 zeigt die Ergebnisse der durchgeführten Tests. Es wird deutlich, dass in allen betrachteten Fällen die Erkennungsgenauigkeiten im Vergleich zu einem Referenzmedian von 80 % signifikant höhere Mediane aufweisen. Die Effektstärke r nach Cohen [154] entspricht für alle betrachteten Hypothesen einem starken Effekt. Die Nullhypothesen FE1 – FE5 können damit nicht abgelehnt werden.

Tabelle 5-32: Einstichproben-Wilcoxon-Test – Formelemente – Rundumsicht

These	Stichprobe* (n=6)	Median**	Z	p	r
FE1	F1x45_F1,5x45; F1x45_F2x45	100	2,449	0,014	0,9998
FE2	R1_R15; R1_R2	100	2,264	0,024	0,924
FE3	N4_N5; N4_N6	100	2,264	0,024	0,924
FE4	T3_T4,5; T3_T6	99,94	2,226	0,026	0,909
FE5	D3_D3,5; D3_D4	99,98	2,226	0,026	0,909

* R = Radius; N = Nut; F = Fase; T = Tiefe; D = Durchmesser; ** in %

Auch die Hauptuntersuchung zur Dimensionierung von Formelementen ergibt im Vergleich zur Rundumsicht höhere Erkennungsgenauigkeiten, wenn für den Trainings- und Validierungsprozess ausschließlich Bilder aus der Draufsicht genutzt werden, siehe Tabelle 5-33. Dabei erreichen alle betrachteten Versuche die maximale mAP von 100 %. Eine statistische Auswertung der Messwerte ist nicht sinnvoll, da diese keine Unterschiede aufweisen. Weil die Messwerte im Vergleich zur Rundumsicht mindestens die gleiche Erkennungsgenauigkeit aufweisen, ist jedoch davon auszugehen, dass sie ebenfalls statistisch signifikant den Grenzwert von 80 % übersteigen, sodass die Nullhypothesen abzulehnen sind.

Tabelle 5-33: Versuchsergebnisse zu Formelementen – Draufsicht

These	Modell*	V1**	V2**	V3**	\bar{x} **	s	s ²	\bar{R} **
FE1	F1x45_F1,5x45	100	100	100	100	0	0	0
	F1x45_F2x45	100	100	100	100	0	0	0
FE2	R1_R15	100	100	100	100	0	0	0
	R1_R2	100	100	100	100	0	0	0
FE3	N4_N5	100	100	100	100	0	0	0
	N4_N6	100	100	100	100	0	0	0
FE4	T3_T4,5	100	100	100	100	0	0	0
	T3_T6	100	100	100	100	0	0	0
FE5	D3_D3,5	100	100	100	100	0	0	0
	D3_D4	100	100	100	100	0	0	0

* R = Radius; N = Nut; F = Fase; T = Tiefe; D = Durchmesser; ** in %

5.7.5 Zusammenfassung der Versuchsergebnisse

Um einen transparenten Überblick zu allen in Abschnitt 4.2 aufgestellten Hypothesen zu geben, werden die zuvor detailliert beschriebenen Versuchsergebnisse abschließend zusammengefasst. Tabelle 5-34 zeigt, welche Hypothesen durch die vorliegende Untersuchung statistisch signifikant nachgewiesen werden. Prinzipiell gilt, dass die aufgestellten Hypothesen durch die Alternativhypothesen repräsentiert werden, welche nur anzunehmen sind, wenn die zugehörige Nullhypothese mit statistischer Signifikanz abzulehnen ist. Es wird deutlich, dass ausschließlich die Hypothese BG1 durch die vorliegende Untersuchung bestätigt wird.

Tabelle 5-34: Versuchsergebnisse – Zusammenfassung

Nr.	Nullhypothese	Alternativhypothese	Ergebnis
Hypothesen zur Bauteilgestaltung			
BT1	$mAP_{BT-Spiegelnd} \geq mAP_{BT-Matt}$	$mAP_{BT-Spiegelnd} < mAP_{BT-Matt}$	Nicht bestätigt
BT2	$mAP_{BT-Dunkel} \geq mAP_{BT-Hell}$	$mAP_{BT-Dunkel} < mAP_{BT-Hell}$	Nicht bestätigt
BT3	$mAP_{BT-Weich} \geq mAP_{BT-Hart}$	$mAP_{BT-Weich} < mAP_{BT-Hart}$	Nicht bestätigt
BT4	$mAP_{BT-Ähnlich} \geq mAP_{BT-NichtÄhnlich}$	$mAP_{BT-Ähnlich} < mAP_{BT-NichtÄhnlich}$	Nicht bestätigt
BT5	$mAP_{BT-Komplex} \geq mAP_{BT-Simpel}$	$mAP_{BT-Komplex} < mAP_{BT-Simpel}$	Nicht bestätigt
Hypothese zur Baugruppengestaltung			
BG1	$mAP_{BG-Gleich} \geq mAP_{BG-NichtGleich}$	$mAP_{BG-Gleich} < mAP_{BG-NichtGleich}$	Bestätigt
Hypothesen zur Auswahl von Normteilen			
N1	$mAP_{Größe} \geq 80 \%$	$mAP_{Größe} < 80 \%$	Nicht bestätigt
N2	$mAP_{Schraubenkopf} \geq 80 \%$	$mAP_{Schraubenkopf} < 80 \%$	Nicht bestätigt
Hypothesen zur Dimensionierung von Formelementen			
FE1	$mAP_{Fase} \geq 80 \%$	$mAP_{Fase} < 80 \%$	Nicht bestätigt
FE2	$mAP_{Radius} \geq 80 \%$	$mAP_{Radius} < 80 \%$	Nicht bestätigt
FE3	$mAP_{Nut} \geq 80 \%$	$mAP_{Nut} < 80 \%$	Nicht bestätigt
FE4	$mAP_{Bohrungstiefe} \geq 80 \%$	$mAP_{Bohrungstiefe} < 80 \%$	Nicht bestätigt
FE5	$mAP_{Bohrungsdurchmesser} \geq 80 \%$	$mAP_{Bohrungsdurchmesser} < 80 \%$	Nicht bestätigt

6 Übertragbarkeit der Ergebnisse auf reale Produkte

6.1 Produktbeispiel I: Explosionsgeschützte Rohrleuchte

6.1.1 Produktbeschreibung

Die Untersuchungsergebnisse zeigen, dass alle im Rahmen der vorliegenden Arbeit betrachteten Bauteile und Baugruppen durch das eingesetzte Objekterkennungsmodell zuverlässig detektiert werden. Um Störeinflüsse zu reduzieren, erfolgt die Untersuchung der Hypothesen dabei jedoch bewusst anhand von sehr simplen Geometrien. Weiterhin werden die Bauteile und Baugruppen aus dem gleichen Grund stets auf neutralem Hintergrund aufgezeichnet.

Im Gegensatz dazu weisen reale Produkte vielfältige Geometrien auf und werden in komplexen Montageumgebungen gefertigt. Um zu überprüfen, ob die Untersuchungsergebnisse auf die Realität übertragbar sind, werden im vorliegenden Abschnitt industrielle Bauteile in einer realitätsnahen Montageumgebung hinsichtlich der in den Hypothesen berücksichtigten Aspekte betrachtet. Dazu werden zwei Produktbeispiele herangezogen.

Bei dem ersten Beispielprodukt handelt es sich um eine explosionsgeschützte Rohrleuchte des Unternehmens R. STAHL Schaltgeräte GmbH, siehe Abbildung 6-1. Diese wurde ausgewählt, da sie aufgrund ihres Einsatzgebietes besonderen Sicherheitsanforderungen unterliegt und der Einsatz einer Objekterkennung auf Grundlage von Deep Learning besonders sinnvoll erscheint.



Abbildung 6-1: Explosionsgeschützte Rohrleuchte [155]

Die Rohrleuchte ist für den Einsatz in explosionsgefährdeten Bereichen konzipiert. Durch eine druckfeste Kapselung ist sichergestellt, dass alle Komponenten, welche potenzielle Zündquellen aufweisen, in einem Gehäuse verbaut sind, welches so ausgelegt ist, dass sich eine Explosion auf sein Inneres beschränkt. Da fehlerhafte Montagen bei explosions-

geschützten Produkten dazu führen können, dass sich eine Explosion auf die Gehäuseumgebung ausweitet, kommt der Qualitätssicherung eine besondere Bedeutung zu. Durch den Einsatz einer Objekterkennung könnten Fehler schon im Montageprozess detektiert werden.

Weiterhin weist die Rohrleuchte Eigenschaften auf, anhand derer sich die im Rahmen der Hauptuntersuchung betrachteten Hypothesen sehr gut überprüfen lassen. Sie wurde für die manuelle Montage konzipiert und ist damit gut für die vorliegende Untersuchung händelbar.

Die Rohrleuchte besteht aus insgesamt 36 Einzelteilen, welche aus einer Vielzahl unterschiedlicher Materialien bestehen. Neben verschiedenen Kunststoffen kommen Aluminium, Stahl und elektrische Komponenten zum Einsatz. Eine Erzeugnisgliederung kann dem Anhang auf Seite 144 entnommen werden.

Neben der Vielzahl unterschiedlicher Werkstoffe finden im Rahmen des Montageprozesses auch viele verschiedene Montage- und Handhabungstätigkeiten Anwendung. So beinhaltet die Montage unter anderem die Handhabungstätigkeiten Positionieren und Spannen. Außerdem kommen verschiedene Fügeverfahren wie das Zusammensetzen, Kleben sowie An- und Einpressen zum Einsatz. Tabelle 6-1 zeigt den vollständigen Arbeitsplan.

Tabelle 6-1: Arbeitsplan zur Montage der explosionsgeschützten Rohrleuchte

Nr.	Tätigkeitsbeschreibung	Tätigkeitsart
1	<u>BG Kabeleinführung</u>	
1.1	Gummikappe Kabeleinführung auf Kabel aufschieben	Zusammensetzen
1.2	Hülse Kabeleinführung auf Kabel aufschieben	Zusammensetzen
1.3	Anschlüsse von Kabel in Kabeleinführungskappe einsetzen	Zusammensetzen
1.4	Kabeleinführungskappe und Kabel verspannen	Spannen
1.5	Kabeleinführungskappe und Kabel verschrauben	An- und Einpressen
2	<u>BG LED-Treiber</u>	
2.1	Kondensatoren von LED-Treiber mit Elektroisierband 32x20 isolieren	Kleben
2.2	Isolierfolie auf Unterseite LED-Treiber positionieren	Positionieren
2.3	Isolierfolie mittels Elektroisierband 20x30 befestigen	Kleben
2.4	Aderleitung rot und Aderleitung schwarz mit LED-Treiber verbinden	An- und Einpressen
3	<u>BG-LED-Träger</u>	
3.1	Typenschild auf Aluprofil kleben	Kleben
3.2	2 x LED-Leiterplatte mittels Verbindungselement verbinden	An- und Einpressen
3.3	2 x LED-Leiterplatte + Verbindungselement auf Aluprofil positionieren	Positionieren
3.4	2 x LED-Leiterplatte + Verbindungselement mit Aluprofil verschrauben	An- und Einpressen
3.5	Kammbrücker mit LED-Leiterplatte verbinden	An- und Einpressen
<i>Fortsetzung auf der nächsten Seite</i>		

Fortsetzung Tabelle 6-1

Nr.	Tätigkeitsbeschreibung	Tätigkeitsart
4	<u>BG Endkappe</u>	
4.1	Endkappe und Schaumpuffer verbinden	Kleben
5	<u>BG Leuchteinheit</u>	
5.1	BG Kabeleinführung und BG LED-Treiber verbinden	An- und Einpressen
5.2	Distanzstück in Aluprofil einschieben	Zusammensetzen
5.3	BG Kabeleinführung + BG LED-Treiber in Aluprofil einschieben	Zusammensetzen
5.4	Isolierschlauch auf Aderleitung rot und Aderleitung schwarz ziehen	Zusammensetzen
5.5	Aderleitung rot und Aderleitung schwarz mit LED-Leiterplatte verbinden	An- und Einpressen
5.6	Streublende auf Aluprofil aufschieben	Zusammensetzen
5.7	BG Endkappe auf Ende von Aluprofil aufschieben	Zusammensetzen
6	<u>Endmontage</u>	
6.1	BG Leuchteinheit in Rohr einschieben	Zusammensetzen
6.2	Hülse Kabeleinführung mit Rohr verschrauben	An- und Einpressen
6.3	Hülse Kabeleinführung mit Drehmomentschlüssel anziehen	An- und Einpressen
6.4	Gewindestift mit Hülse Kabeleinführung verschrauben	An- und Einpressen
6.5	Gummikappe Kabeleinführung auf Hülse Kabeleinführung aufschieben	Zusammensetzen
6.6	Gummikappe Kopfseite auf Kopfseite von Rohr aufschieben	Zusammensetzen

6.1.2 Montage mittels digitalem Assistenzsystem

Ein weiterer Grund für die Auswahl der explosionsgeschützten Rohrleuchte besteht durch den Umstand, dass diese bereits als Referenzprodukt in einem Forschungsprojekt Verwendung findet, welches den Einsatz von Objekterkennungsalgorithmen auf Grundlage von Deep Learning in der manuellen Montage untersucht. Im Rahmen des Projektes „Smart Assembly – Werkerassistenzsysteme mittels Augmented Reality“ wird der Einfluss verschiedener digitaler Assistenzsysteme auf die manuelle Montage untersucht. Dazu wurden drei Montagearbeitsplätze aufgebaut, auf welchen die Rohrleuchte mithilfe verschiedener Assistenzsysteme montiert wird.

Einer der beschriebenen Arbeitsplätze nutzt eine auf Deep Learning basierende Objekterkennung, um den Montageablauf der Rohrleuchte zu erfassen. Dazu ist der Arbeitsplatz mit einem digitalen Assistenzsystem ausgestattet, welches verschiedene Bestandteile aufweist. Neben dem Objekterkennungs-Modul beinhaltet der Arbeitsplatz ein Pick-by-Light-Modul, ein Schrauber-Modul sowie ein digitales Visualisierungs-Modul, siehe Abbildung 6-2. Die Speicherprogrammierbare Steuerung (SPS) ist Bestandteil des Pick-by-Light-Moduls.

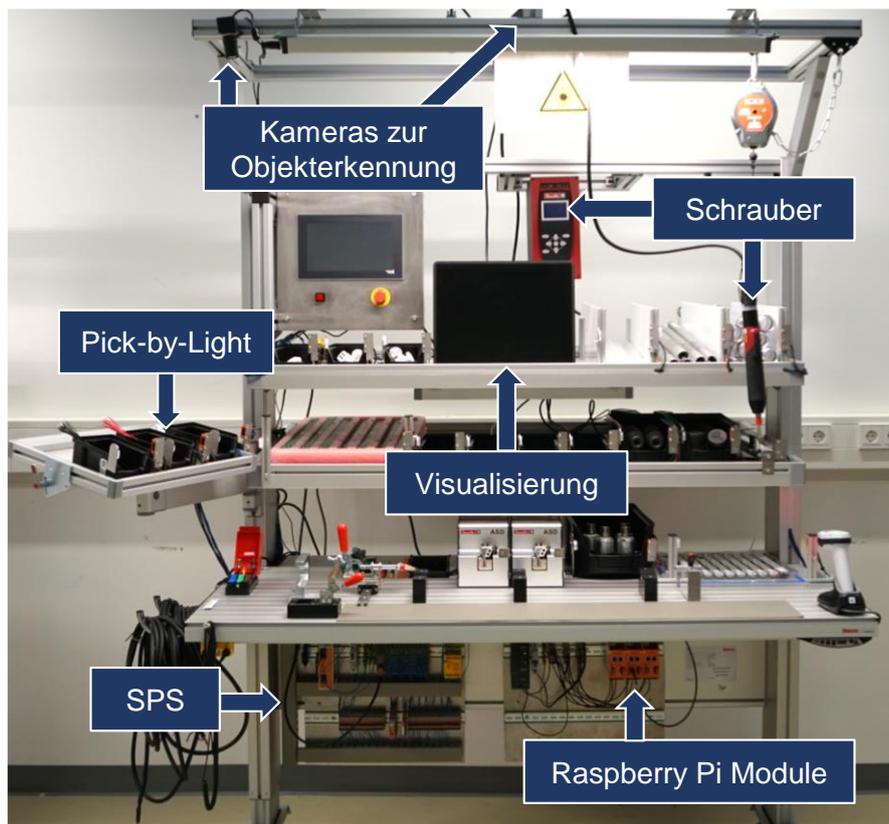


Abbildung 6-2: Montagearbeitsplatz mit digitalem Assistenzsystem

Die einzelnen Module sind jeweils mit einem industriellen Raspberry Pi verknüpft und kommunizieren über das MQTT-Netzwerkprotokoll miteinander. Dabei wird der vom Objekterkennungs-Modul identifizierte, aktuelle Werkstückstatus als Schlüsselinformation genutzt. Dieser kann verschiedene Zustände annehmen. Neben den Status „Werkstück vorhanden“ und „Werkstück nicht vorhanden“ kann er den Zustand „Fehlerhafte Baugruppe“ aufweisen. Diese Informationen werden genutzt, um auf den aktuellen Arbeitsschritt zu schließen und eine zugehörige Arbeitsanweisung bzw. im Fall eines auftretenden Fehlers einen Warnhinweis über das Visualisierungs-Modul darzustellen. Weiterhin werden Schrauber- und Pick-by-Light-Modul basierend auf dem aktuellen Arbeitsschritt aktiviert. Indem für jeden Arbeitsschritt definiert wird, welche Schlüsselobjekte vorhanden sein oder fehlen müssen, lassen sich die Arbeitsschritte aus den detektierten Bauteilen und Baugruppen ableiten. Durch diesen Ansatz kann das Assistenzsystem flexibel auf sich ändernde Arbeitsabläufe angepasst werden.

Da das eingesetzte Assistenzsystem auf dem Einsatz einer Objekterkennung auf Grundlage von Deep Learning basiert, wurde der Montagearbeitsplatz entsprechend gestaltet. Dies sind ideale Voraussetzungen für die vorliegende Untersuchung, da die Montageumgebung die zukünftigen Einsatzbedingungen sehr gut abbildet. Weiterhin liegen durch das Forschungsprojekt bereits Erkenntnisse zum Einsatz einer Objekterkennung in der manuellen Montage vor, siehe [156] und [157]. Im Gegensatz zur vorliegenden Untersuchung wurde im beschriebenen Forschungsprojekt jedoch kein Faster R-CNN Modell verwendet. Stattdessen kam der One-Stage-Detektor YOLOv4 zum Einsatz, da dieser sich

besonders gut für eine Echtzeit-Anwendung eignet [158]. Um das Modell zu trainieren, wurden insgesamt 1512 Trainingsbilder mit 5485 Objektinstanzen und 790 Validierungsbilder mit 1468 Objektinstanzen genutzt. Dabei bestanden die Trainingsdaten aus 41 einzelnen Objektklassen, sodass jede Klasse im Durchschnitt durch 134 Trainingsinstanzen repräsentiert wurde. Das Modell wurde auf einer Nvidia GeForce 1650 Super GPU unter Verwendung des Darknet Frameworks [159] trainiert und erzielte nach 6 Trainingsepochen eine mAP von 71,96 %. Durch die praktische Anwendung des trainierten Modells im Assistenzsystem am Montagearbeitsplatz wurde deutlich, dass das Modell in der Lage ist, alle Objekte auf einem zufriedenstellenden Niveau zu detektieren.

6.1.3 Übertragbarkeit der Versuchsergebnisse

Während das zuvor beschriebene Objekterkennungsmodell für die gesamte Montage der Rohrleuchte trainiert wurde, beziehen sich die in Abschnitt 5 beobachteten Versuchsergebnisse auf bestimmte Bauteil- bzw. Baugruppeneigenschaften. Um zu überprüfen, ob diese auf die explosionsgeschützte Rohrleuchte übertragbar sind, kann nicht die gesamte Montage in einem Objekterkennungsmodell berücksichtigt werden, sondern es sind bestimmte, den Hypothesen entsprechende Produktmerkmale auszuwählen. Aus diesem Grund werden im Rahmen der Untersuchung zur Übertragbarkeit der Ergebnisse einzelne, zu den Hypothesen passende Bestandteile und Montageschritte der Rohrleuchte bestimmt und jeweils in ein Objekterkennungsmodell integriert. Über die erzielten Erkennungsgenauigkeiten lassen sich schließlich Erkenntnisse bezüglich einzelner Hypothesen ableiten.

In Bezug auf den Versuchsablauf wird die Vorgehensweise der Hauptuntersuchung übernommen. Dementsprechend werden zunächst Trainings- und Validierungsbilder der relevanten Objekte aufgezeichnet und im Rahmen eines Annotationsprozesses mit Informationen zur Objektklasse und -position versehen. Im Sinne einer praxisnahen Umsetzung erfolgt die Aufnahme der Bilder mithilfe der am Montagearbeitsplatz installierten Kameras (siehe Abbildung 6-2). Da die Echtzeit-Objekterkennung im Assistenzsystem aus dieser Perspektive erfolgt, ist es sinnvoll, auch die Trainings- und Validierungsbilder aus diesem Blickwinkel aufzuzeichnen.

Um den Aufwand zu reduzieren, werden wie in der Hauptuntersuchung zunächst Videos aufgezeichnet und im Anschluss einzelne Bilder extrahiert. Im Gegensatz zur Hauptuntersuchung wird dabei jedoch der reale Montageprozess aufgezeichnet, welcher verschiedene Störeinflüsse wie Hände oder Werkzeuge im Bild beinhaltet. Dieses Vorgehen ist für die praktische Anwendung von Objekterkennungsalgorithmen von großer Bedeutung, da diese deutlich robuster gegenüber Störeinflüssen sind, wenn diese bereits in den Trainings- und Validierungsdaten enthalten sind. Im Anschluss an die Aufzeichnung der Videos erfolgt die Extraktion der Bilder, sowie die manuelle Annotation und die Einteilung in Trainings- und Validierungsbilder. Wie in der Hauptuntersuchung findet dabei ein Verhältnis von 80 % Trainings- und 20 % Validierungsbildern Anwendung.

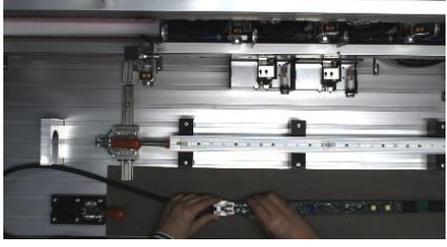
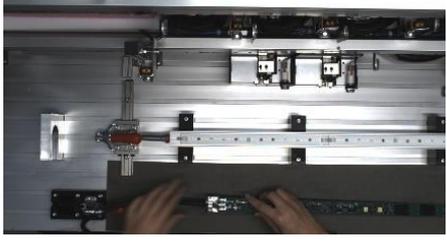
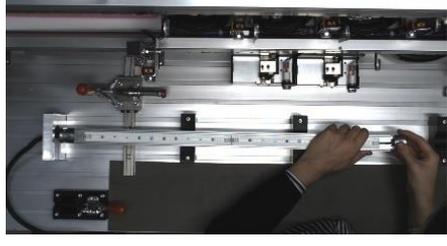
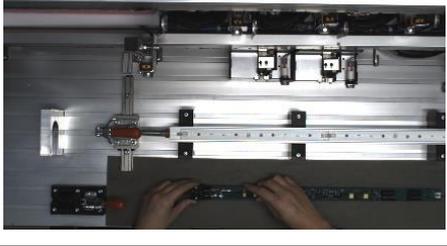
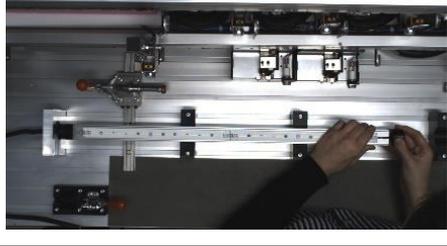
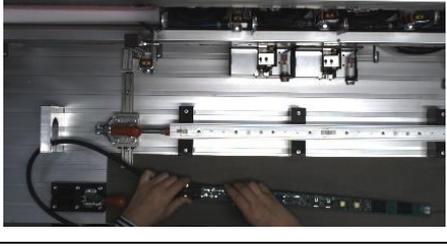
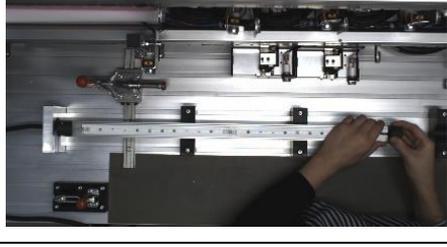
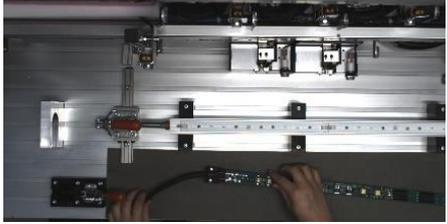
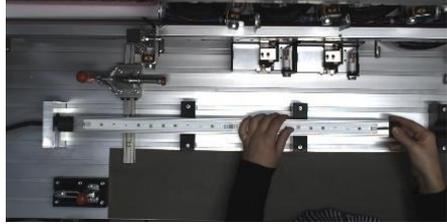
6.1.3.1 Bauteilgestaltung

Um die Hypothesen zur Bauteilgestaltung anhand der explosionsgeschützten Rohrleuchte zu untersuchen, kommen die in Tabelle 6-2 aufgeführten Bauteile zum Einsatz. Für die Hypothesen Spiegelnde Oberflächen (BT1), Dunkle Oberflächen (BT2) und Weiche Konturen (BT3) finden die Bauteile Kabeleinführungskappe und Endkappe Verwendung, da ihr optisches Erscheinungsbild mit wenig Aufwand angepasst werden kann und beide Bauteile aus Kunststoff gefertigt sind. Dadurch lassen sie sich problemlos durch 3D-gedruckte Bauteile ersetzen, deren Konturen einfach zu variieren sind.

Um eine spiegelnde Oberfläche zu erzeugen, werden Kabeleinführungs- und Endkappe durch das gleiche Airbrush-Verfahren wie in der Hauptuntersuchung mit silbernem Chromlack versehen. Auch für die Hypothese BT2 werden beide Bauteile lackiert. Dabei kommt jedoch mattschwarzer Lack zum Einsatz. Da im Rahmen der Hypothese BT3 untersucht wird, ob Bauteile mit weichen Konturen mit einer anderen Genauigkeit detektiert werden als Bauteile mit harten Konturen, ist die Geometrie der Bauteile anzupassen. Zwar sind die Kanten der Originalbauteile bereits mit einem sehr kleinen Radius versehen. Dieser ist jedoch zu wenig ausgeprägt, um ihn als weiche Kontur einzustufen. Da die Kanten der Bauteile aufgrund des vorhandenen Radius jedoch auch keine harte Kontur aufweisen, kommen jeweils zwei neue Versionen der Kabeleinführungs- und Endkappe zum Einsatz. Während die Erste keinerlei Rundungen aufweist, werden bei der zweiten Version alle Kanten mit dem maximal möglichen Radius versehen.

Prinzipiell wird wie in der Hauptuntersuchung für jedes Modell ein Referenzmodell definiert, welches Bauteile mit entgegengesetzten Eigenschaften enthält. Dies ermöglicht den Vergleich der Erkennungsgenauigkeiten beider Modelle. Während für die Hypothese BT3 Kabeleinführungs- und Endkappe sowohl für das Referenz- als auch das Untersuchungsmodell angepasst werden, kommen für die Hypothesen BT1 und BT2 die Originalbauteile als Referenz zum Einsatz. Diese bestehen aus einem weißen, matten Kunststoff und weisen damit die passenden, entgegengesetzten Eigenschaften auf.

Tabelle 6-2: Modelle der Rohrleuchte zur Bauteilgestaltung – Teil 1

Modell	Enthaltene Objektklassen	
Originalbauteile	<p data-bbox="539 376 826 409">Kabeleinführungskappe</p> 	<p data-bbox="1098 376 1225 409">Endkappe</p> 
Spiegelnde Oberfläche	<p data-bbox="491 689 890 723">Kabeleinführungskappe_Spiegelnd</p> 	<p data-bbox="1034 689 1289 723">Endkappe_Spiegelnd</p> 
Dunkle Oberfläche	<p data-bbox="491 1003 890 1037">Kabeleinführungskappe_Schwarz</p> 	<p data-bbox="1050 1003 1289 1037">Endkappe_Schwarz</p> 
Weiche Konturen	<p data-bbox="507 1317 874 1350">Kabeleinführungskappe_Weich</p> 	<p data-bbox="1066 1317 1273 1350">Endkappe_Weich</p> 
Harte Konturen	<p data-bbox="523 1630 858 1664">Kabeleinführungskappe_Hart</p> 	<p data-bbox="1082 1630 1257 1664">Endkappe_Hart</p> 

Für die Hypothese Ähnliche Bauteile (BT4) findet das Bauteil LED-Leiterplatte Verwendung, da dieses zweimal im Produkt verbaut ist und eine weitere Variante des Bauteils zur Verfügung steht, siehe Tabelle 6-3. So beinhaltet die Rohrleuchte zwar keine Bauteile, welche sich ähneln, jedoch lässt sich eine solche Ähnlichkeit durch die leichte Modifikation einer der beiden montierten LED-Leiterplatten erreichen. Dazu wird jede zweite LED der Originalleiterplatte (LED_Typ 2) entfernt, sodass eine dem Originalbauteil sehr ähnliche LED-Leiterplatte (LED_Typ 1) entsteht. Im Rahmen des aufgezeichneten Montageprozesses werden anschließend verschiedene Leiterplatten-Kombinationen verbaut. Neben der Montage von jeweils zwei LED-Leiterplatten desselben Typs werden die einzelnen Varianten der Leiterplatte kombiniert montiert. Um die erzielte Erkennungsgenauigkeit mit einem Referenzmodell zu vergleichen, wird außerdem ein Modell trainiert, welches neben der Originalleiterplatte (LED_Typ 2) eine optisch deutlich unterscheidbare Bauteilvariante (LED_Typ 3) beinhaltet. Für die beschriebenen Objekte werden im Rahmen der Erstellung der Trainings- und Validierungsbilder alle Montageschritte aufgezeichnet, in welchen die entsprechenden Bauteile Verwendung finden.

Tabelle 6-3: Modelle der Rohrleuchte zur Bauteilgestaltung – Teil 2

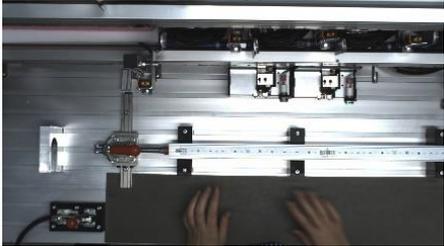
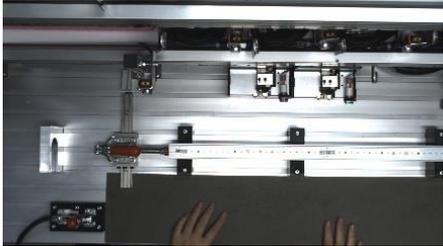
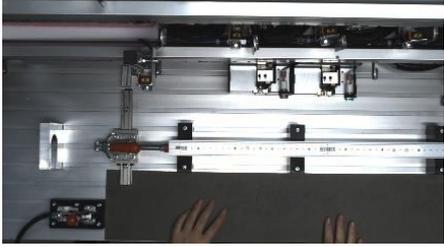
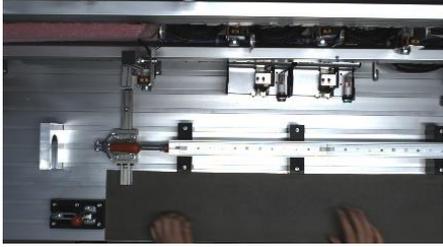
Modell	Enthaltene Objektklassen	
Ähnliche Bauteile	<p style="text-align: center;">LED_Typ1</p> 	<p style="text-align: center;">LED_Typ2</p> 
Nicht ähnliche Bauteile	<p style="text-align: center;">LED_Typ2</p> 	<p style="text-align: center;">LED_Typ3</p> 

Tabelle 6-4 zeigt die erzielten, maximalen Erkennungsgenauigkeiten der Modelle zur Bauteilgestaltung. Es wird deutlich, dass das Modell mit spiegelnden Bauteilen eine ähnlich hohe Erkennungsgenauigkeit erzielt wie das Referenzmodell mit den Originalbauteilen. Darüber hinaus ist zu erkennen, dass die mAP der Modelle, welche weich bzw. hart konturierte Bauteile enthalten, ähnliche Werte annimmt. Die Ergebnisse der Hauptuntersuchung in Bezug auf die Hypothesen BT1 und BT3 lassen sich damit auch für reale Produkte bestätigen.

Im Gegensatz dazu zeigt sich für die Modelle, welche ähnliche bzw. nicht ähnliche Bauteile beinhalten, ein deutlicher Unterschied hinsichtlich der erzielten mAP, siehe Tabelle 6-4. Zwar lassen sich die Ergebnisse der Hauptuntersuchung damit nicht eindeutig bestätigen, jedoch erzielt auch das Modell mit ähnlichen Bauteilen eine mAP, welche über dem Grenzwert von 80 % liegt. Eine Einschränkung in Bezug auf die Gestaltung von Bauteilen ist damit trotz der Abweichung zur Hauptuntersuchung nicht erforderlich.

Tabelle 6-4: Versuchsergebnisse zur Rohrleuchte – Bauteilgestaltung

Modell	mAP [%]
Originalbauteile	93,98
Spiegelnde Oberfläche	90,96
Dunkle Oberfläche	54,03
Weiche Konturen	85,36
Harte Konturen	81,80
Nicht ähnliche Bauteile	92,13
Ähnliche Bauteile	83,37

Eine noch geringere mAP weist das Modell Dunkle Oberflächen auf. Mit 54,03 % erzielt es eine deutlich niedrigere Erkennungsgenauigkeit als der in Abschnitt 4.2.3 definierte Grenzwert von 80 %. Die Analyse der Validierungsergebnisse zeigt, dass die Bauteile insbesondere dann nicht erkannt werden, wenn sie sich nicht deutlich von ihrem Hintergrund abheben. Da der Montagearbeitsplatz und die Vorrichtungen dunkle Farben aufweisen, lassen sich die schwarz lackierten Bauteile deutlich schlechter vom Hintergrund unterscheiden, als die weißen Originalbauteile. Diese Beobachtung wird durch die Ergebnisse der Modelle mit weichen und harten Konturen bestätigt. Beide Modelle enthalten graue Bauteile, welche sich schlechter als die Originalbauteile vom Hintergrund abheben und erzielen geringere Erkennungsgenauigkeiten als diese.

Um die Beobachtung näher zu untersuchen, wird ein weiteres Experiment durchgeführt. Dazu werden die Untersuchungsobjekte aus der Hauptuntersuchung genutzt und der Bauteilhintergrund variiert. Die Modelle detektieren jeweils die Formen 1-4 in weiß bzw. schwarz auf weißem bzw. schwarzem Hintergrund. Tabelle 6-5 zeigt die erzielten Erkennungsgenauigkeiten. Es wird deutlich, dass höhere Erkennungsgenauigkeiten erreicht werden, wenn Bauteile sich vom Hintergrund abheben.

Tabelle 6-5: mAP in Abhängigkeit des Hintergrundes

Modell	Weiß auf Weiß	Weiß auf Schwarz	Schwarz auf Weiß	Schwarz auf Schwarz
mAP [%]	12,40	97,40	99,80	90,07

In Bezug auf die Übertragbarkeit der Ergebnisse der Hauptuntersuchung auf die Rohrleuchte kann die Hypothese BT2 damit zwar nicht eindeutig bestätigt werden, jedoch ist anzunehmen, dass die geringe Erkennungsgenauigkeit auf die Ähnlichkeit zur Montageumgebung zurückzuführen ist. Die Ergebnisse stehen damit nicht zwangsläufig im Widerspruch zur Hauptuntersuchung.

6.1.3.2 Baugruppengestaltung

Um die Hypothese zur Baugruppengestaltung (BG1) anhand der Rohrleuchte zu untersuchen, wird die in Tabelle 6-6 aufgeführte Baugruppe genutzt und Modelle für unterschiedliche optische Ausprägungen der Bauteile betrachtet. Neben der ursprünglichen Baugruppe, welche aus einem aluminiumfarbigen Profil und zwei weiß lackierten LED-Leiterplatten besteht (Einfarbig_Aluminium), erfolgt die Untersuchung einer Baugruppe, deren Bauteile mit schwarzem Lack überzogen wurden (Einfarbig_Schwarz). Außerdem wird ein Modell trainiert, welches die Kombination von schwarzem bzw. weißem Aluprofil und weißen bzw. schwarzen LED-Leiterplatten beinhaltet (Mehrfarbig).

Tabelle 6-6: Modelle der Rohrleuchte zur Baugruppengestaltung

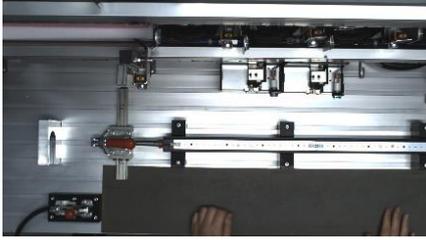
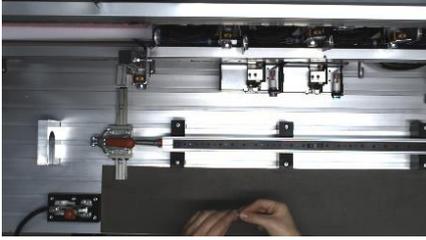
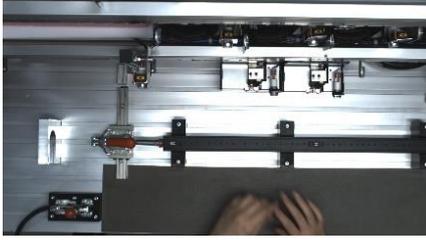
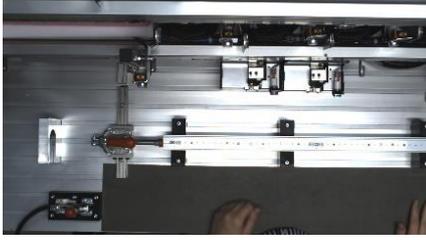
Modell	Enthaltene Objektklassen	
Mehrfarbig	<p style="text-align: center;">Aluprofil_Schwarz + LED</p> 	<p style="text-align: center;">Aluprofil + LED_Schwarz</p> 
Einfarbig_Schwarz	<p style="text-align: center;">Aluprofil_Schwarz + LED_Schwarz</p> 	
Einfarbig_Aluminium	<p style="text-align: center;">Aluprofil + LED</p> 	

Tabelle 6-7 zeigt die erzielten Erkennungsgenauigkeiten der Modelle zur Baugruppengestaltung. Es wird deutlich, dass alle Modelle sehr hohe Erkennungsgenauigkeiten erzielen, welche über dem Grenzwert von 80 % liegen. Außerdem weist das Modell Einfarbig_Aluminium mit 90,64 % eine geringere mAP auf als die anderen beiden Modelle (>92%). Zwar ist der Unterschied zwischen den einfarbigen und mehrfarbigen Modellen sehr gering, dennoch lässt sich das Ergebnis der Hauptuntersuchung in Bezug auf die Hypothese BG1 damit auch für reale Produkte bestätigen. So beträgt die Differenz zwischen einfarbigen und mehrfarbigen Modellen auch in der Hauptuntersuchung nur 2 %P, siehe Abschnitt 5.7.2.

Tabelle 6-7: Versuchsergebnisse zur Rohrleuchte – Baugruppengestaltung

Modell	mAP [%]
Mehrfarbig	92,48
Einfarbig_Schwarz	92,20
Einfarbig_Aluminium	90,64

6.1.3.3 Auswahl von Normteilen

Auch für die Untersuchung der Ergebnisübertragbarkeit der Hypothesen zur Auswahl von Normteilen, kommt die Baugruppe bestehend aus Aluprofil und LED-Leiterplatten zum Einsatz. Dabei werden jedoch ausschließlich die acht Schrauben betrachtet, welche für die Verbindung der beiden Bauteile Anwendung finden. Um die Hypothesen Ähnlicher Nenndurchmesser (N1) und Ähnlicher Schraubenkopf (N2) in Bezug auf die Rohrleuchte zu prüfen, werden die in Tabelle 6-8 aufgeführten Schraubenkombinationen eingesetzt. Während bei den Modellen M3Linse, M3Zylinder und M2Linse jeweils ein Schraubentyp und Nenndurchmesser zum Einsatz kommt, wird der Nenndurchmesser im Modell Durchmesser bzw. der Schraubenkopf im Modell Schraubenkopf variiert.

Dabei werden verschiedene Kombinationen der Schraubenpositionen berücksichtigt, sodass die Schrauben nicht durch das bloße Lernen ihrer Position detektierbar sind. Die Ergebnisse der Modelle mit einheitlichen Schrauben dienen als Referenz für die Auswertung.

Tabelle 6-8: Modelle der Rohrleuchte zur Auswahl von Normteilen

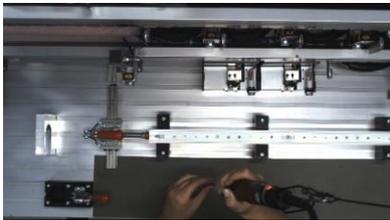
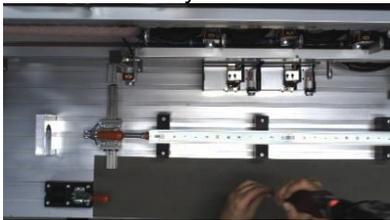
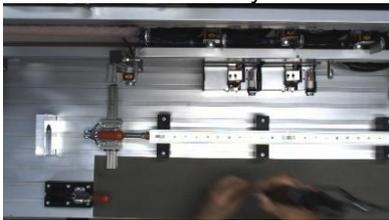
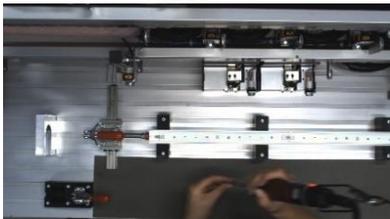
Modell	Enthaltene Objektklassen	Modell	Enthaltene Objektklassen
M3Linse	M3Linse 	Durchmesser	M3Linse + M2Linse 
M3Zylinder	M3Zylinder 	Schraubenkopf	M3Linse + M3Zylinder 
M2Linse	M2Linse 		

Tabelle 6-9 zeigt, dass die Modelle mit variierendem Schraubenkopf und Nenndurchmesser eine Erkennungsgenauigkeit aufweisen, welche unter dem in Abschnitt 4.2.3 definierten Grenzwert von 80 % liegt. Obwohl dies auf eine unzureichende Objekterkennung hindeutet, lässt sich aus dieser Beobachtung nicht ableiten, dass die einzelnen Objektinstanzen verwechselt werden. So zeigen die Ergebnisse der Referenzmodelle M3Linse, M3Zylinder und M2 Linse, dass ähnlich niedrige Erkennungsgenauigkeiten erzielt werden, wenn nur eine Objektinstanz zu identifizieren ist. Dies deutet auf eine insgesamt schlechte Objekterkennung z. B. aufgrund einer zu geringen Auflösung, eines ungünstigen Bildausschnittes oder anderer Störeinflüsse hin. Die Ergebnisse der Hauptuntersuchung in Bezug auf die Hypothesen N1 und N2 lassen sich damit weder bestätigen noch widerlegen.

Im Gegensatz dazu kann der im Rahmen der Hauptuntersuchung für aus der Draufsicht erfasste Objekte beobachtete Effekt, dass Normteile mit großen Nennweiten besser detektiert werden als solche mit kleinen, durch die Ergebnisse bestätigt werden. So weist das Modell M2Linse mit 52,23 % eine deutlich niedrigere mAP auf, als das Modell M3Linse mit 73,18 % und auch die einzelnen AP-Werte im Modell Durchmesser zeigen diesen Effekt (71,99 % vs. 60,22 %).

Tabelle 6-9: Versuchsergebnisse zur Rohrleuchte – Normteile

Modell	mAP [%]	AP M3Linse [%]	AP M3Zylinder [%]	AP M2Linse [%]
M3Linse	73,18	73,18	-	-
M3Zylinder	69,90	-	69,90	-
M2Linse	52,23	-	-	52,23
Schraubenkopf	73,68	75,63	71,74	-
Durchmesser	66,11	71,99	-	60,22

6.1.3.4 Dimensionierung von Formelementen

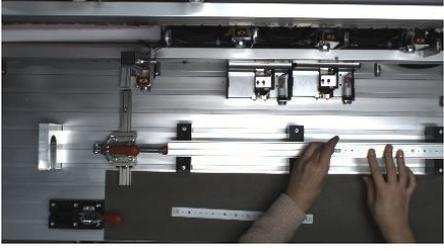
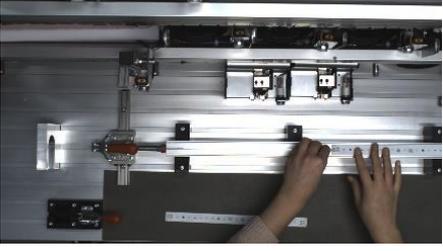
Um die Hypothesen zur Dimensionierung von Formelementen anhand der Rohrleuchte zu untersuchen, kommen schließlich die in Tabelle 6-10 aufgeführten Baugruppen zum Einsatz. Zur Prüfung der Übertragbarkeit der Hauptuntersuchungsergebnisse in Bezug auf die Hypothese FE1 wird die Hülse zur Kabeleinführung herangezogen. Diese weist im Original eine Fase von 5 x 30° auf. Für die Untersuchung, ob Bauteile verwechselt werden, wenn die Maße ihrer Fasen sich nur geringfügig unterscheiden, wird eine identische Hülse mit einer Fase von 6 x 45° versehen. Im Rahmen des Trainings- und Validierungsprozesses werden Bilder beider Varianten in einem Modell berücksichtigt.

Ein ähnliches Vorgehen wird zur Überprüfung der Hypothese FE3 genutzt. Hierzu kommt das Aluprofil zum Einsatz, welches im Original eine Nut mit einer Breite von 3 mm aufweist. Zusätzlich zu diesem Originalbauteil wird eine Variante erstellt, bei welcher die Nut auf eine Breite von 4 mm erweitert ist. Im Rahmen der Bilderaufnahme wird neben dem Positionieren des Aluprofils in der Montagevorrichtung auch das Positionieren der LED-Leiterplatten auf dem Aluprofil berücksichtigt.

Die Überprüfung der Ergebnisse der Hauptuntersuchung bezüglich der Hypothese FE5 erfolgt schließlich unter Verwendung der LED-Leiterplatten. Diese weisen im Original jeweils vier Bohrungen mit einem Durchmesser von 3 mm auf. Zusätzlich kommen jeweils zwei LED-Leiterplatten mit einem Bohrungsdurchmesser von 4 mm bzw. 4,5 mm zum Einsatz. Wie für die Überprüfung der Hypothese BT4 (siehe Abschnitt 6.1.3.1) werden im Rahmen des aufgezeichneten Montageprozesses verschiedene Leiterplatten-Kombinationen verbaut. Neben der Montage von jeweils zwei LED-Leiterplatten desselben Bohrungsdurchmessers werden jeweils zwei verschiedene Varianten der Leiterplatte kombiniert montiert. Um wie in der Hauptuntersuchung zwischen unterschiedlichen Größenabständen differenzieren zu können, werden dabei verschiedene Durchmesserkombinationen berücksichtigt, siehe Tabelle 6-10.

Die Hypothesen FE2 und FE4 waren nicht durch die Bauteile der Rohrleuchte abbildbar. Diese Einschränkung ist jedoch zu vernachlässigen, da die Untersuchung zur Übertragbarkeit der Ergebnisse auf reale Produkte nur einen ersten Hinweis auf die Anwendbarkeit liefern soll. Um den Aspekt der Übertragbarkeit zuverlässig zu untersuchen, sind umfangreiche, an die vorliegende Arbeit anknüpfende Untersuchungen notwendig.

Tabelle 6-10: Modelle der Rohrleuchte zur Dimensionierung von Formelementen

Modell	Enthaltene Objektklassen	
	I	II
Fase	Hülse Kabeleinführung_F5x30 	Hülse Kabeleinführung_F6x45 
Nut	Aluprofil_N3 	Aluprofil_N4 
<i>Fortsetzung auf der nächsten Seite</i>		

Fortsetzung Tabelle 6-10

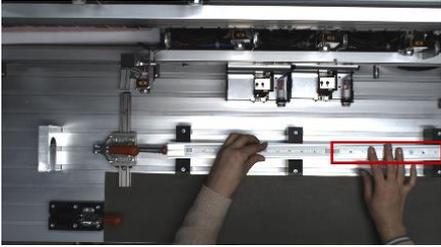
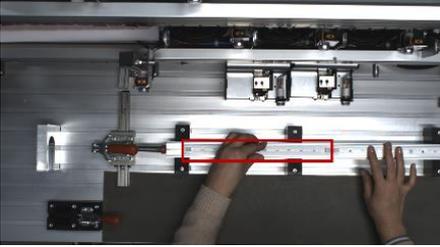
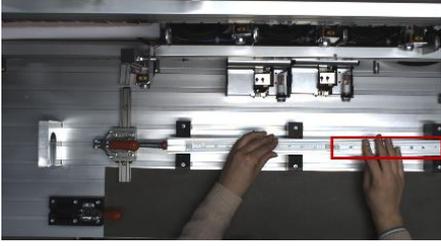
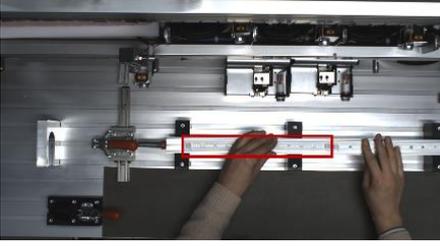
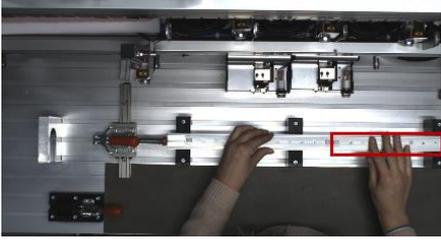
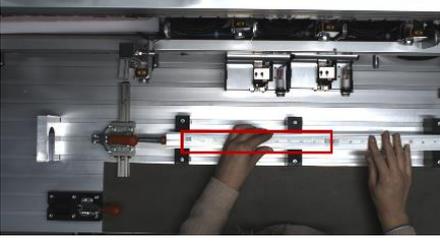
Modell	Enthaltene Objektklassen	
	I	II
Bohrungs- durchmesser	LED-Leiterplatte_D3 	LED-Leiterplatte_D4 
	LED-Leiterplatte_D3 	LED-Leiterplatte_D4,5 
	LED-Leiterplatte_D4 	LED-Leiterplatte_D4,5 

Tabelle 6-11 zeigt die erzielten Erkennungsgenauigkeiten aller in Bezug auf die Hypothesen zur Dimensionierung von Formelementen trainierten Modelle. Es wird deutlich, dass drei von fünf Modellen den Grenzwert von 80 % übersteigen. Darüber hinaus erzielen auch die übrigen zwei Modelle mit mAP-Werten von 75,86 % und 77,87 % hohe Erkennungsgenauigkeiten. Auch der Durchschnitt der Ergebnisse aller Modelle zum Bohrungsdurchmesser (Hypothese FE5) liegt mit 82,57 % über dem in Abschnitt 4.2.3 hergeleiteten Grenzwert. In Bezug auf die Übertragbarkeit der Hauptuntersuchungsergebnisse auf reale Produkte lassen sich damit die Erkenntnisse zu den Hypothesen FE3 und FE5 bestätigen. Im Gegensatz dazu kann das Ergebnis der Hauptuntersuchung zur Dimensionierung von Fasen (Hypothese FE1) nicht bestätigt werden. Da die Differenz der erzielten Erkennungsgenauigkeit zum Grenzwert mit ca. 5 %P jedoch gering ist und sich auf einen einzigen Versuchsdurchlauf bezieht, steht die Beobachtung nicht zwangsläufig im Widerspruch zum Resultat der Hauptuntersuchung. Vielmehr sollten im Anschluss an der vorliegende Arbeit weitere Untersuchungen zur Generierung statistisch gesicherter Ergebnisse durchgeführt werden.

Tabelle 6-11: Versuchsergebnisse zur Rohrleuchte – Formelemente

Modell	mAP [%]	AP I [%]	AP II [%]
Fase	75,86	74,26	77,45
Nut	100	100	100
Bohrungsdurchmesser (3vs4)	77,87	86,37	69,36
Bohrungsdurchmesser (3vs4,5)	85,58	86,54	84,62
Bohrungsdurchmesser (4vs4,5)	84,27	81,57	86,97

6.2 Produktbeispiel II: Würfel-Regal

6.2.1 Produktbeschreibung

Neben der explosionsgeschützten Rohrleuchte, welche für den Einsatz in einem industriellen Umfeld konzipiert ist, wird ein weiteres Beispielprodukt herangezogen, um die Übertragbarkeit der Ergebnisse aus der Hauptuntersuchung auf reale Produkte zu untersuchen. Dabei wird mit dem in Abbildung 6-3 dargestellten IKEA Würfel-Regal ein Erzeugnis ausgewählt, welches in der Regel im privaten Umfeld Anwendung findet. Das Regal kann in verschiedenen Farbausführungen bezogen werden, wodurch es sich ideal für die Überprüfung der Hypothesen eignet. Konkret kommt das Produkt in der vorliegenden Arbeit in weißer und schwarzer Ausführung zum Einsatz.



Abbildung 6-3: Würfel-Regal in Weiß und Schwarz [160] [161]

Das Regal ist für den Einsatz von Objekterkennungsalgorithmen besonders interessant, da es durch den Käufer selbstständig zu montieren ist. Eine intelligente Montageanweisung könnte den Montageprozess deutlich erleichtern. Weiterhin bietet der Hersteller bereits eine Applikation an, welche es ermöglicht, Möbel durch die Nutzung von Augmented Reality (AR) virtuell in einem realen Raum zu platzieren [162] und es wurde eine Applikation entwickelt, welche den Montageprozess verschiedener IKEA Produkte mittels AR visualisiert [163].

Die beschriebenen Applikationen zeigen das große Einsatzpotential digitaler Technologien zur Unterstützung der Möbelmontage. Ob die Anwendungen bereits Objekterkennungsalgorithmen nutzen, ist nicht bekannt. Es besteht jedoch in jedem Fall ein großes Potenzial durch die Implementierung von Objekterkennungsalgorithmen, da sich die Montageanweisungen mit den Echtzeit-Montageschritten verknüpfen ließen.

Im Vergleich zur explosionsgeschützten Rohrleuchte weist das zweite Beispielprodukt eine deutlich geringere Komplexität auf. Sowohl das schwarze als auch das weiße Regal besteht aus drei Bauteilen, welche in mehrfacher Stückzahl Verwendung finden. Tabelle 6-12 zeigt die vollständige Stückliste eines Würfel-Regals.

Tabelle 6-12: Stückliste des Produktbeispiels Würfel-Regal

Nr.	Bezeichnung	Stückzahl
1	Deckel	2
2	Seitenwand	2
3	Holzschraube Senkkopf mit Innensechskant 5x70	8

Die Montage des Regals besteht aus drei Schritten. Zunächst werden ein Deckel und eine Seitenwand durch zwei Schrauben miteinander verbunden. Anschließend erfolgt die Montage der zweiten Seitenwand, ebenfalls unter Verwendung zweier Schrauben. Schlussendlich wird der verbleibende Deckel auf die Baugruppe aufgesetzt und mithilfe von vier Schrauben befestigt. Die vollständige Montageanleitung kann [164] entnommen werden.

6.2.2 Übertragbarkeit der Versuchsergebnisse

In Bezug auf den Versuchsablauf wird auch für die Übertragbarkeitsuntersuchung mittels Würfel-Regalen die Vorgehensweise der Hauptuntersuchung übernommen. So werden zunächst Trainings- und Validierungsbilder der relevanten Objekte aufgezeichnet und im Rahmen eines Annotationsprozesses mit Informationen zur Objektklasse und -position versehen. Im Gegensatz zur Montage der Rohrleuchte werden die Regale in der Praxis jedoch nicht auf professionellen Montagearbeitsplätzen montiert. Die Montage erfolgt in der Regel in einem privaten Umfeld, dessen Erscheinungsbild (Bodenbeschaffenheit, weitere Möbel im Blickfeld, Beleuchtung, etc.) extrem stark variiert. Um die reale Produktumgebung abzubilden, erfolgt die Aufnahme der Bilder in einem privaten Raum mit diversem Mobiliar.

Ein weiterer Unterschied zur explosionsgeschützten Rohrleuchte besteht durch die Perspektive, aus welcher die Montage im realen Anwendungsfall erfasst wird. Während der Montageprozess der Rohrleuchte aus der Draufsicht erfasst wird, steht ein solches Kamera-System im privaten Umfeld in der Regel nicht zur Verfügung. Mögliche Anwendungsformen einer Echtzeit-Objekterkennung bestehen hier z. B. durch den Einsatz von Smartphones oder Smart Glasses. Diese können die Montage erfassen und relevante Informationen

bereitstellen. In beiden Fällen wird das zu montierende Produkte aus der Ich-Perspektive aufgezeichnet. Um diese Perspektive nachzubilden, erfolgt die Aufnahme der Trainings- und Validierungsbilder mithilfe einer Kamera des Typs GoPro Hero7, welche durch einen Kopfgurt getragen wird und die Montage aus dem Blickwinkel des Montierenden aufzeichnet.

Um den Aufwand zu reduzieren, werden Videos erstellt und im Anschluss einzelne Bilder extrahiert. Dabei wird der reale Montageprozess aufgezeichnet, welcher verschiedene Störeinflüsse wie Hände oder Werkzeuge im Bild beinhaltet, um die Objekterkennungsmodelle robuster gegenüber Störeinflüssen zu machen. Im Anschluss erfolgt die manuelle Annotation der Bilder und ihre Einteilung in Trainings- und Validierungsdatensätze. Wie in der Hauptuntersuchung wird dabei ein Verhältnis von 80 % Trainings- und 20 % Validierungsbilder angewendet.

Aufgrund der geringen Komplexität der Baugruppe sind nicht alle im Rahmen der Hauptuntersuchung berücksichtigten Hypothesen durch die Regale abbildbar. Neben der Hypothese BT2 lässt sich insbesondere die Hypothese BG1 durch das ausgewählte Produkt überprüfen. Diese These ist von besonderer Bedeutung, da sie die einzige Hypothese darstellt, welche sich durch die Hauptuntersuchungsergebnisse statistisch signifikant bestätigen lässt. Darüber hinaus wird ein Großteil der Hypothesen bereits durch die in Abschnitt 6.1 betrachtete Rohrleuchte hinsichtlich ihrer Übertragbarkeit untersucht.

Tabelle 6-13 zeigt die für die Regale erstellten Modelle sowie die in ihnen enthaltenen Objektklassen. Um die Übertragbarkeit der Hauptuntersuchung in Bezug auf die Hypothese BT2 zu untersuchen, wird für die Montage des weißen und des schwarzen Regals jeweils ein Objekterkennungsmodell erstellt. Über den Vergleich der Erkennungsgenauigkeiten dieser Modelle lässt sich feststellen, ob Bauteile, welche dunkle Oberflächen aufweisen, schlechter als Bauteile mit hellen Oberflächen detektiert werden.

Die Baugruppe beinhaltet jeweils zwei Objektklassen. Neben einem Deckel, von welchem zwei Stück zur Anwendung kommen, besteht die Baugruppe aus zwei identischen Seitenwänden. Die Schrauben, welche zur Verbindung der einzelnen Teile eingesetzt werden, sind nicht Bestandteil der Objekterkennungsmodelle. Für die Untersuchung der Hypothese BG1 wird zusätzlich zu den beschriebenen, einfarbigen Modellen, ein Objekterkennungsmodell trainiert, welches mehrfarbige Baugruppen beinhaltet. Dazu wird die Montage eines Regals mit weißen Deckeln und schwarzen Seitenwänden, sowie eines Regals bestehend aus schwarzen Deckeln und weißen Seitenwänden erfasst und Trainings- sowie Validierungsdatensätze abgeleitet. Über den Vergleich der erzielten Erkennungsgenauigkeiten mit den zuvor beschriebenen, einfarbigen Modellen, lässt sich feststellen, ob die im Rahmen der Hauptuntersuchung verifizierte Hypothese BG1 auf reale Produkte übertragbar ist.

Tabelle 6-13: Modelle der Würfel-Regale

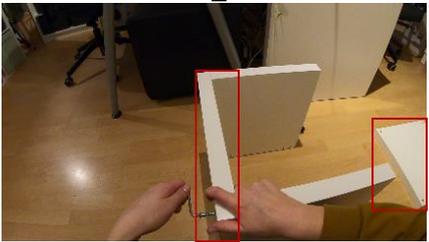
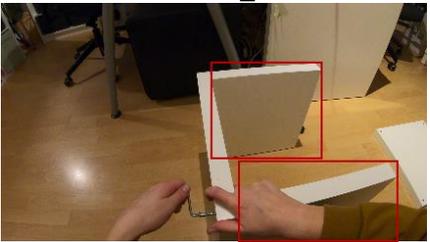
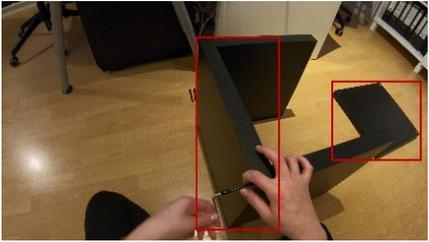
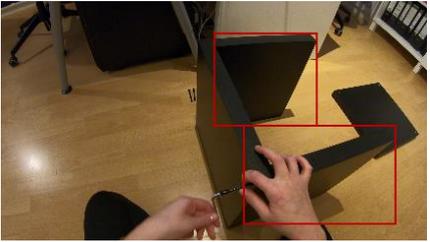
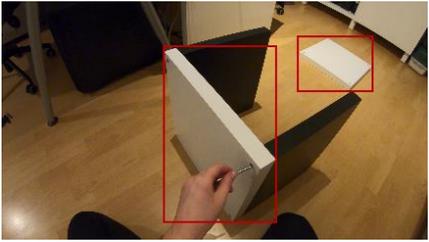
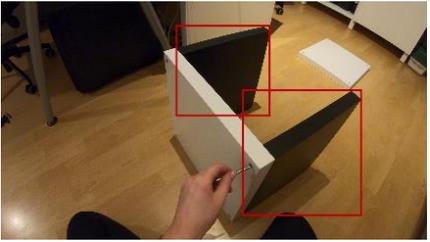
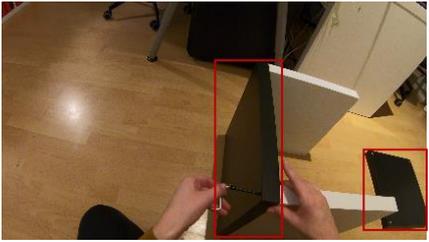
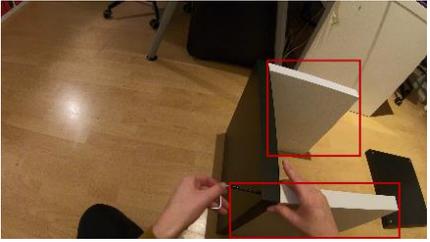
Modell	Enthaltene Objektklassen	
Einfarbig_Weiss	<p style="text-align: center;">Deckel_Weiss</p> 	<p style="text-align: center;">Seitenwand_Weiss</p> 
Einfarbig_Schwarz	<p style="text-align: center;">Deckel_Schwarz</p> 	<p style="text-align: center;">Seitenwand_Schwarz</p> 
Mehrfarbig	<p style="text-align: center;">Deckel_Weiss</p> 	<p style="text-align: center;">Seitenwand_Schwarz</p> 
	<p style="text-align: center;">Deckel_Schwarz</p> 	<p style="text-align: center;">Seitenwand_Weiss</p> 

Tabelle 6-14 zeigt die erzielten Erkennungsgenauigkeiten der trainierten Modelle. In Bezug auf die Hypothese BT2 wird deutlich, dass das einfarbig weiße Regal eine höhere mAP aufweist, als das einfarbig schwarze. Das Ergebnis der Hauptuntersuchung, welches keinen signifikanten Unterschied zwischen der Detektierbarkeit heller und dunkler Bauteile nachweisen konnte, lässt sich mit dieser Beobachtung nicht bestätigen. Da die Differenz mit ca. 5 %P jedoch gering ist und sich auf einen einzigen Versuchsdurchlauf bezieht, steht die Beobachtung nicht zwangsläufig im Widerspruch zum Resultat der Hauptuntersuchung.

Im Gegensatz dazu, lässt sich das Ergebnis der Hauptuntersuchung in Bezug auf die Hypothese BG1 durch die erzielten Erkennungsgenauigkeiten bestätigen. So erzielt das mehrfarbige Modell mit einem Wert von 78,60 % eine höhere mAP, als die beiden

einfarbigen Modelle, welche im Durchschnitt eine mAP von 73,25 % aufweisen. Insbesondere die schwarzen Deckel und Seitenwände werden im mehrfarbigen Modell mit einer deutlich höheren Genauigkeiten detektiert, als im einfarbig schwarzen Modell. Die Ergebnisse deuten darauf hin, dass die Einzelteile von Baugruppen mit einer höheren Genauigkeit detektiert werden, wenn sie unterschiedliche optische Eigenschaften aufweisen.

Tabelle 6-14: Versuchsergebnisse zu den Würfel-Regalen

Modell	mAP [%]	AP D_W* [%]	AP S_W* [%]	AP D_S* [%]	AP S_S* [%]
Mehrfarbig	78,60	84,25	68,59	88,63	72,93
Einfarbig_Weiss	75,62	85,48	65,76		
Einfarbig_Schwarz	70,88			75,67	66,08
D_W = Deckel_Weiss; S_W = Seitenwand_Weiss; D_S = Deckel_Schwarz; S_S = Seitenwand_Schwarz					

7 Ableitung konstruktiver Gestaltungsrichtlinien

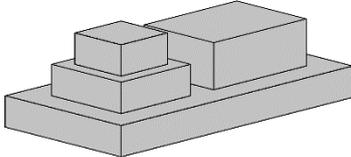
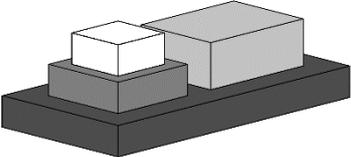
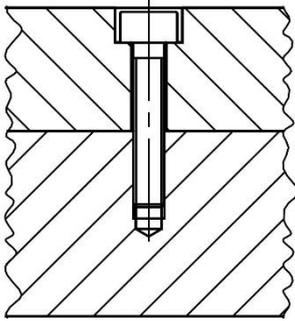
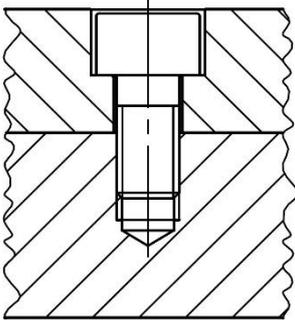
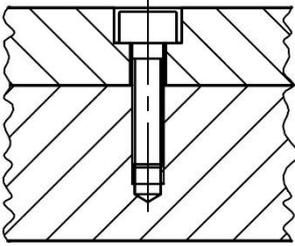
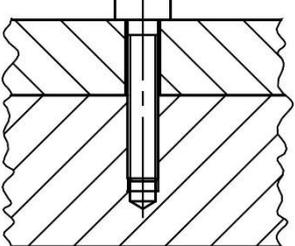
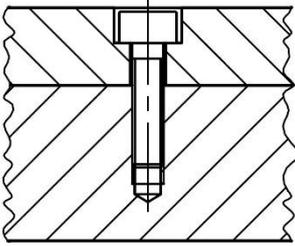
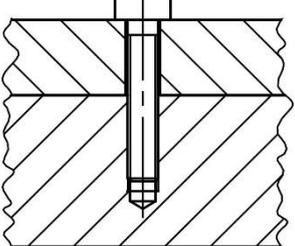
Ziel der vorliegenden Arbeit ist die Definition von Gestaltungsrichtlinien, welche dem Konstrukteur der Zukunft helfen, Produkte für den Einsatz von Deep Learning in der Montage effektiv und effizient zu gestalten. Als Grundlage dienen die Ergebnisse der vorliegenden Arbeit. So zeigt die Hauptuntersuchung, dass Baugruppen besser detektiert werden, wenn Einzelteile unterschiedliche optische Eigenschaften aufweisen. Diese Beobachtung wird zum einen durch die Experimente unter Laborbedingungen erzielt, siehe Kapitel 5. Darüber hinaus bestätigen die Versuche zur Übertragbarkeit auf reale Produkte, dass der Effekt auch in realen Montageumgebungen zu beobachten ist, siehe Kapitel 6.

Weitere relevante Beobachtungen ergeben sich in Bezug auf die Auswahl von Normteilen. Diese sind zwar nicht Bestandteil der in Kapitel 4 definierten Hypothesen, sie können jedoch als Zusatzbeobachtungen angesehen werden. So zeigen die Versuchsergebnisse, dass Normteile im Nenndurchmesserbereich von M3 bis M10 aus der Draufsicht mit höherer Genauigkeit detektiert werden, je größer ihr Nenndurchmesser ist. Auch diese Beobachtung wird durch die Versuche zur Übertragbarkeit auf reale Produkte bestätigt.

Schließlich zeigt die Hauptuntersuchung, dass Schrauben besser erkannt werden, wenn sie nicht vollständig im Bauteil versinken. Diese Beobachtung wird durch die Versuche zur Übertragbarkeit auf reale Produkte weder bestätigt noch widerlegt, da die ausgewählten Beispielprodukte keine geeigneten Elemente zur Untersuchung aufweisen. Da die vorliegende Arbeit das Ziel verfolgt, grundlegende Erkenntnisse zur montagegerechten Gestaltung für den Einsatz von Deep Learning zu gewinnen, welche zukünftig durch weitere, ausführliche Versuche anhand von realen Fallbeispielen zu überprüfen sind, wird die Beobachtung dennoch für die Ableitung der Gestaltungsrichtlinien herangezogen.

Zusammenfassend lassen sich vier konkrete Empfehlungen für die Gestaltung von Produkten zur Anwendung von auf Deep Learning basierenden Objekterkennungsalgorithmen aus den Ergebnissen der vorliegenden Untersuchung ableiten, siehe Tabelle 7-1. Als Vorbild für die Struktur der Tabelle dienen die Gestaltungsrichtlinien für montagegerechte Konstruktionen von Pahl und Beitz [30]. Diese liefern neben einer allgemeinen Empfehlung Beispiele für montagegerechte und nicht montagegerechte Konstruktionen.

Tabelle 7-1: Montagegerechte Gestaltungsrichtlinien zum Einsatz von Deep Learning

Nr.	Gestaltungsrichtlinie	Nicht montagegerecht	Montagegerecht
1	Gestaltung von Baugruppen so, dass angrenzende Einzelteile unterschiedliche optische Eigenschaften aufweisen		
2	Auswahl von Schrauben großer Nennweiten		
3	Auswahl von Schrauben, welche nicht im Bauteil versinken		
4	Gestaltung von Bauteilen und Baugruppen, sodass Schrauben nicht im Bauteil versinken		

Die abgeleiteten Gestaltungsrichtlinien stehen teilweise im Widerspruch zu anderen konstruktiven Empfehlungen. So steht dem Gebot, angrenzende Einzelteile mit unterschiedlichen optischen Eigenschaften zu gestalten (Gestaltungsrichtlinie Nr. 1), der Gedanken der Vereinheitlichung von Bauteilen entgegen. Darüber hinaus widerspricht die Nutzung von Schrauben großer Nennweiten (Gestaltungsrichtlinie Nr. 2) den Zielen des Leichtbaus.

Trotz der beobachteten Unterschiede der Detektierbarkeit verschiedener Konstruktionen zeigt die vorliegende Arbeit, dass alle betrachteten Untersuchungsobjekte unabhängig von ihrer konstruktiven Gestaltung zuverlässig und mit sehr hoher Genauigkeit durch auf Deep Learning basierende Objekterkennungsalgorithmen detektiert werden. Die abgeleiteten Gestaltungsrichtlinien sind daher nicht zwingend für eine zuverlässige Objekterkennung durch Deep Learning notwendig. Sie sind eher als Möglichkeit zu verstehen, die Erkennungsgenauigkeit zu optimieren. Stehen den vorgeschlagenen Empfehlungen andere Interessen (z. B. Kostenfaktoren, Designfaktoren, etc.) oder Gestaltungsrichtlinien entgegen, ist zu prüfen, ob eine Optimierung der Detektierbarkeit erforderlich ist. Für jeden Einzelfall muss dann zwischen den konträren Zielen abgewogen werden.

Die Ergebnisse der Hauptuntersuchung sowie der Untersuchung zur Übertragbarkeit auf reale Produkte zeigen außerdem, dass ein Großteil der betrachteten konstruktiven Eigenschaften die Erkennungsgenauigkeit durch eine auf Deep Learning basierende Objekterkennung nicht beeinflussen. So werden Bauteile mit spiegelnden und dunklen Oberflächen mit der gleichen Genauigkeit detektiert, wie matte bzw. helle Bauteile und auch ihre Komplexität sowie die Ausprägung von Kanten führen zu keinen Unterschieden in der Objekterkennung. Selbst die Detektion von sehr ähnlichen Bauteile erfolgt mit der gleichen Genauigkeit wie für Bauteile, welche deutliche Unterschiede aufweisen. Diese Beobachtung wird zudem für Bauteile nachgewiesen, welche Formelemente mit ähnlichen Abmaßen aufweisen. Konkret zeigen die Ergebnisse, dass Bauteile mit Fasen, Radien, Nuten sowie Bohrungen ähnlicher Abmessung zuverlässig durch die Objekterkennung unterschieden werden.

Für Konstrukteure resultiert aus den beschriebenen Ergebnissen eine umfangreiche Freiheit in der Gestaltung von Produkten. Zwar werden durch Tabelle 7-1 neue konstruktive Empfehlungen für eine montagegerechte Produktgestaltung abgeleitet, diese sind wie zuvor beschrieben, jedoch eher als Möglichkeit zu verstehen, die Erkennungsgenauigkeit zu optimieren. Da nachgewiesen wird, dass alle untersuchten Objekte unabhängig von ihren konstruktiven Eigenschaften zuverlässig detektiert werden, besteht außerdem das Potential, etablierte, montagegerechte Gestaltungsrichtlinien zu überarbeiten.

Indem sich die Montage durch den Einsatz von Objekterkennungsalgorithmen in Echtzeit nachvollziehen lässt, können Mitarbeiter sofort auf entstehende Fehler hingewiesen werden. Da sich die Gefahr auftretender Montagefehler damit deutlich reduzieren lässt, besteht die Möglichkeit, konstruktive Einschränkungen zur Vermeidung dieser Fehler zu reduzieren. Konstrukteuren wird damit eine größere Freiheit in der Produktgestaltung gewährt, sodass sie andere Ziele wie z. B. eine Kosten- oder Gewichtsoptimierung in den Vordergrund stellen können.

Um Optimierungspotenziale aufzudecken, werden alle in Abschnitt 2.1.3 aufgeführten Gestaltungsrichtlinien hinsichtlich ihrer Notwendigkeit in einem Montagesystem analysiert, welches Verwechslungs- und Ausrichtungsfehler durch eine Objekterkennung zuverlässig verhindert. Dabei zeigt sich, dass insgesamt fünf bestehende konstruktive Empfehlungen durch den Einsatz eines überwachenden Assistenzsystems entfallen können. Tabelle 7-2 zeigt die identifizierten Gestaltungsrichtlinien sowie ihre Autoren.

Tabelle 7-2: Entfallende montagegerechte Gestaltungsrichtlinien

Nr.	Gestaltungsrichtlinie	Autoren
1	Anstreben symmetrischer Konturen	Pahl/Beitz, Gairola, Bäßler, Hesse
2	Vermeiden von Fast-Symmetrien	Pahl/Beitz, Gairola, Bäßler, Hesse
3	Anstreben geometrischer Erkennungsmerkmale / Verwechslungen vermeiden	Pahl/Beitz, Gairola, Bäßler, Hesse
4	Bevorzugen von Erkennungsmerkmalen an Außenkonturen	Pahl/Beitz, Gairola
5	Anstreben einer erkenntungsgerechten Gestaltung für den Einsatz sensorisierter Handhabungseinrichtungen	Hesse

Darüber hinaus kann die Anwendung von Poka-Yoke in der Produktgestaltung entfallen. Das Prinzip basiert auf dem Grundsatz, Bauteile so zu gestalten, dass ein falscher Einbau durch die geometrischen Eigenschaften und Maße ausgeschlossen ist. Es bedient sich dabei ähnlichen Gestaltungsrichtlinien wie in Tabelle 7-2 beschrieben und empfiehlt die Verwendung von Gleichteilen sowie die Gestaltung von symmetrischen oder deutlich asymmetrischen Bauteilen, um deren fehlerfreie Positionierung zu gewährleisten. Außerdem ist die Unterscheidbarkeit einzelner Teile sowie eine eindeutige Bauteilkennzeichnung zu beachten.

Die Nicht-Anwendung der aufgeführten Gestaltungsrichtlinien ist jedoch nur sinnvoll, wenn das montageüberwachende Assistenzsystem konsequent durch die Anwender genutzt wird. Übergehen Mitarbeiter das System, besteht durch das Entfallen der konstruktiven Empfehlungen ein erhöhtes Fehlerpotenzial, da fehlervermeidende Produkteigenschaften reduziert werden. Ob die Vereinfachung der etablierten Gestaltungsrichtlinien in der Praxis umsetzbar und sicher ist, sollte daher durch weitere Untersuchungen eruiert werden. Bestandteil dieser Studie sollte außerdem sein, ob der Verzicht auf die genannten Gestaltungsrichtlinien erhöhte Montagezeiten zur Folge hat. Da die konstruktiven Empfehlungen neben der Fehlervermeidung auch die Bauteilorientierung im Handhabungsprozess beeinflussen, besteht die Gefahr, dass Orientierungsvorgänge erschwert werden. Insgesamt ist daher zu prüfen, ob das Entfallen der Gestaltungsrichtlinien und die damit einhergehende konstruktive Freiheit wirtschaftlich ist.

8 Zusammenfassung

Durch den Einsatz von Deep Learning in der manuellen Montage bestehen große Potenziale, Montagezeiten zu reduzieren und Montagefehler zu vermeiden. So können Position, Lage und Art von Bauteilen durch die Erfassung des Montageprozesses mittels Kamera und anschließender Objektdetektion bestimmt und Informationen über Arbeitsschritte, Montagefehler oder den aktuellen Produktzustand abgeleitet werden.

Der Einsatz von Deep Learning ist jedoch nur zielführend, wenn die zu montierenden Produkte zuverlässig erfasst werden. Zum aktuellen Zeitpunkt bestehen zwar zahlreiche Empfehlungen zur montagegerechten Produktgestaltung, diese behandeln den Aspekt des Einsatzes einer Objekterkennung jedoch an keiner Stelle. Es fehlen Erkenntnisse zum Einfluss verschiedener Produkteigenschaften auf die Detektierbarkeit durch Deep Learning.

Die vorliegende Arbeit greift diese Problemstellung auf und liefert Gestaltungsrichtlinien, welche eine optimierte Objekterkennung ermöglichen. Basierend auf einer umfangreichen Analyse der Deep Learning Technologie werden insgesamt 13 Hypothesen zum Einfluss verschiedener Produkteigenschaften auf die Objekterkennung aufgestellt und durch Experimente unter Laborbedingungen untersucht. Die Hypothesen gliedern sich in die Bereiche Bauteilgestaltung, Baugruppengestaltung, Auswahl von Normteilen und Dimensionierung von Formelementen. Neben dem Einfluss spiegelnder und dunkler Oberflächen, wird die Auswirkung weicher, komplexer und ähnlicher Konturen betrachtet. Darüber hinaus wird ermittelt, wie sich die Kombination von Bauteilen ähnlicher und nicht-ähnlicher Optik auf die Erkennungsgenauigkeit auswirkt. Weiterhin wird untersucht, ob Normteile ähnlicher Nennweiten, Schrauben mit gleichartigen Köpfen sowie Formelemente mit ähnlichen Abmessungen zuverlässig detektiert werden. Neben Fasen werden dabei Radien, Nuten sowie Bohrungen berücksichtigt.

Zur Verifizierung der Hypothesen werden Untersuchungsobjekte mit unterschiedlichen konstruktiven Eigenschaften erstellt und mit Hilfe von Deep Learning detektiert. Dazu wird zunächst ein Versuchsaufbau entwickelt und eine umfangreiche statistische Versuchsplanung durchgeführt. Bestandteil dieser Planung sind ausführliche Voruntersuchungen zur Festlegung des Stichprobenumfangs sowie zur Definition der konstanten Einflussfaktoren. Neben der Iterationenanzahl im Trainingsprozess wird dabei der optimale Umfang der Trainings- und Validierungsdatensätze ermittelt. Basierend auf den zu verifizierenden Hypothesen werden Untersuchungsobjekte konzipiert und gefertigt bzw. beschafft. Im Anschluss erfolgt die Aufnahme von Trainings- und Validierungsbildern sowie ihre Annotation. Dazu wird jedes Bild mit Informationen zur Position und Klasse der enthaltenen Objekte versehen. Unter Nutzung dieser Daten werden im Rahmen der Hauptuntersuchung insgesamt 330 Objekterkennungsmodelle trainiert und validiert. Als Ergebnis der Versuche wird die Genauigkeit der Detektion in Abhängigkeit der konstruktiven Eigenschaften erfasst.

Die statistische Auswertung der Versuchsergebnisse zeigt, dass Baugruppen mit einer höheren Genauigkeit detektiert werden, wenn ihre Einzelteile unterschiedliche optische Eigenschaften aufweisen. Baugruppen mit Einzelteilen identischer Optik erzielen hingegen

niedrigere Erkennungsgenauigkeiten. Die weiteren zwölf Hypothesen können nicht statistisch signifikant verifiziert werden. Neben den Aussagen zu den aufgestellten Hypothesen lassen sich jedoch weitere Effekte aus den Versuchsergebnissen ableiten. So zeigt sich, dass Schrauben mit einer höheren Genauigkeiten erkannt werden, wenn diese nicht in einem anderen Bauteil versinken. Darüber hinaus ergibt sich eine höhere Erkennungsgenauigkeit, je größer der Nenndurchmesser einer Schraube ist. Diese Beobachtung wird durch die vorliegende Arbeit für einen Nenndurchmesserbereich von M3 bis M10 nachgewiesen.

Um Störeinflüsse zu reduzieren, erfolgt die empirische Untersuchung unter Laborbedingungen. Durch anschließende Versuche, welche die Montage einer industriellen, explosionsgeschützten Leuchte und eines kommerziell vertriebenen Regals unter realen Umgebungsbedingungen beinhalten, wird die Übertragbarkeit der Ergebnisse auf reale Anwendungen überprüft. Dabei zeigt sich, dass die Ergebnisse der empirischen Untersuchung zu einem überwiegenden Anteil auf reale Anwendungen übertragbar sind. Insbesondere die statistisch signifikanten Beobachtungen werden durch die Übertragbarkeitsprüfung bestätigt, sodass als Ergebnis der vorliegenden Arbeit insgesamt vier Gestaltungsrichtlinien für montagegerechte Konstruktionen mittels Deep Learning abgeleitet werden.

Insgesamt zeigt die vorliegende Arbeit, dass alle betrachteten Untersuchungsobjekte unabhängig von ihren konstruktiven Eigenschaften zuverlässig und mit sehr hoher Genauigkeit detektiert werden. Die abgeleiteten Gestaltungsrichtlinien sind daher nicht zwangsläufig erforderlich, um eine zuverlässige Objekterkennung durch Deep Learning zu erzielen. Vielmehr dienen sie als Optimierungsoptionen für die Erkennungsgenauigkeit. Stehen den vorgeschlagenen Empfehlungen andere Interessen wie Kostenfaktoren, Designfaktoren oder konträre Gestaltungsrichtlinien entgegen, ist die Notwendigkeit der Erkennungsgenauigkeitsoptimierung zu prüfen.

Neben den neu abgeleiteten konstruktiven Empfehlungen, werden außerdem bestehende Gestaltungsrichtlinien hinsichtlich ihrer Erfordernisse überprüft. Durch eine Analyse etablierter Richtlinien für eine montagegerechte Produktgestaltung werden Regeln identifiziert, welche durch den Einsatz einer Objekterkennung im Montageprozess entfallen können, da Montagefehler (z. B. Verwechslungen oder Falschrichtungen) durch das Assistenzsystem verhindert werden. Sie sind damit nicht mehr im Rahmen der Produktgestaltung zu berücksichtigen, sodass Konstrukteuren eine größere Freiheit gewährt werden kann. Da durch die ausgemachten Gestaltungsrichtlinien auch die Orientierung im Handhabungsprozess und damit die Montagezeit beeinflusst werden könnte, sollte die Wirkung der reduzierten konstruktiven Empfehlungen auf den Montageprozess durch weitere Untersuchungen im Detail analysiert werden.

Die Versuche in realen Produktumgebungen machen deutlich, dass die Erkennungsgenauigkeit nicht nur von den zu detektierenden Objekten, sondern auch von äußeren Einflüssen wie der Montageumgebung abhängt. So werden dunkle Bauteile auf dunklem Hintergrund und helle Bauteile auf hellem Hintergrund mit einer geringeren Genauigkeit

detektiert, als helle bzw. dunkle Bauteile auf dunklem bzw. hellem Grund. Der Fokus der vorliegenden Arbeit liegt jedoch auf der Gestaltung der zu montierenden Objekte. Um den Einfluss der Montageumgebung zu erfassen und wissenschaftliche Aussagen bezüglich ihrer optimalen Gestaltung abzuleiten, sollten sich weitere Untersuchungen an die vorliegende Arbeit anschließen.

Darüber hinaus ist die Übertragbarkeit der Erkenntnisse auf reale Produkte in realen Montageumgebungen durch die Anwendung auf unterschiedliche Fallbeispiele weiter zu untersuchen. Im Rahmen der vorliegenden Arbeit wurde die Anwendbarkeit zwar anhand zweier Beispielprodukte überprüft, diese können jedoch nicht die sehr hohe Variabilität aller am Markt verfügbaren Produkte abbilden.

Schlussendlich schafft die vorliegende Arbeit die konstruktionsseitige Grundlage für eine optimale Anwendung von Objekterkennungsalgorithmen basierend auf Deep Learning in der manuellen Montage. Sie liefert dem Konstrukteur der Zukunft durch die erarbeiteten Gestaltungsrichtlinien ein Werkzeug, mit welchem er den sich ändernden Rahmenbedingungen in der Montage begegnen kann.

9 Literatur

- [1] C. Schröder, „IW-Trends 2/2019. Industrielle Arbeitskosten im internationalen Vergleich,“ *Institut der deutschen Wirtschaft Köln e. V.*, Jg. 46, Nr. 2, Juni 2019. [Online]. Verfügbar unter: https://www.iwkoeln.de/fileadmin/user_upload/Studien/IW-Trends/PDF/2019/IW-Trends_2019-02-05_Industrielle_Arbeitskosten.pdf. [Zugriff am 02. Juni 2022].
- [2] E. Bogner, U. Löwen, und J. Franke, „Bedeutung der zukünftigen Produktion kundenindividueller Produkte in Losgröße 1,“ In *Interdisziplinäre Perspektiven zur Zukunft der Wertschöpfung*, T. Redlich, M. Moritz, und J. Wulfsberg, Hrsg. Wiesbaden: Springer Gabler, 2018, S. 63-75.
- [3] Promotorengruppe Kommunikation der Forschungsunion Wirtschaft – Wissenschaft, Deutsche Akademie der Technikwissenschaften e.V., „Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0,“ April 2013. [Online]. Verfügbar unter: https://www.bmbf.de/bmbf/shareddocs/downloads/files/umsetzungsempfehlungen_industrie4_0.pdf. [Zugriff am 02. Juni 2022].
- [4] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, und M. Hoffmann, „Industrie 4.0,“ In *Wirtschaftsinformatik*, Ausgabe 4/2014, Wiesbaden: Springer Gabler, August 2014, S. 261-264.
- [5] Informationsdienst des Instituts der deutschen Wirtschaft, „Fluktuationsquote: Starke Wirtschaft führt zu mehr Jobwechseln,“ November 2021. [Online]. Verfügbar unter: <https://www.iwd.de/artikel/fluktuation-starke-wirtschaft-fuehrt-zu-mehr-jobwechseln-401583/>. [Zugriff am 02.06.2022].
- [6] Der Informationsdienst des Instituts der deutschen Wirtschaft, „Zuwanderer: Sprachniveau ist entscheidend,“ Oktober 2019. [Online]. Verfügbar unter: <https://www.iwd.de/artikel/zuwanderer-sprachniveau-ist-entscheidend-447362/>. [Zugriff am 02.06.2022].
- [7] S. Khan, H. Rahmani, S. Shah, und M. Bennamoun, *A Guide to Convolutional Neural Networks for Computer Vision*. Williston, VT: Morgan & Claypool, 2018.
- [8] Verein Deutscher Ingenieure e.V., „Wirtschaftliche Entscheidungen beim Konstruieren,“ VDI-Richtlinie 2235, Oktober 1987.
- [9] Verein Deutscher Ingenieure e.V., „Entwicklung technischer Produkte und Systeme - Modell der Produktentwicklung,“ VDI-Richtlinie 2221 Blatt 1, November 2019.
- [10] F. Reuleaux, *Lehrbuch der Kinematik*. Braunschweig: Vieweg, 1875.
- [11] G. Pahl und W. Beitz, *Konstruktionslehre*. Berlin: Springer, 1977.
- [12] H. Wögerbauer, *Die Technik des Konstruierens*. München: Oldenbourg, 1942.
- [13] F. Kesselring, *Technische Konstruktionslehre*. Berlin: Springer, 1954.

-
- [14] F. Hansen, *Konstruktionssystematik: Grundlagen für eine allgemeine Konstruktionslehre*. Berlin: Verlag Technik, 1965.
- [15] V. Hubka, *Theorie der Konstruktionsprozesse: Analyse der Konstruktionstätigkeit*. Berlin: Springer, 1976.
- [16] R. Koller, *Konstruktionsmethode für den Maschinen-, Geräte- und Apparatebau*. Berlin: Springer, 1976.
- [17] K. Roth, *Konstruieren mit Konstruktionskatalogen*. Berlin: Springer, 1982.
- [18] W. Rodenacker, *Methodisches Konstruieren: Grundlagen, Methodik, praktische Beispiele*, 4., überarbeitete Auflage, Berlin: Springer, 1991.
- [19] Verein Deutscher Ingenieure e.V., „Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte,“ VDI-Richtlinie 2221, November 1986.
- [20] Verein Deutscher Ingenieure e.V., „Konstruktionsmethodik. Methodisches Entwickeln von Lösungsprinzipien,“ VDI-Richtlinie 2222 Blatt 1, Juni 1997.
- [21] Verein Deutscher Ingenieure e.V., „Konstruktionsmethodik. Erstellung und Anwendung von Konstruktionskatalogen,“ VDI-Richtlinie 2222 Blatt 2, Februar 1982.
- [22] Verein Deutscher Ingenieure e.V., „Methodisches Entwerfen technischer Produkte,“ VDI-Richtlinie 2223, Januar 2004.
- [23] Verein Deutscher Ingenieure e.V., „Entwicklung mechatronischer und cyber-physischer Systeme,“ VDI/VDE-Richtlinie 2204, November 2021.
- [24] B. Lotter und H.-P. Wiendahl, Hrsg., *Montage in der industriellen Produktion. Ein Handbuch für die Praxis*. Berlin: Springer, 2006.
- [25] Verein Deutscher Ingenieure e.V., „Montage- und Handhabungstechnik - Handhabungsfunktionen, Handhabungseinrichtungen, Begriffe, Definitionen, Symbole,“ VDI-Richtlinie 2860 Blatt 1, Mai 1990.
- [26] B. Bender und K. Gericke, Hrsg., *Pahl/Beitz Konstruktionslehre. Methoden und Anwendung erfolgreicher Produktentwicklung*, 9. Auflage, Berlin: Springer, 2021.
- [27] Deutsches Institut für Normung, „Fertigungsverfahren Fügen – Teil 0: Allgemeines. Einordnung, Unterteilung, Begriffe,“ DIN 8593-0, September 2003.
- [28] U. Kurz, H. Hintzen und H. Laufenberg, *Konstruieren, Gestalten, Entwerfen. Ein Lehr und Arbeitsbuch für das Studium der Konstruktionstechnik*, 4. Auflage, Wiesbaden: Vieweg & Teubner, 2009.
- [29] G. Stoll, „Montagegerechte Produkte mit feature-basiertem CAD,“ Dissertation, Fakultät für Maschinenwesen, Technische Universität München, 1995.
- [30] G. Pahl und W. Beitz, *Konstruktionslehre. Handbuch für Studium und Praxis*, 2. Auflage, Berlin: Springer, 1986.

-
- [31] M. Andreasen, S. Kähler und T. Lund, *Montagegerechtes Konstruieren*, Berlin, Heidelberg: Springer, 1985.
- [32] S. Hesse, *Montage-Atlas. Montage- und automatisierungsgerecht konstruieren*, Darmstadt: Hoppenstedt Technik Tabellen Verlag, 1994.
- [33] R. Bäßler, *Integration der montagegerechten Produktgestaltung in den Konstruktionsprozess*, Berlin, Heidelberg: Springer, 1988.
- [34] A. Gairola, „Montagegerechtes Konstruieren – Ein Beitrag zur Konstruktionsmethodik,“ Dissertation, Technische Hochschule Darmstadt, 1981.
- [35] G. Boothroyd, P. Dewhurst und W.A. Knight, *Product design for manufacture and assembly*, New York: Dekker, 1994.
- [36] K. Ehrlenspiel, *Integrierte Produktentwicklung. Methoden für Prozessorganisation, Produkterstellung und Konstruktion*, München: Hanser, 1995.
- [37] K. Ehrlenspiel und H. Meerkamm, *Integrierte Produktentwicklung. Denkabläufe, Methodeneinsatz, Zusammenarbeit*, 6., überarbeitete und erweiterte Auflage, München: Hanser, 2017.
- [38] A. Turing, „Computing machinery and intelligence,“ In *Mind*, Ausgabe 59, Oktober 1950, S. 433–460.
- [39] C. Gethmann, P. Buxmann, J. Distelrath, B. Humm, S. Linger, V. Nitsch, J. Schmidt und I. Spiecker, *Künstliche Intelligenz in der Forschung. Neue Möglichkeiten und Herausforderungen für die Wissenschaft*. Berlin: Springer, 2022.
- [40] High-Level Expert Group on Artificial Intelligence, „A definition of AI: Main capabilities and scientific disciplines,“ April 2019. [Online]. Verfügbar unter: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341 [Zugriff am 29. Juni 2022].
- [41] K. Murphy, *Machine Learning. A Probabilistic Perspective*. Cambridge, MA: MIT Press, 2012.
- [42] J. Wang, Y. Ma, L. Zhang, und R. Gao, „Deep learning for smart manufacturing: Methods and applications,“ In *Journal of Manufacturing Systems*, Ausgabe 48, Juli 2018, S. 144-156.
- [43] W. McCulloch und W. Pitts, „A logical calculus of ideas immanent in nervous activity,“ In *Bulletin of Mathematical Biophysics*, Ausgabe 5, Heft 4, 1943, S. 115–133.
- [44] F. Rosenblatt, „The perceptron: A probabilistic model for information storage and organization in the brain,“ In *Psychological Review*, Ausgabe 65, Heft 6, 1958, S. 386–408.
- [45] D. O. Hebb, *The Organization of Behavior*. Wiley, NY: John Wiley & Sons, Inc., 1949.

-
- [46] B. Widrow und M. E. Hoff, "Adaptive switching circuits", In *IRE WESCON Convention Record 1960*, Ausgabe 4, 1960, S. 96–104.
- [47] M. L. Minsky und S. A. Papert, *Perceptrons*, Cambridge, MA: MIT Press, 1969.
- [48] I. Goodfellow, Y. Bengio und A. Courville, *Deep Learning*, Cambridge, MA: MIT Press, 2016.
- [49] G. E. Hinton, J. McClelland und D. Rumelhart, "Distributed Representations" In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D. E. Rumelhart und J. L. McClelland, Hrsg. Cambridge, MA: MIT Press, 1987, S. 77-109.
- [50] D. Rumelhart, G. Hinton und R. Williams, "Learning re representations by back-propagating errors," In *Nature*, Ausgabe 323, 1986, S. 533-536.
- [51] Y. LeCun und F.-S. Fogelman, "Modèles connexionnistes de l'apprentissage," In *Intellectia. Revue de l'Association pour la Recherche Cognitive*, Ausgabe 2-3, Heft 1, 1987, S. 114-143.
- [52] S. Hochreiter und J. Schmidhuber, "Long short-term memory," In *Neural Computation*, Ausgabe 9, Heft 8, 1997, S. 1735-1780.
- [53] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard und L. Jackel, "Handwritten Digit Recognition with a Back-Propagation Network," In *Advances in Neural Information Processing Systems 2*, D. S. Touretzky, Hrsg. San Francisco, CA: Morgan Kaufmann Publishers Inc., 1990, S. 396-404.
- [54] G. E. Hinton, S. Osindero und Y.-W. The, "A fast learning algorithm for deep belief nets," In *Neural Computation*, Ausgabe 18, Heft 7, 2006, S. 1527-1554.
- [55] Y. Bengio, P. Lamblin, D. Popovici und H. Larochelle, "Greedy Layer-Wise Training of Deep Networks" In *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, B. Schölkopf, J. Platt und T. Hofmann, Hrsg. Cambridge, MA: MIT Press, 2007, S. 153-160.
- [56] M. Ranzato, C. Poultney, S. Chopra und Y. LeCun, "Efficient Learning of Sparse Representations with an Energy-Based Model" In *Advances in Neural Information Processing Systems 19: Proceedings of the 2006 Conference*, B. Schölkopf, J. Platt und T. Hofmann, Hrsg. Cambridge, MA: MIT Press, 2007, S. 1137-1144.
- [57] G. E. Dahl, M. A. Ranzato, A. Mohamed und G. Hinton, "Phone recognition with the mean-covariance restricted Boltzmann machine" In *Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010*, J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel und A. Culotta, Red Hook, Hrsg. NY: Curran Associates Inc, 2010, S. 469-477.
- [58] F. Seide, G. Li und D. Yu, "Conversational speech transcription using context-dependent deep neural networks," In *Proc. Interspeech 2011*, 2011, S. 437-440.

-
- [59] G. Hinton, L. Deng, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. Sainath und B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition," In *IEEE Signal Processing Magazine*, Ausgabe 29, Heft 6, 2012, S. 82-97.
- [60] P. Sermanet, K. Kavukcuoglu, S. Chintala und Y. LeCun, "Pedestrian detection with unsupervised multi-stage feature learning," In Proc. International Conference on Computer Vision and Pattern Recognition 2013, 2013, S. 3626-3633.
- [61] C. Couprie, C. Farabet, L. Najman und Y. LeCun, "Indoor semantic segmentation using depth information," In Proc. International Conference on Learning Representations (ICLR) 2013, 2013, S. 1-8.
- [62] D. Ciresa, U. Meier, J. Masci und J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," In *Neural Networks: the official journal of the International Neural Network Society*, Ausgabe 32, 2012, S. 333-338.
- [63] P. Baldi, P. Sadowski und D. Whiteson, "Searching for exotic particles in high-energy physics with deep learning," In *Nature Communications*, Ausgabe 5, Heft 1, 2014.
- [64] S. Roy, W. Menapace, S. Oei, B. Luijten, E. Fini, C. Saltori, I. Huijben, N. Chennakeshava, F. Mento, A. Sentelli, E. Peschiera, R. Trevisan, G. Maschietto, E. Torri, R. Inchigolo, A. Smargiassi, G. Soldati, P. Rota, A. Passerini, R. J. G. van Sloun, E. Ricci und L. Demi, "Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound," In *IEEE Transactions on Medical Imaging*, Ausgabe 39, Heft 8, 2020, S. 2676-2687.
- [65] L. Kouhalvandi, L. Matekovits und I. Peter, "Deep Learning and its Benefits in Prediction of Patients Through Medical Images," In 2021 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2021, S. 139-142.
- [66] C. Tang, C. Yu, Y. Gao, J. Chen, J. Yang, J. Lang, C. Liu, L. Zhong, Z. He und J. Lv, "Deep learning in nuclear industry: A survey," In *Big Data Mining and Analytics*, Ausgabe 5, Heft 2, 2022, S. 140-160.
- [67] J. -S. Kim und D. M. Lee, "A Deep Learning Module Design for Workspace Identification in Manufacturing Industry," In 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), 2021, S. 390-393.
- [68] F. – F. Li, J. Johnson und S. Yeung. CS231n, Class Lecture, Topic: "Convolutional Neural Networks for Visual Recognition." Stanford Vision and Learning Lab, Stanford University, Stanford, CA, Frühling 2017.
- [69] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn und A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge," In *International Journal of Computer Vision*, Ausgabe 88, 2010, S. 303-338.

- [70] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg und F. – F. Li, "ImageNet Large Scale Visual Recognition Challenge," In *International Journal of Computer Vision*, Heft 115, 2015, S. 211-252.
- [71] Google API, "Overview of the Open Images Challenge," Google LLC, 2022. [Online]. Verfügbar unter: https://storage.googleapis.com/openimages/web/challenge_overview.html. [Zugriff am 20. Juli 2022].
- [72] R. Padilla, S. L. Netto und E. A. B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," In *International Conference on Systems, Signals und Image Processing (IWSSIP) 2020*, 2020, S. 237-242.
- [73] M. Salikan, M. Rosli, N. Mohammad, I. Rahim und M. Fadzly, "Design and Analysis of Grass Cutting Machine By Using DFMA Method," In *IOP Conference Series: Materials Science and Engineering*, Ausgabe 864, 2020.
- [74] J. Butt und S. Jedi, "Redesign of an In-Market Conveyor System for Manufacturing Cost Reduction and Design Efficiency Using DFMA Methodology," In *Designs*, Ausgabe 4, Heft 1, 2020, S. 6.
- [75] B. Mora, I. Retolaza, M. Campos, A. Ramirez, M. Cabello und F. Martinez, "DEVELOPMENT OF A NEW DESIGN METHODOLOGY FOR LARGE SIZE PRODUCTS BASED ON DSM AND DFMA," In *Proceedings of the Design Society: DESIGN Conference*, Ausgabe 1, 2020, S. 2315–2324.
- [76] G. Formentini, N. Rodríguez und C. Favi, "Design for manufacturing and assembly methods in the product development process of mechanical products: a systematic literature review," In *The International Journal of Advanced Manufacturing Technology*. Ausgabe 120, 2022, S. 4307–4334.
- [77] A. Ramirez, A. Ramos, I. Retolaza, M. Cabello, M. Campos und F. Martinez, "New design for assembly methodology adapted to large size products: Application on a solar tracker design," In *Procedia CIRP*, Ausgabe 84, 2019, S. 468-473.
- [78] C.D. Naiju, "DFMA for product designers: A review," In *Materials Today: Proceedings*, Ausgabe 46, Heft 17, 2021, S. 7473-7478.
- [79] F. Bouissiere, C. Cuiller, P.-E. Dereux, C. Malchair, C. Favi, und G. Formentini, "Conceptual Design for Assembly in Aerospace Industry: A Method to Assess Manufacturing and Assembly Aspects of Product Architectures," In *Proceedings of the Design Society: International Conference on Engineering Design*, Ausgabe 1, Heft 1, 2019, S. 2961–2970.
- [80] V. Pomazan und S. Sinteia, "Generative design for assembly wrapping up," In *IOP Conference Series: Materials Science and Engineering*, Ausgabe 916, 2020, S. 23-27.

- [81] T. Robinson, I. Friel, C. Armstrong, A. Murphy, J. Butterfield, M. Price und A. Marzano, "Computer-aided design model parameterisation to derive knowledge useful for manufacturing design decisions," In *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, Ausgabe 232, Heft 4, 2018, S. 621-628.
- [82] A. Atzberger, S. Nicklas, J. Schrof, S. Weiss und K. Paetzold, *Agile Entwicklung physischer Produkte: Eine Studie zum aktuellen Stand in der industriellen Praxis*. Neubiberg: Universität der Bundeswehr München, 2020.
- [83] J. Heimicke, T. Rösel und A. Albers, "Analyse des Einflusses von Faktoren auf die agilen Fähigkeiten von Organisationseinheiten in der Entwicklung physischer Systeme," In *Entwickeln Entwerfen Erleben in Produktentwicklung und Design 2021*, 2021, S. 691-702.
- [84] H. Anacker, R. Dumitrescu, N. Michels, und R. Rasor, "EngOps – Ein Leitbild für das agile Zusammenwirken von Engineering und Operations," In *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, Ausgabe 117, Heft 6, 2022, S. 405-409.
- [85] E. Windisch, C. Mandel, S. Rapp, N. Bursac und A. Albers, "Approach for model-based requirements engineering for the planning of engineering generations in the agile development of mechatronic systems," In *Procedia CIRP*, Ausgabe 109, 2022, S. 550-555.
- [86] B. Corves, K. Gericke, K.-H. Grote, A. Lohrengel, N. Müller, A. Nagarajah, F. Rieg, G. Scharr und R. Stelzer (Hrsg.), *Tagungsband - 17. Gemeinsames Kolloquium Konstruktionstechnik – Agile Entwicklung physischer Produkte*. Aachen: RWTH Aachen Universität, 2019.
- [87] J. Miehl, B. Schleich, und S. Wartzack, "Digitale Zwillinge und Digitale Zwillingspaare im Kontext des Digital Engineerings: Mögliche Anwendungen und Herausforderungen," In *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, Ausgabe 115, Heft S1, 2020, S. 55-57.
- [88] C. Jung, B. Klöppel, S. Leidig und W. Holz, "Signalverarbeitung - Big Data und Digitale Zwillinge beschleunigen die Erprobungsphase," In *ATZ Elektron*, Ausgabe 14, 2019, S. 38–43.
- [89] L. Bauer, M. Bauer und M. Kley, "Modellbasierte Validierung der Prüfstandsdynamik zur Erprobung von Komponenten elektrifizierter Antriebsstränge mithilfe eines digitalen Zwillings," In *Stuttgarter Symposium für Produktentwicklung 2021*, 2021.
- [90] IEEE Xplore, "Home Page," *IEEE*, 2022. [Online]. Verfügbar unter: <https://ieeexplore.ieee.org/Xplore/home.jsp>. [Zugriff am 20. Juli 2022].
- [91] IEEE, "IEEE-History," *IEEE*, 2022. [Online]. Verfügbar unter: <https://www.ieee.org/about/ieee-history.html>. [Zugriff am 20. Juli 2022].

- [92] IEEE, "IEEE Annual Reports," *IEEE*, 2022. [Online]. Verfügbar unter: <https://www.ieee.org/about/annual-report.html>. [Zugriff am 20. Juli 2022].
- [93] Kimoknow, "Home Page," *Kimoknow UG*, 2022. [Online]. Verfügbar unter: <https://www.kimoknow.de/>. [Zugriff am 20. Juli 2022].
- [94] E. Monari und M. Kayser, "Mit Montageassistenz zur perfekten Bauanleitung," *Robert Bosch GmbH*, 2022. [Online]. Verfügbar unter: <https://www.bosch.com/de/stories/montageassistenz/>. [Zugriff am 20. Juli 2022].
- [95] S-Tec, "NeuroCAD: Intelligentes-Bauteile-Assessment," *S-Tec*, 2022. [Online]. Verfügbar unter: <https://s-tec.de/projekte/neurocad-intelligentes-cad-bauteile-assessment/>. [Zugriff am 20. Juli 2022].
- [96] Stanford, "Tutorial 1: Image Filtering," *Stanford University*, 2022. [Online]. Verfügbar unter: <https://ai.stanford.edu/~syyeung/cvweb/tutorial1.html> [Zugriff am 21. Juli 2022].
- [97] S. Ravi und A. M. Khan, "Operators used in edge detection computation: A case study," In *International Journal of Applied Engineering Research*, Ausgabe 7, Heft 11, 2012, S. 2022-2027.
- [98] G. Kramper, „Squad Auto – dieses Micro-Car aus Holland ist billig und lädt sich in der Sonne selbst auf,“ *Stern*, 18. Mai 2022. [Online]. Verfügbar unter: <https://www.stern.de/auto/fahrberichte/squad-auto---dieses-micro-car-aus-holland-ist-billig-und-laedt-sich-in-der-sonne-selbst-aus-31873432.html>. [Zugriff am 04. Januar 2023].
- [99] G. Kramper, „In China ist die 1-Million-Kilometer-Batterie bereits Realität,“ *Stern*, 12. April 2021. [Online]. Verfügbar unter: <https://www.stern.de/auto/news/auto-industrie--1-million-kilometer-batterie-in-china-bereits-realitaet-30456448.html>. [Zugriff am 04. Januar 2023].
- [100] Jenaer Nahverkehr, „Neue Busse im Jenaer Nahverkehr,“ *Jenaer Nachrichten*, 02. September 2016. [Online]. Verfügbar unter: <https://www.jenaer-nachrichten.de/stadtleben/4746-neue-busse-im-jenaer-nahverkehr>. [Zugriff am 04. Januar 2023].
- [101] BBC, „London bus cuts: Sadiq Khan pulls handbrake on planned changes,“ *BBC*, 23. November 2022. [Online]. Verfügbar unter: <https://www.bbc.com/news/uk-england-london-63727784>. [Zugriff am 04. Januar 2023].
- [102] COCO, "Home Page," *COCO Consortium*, 2022. [Online]. Verfügbar unter: <https://cocodataset.org/#home>. [Zugriff am 25. Juli 2022].
- [103] ImageNet, "Home Page," *Stanford Vision Lab, Stanford University, Princeton University*, 2021. [Online]. Verfügbar unter: <https://www.image-net.org/>. [Zugriff am 27. Juli 2022].
- [104] M. Kornmeier, *Wissenschaftstheorie und wissenschaftliches Arbeiten. Eine Einführung für Wirtschaftswissenschaftler*, Heidelberg, BW: Physica Verlag, 2007.

- [105] M. Janczyk und R. Pfister, *Interferenzstatistik verstehen. Von A wie Signifikanztest bis Z wie Konfidenzintervall*, 3. Auflage, Berlin: Springer Spektrum, 2020.
- [106] K. Siebertz, D. Bebbler und T. Hochkirchen, *Statistische Versuchsplanung. Design of Experiments (DoE)*, 2. Auflage, Berlin: Springer Vieweg, 2017.
- [107] U. Iqbal, M. Z. B. Riaz, J. Barthelemy, N. Hutchison und P. Perez, "Floodborne Objects Type Recognition Using Computer Vision to Mitigate Blockage Originated Floods," In *Water*, Ausgabe 14, Heft 7, August 2022.
- [108] M. Saleem, S. Khanchi, J. Potgieter und K. M. Arif, "Image-Based Plant Disease Identification by Deep Learning Meta Architectures," In *Plants*, Ausgabe 9, Heft 11, Oktober 2020.
- [109] C. Qiu, G. Tian, J. Zhao, Q. Liu, S. Xie und K. Zheng, "Grape Maturity Detection and Visual Pre-Positioning Based on Improved YOLOv4," In *Electronics*, Ausgabe 11, Heft 17, August 2022.
- [110] H. Ji, J. Yu, F. Lao, Y. Zhuan, Y. Wen und G. Teng, "Automatic Position Detection and Posture Recognition of Grouped Pigs Based on Deep Learning," In *Agriculture*, Ausgabe 12, Heft 9, August 2022.
- [111] M. Humayun, F. Ashfaq, N. Z. Jhanji und M. K. Alsadun, "Traffic Management: Multi-Scale Vehicle Detection in Varying Weather Conditions Using YOLOv4 and Spatial Pyramid Pooling Network," In *Electronics*, Ausgabe 11, Heft 7, September 2022.
- [112] S. Rajaraman, P. Guo, Z. Xue und S. K. Antani, "A Deep Modality-Specific Ensemble for Improving Pneumonia Detection in Chest X-rays," In *Diagnostics*, Ausgabe 12/6, Juni 2022.
- [113] J. Yep, W. Yolland und P. Tschandl, "Multimodal skin lesion classification using deep learning," In *Exp Dermatol.*, Ausgabe 27, Heft 11, September 2018.
- [114] S. Alkentar, B. Alsahwa, A. Assalem und D. Karakoll, "Practical comparison of the accuracy and speed of YOLO, SSD and Faster RCNN for drone detection," In *Journal of Engineering*, Ausgabe 27, Heft 8, August 2021.
- [115] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Fathi, Z. Wojna, Y. Song, S. Guadarrama und K. Murphy, "Speed-accuracy trade-offs for modern convolutional object detectors", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, S. 3296-3297.
- [116] L. Tan, T. Huangfu, L. Wu und W. Chen, "Comparison of YOLO v3, Faster R-CNN, and SSD for Real-Time Pill Identification," In *BMC Medical Informatics and Decision Making*, Ausgabe 21, Heft 1, Juli 2021.
- [117] S. Srivastava, A. V. Divekar, C. Anilkumar, I. Naik, V. Kulkarni und V. Pattabiraman, "Comparative analysis of deep learning image detection algorithms," In *Journal of Big Data*, Ausgabe 8, Heft 1, Nr. 66, Mai 2021.

- [118] Google, Crash Course, "Training and Test Sets: Splitting Data" Google Foundational courses, Google, Mountain View, CA: 18. September 2022.
- [119] V. R. Joseph, "Optimal ratio for data splitting," In *Statistical Analysis and Data Mining: The ASA Data Science Journal*, Ausgabe 15, Heft 4, 2022, S. 531-538.
- [120] S. Ren, K. He, R. Girshick und J. Sun, "Faster R-CNN: Towards RealTime Object Detection with Regional Proposal Networks," In Proc. Advances in Neural Information Processing Systems 28, 2015, S. 91-99.
- [121] K. He, G. Gkioxari, P. Dollar und R. Girshick, "Mask R-CNN," In Proc. IEEE International Conference on Computer Vision (ICCV) 2017, 2017, S. 2961-2969.
- [122] P. Soviany und R. T. Ionescu, "Optimizing the Trade-off between Single-Stage and Two-Stage Deep Object Detectors using Image Difficulty Prediction," *Cornell University*, 31. August 2018. [Online]. Verfügbar unter: <https://arxiv.org/pdf/1803.08707.pdf>. [Zugriff am 09. September 2022]
- [123] L. Jiao, F. Zheng, F. Liu, S. Yang, L. Li, Z. Feng und R. Qu, "A Survey of Deep Learning-based Object Detection," *Cornell University*, 10. Oktober 2019. [Online]. Verfügbar unter: <https://arxiv.org/pdf/1907.09408.pdf>. [Zugriff am 09. September 2022].
- [124] J. Redmon, S. Divvala, R. Girshick und A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016, 2016, S. 779-788.
- [125] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu und A. C. Berg, "SSD: Single Shot MultiBox Detector," In Proc. European Conference on Computer Vision 2016, 2016, S. 21-37.
- [126] Tensorflow, "Home Page," *Google*, 2022. [Online]. Verfügbar unter: <https://www.tensorflow.org/>. [Zugriff am 09. September 2022].
- [127] PyTorch, "Home Page," *Meta AI*, 2022. [Online]. Verfügbar unter: <https://pytorch.org/>. [Zugriff am 09. September 2022].
- [128] M. C. Chirodea, O. C. Novac, C. M. Novac, N. Bizon, M. Oproescu und C. E. Gordan, "Comparison of Tensorflow and PyTorch in Convolutional Neural Network - based Applications," In International Conference on Electronics, Computers and Artificial Intelligence (ECAI) 13, 2021, S. 1-6.
- [129] Y. Wu, A. Kirillov, F. Massa, W. Y. Lo und R. Girshick, "Detectron2: A PyTorch-based modular object detection library," *ai.facebook.com*, 10. Oktober 2019. [Online]. Verfügbar unter: <https://ai.facebook.com/blog/-detectron2-a-pytorch-based-modular-object-detection-library/>. [Zugriff am 09. September 2022].
- [130] Y. Wu, A. Kirillov, F. Massa, W. Y. Lo und R. Girshick, "Detectron2," *github.com*, 15. November 2021. [Online]. Verfügbar unter: <https://github.com/facebookresearch/detectron2>. [Zugriff am 09. September 2022].

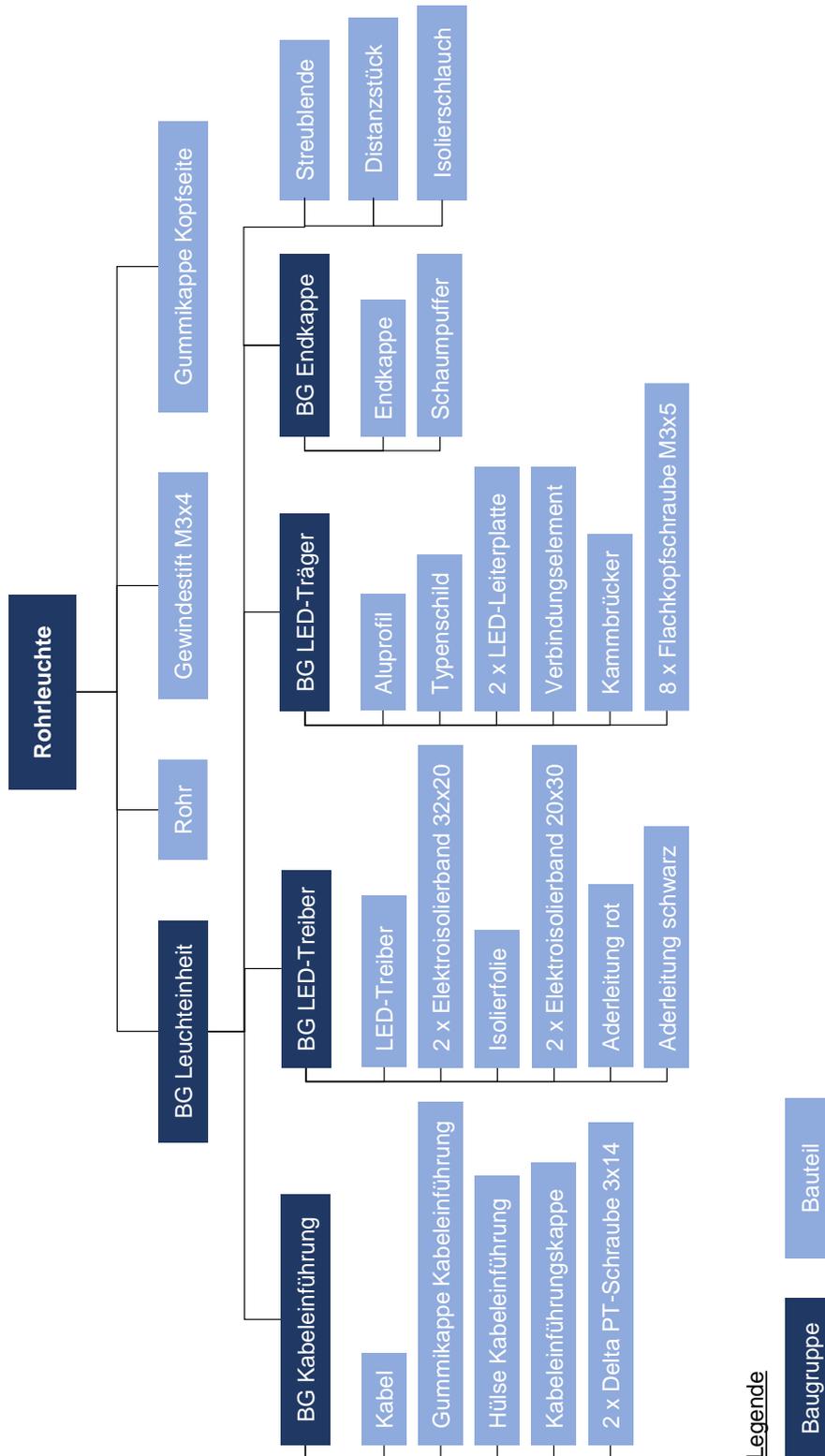
- [131] Y. Wu, A. Kirillov, F. Massa, W. Y. Lo und R. Girshick, "MODEL_ZOO," *github.com*, 14. November 2021. [Online]. Verfügbar unter: https://github.com/facebookresearch/detectron2/blob/main/MODEL_ZOO.md. [Zugriff am 09. September 2021].
- [132] K. He, X. Zhang, S. Ren und J. Sun, "Deep Residual Learning for Image Recognition," In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016, 2016, S. 770-778.
- [133] G. Aurélien, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems*, 2. Auflage, Sebastopol, CA: O'Reilly Media Inc, 2019.
- [134] T.-Y. Lin, P. Dollár, R. Girshick, K. He B. Hariharan und S. Belongie, "Feature Pyramid Networks for Object Detection," In Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017, 2017, S. 936-944.
- [135] R. Rothe, M. Guillaumin und L. Van Gool, "Non-Maximum Suppression for Object Detection by Passing Messages between Windows," In Proc. Asian Conference on Computer Vision (ACCV) 2014, 2015, S. 290-306.
- [136] W. Kleppmann, *Versuchsplanung – Produkte und Prozesse optimieren*, 10., überarbeitete Auflage, München: Hanser, 2020.
- [137] J. Gerlach, *Technische Zeichnungen zur Dissertation*, Jena: Ernst-Abbe-Hochschule Jena, Januar 2023, unveröffentlicht.
- [138] Ultimaker, "Ultimaker S5," *Ultimaker*. [Online]. Verfügbar unter: https://ultimaker.com/3d-printers/ultimaker-s5?utm_medium=cpc&utm_source=google&utm_campaign=2022_Alwayson_srengineer_traffic_do_DE. [Zugriff am 10. Oktober 2022].
- [139] A. Riedel, Video2Image. [Programmcode]. Jena: Ernst-Abbe-Hochschule Jena, 2020, unveröffentlicht.
- [140] J. Gerlach, *Messwerte zur Dissertation – Lichtstärke*, Jena: Ernst-Abbe-Hochschule Jena, Januar 2023, unveröffentlicht.
- [141] A. Riedel, Image2Coco. [Programmcode]. Jena: Ernst-Abbe-Hochschule Jena, 2020, unveröffentlicht.
- [142] A. Riedel, Labelme2Coco. [Programmcode]. Jena: Ernst-Abbe-Hochschule Jena, 2020, unveröffentlicht.
- [143] conansherry, "detectron2 windows build," *github.com*, 30. November 2019. [Online]. Verfügbar unter: <https://github.com/conansherry/detectron2>. [Zugriff am 13. Oktober 2022].
- [144] Anaconda, "Home Page," *Anaconda*. [Online]. Verfügbar unter: <https://www.anaconda.com/>. [Zugriff 13. Oktober 2022].

- [145] Nvidia, "CUDA Toolkit – Free Tools and Training," *Nvidia*. [Online]. Verfügbar unter: <https://developer.nvidia.com/cuda-toolkit>. [Zugriff am 13. Oktober 2022].
- [146] Visual Studio, "Downloads," *Microsoft*. [Online]. Verfügbar unter: <https://visualstudio.microsoft.com/de/downloads/>. [Zugriff am 13. Oktober 2022].
- [147] OpenCV, "Home Page," *OpenCV*. [Online]. Verfügbar unter: <https://opencv.org/>. [Zugriff am 13. Oktober 2022].
- [148] K. Wada, "labelme: Image Polygonal Annotation with Python," *github.com*, 26. September 2022. [Online]. Verfügbar unter: <https://github.com/wkentaro/labelme>. [Zugriff am 13. Oktober 2022].
- [149] Visual Studio, "Code," *Microsoft*. [Online]. Verfügbar unter: <https://code.visualstudio.com/>. [Zugriff am 13. Oktober 2022].
- [150] GitHub, "facebookresearch/ detectron2/ config/ defaults.py," *github.com*. [Online]. Verfügbar unter: <https://github.com/facebookresearch/detectron2/blob/main/detectron2/config/defaults.py>. [Zugriff am 13. Oktober 2022].
- [151] GitHub, "facebookresearch/ detectron2/ configs/ COCO-Detection/ fast_rcnn_R_50_FPN_1x.yaml," *github.com*. [Online]. Verfügbar unter: https://github.com/facebookresearch/detectron2/blob/main/configs/COCO-Detection/fast_rcnn_R_50_FPN_1x.yaml. [Zugriff am 13. Oktober 2022].
- [152] IBM, „IBM SPSS-Software,“ *IBM*. [Online]. Verfügbar unter: <https://www.ibm.com/de-de/spss>. [Zugriff am 13. Oktober 2022].
- [153] L. Sachs, *Angewandte Statistik*. Berlin, Heidelberg: Springer, 2013.
- [154] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Hoboken: Taylor and Francis, 1988.
- [155] R. STAHL AG, „Rohrleuchte mit LED Reihe 6036,“ *R. STAHL AG*. [Online]. Verfügbar unter: <https://r-stahl.com/de/global/produkte/leuchten/rohrleuchten/rohrleuchte-mit-led-reihe-6036-33916/274247-56306/>. [Zugriff am 11. November 2022].
- [156] A. Riedel, „Transformation von Montageprozessen in die digitale Welt: Erfassen von manuellen Montageschritten durch Objekterkennungsalgorithmen,“ Masterarbeit, Ernst-Abbe-Hochschule Jena, Jena, 2020.
- [157] A. Riedel, J. Gerlach, M. Dietsch, S. Herbst, F. Engelmann, N. Brehm und T. Pfeifroth, „A deep learning-based worker assistance system for error prevention: Case study in a real-world manual assembly,“ In *Advances in Production Engineering & Management*, Ausgabe 16, Heft 4, 2021, S. 393-404.
- [158] A. Bochkovskiy, C.-Y. Wang und H.-Y. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *Cornell University*, 23. April 2020. [Online].

- Verfügbar unter: <https://arxiv.org/pdf/2004.10934.pdf>. [Zugriff am 02. Februar 2022].
- [159] J. Redmon, "Darknet: Open Source Neural Networks in C," [Online]. Verfügbar unter: <http://pjreddie.com/darknet/> [Zugriff am 02. Februar 2022].
- [160] IKEA, "KALLAX, Regal, weiß, 42x42 cm," *IKEA*. [Online]. Verfügbar unter: <https://www.ikea.com/de/de/p/kallax-regal-weiss-20301554/#content>. [Zugriff am 10. Dezember 2022].
- [161] IKEA, "KALLAX, Regal, schwarzbraun, 42x42 cm," *IKEA*. [Online]. Verfügbar unter: <https://www.ikea.com/de/de/p/kallax-regal-schwarzbraun-70309840/#content>. [Zugriff am 10. Dezember 2022].
- [162] K. Steuerwald, "Mit der IKEA App per Augmented Reality einrichten," *IKEA*, 23. September 2019. [Online]. Verfügbar unter: <https://www.ikea.com/de/de/this-is-ikea/corporate-blog/ikea-place-app-augmented-reality-puba55c67c0>. [Zugriff am 10. Dezember 2022].
- [163] A. Pickard, "AssemblyAR," [Online]. Verfügbar unter: <https://adampickard.com/AssembleAR>. [Zugriff am 11. Dezember 2022].
- [164] IKEA, "KALLAX," *IKEA*. [Online]. Verfügbar unter: https://www.ikea.com/de/de/assembly_instructions/kallax-regal-weiss__AA-2135685-3-1.pdf. [Zugriff am 11. Dezember 2022].

A Anhang

A.1 Erzeugnisgliederung der explosionsgeschützten Rohrleuchte



A.2 Programmcode – Video2image

```

"""
Author: Alexander Riedel
Date: 10/07/2020
Name: video2image

Choose a framerate, choose multiple videos, choose a destination path, choose a resize factor
The script will split the videos and produce images with a name structure of: video name + "_frame_"
+ number of frame

"""

import cv2
from tkinter import Tk
from tkinter import filedialog
import ntpath
from tkinter.simpledialog import askfloat

Tk().withdraw()

framerate = askfloat("Choose a Framerate", "Capture Frame every n seconds", initialValue="1")
resize = askfloat("Choose a Resize Factor", "Resize Factor", initialValue="1")

video_path = filedialog.askopenfilenames(title='Choose Video Files')
image_dest = filedialog.askdirectory(title='Choose a Destination Path')
j = 0

def getFrame(sec):
    cap.set(cv2.CAP_PROP_POS_MSEC, sec*1000)
    hasFrames, image = cap.read()
    if hasFrames:

        image = cv2.copyMakeBorder(image, 0, 0, 0, 0, cv2.BORDER_CONSTANT,value=[255,255,255])
        image = cv2.resize(image, (0,0), fx = resize, fy = resize)
        height, width = image.shape[0:2]
        cv2.imwrite(image_dest + "/" + vid_name + "_frame_" + str(count) + "_res_" + str(width) + "x"
+ str(height) + ".jpg", image)

    return hasFrames

for i in range(len(video_path)):
    vid_name=ntpath.splitext(ntpath.basename(video_path[i]))[0]
    cap = cv2.VideoCapture(video_path[i])
    sec = 0
    count = 1
    success = getFrame(sec)

    while success:
        count +=1
        sec += framerate
        sec = round(sec, 2)
        success = getFrame(sec)
        if sec%20 == 0:
            print(str(sec) + " Frames converted to Images")

```

A.3 Programmcode – Image2coco

```
"""
```

```
Author: Alexander Riedel
Date: 10/07/2020
Name: image2coco
```

```
Choose a choose multiple Pictures, choose a destination path
The script will create a coco annotation JSON with the category structure of:
Supercategory_Category_anything.jpg
Supercategory and category are the same in this.
Because of this, only one object per Image is possibly!
```

```
It will only get the bbox and the area of the bbox. The bbox is created through maximum contours,
so take care that only the object of interest is in the input picture and nothing else! Check after
the process!
```

```
"""
```

```
from PIL import Image
from tkinter import filedialog
from tkinter import Tk
import ntpath
import random as rng
import cv2
import imutils
import numpy as np
import json
import ntpath
import os
from tkinter import *
from PIL import ImageTk, Image

Tk().withdraw()

image_files = filedialog.askopenfilenames(title='Choose Images')
dest = filedialog.askdirectory(title='Choose a Destination Path')

def uniquify(seq):
    checked = []
    for e in seq:
        if e not in checked:
            checked.append(e)
    return checked

coco={}
coco["images"] = []
coco["categories"] = []
coco["annotations"]=[]

im_id=0
ann_id=1

cats_dict = {}
cats_list=[]

#show how many occurrences of each category
for images in image_files:
    cats_list.append(ntpath.basename(images).split('_')[0])
cats_list_unique = uniquify(cats_list)
for e in cats_list_unique:
    print(e, cats_list.count(e))

#write categories and according supercategory in dict
```

```

for images in image_files:
    cats_dict[ntpath.basename(images).split('_')[0]] = ntpath.basename(images).split('_')[0]

#write cats und sup cats to json file
for cat in cats_dict:
    coco["categories"].append({
        "supercategory": cats_dict[cat],
        "id": list(cats_dict.keys()).index(cat),
        "name" : cat,
    })

for images in image_files:
    image = cv2.imread(images)

    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    blurred = cv2.GaussianBlur(gray, (5, 5), 0)
    canny = cv2.Canny(blurred, 1, 200) #play with these values

    # find contours in the thresholded image and initialize the
    # shape detector
    cnts = cv2.findContours(canny.copy(), cv2.RETR_EXTERNAL,
        cv2.CHAIN_APPROX_SIMPLE)
    cnts = imutils.grab_contours(cnts)

    height, width, channels = image.shape

    extLeft = width
    extRight = 0
    extTop = height
    extBot = 0
    #get maximum dimensions of the contours (this will be the bbox later!)
    for c in cnts:
        extLeft1 = tuple(c[c[:, :, 0].argmin()][0])[0]
        extRight1 = tuple(c[c[:, :, 0].argmax()][0])[0]
        extTop1 = tuple(c[c[:, :, 1].argmin()][0])[1]
        extBot1 = tuple(c[c[:, :, 1].argmax()][0])[1]
        cv2.drawContours(image, [c], 0, 255, -1)

        if extLeft1 < extLeft:
            extLeft = extLeft1
        if extRight1 > extRight:
            extRight = extRight1
        if extTop1 < extTop:
            extTop = extTop1
        if extBot1 > extBot:
            extBot = extBot1

    x = extLeft
    y = extTop
    w = extRight-extLeft
    h = extBot-extTop

    coco["images"].append({
        "height": image.shape[0],
        "width": image.shape[1],
        "id": im_id,
        "file_name": ntpath.basename(images)
    })

for categories in coco["categories"]:
    if categories["name"] == ntpath.basename(images).split('_')[0]:
        coco["annotations"].append({
            "segmentation": [],
            "iscrowd": 0,
            "area": float(w) * float(h),

```

```
        "image_id": im_id,
        "bbox": [float(x),float(y),float(w),float(h)],
        "category_id": categories["id"],
        "id": ann_id
    })

    im_id += 1
    ann_id += 1

    #uncomment to verify bounding boxes

    cv2.rectangle(image,(x,y),(x+w,y+h),(0,255,0),2)
    print(x,y,w,h)

    cv2.imshow(ntpath.basename(images), image)
    cv2.waitKey(0)
    cv2.destroyWindow(ntpath.basename(images))

cv2.destroyAllWindows()
with open(os.path.join(dest, 'coco.json'), 'w') as outfile:
    json.dump(coco, outfile)
```

A.4 Programmcode – Labelme2coco

```

"""
author: Alexander Riedel
date: 10/07/2020
name: labelme2coco

based on: https://github.com/Tony607/labelme2coco

"""

import os
import argparse
import json

from labelme import utils
import numpy as np
import glob
import PIL.Image

from tkinter import Tk
from tkinter import filedialog
import ntpath
from tkinter.simpledialog import askfloat

Tk().withdraw()
class labelme2coco(object):
    def __init__(self, labelme_json=[], save_json_path="./coco.json"):
        """
        :param labelme_json: the list of all labelme json file paths
        :param save_json_path: the path to save new json
        """
        self.labelme_json = labelme_json
        self.save_json_path = save_json_path
        self.images = []
        self.categories = []
        self.annotations = []
        self.label = []
        self.annID = 1
        self.height = 0
        self.width = 0

        self.save_json()

    def data_transfer(self):
        for num, json_file in enumerate(self.labelme_json):
            with open(json_file, "r") as fp:
                data = json.load(fp)
                self.images.append(self.image(data, num))
                for shapes in data["shapes"]:
                    label = shapes["label"].split("_")
                    if label not in self.label:
                        self.label.append(label)
                    points = shapes["points"]
                    self.annotations.append(self.annotation(points, label, num))
                    self.annID += 1

        # Sort all text labels so they are in the same order across data splits.
        self.label.sort()
        for label in self.label:
            self.categories.append(self.category(label))
        for annotation in self.annotations:
            annotation["category_id"] = self.getcatid(annotation["category_id"])

```

```

def image(self, data, num):
    image = {}
    img = utils.img_b64_to_arr(data["imageData"])
    height, width = img.shape[:2]
    img = None
    image["height"] = height
    image["width"] = width
    image["id"] = num
    image["file_name"] = data["imagePath"].split("/")[-1]

    self.height = height
    self.width = width

    return image

def category(self, label):
    category = {}
    category["supercategory"] = label[0]
    category["id"] = len(self.categories)
    category["name"] = label[0]
    return category

#deleted segmentation, made area only for bbox! (not segmentation)
def annotation(self, points, label, num):
    annotation = {}
    contour = np.array(points)
    x = contour[:, 0]
    y = contour[:, 1]
    #area = 0.5 * np.abs(np.dot(x, np.roll(y, 1)) - np.dot(y, np.roll(x, 1)))
    area = list(map(float, self.getbbox(points)))[2] * list(map(float, self.getbbox(points)))[3]
    #annotation["segmentation"] = [list(np.asarray(points).flatten())]
    annotation["segmentation"] = []
    annotation["iscrowd"] = 0
    annotation["area"] = area
    annotation["image_id"] = num

    annotation["bbox"] = list(map(float, self.getbbox(points)))

    annotation["category_id"] = label[0] # self.getcatid(label)
    annotation["id"] = self.annID
    return annotation

def getcatid(self, label):
    for category in self.categories:
        if label == category["name"]:
            return category["id"]
    print("label: {} not in categories: {}".format(label, self.categories))
    exit()
    return -1

def getbbox(self, points):
    polygons = points
    mask = self.polygons_to_mask([self.height, self.width], polygons)
    return self.mask2box(mask)

def mask2box(self, mask):
    index = np.argwhere(mask == 1)
    rows = index[:, 0]
    clos = index[:, 1]

    left_top_r = np.min(rows) # y
    left_top_c = np.min(clos) # x

    right_bottom_r = np.max(rows)

```

```
right_bottom_c = np.max(clos)

return [
    left_top_c,
    left_top_r,
    right_bottom_c - left_top_c,
    right_bottom_r - left_top_r,
]

def polygons_to_mask(self, img_shape, polygons):
    mask = np.zeros(img_shape, dtype=np.uint8)
    mask = PIL.Image.fromarray(mask)
    xy = list(map(tuple, polygons))
    PIL.ImageDraw.Draw(mask).polygon(xy=xy, outline=1, fill=1)
    mask = np.array(mask, dtype=bool)
    return mask

def data2coco(self):
    data_coco = {}
    data_coco["images"] = self.images
    data_coco["categories"] = self.categories
    data_coco["annotations"] = self.annotations
    return data_coco

def save_json(self):
    print("save coco json")
    self.data_transfer()
    self.data_coco = self.data2coco()

    print(self.save_json_path)
    os.makedirs(
        os.path.dirname(os.path.abspath(self.save_json_path)), exist_ok=True
    )
    json.dump(self.data_coco, open(self.save_json_path, "w"), indent=4)

if __name__ == "__main__":
    path = filedialog.askdirectory(title='Choose a Destination Path')
    labelme_json = glob.glob(os.path.join(path, "*.json"))
    output = path + "/coco.json"

    labelme2coco(labelme_json, output)
```