

# Montagegerechte Gestaltungsrichtlinien mittels Deep Learning

Johanna Gerlach, Alexander Riedel, Seyyid Uslu, Frank Engelmann, Nico Brehm

*Die Anwendung von Deep Learning in der manuellen Montage birgt großes Potenzial, Montagezeiten zu reduzieren und Montagefehler zu vermeiden. Indem der Montageablauf mithilfe einer Kamera erfasst und die aufgezeichneten Bilder durch einen Objekterkennungsalgorithmus analysiert werden, lassen sich Position, Lage und Art der montierten Bauteile bestimmen. Daraus lassen sich wiederum Informationen über Arbeitsschritte, Montagefehler oder den aktuellen Zustand des Produkts ableiten, sodass die Mitarbeiter bei der Montage durch entsprechende Anweisungen unterstützt werden können.*

*Es stellt sich jedoch die Frage, inwieweit gegenwärtige Produkte für den Einsatz von Deep Learning geeignet sind. Nur wenn die zu montierenden Bauteile sicher erkannt werden, ist der Einsatz in der manuellen Montage sinnvoll. Bestehende Gestaltungsrichtlinien adressieren diesen Aspekt bislang nicht. Im Forschungsprojekt wurde daher untersucht, welche Eigenschaften Produkte aufweisen sollten, um eine optimale Objekterkennung zu ermöglichen.*

*Dazu wurden Hypothesen zu positiven und negativen Bauteileigenschaften hinsichtlich der Erkennungsgenauigkeit formuliert und in praktischen Versuchen überprüft. Dabei konnte gezeigt werden, dass alle untersuchten Objekte durch den eingesetzten Objekterkennungsalgorithmus sehr gut detektiert werden. Aus den vorliegenden Forschungsergebnissen lassen sich daher keine Einschränkungen in der Produktgestaltung ableiten.*

*Keywords: Computer Vision, Deep Learning, Produktentwicklung, Montagegerechte Gestaltungsrichtlinien*

## Einleitung

---

Trotz des anhaltenden Trends zur Automatisierung stellt die manuelle Montage in den meisten Unternehmen noch immer einen wichtigen Fertigungsschritt dar. Aus dem Prozess der manuellen Montage resultiert jedoch ein hoher Zeit- und Kostenaufwand. Neben den hohen Arbeitskosten im europäischen Raum stellt die wachsende Variantenvielfalt Unternehmen vor große Herausforderungen. Mitarbeiter müssen immer flexibler und schneller für neue Produkte und Varianten angelernt werden. Durch den zunehmenden Fachkräftemangel in Deutschland wird die Situation weiter verschärft.

Um diesen Herausforderungen zu begegnen, werden in laufenden Forschungsprojekten potenzielle Optimierungsansätze für die manuelle Montage untersucht (vgl. Schreiber/Zürl/ Zimmermann: 2017; BMAS: 2018; KIT: 2020). 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.

Die Objekterkennung als Teilgebiet der Computer Vision befasst sich damit, den Inhalt und die Position verschiedener Objekte anhand von bestimmten Bildeigenschaften (Features) zu identifizieren. Deep Learning liegt vor, wenn diese Features in komplexen Strukturen von künstlichen neuronalen Netzen (KNN) erkannt werden (vgl. Khan/Rahmani/Shah/Bennamoun: 2018). Zu diesem Zweck können Convolutional Neural Networks (CNNs) eingesetzt werden. Ein CNN wird mit Trainingsbildern, auf denen die Position und Klasse der Objekte von Interesse markiert sind, für eine bestimmte Anzahl von Iterationen trainiert und über einen Validierungsdatensatz validiert, bis die Erkennungsqualität optimal ist. Auf diese Weise lassen sich Bilder aller Bauteile und typische Montagefehler in das Objekterkennungsmodell integrieren.

Die Leistungsfähigkeit von Deep Learning Modellen ist weitreichend für Datensätze wie COCO, CIFAR oder ImageNet untersucht (vgl. Bochkovskiy/Wang/Liao: 2020), deren Objektklassen Tiere, Gebäude, Blumen oder Einrichtungsgegenstände beinhalten. Im Gegensatz dazu weisen Objekte aus dem Bereich der industriellen Montage andere Eigenschaften auf. Um diese Objekte ebenso zuverlässig zu erkennen, kann es notwendig sein, die Architektur des neuronalen Netzes zu optimieren oder die Form der Objekte zu modifizieren. Im Folgenden soll der zweite Ansatz untersucht werden. Dieser Ansatz wurde in der Literatur bisher selten berücksichtigt, da Gestaltungsrichtlinien für industrielle Produkte leichter anpassbar sind als die Form von Tieren, Pflanzen oder Gebäuden.

Insgesamt birgt der Einsatz von Deep Learning in der manuellen Montage große Potenziale, Montagezeiten zu reduzieren sowie Montagefehler zu vermeiden. Es stellt sich jedoch die Frage, inwieweit gegenwärtige Produkte für den Einsatz von Deep Learning Modellen geeignet sind, ohne dass diese optimiert werden müssen. 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.

Um das Potential der Deep Learning Technologien in der manuellen Montage in vollem Umfang auszuschöpfen, sollte die Produktgestaltung auf diesen Anwendungsfall ausgerichtet werden. Aus diesem Grund wird im Rahmen des Forschungsvorhabens untersucht, welche konstruktiven Eigenschaften Produkte und ihre Bestandteile aufweisen müssen, um eine optimale Objekterkennung zu ermöglichen.

## **Methode**

---

### **Null- und Alternativhypothesen**

Im ersten Schritt wurden basierend auf der grundlegenden Funktionsweise von Objekterkennungsalgorithmen verschiedene Thesen zur optimalen Bauteilgestaltung formuliert. Die Basis der zum Untersuchungszeitpunkt verfügbaren Objekterkennungsalgorithmen bilden CNNs, welche in mehreren Schichten hintereinander angeordnet sind. Das rezeptive Feld der einzelnen Schichten nimmt zu, je tiefer sie im Netz liegen. Während des Trainingsprozesses lernen die ersten Schichten, einfache Bildmerkmale wie Ecken und Kanten zu erkennen. Tiefere Schichten reagieren auf komplexere Konturen und höherdimensionale Merkmale. Auf dieser Grundlage wurden Hypothesen entwickelt (siehe Tabelle 1), welche die Grenzen des Objekterkennungsalgorithmus in Bezug auf kleine Unterschiede zwischen zwei Objekten oder schwache Ausprägung von Merkmalen wie Konturen aufzeigen.

Da eine Hypothese statistisch nicht bewiesen, sondern lediglich widerlegt werden kann, wurden Null- und Alternativhypothesen formuliert. Lässt sich eine Nullhypothese durch die nachfolgend beschriebenen Versuche widerlegen, kann die jeweilige Alternativhypothese angenommen werden (vgl. Siebertz/Bebber/Hochkirchen: 2017, 101). Die Mean Average Precision (mAP) ist eine quantitative Angabe für die Erkennungsgenauigkeit eines Objekterkennungsmodells und soll in den Versuchen als Zielgröße dienen. Wenn sich die Werte der mAP zwischen zwei Modellen nicht signifikant unterscheiden, kann die Nullhypothese nicht widerlegt und die Alternativhypothese nicht angenommen werden.

### **Abhängige und unabhängige Variablen**

Die aufgestellten Hypothesen wurden in diesem Forschungsprojekt experimentell untersucht. Grundlage des Versuchsplans war die Bestimmung der unabhängigen und abhängigen Variablen. Wie in weiter unten beschrieben, wurde die mAP als abhängige Variable verwendet.

Tabelle 1: Null- und Alternativhypothesen

**Nr. Nullhypothese**

1 Bauteile, welche spiegelnde Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche matte Oberflächen aufweisen.

$$mAP_{\text{Spiegelnd}} = mAP_{\text{Matt}}$$

2 Bauteile, welche dunkle Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche helle Oberflächen aufweisen.

$$mAP_{\text{Dunkel}} = mAP_{\text{Hell}}$$

3 Bauteile, welche komplexe Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche simple Konturen aufweisen.

$$mAP_{\text{Komplex}} = mAP_{\text{Simpel}}$$

4 Bauteile, welche weiche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche harte Konturen aufweisen.

$$mAP_{\text{Weich}} = mAP_{\text{Hart}}$$

5 Bauteile, welche ähnliche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche sich hinsichtlich ihrer Konturen nicht ähneln.

$$mAP_{\text{Ähnlich}} = mAP_{\text{NichtÄhnlich}}$$

6 Bauteile, welche die gleiche Form, jedoch eine unterschiedliche Größe aufweisen, werden durch einen Objekterkennungsalgorithmus genauso gut unterschieden, wie Bauteile, welche nicht die gleiche Form aufweisen.

$$mAP_{\text{Gleich}} = mAP_{\text{NichtGleich}}$$

**Alternativhypothese**

Bauteile, welche spiegelnde Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche matte Oberflächen aufweisen.

$$mAP_{\text{Spiegelnd}} < mAP_{\text{Matt}}$$

Bauteile, welche dunkle Oberflächen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche helle Oberflächen aufweisen.

$$mAP_{\text{Dunkel}} < mAP_{\text{Hell}}$$

Bauteile, welche komplexe Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche simple Konturen aufweisen.

$$mAP_{\text{Komplex}} < mAP_{\text{Simpel}}$$

Bauteile, welche weiche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche harte Konturen aufweisen.

$$mAP_{\text{Weich}} < mAP_{\text{Hart}}$$

Bauteile, welche ähnliche Konturen aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche sich hinsichtlich ihrer Konturen nicht ähneln.

$$mAP_{\text{Ähnlich}} < mAP_{\text{NichtÄhnlich}}$$

Bauteile, welche die gleiche Form, jedoch eine unterschiedliche Größe aufweisen, werden durch einen Objekterkennungsalgorithmus schlechter unterschieden, als Bauteile, welche nicht die gleiche Form aufweisen.

$$mAP_{\text{Gleich}} < mAP_{\text{NichtGleich}}$$

Die in dieser Arbeit verwendete Definition der mAP ist analog zu der im COCO-Datensatz verwendeten Metrik  $mAP@[50\%;95\%]$ . Sie beschreibt den Durchschnitt der verschiedenen Average Precision Werte unter der Annahme, dass eine positive Erkennung bei 50%, 55%, ..., bis 95% Übereinstimmung zwischen Vorhersage und tatsächlicher Position des Objekts bewertet wird. Die Average Precision ergibt sich aus der Fläche unter der Funktion Precision (Recall), wobei der Recall als die richtige Positivquote (Sensitivität) und die Precision als der positive Vorhersagewert (Genauigkeit) definiert ist (vgl. Henderson/ Ferrari: 2017).

Diese Metrik bewertet sowohl die Qualität der Klassifikation als auch der Positionsbestimmung durch das Vorhersagemodell. Im Hinblick auf die zu untersuchenden Themen scheint sie somit geeignet, einen aussagekräftigen Wert zum Vergleich der Erkennungsgenauigkeit verschiedener Bauteile zu liefern. (vgl. Henderson/Ferrari: 2017)

Bei den unabhängigen Variablen ist zwischen Einflussgrößen, welche im Rahmen der Versuche bewusst variiert werden sollen (Faktoren) und Einflussgrößen, welche konstant gehalten werden (Steuergrößen), zu unterscheiden. Da im Rahmen der Untersuchung die optimale Bauteilbeschaffenheit für die Erkennbarkeit durch Objekterkennungsalgorithmen untersucht werden soll, wurde als zu variiender Faktor die Geometrie und das optische Erscheinungsbild der Untersuchungsgegenstände definiert.

Objekterkennungsmodelle basieren meist auf einer Datenbank, die aus Bildern der zu erkennenden Objekte besteht. Da diese ausschlaggebend für die Erkennungsgenauigkeit des Modells sind, müssen alle Größen, welche Einfluss auf die Eigenschaften der Bilder nehmen, als unabhängige Variablen definiert und im Rahmen der Versuche konstant gehalten werden. Dadurch wird sichergestellt, dass nur der Einfluss der Bauteileigenschaften auf die mAP gemessen wird. Neben der Beleuchtungsintensität müssen Bildhintergrund, Bildqualität, Bildperspektive und die Anzahl der verwendeten Trainings- und Validierungsbilder konstant gehalten werden. Schlussendlich gibt es eine Vielzahl unterschiedlicher Objekterkennungsalgorithmen, welche zur Objekterkennung eingesetzt werden können. Da die Art des eingesetzten Modells ebenfalls einen Einfluss auf die mAP hat, ist für die Versuche ein Algorithmus auszuwählen und anzuwenden.

## **Versuchsaufbau**

Nach der Definition der unabhängigen und abhängigen Variablen wurde ein Versuchsaufbau geplant, welcher es ermöglicht, die Steuergrößen konstant zu halten, während die Geometrie bzw. das optische Erscheinungsbild der Untersuchungsgegenstände variiert werden.

Der prinzipielle Ablauf der Untersuchung ist in Abbildung 1 dargestellt. Zunächst wurden Untersuchungsgegenstände konstruiert und beschafft, welche zur Überprüfung der aufgestellten Hypothesen dienen. Prinzipiell sind mehrere Modelle zu trainieren, welche jeweils andere Objekte enthalten und diese detektieren sollen. Über den Vergleich der mAPs kann dann eine Aussage getroffen werden, welche Objekte durch den Objekterkennungsalgorithmus besser unterschieden werden können.

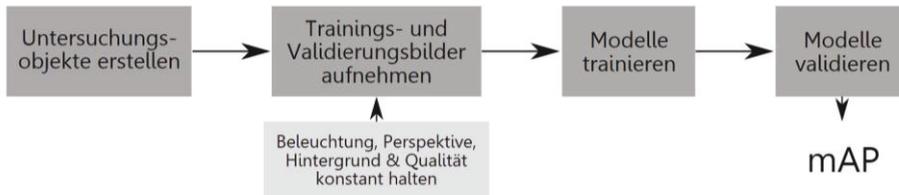


Abbildung 3: Versuchsablauf

Das Referenzmodell enthält die Formen 1-4. Diese weisen simple, nicht ähnliche und harte Konturen sowie matte und helle Oberflächen auf. Für die Testmodelle wurden Objekte konstruiert, welche sich entweder hinsichtlich ihrer Form oder ihrem optischen Erscheinungsbild von den Formen 1-4 unterscheiden. Beispielsweise weisen die Formen 1-1.3, 2-2.3, 3-3.3 und 4-4.3 nur sehr geringe Unterschiede auf und dienen damit zur Überprüfung der Annahme, dass Bauteile mit ähnlichen Konturen verwechselt werden. Tabelle 2 zeigt alle im Rahmen des Forschungsprojekts erstellten Testmodelle sowie die darin enthaltenen Untersuchungsgegenstände.

Anschließend wurden Bilder von den Untersuchungsobjekten erstellt. Um die Beleuchtungsstärke konstant zu halten, wurden diese in einem ausschließlich mit künstlichem Licht beleuchteten Raum aufgezeichnet und die Lichtstärke kontinuierlich mittels kalibriertem Luxmeter kontrolliert. Die Objekte wurden auf einem elektrischen Drehteller, welcher eine Umdrehungsgeschwindigkeit von 60 Sekunden je Umdrehung sowie eine weiße Oberfläche aufweist, an jeweils der gleichen Stelle positioniert. Parallel wurde eine Kamera (Sony Alpha 6300 mit SELP1650 Objektiv) auf den jeweiligen Untersuchungsgegenstand ausgerichtet. Um Bilder von den Objekten aufzunehmen, wurde der Drehteller in Rotation versetzt und beginnend bei einer definierten Startmarkierung jede Sekunde ein Bild aufgezeichnet. Die Einstellungen sowie die Position der Kamera zum Objekt wurden während der Aufnahme der einzelnen Untersuchungsgegenstände konstant gehalten.

Tabelle 2: Untersuchungsobjekte

Modell	Beinhaltete Objekte			
Referenzmodell	Form1 	Form2 	Form3 	Form4 
Modell zur Hypothese 1 (Spiegelnde Bauteile)	Form1Spiegelnd 	Form2Spiegelnd 	Form3Spiegelnd 	Form4Spiegelnd 
Modell zur Hypothese 2 (Dunkle Bauteile)	Form1Schwarz 	Form2Schwarz 	Form3Schwarz 	Form4Schwarz 
Modell zur Hypothese 3 (Komplexe Konturen)	Form5 	Form6 	Form7 	Form8 
Modell zur Hypothese 4 (Weiche Konturen)	Form1Weich 	Form2Weich 	Form3Weich 	Form4Weich 
Modell 1 zur Hypothese 5 (Ähnliche Konturen)	Form1 	Form1.1 	Form1.2 	Form1.3 
Modell 2 zur Hypothese 5 (Ähnliche Konturen)	Form2 	Form2.1 	Form2.2 	Form2.3 
Modell 3 zur Hypothese 5 (Ähnliche Konturen)	Form3 	Form3.1 	Form3.2 	Form3.3 
Modell 4 zur Hypothese 5 (Ähnliche Konturen)	Form4 	Form4.1 	Form4.2 	Form4.3 
Modell zur Hypothese 6 (Verschiedene Größe)	Form1 (siehe oben) & Form1Klein 	Form2 (siehe oben) & Form2Klein 	Form3 (siehe oben) & Form3Klein 	Form4 (siehe oben) & Form4Klein 

Nachdem alle Objekte aus einer Perspektive aufgenommen wurden, erfolgte ein Positionswechsel der Kamera, um alle Objekte aus einer weiteren Perspektive aufzuzeichnen und damit eine größere Bilderanzahl zu generieren. Durch das beschriebene Vorgehen wurde sichergestellt, dass sich die Objektbilder für die einzelnen Modelle lediglich hinsichtlich der dargestellten Objektgeometrie oder dem optischen Erscheinungsbild unterscheiden.

Insgesamt wurden jeweils 360 Bilder von den Objekten aufgezeichnet. Anschließend wurden die Bilder in Trainings- und Validierungsbilder unterteilt. Dabei wurde jedes fünfte Bild als Validierungsbild definiert, während die übrigen Bilder als Trainingsbilder genutzt wurden. Unter Verwendung dieser Bilder wurde das Modell trainiert.

Zu diesem Zweck wurde ein Faster-R-CNN-Modell mit ResNet-50-Backbone und einem Feature-Pyramid-Network-Head eingesetzt, das zum Zeitpunkt der Studie einen guten Kompromiss zwischen Erkennungsgenauigkeit, Trainingszeit, Inferenzzeit und Modellgröße darstellt. Die Modellarchitektur, sowie ein über ~37 Epochen auf dem COCO-Datensatz vortrainiertes Modell wurden dem von FAIR entwickelten PyTorch Framework detectron2 entnommen. (vgl. Wu/Kirillov/Massa/Lo/Girshick: 2021; Ren/He/Girshick/Sun: 2016)

Das detectron2 Framework ermöglicht einen geringen Entwicklungsaufwand für Objekterkennungsmethoden, bietet eine Vielzahl unterschiedlicher Modellarchitekturen und ist hinsichtlich der Trainings- und Inferenzzeit performanter als eine reine PyTorch Umsetzung. Die Verwendung eines vortrainierten Modells ist gängige Praxis, weil dadurch das Erlernen einfacher Konturen der ersten Netzwerkschichten entfallen kann. Dies wirkt sich positiv auf die Trainingsdauer und die Zahl der zum Training benötigten Bilder aus.

Auf dem vortrainierten Modell wurden je Hypothese 288 Trainingsbilder mit gleichmäßiger Verteilung der Objektklassen für 40.000 Iterationen lokal trainiert und anhand des Validierungsdatensatzes (72 Bilder) validiert. Die erzielten Ergebnisse werden im folgenden Abschnitt erläutert.

## Ergebnisse

---

Als Untersuchungsergebnis wurde für jedes Objekterkennungsmodell ein mAP erzeugt. Je höher dieser Wert, desto besser wurden die im Modell enthaltenen Objekte durch den eingesetzten Algorithmus unterschieden. Tabelle 3 zeigt die Ergebnisse aller trainierten Modelle. Während sich die  $mAP_{\text{Spiegelnd}}$ ,  $mAP_{\text{Komplex}}$ ,  $mAP_{\text{Weich}}$  und  $mAP_{\text{Gleich}}$  jeweils auf ein trainiertes Modell beziehen, stellt der  $mAP_{\text{Ähnlich}}$  einen Mittelwert der

mAPs der einzelnen Modelle für die These „Ähnliche Konturen“ dar. Hierfür wurden die mAPs der Modelle 1-1.3 (82.16), 2-2.3 (85.34), 3-3.3 (80.34) und 4-4.3 (88.01) zusammengefasst.

Aus den Ergebnissen wird deutlich, dass die Untersuchungsobjekte, welche spiegelnde Oberflächen, komplexe Konturen, weiche Konturen, ähnliche Konturen und gleiche Formen bei unterschiedlicher Größe aufweisen, genauso gut wie die Referenzobjekte detektiert werden. Die auftretenden Differenzen der mAP (z.B. 86,04 vs. 83,96) können nicht als signifikanter Unterschied eingestuft werden. Die hierzu aufgestellten Nullhypothesen lassen sich demnach nicht widerlegen, sodass die Alternativhypothesen nicht angenommen werden können.

*Tabelle 3: Versuchsergebnisse*

<b>mAP<sub>Spiegelnd</sub></b>	<b>mAP<sub>Dunkel</sub></b>	<b>mAP<sub>Komplex</sub></b>	<b>mAP<sub>Weich</sub></b>	<b>mAP<sub>Ähnlich</sub></b>	<b>mAP<sub>Gleich</sub></b>
84,02	99,80	86,52	86,85	83,96	86,67
<b>mAP<sub>Matt</sub> / mAP<sub>Hell</sub> / mAP<sub>Simplel</sub> / mAP<sub>Hart</sub> / mAP<sub>NichtÄhnlich</sub> / mAP<sub>NichtGleich</sub></b>					
86,04					

Weiterhin zeigt sich, dass die mAP für das Modell, welches schwarz eingefärbte Untersuchungsobjekte enthielt, einen deutlich höheren Wert als das Referenzmodell annimmt. Die Nullhypothese bezüglich der dunklen Oberflächen kann demnach zwar widerlegt, die Alternativhypothese ( $mAP_{Dunkel} < mAP_{Hell}$ ) jedoch nicht bestätigt werden. Es wurde ein genau entgegengesetzter Zusammenhang beobachtet. Entgegen der Hypothese wurden Bauteile mit dunklen Oberflächen besser erkannt als Bauteile mit hellen Oberflächen. Um diese Beobachtung zu erklären, wurden weitere Untersuchungen durchgeführt. Dabei wurde untersucht, inwieweit der Bauteilhintergrund einen Einfluss auf die Erkennungsgenauigkeit hat und ob sich dunkle Bauteile besser von einem weißen Hintergrund abheben, als helle Bauteile, sodass Ihre Außenkonturen besser detektiert werden. Dazu wurden vier weitere Modelle erzeugt, welche Bilder von den Formen 1-4 in schwarz bzw. weiß auf schwarzem bzw. weißen Hintergrund enthalten. Die Bilder und Modelle wurden nach dem gleichen Verfahren wie oben beschrieben aufgenommen und trainiert.

Abbildung 2 zeigt die erreichten Erkennungsgenauigkeiten (mAP) der einzelnen Modelle über 40.000 Iterationen. Es wird deutlich, dass Bauteile, welche sich vor einem andersfarbigen Hintergrund befinden, besser erkannt werden, als Bauteile, deren Hintergrund die gleiche Farbe wie sie selbst aufweist. Insbesondere weiße Bauteile vor weißem Hintergrund werden durch die Objekterkennung nur sehr eingeschränkt detektiert.

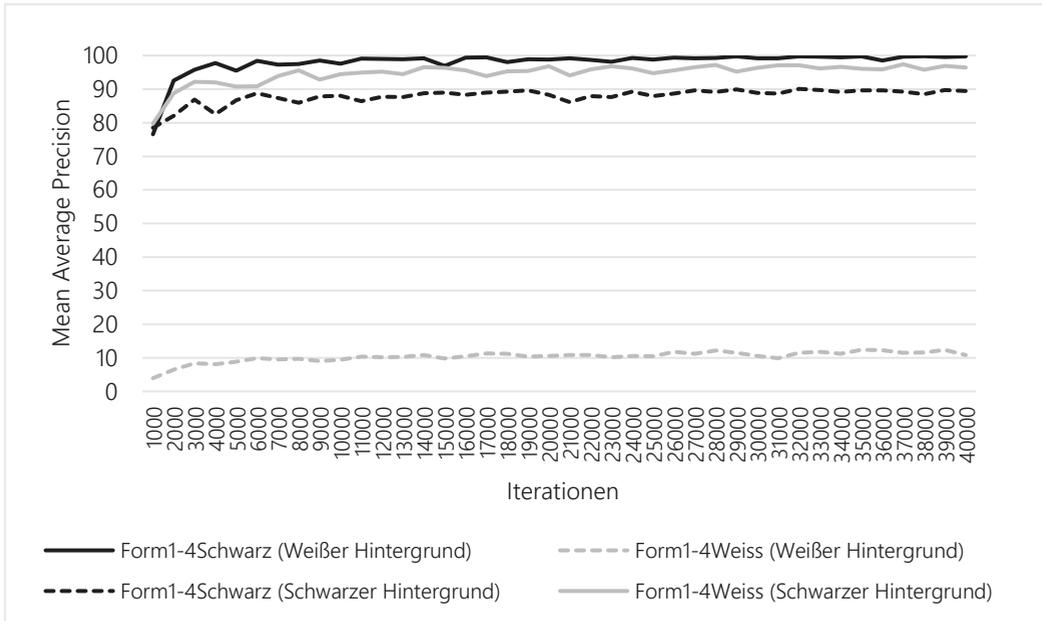


Abbildung 2: Mean Average Precision für Modelle zur Hypothese „Dunkle Oberflächen“

Insgesamt war das verwendete Objekterkennungsmodell in der Lage, alle getesteten Untersuchungsobjekte sehr gut zu erkennen, unabhängig von ihrer Geometrie und ihrem optischen Erscheinungsbild. Dementsprechend lassen sich aus der durchgeführten Untersuchung keine konstruktiven Einschränkungen in Bezug auf die Produktgestaltung ableiten. Es wird jedoch deutlich, dass der Bauteilhintergrund einen nicht zu vernachlässigenden Einfluss auf die Erkennungsqualität hat, sodass dieser Zusammenhang bei der Gestaltung von Montagearbeitsplätzen berücksichtigt werden sollte.

## Ausblick

Aus den Ergebnissen der Untersuchung lassen sich keine Einschränkungen für Konstrukteure in Bezug auf die Produktgestaltung ableiten. Alle in der Untersuchung betrachteten Objekte konnten durch den eingesetzten Algorithmus zuverlässig erkannt werden. Dies zeigt, dass ein konventionelles Faster-R-CNN-Modell keine Probleme hat, die Objekte in Bezug auf die untersuchten Gestaltungsmerkmale zu erkennen. Es ist also weder eine Optimierung der Architektur des Modells noch der Objektgestalt notwendig.

Neben den betrachteten Thesen gibt es jedoch eine Vielzahl weiterer Produktmerkmale, welche in der vorliegenden Untersuchung nicht behandelt werden. So ist bspw. zu prüfen, inwieweit sich Normteile (z. B. Schrauben, Sicherungsringe, Unterlegscheiben, etc.) verschiedener Größe oder Ausprägung durch Objekterkennungsalgorithmen unterscheiden lassen. Da die vorliegende Untersuchung bewusst anhand vereinfachter Geometrien durchgeführt wurde, stellt sich außerdem die Frage, ob die Ergebnisse auf reale Produkte übertragbar sind.

Insgesamt stellt das Forschungsvorhaben einen ersten Schritt zur Ableitung von Gestaltungsrichtlinien für den Einsatz von Objekterkennungsalgorithmen in der manuellen Montage dar, welchem weitere Untersuchungen folgen müssen. Erst nach einer umfassenden Betrachtung der Thematik kann eine Aussage darüber getroffen werden, ob sich durch den Einsatz von Objekterkennungsalgorithmen in der manuellen Montage Einschränkungen hinsichtlich der Produktgestaltung ergeben.

## Literaturverzeichnis

---

Schreiber, W.; Zürl, K.; Zimmermann, P. (Hrsg.) (2017): *Web-basierte Anwendungen Virtueller Techniken – Das ARVIDA-Projekt – Dienste-basierte Software-Architektur und Anwendungsszenarien für die Industrie*. Heidelberg; Berlin: Springer Vieweg

Bundesministerium für Arbeit und Soziales (BMAS) (2018): *Forschungsbericht 502 – Einsatz von digitalen Assistenzsystemen im Betrieb*. [https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/Forschungsberichte/fb502-einsatz-von-digitalen-assistenzsystemen-im-betrieb.pdf?\\_\\_blob=publicationFile&v=1](https://www.bmas.de/SharedDocs/Downloads/DE/PDF-Publikationen/Forschungsberichte/fb502-einsatz-von-digitalen-assistenzsystemen-im-betrieb.pdf?__blob=publicationFile&v=1), veröffentlicht 2018, abgerufen am 03.10.2020

Karlsruher Institut für Technologie (KIT) (2020): *KI unterstützt Fachkräfte bei der Montage*. [https://www.kit.edu/kit/pi\\_2020\\_047\\_ki-unterstutzt-fachkraefte-bei-der-montage.php](https://www.kit.edu/kit/pi_2020_047_ki-unterstutzt-fachkraefte-bei-der-montage.php), veröffentlicht 2020, abgerufen am 14.10.2020

Khan, S.; Rahmani, H.; Shah, S. A.; Bennamoun, M. (2018): *A Guide to Convolutional Neural Networks for Computer Vision*. Williston, VT: Morgan & Claypool

Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.: YOLOv4 (2020): *Optimal Speed and Accuracy of Object Detection*. <https://arxiv.org/pdf/2004.10934.pdf>, veröffentlicht 2020, abgerufen am 05.02.21

Siebertz, K.; Bebbler, D.; Hochkirchen, T. (2017): *Statistische Versuchsplanung, Design of Experiment (DoE)*. 2. Auflage. Heidelberg; Berlin: Springer Vieweg

Henderson, P.; Ferrari, V. (2017): *End-to-end training of object class detectors for mean average precision*. <https://arxiv.org/pdf/1607.03476.pdf>, veröffentlicht 2017, abgerufen am 05.02.21

Wu, Y.; Kirillov, A.; Massa, F.; Lo, W.-Y.; Girshick, R. (2020): *facebookresearch / detectron2*. <https://github.com/facebookresearch/detectron2>, abgerufen am 10.10.2020

Ren, S.; He, K.; Girshick, R.; Sun, J. (2016): *Faster R-CNN: Towards Real-Time Objekt Detection with Region Proposal Networks*. <https://arxiv.org/pdf/1506.01497.pdf>, veröffentlicht 2016, abgerufen am 05.02.21

## Kontakt

---

Johanna Gerlach, M. Sc.  
Alexander Riedel, M. Sc.  
Seyyid Uslu, M. Sc.  
Prof. Dr.-Ing. habil. Frank Engelmann  
Prof. Dr.-Ing. Nico Brehm  
Ernst-Abbe-Hochschule Jena  
Fachbereich Wirtschaftsingenieurwesen  
Carl-Zeiss-Promenade 2  
07745 Jena  
[www.eah-jena.de/wi](http://www.eah-jena.de/wi)