Bei diesem Beitrag handelt es sich um einen wissenschaftlich begutachteten und freigegebenen Fachaufsatz („reviewed paper“).

doi.org/10.37544/1436-4980-2024-09-34

Sensorrealistische Bildsimulation für die KI-basierte Sichtprüfung und Objektlageerkennung

Synthetische Daten für die Automatisierung mit KI

I. Effenberger, F. Seiler, V. Eichinger

ZUSAMMENFASSUNG Die KI-basierte Bildverarbeitung ist eine Schlüsseltechnologie für die Digitalisierung und Automatisierung in der Produktion. Da erforderliche Bilddatensätze zum Beispiel für die Qualitätssicherung unter hohem Zeitaufwand an realen Aufbauten aufgenommen und annotiert werden müssen, bleibt erhebliches Automatisierungspotenzial bislang ungenutzt. Einen großen Zeit- und Kostenvorteil bietet die sensorrealistische Bildsimulation, die repräsentative Bilddatensätze für das Training KI-basierter Inspektions- oder Handhabungssysteme synthetisch erzeugt.

STICHWÖRTER

Simulation, Künstliche Intelligenz, Qualitätssicherung

Synthetic data for AI-based automation

ABSTRACT AI-based image processing is considered as key technology for digitization and automation in production. Since the necessary image data sets, e.g. for quality assurance, have to be recorded and annotated on real structures in a time-consuming manner, considerable automation potential remains unused. Sensor-realistic image simulation, which synthetically generates representative image data sets for training AI-based inspection or handling systems, offers major time and cost advantages.

1 Einführung

In den letzten Jahren wurden durch den Einsatz neuartiger Verfahren des Maschinellen Lernens (ML) im Bereich der Bildverarbeitung enorme Durchbrüche erzielt. Vielversprechende Einsatzgebiete für entsprechende Verfahren mit Künstlicher Intelligenz (KI) liegen einerseits in der industriellen Qualitätsprüfung, andererseits in der Bauteilzuführung und Bauteilhandhabung in Produktionslinien. Aktuell größtes Hindernis für die Umsetzung von KI-basierten Automatisierungslösungen für diese Anwendungen stellt die Verfügbarkeit oder Bereitstellung geeigneter Daten dar. Denn für das Training von KI-Modellen werden repräsentative und große Datensätze benötigt, die in Form von Bilddaten der zu prüfenden oder zu handhabenden Bauteile vorhanden sein müssen. Die Aufnahme solcher Datensätze erfordert einen realen Aufbau mit passenden Kameramodellen, Objektiven und geeigneter Beleuchtung. Da zusätzlich das Vorbereiten des Bilddatensatzes für das ML-Training einschließlich des Annotierens der Bilder durchgeführt werden muss, ist dieses Vorgehen nicht nur zeitaufwendig, sondern auch sehr kostenintensiv.

Ein zusätzliches Problem für den Anwendungsfall der Defekterkennung ist, dass oft nur sehr wenige Schlechtteile vorliegen, die meist nicht die Bandbreite möglicher Defekte abdecken. Sind bestimmte Fehlertypen in den Trainingsdaten unterrepräsentiert oder ist die Anzahl der Fehlerbilder für die einzelnen Defektklassen sehr unterschiedlich, werden die Defekte vom KI-Modell nicht zuverlässig erkannt oder der falschen Defektklasse zugeordnet. Für die Bauteilzuführung und Bauteilhandhabung sind für die Erstellung von Trainingsdaten teilweise auch umgebende Elemente wie Förderbänder oder Robotergreifer relevant, die für die Bildaufnahme in den Aufbau integriert werden müssen.

Die Verwendung synthetischer Daten bietet eine gute Möglichkeit, um diesen Aufwand zu reduzieren und damit schneller und kostengünstiger KI-basierte Automatisierungslösungen in die industrielle Anwendung zu bringen. Sensorrealistische, synthetische Daten, wie sie am Fraunhofer IPA entwickelt werden, sind in vielerlei Hinsicht von großem Vorteil. Aufgrund des häufigen Mangels an fehlerhaften Bauteilen bei optischen Prüfaufgaben herrscht ein Mangel an Fehlerbildern, die für das Training der KI genutzt werden können. Dies erschwert die Erstellung robuster ML-Modelle für die zuverlässige Fehlerdetektion an den Bauteilen. Die Einbringung synthetischer Defekte in das 3D-Modell des Bauteils ermöglicht die Generierung von Datensätzen für die Qualitätskontrolle, auch in Abwesenheit defekter Teile. Zudem ist die manuelle Annotation großer Datensätze für maschinelle Lernverfahren eine zeitaufwendige und sich wiederholende Aufgabe, die fehleranfällig ist und oft zu Inkonsistenzen im annotierten Datensatz führt. Bei der Verwendung synthetisch erzeugter Fehlerbilder wird hingegen die Annotation automatisch mitgeliefert.

Rendering-Techniken haben in den letzten Jahren erhebliche Fortschritte erzielt und erschaffen immer realistischere Visualisierungen und immersive Erlebnisse. Davon profitiert nicht nur die Spiele- und Unterhaltungsindustrie, sondern auch die Bildsimulation, die mithilfe von 3D-Modellen und physikalisch korrekten Rendering-Verfahren synthetische Bilddaten generiert. Das Ziel der am Fraunhofer IPA entwickelten Software zur sensorrealistischen Bildsimulation ist es, synthetische Datensätze zu erstellen, die zum Trainieren eines maschinellen Lernmodells verwendet

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - CC BY

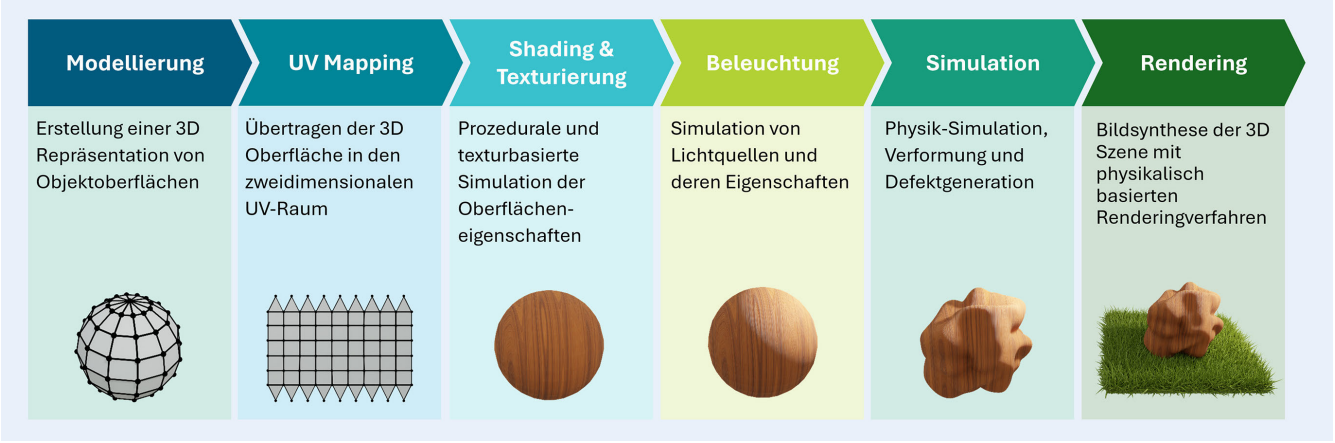

Bild 1. Prozesskette zur sensorrealistischen Bildsimulation auf Basis von 3D-Modellen und physikalisch korrektem Rendering. *Grafik: Fraunhofer IPA*

werden können, das anschließend auf reale Kamerabilder oder 3D-Sensordaten angewendet werden kann.

Durch Verwendung virtueller Szenen und die automatische Erstellung von Ground-Truth-Daten kann die IPA-Methode den manuellen Aufwand für die Annotation erheblich reduzieren und so eine effiziente und skalierbare Datengenerierung ermöglichen. Eine große Herausforderung bei dieser Technik ist die Domänenlücke („Domain Gap"), die zwischen realen Bildern und synthetischen Daten besteht. Modelle, die nur mit synthetischen Daten trainiert wurden, neigen dazu, einen erheblichen Leistungsabfall zu erfahren, wenn sie auf reale Kamerabilder angewendet werden. Um dieses Problem zu lösen, werden synthetische Daten sehr genau modelliert und ein hohes Maß an Vielfalt in die synthetischen Daten eingeführt. Anhand von zwei realen Anwendungen, der automatischen Sichtprüfung von Metallscheiben und der Bauteilerkennung für die Handhabung, wird die Leistungsfähigkeit der sensorrealistischen Bildsimulation demonstriert. Es wird gezeigt, dass ein KI-Modell, das ausschließlich mit synthetischen Daten trainiert wurde, bei realen Daten sehr gute Ergebnisse erzielen kann.

2 Stand der Technik

Um automatisch synthetische Datensätze von zufälligen Objektanordnungen zu generieren, wurden in den letzten Jahren mehrere Pipelines eingeführt. Diese Pipelines basieren meist auf 3D-Grafiksoftware wie „Blender" [1, 2] oder „Unity3D" [3]. Innerhalb dieser Frameworks wird eine Vielzahl von festen Objektinstanzen erzeugt. Anschließend werden diese Objekte einer Starrkörpersimulation unterzogen, die ihr Fallverhalten genau darstellt und die Physik von Bewegung und Kollision nachahmt. Diese Simulation stellt sicher, dass ein vielfältiges und repräsentatives Spektrum an Szenarien generiert wird. Die resultierenden Daten sind wertvoll für Bildverarbeitungsaufgaben, wie etwa die Objekterkennung für den Griff in die Kiste oder den Griff vom Band. Zusätzlich zu den simulierten Bildern werden Tiefenkarten, Objektposen und semantische Segmentierung automatisch generiert. Den Daten fehlen jedoch Informationen zur Greifbarkeit der Objekte in der Szene.

Anstelle einer Starrkörpersimulation präsentieren *Raistrick et al.* [4] das Infinigen-Framework zur Generierung prozeduraler Szenen der natürlichen Welt. Die generierten Bilder weisen eine hohe Vielfalt auf, sind aber nicht fotorealistisch. Auch wurden keine Experimente mit realen Datensätzen durchgeführt.

Eine weitere Anwendung konzentriert sich auf das Generieren synthetischer Bilder für die visuelle Inspektion und Objekterkennung. *Napier et al.* [5] zeigen, wie ein Netzwerk zur Segmentierung ausschließlich mit prozedural generierten synthetischen Daten trainiert werden kann. Aufgrund einer Verzerrung in den synthetischen Daten liefert nur ein Teil der realen Testdaten gute Ergebnisse. *Schmedemann et al.* [6] zeigen, dass die Generierung prozeduraler Defekte zur Erweiterung eines kleinen realen Datensatzes verwendet werden kann. Hier führt die Kombination von synthetischen und realen Daten zu besseren Ergebnissen als die Verwendung ausschließlich realer Daten. Ergebnisse, die ausschließlich auf synthetischen Daten basieren, werden in der Arbeit nicht berücksichtigt.

Frühere Arbeiten identifizierten die Domänenlücke als häufige Herausforderung bei der Verwendung synthetischer Trainingsdaten für Aufgaben des maschinellen Sehens [7]. Modelle gehen normalerweise davon aus, dass Trainings- und Testdaten sehr ähnlich sind. Synthetische Daten unterscheiden sich jedoch von realen Daten. Es wurden verschiedene Techniken ausgearbeitet, um den Transfer zwischen synthetischen Trainingsdaten und realen Testdaten zu erleichtern. Dabei gibt es im Wesentlichen drei grundlegende Strategien:

1. Domänenanpassung: Ziel der Domänenanpassung ist es, die statistische Abweichung zwischen Quell- und Zieldomäne zu verringern [8]. Die Anpassung kann durch die KI-Architektur erfolgen, beispielsweise durch Verwendung einer bestimmten Verlustfunktion, wie „Adversarial Loss". Hierbei gleichen sich die Merkmale, die aus synthetischen Daten gewonnen werden, den Merkmalen von realen Bildern an. Jedoch bedeutet ein überlappender Merkmalsraum nicht automatisch, dass die Merkmale der synthetischen Bilder für das zu lösende Problem geeignet sind.
2. Sensor-Realistic Rendering: Dieser Ansatz strebt Fotorealismus in den generierten Bildern an. Dies wird durch realistische Objektgeometrie, physikalisch korrekte Positionierung von Objekten in einer Szene, präzise Materialmodelle und Beleuchtung erreicht. Eine wichtige Technik für diesen Ansatz ist Physically Based Rendering (PBR) [9]. Der manuelle Aufwand der Szenenmodellierung steigt jedoch mit zunehmendem Anspruch an Realismus und Detailgrad.

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - CC BY

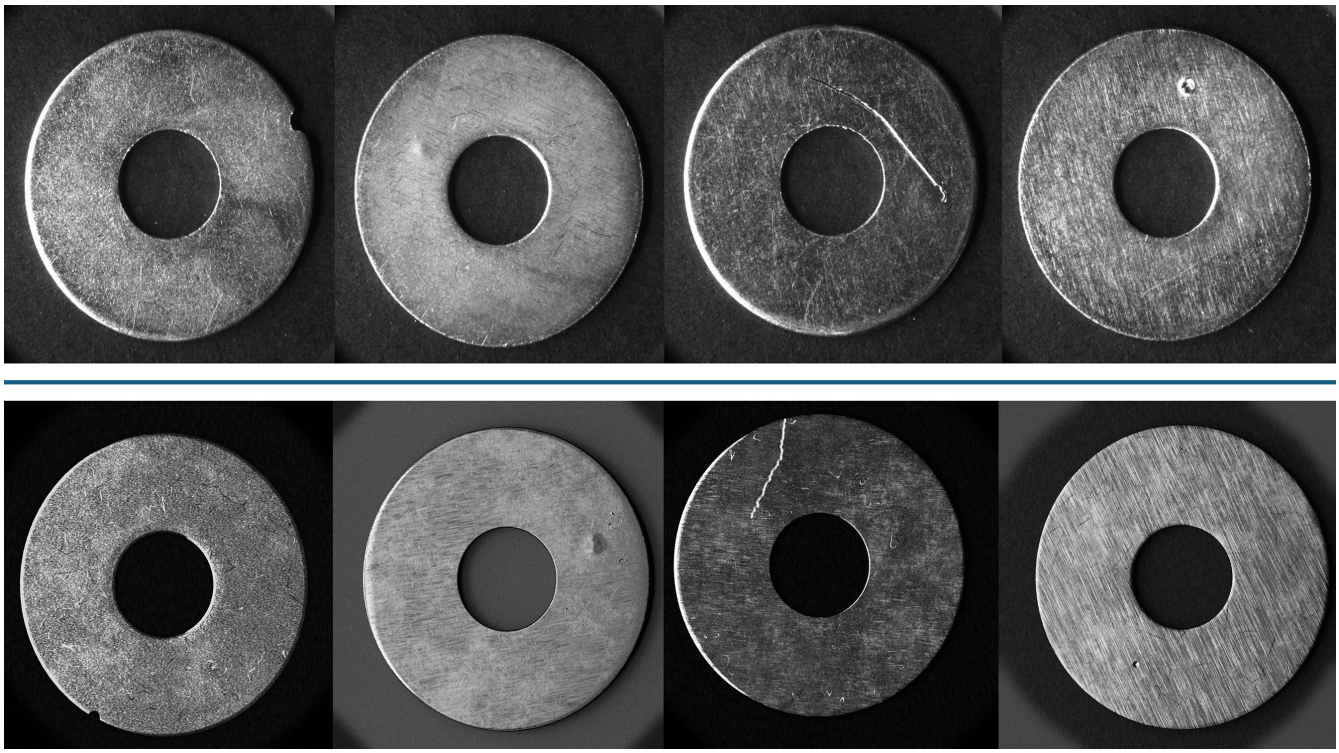

Bild 2. Vergleich realer Kamerabilder von Unterlegscheiben mit Defekten (obere Reihe) und synthetisch erzeugter Defektbilder der Unterlegscheiben (untere Reihe). *Foto: Fraunhofer IPA*

3. Domänenrandomisierung: Dieses von *Tobin et al.* [10] entwickelte Konzept umfasst die Randomisierung synthetischer Daten in einer Weise, dass die synthetische Domäne eine hohe Variation aufweist. Dies erlaubt es, ein Modell zu trainieren, das robust gegenüber verschiedenen Szenarien ist. Die Domänenrandomisierung erlaubt die Verwendung weniger realistischer Renderings, deren Erstellung weniger Aufwand erfordert. Bei dieser Technik werden Parameter für das Rendering von Daten zufällig variiert. Beispiele für solche Parameter sind Materialparameter, Texturen, Beleuchtungseinstellungen, Größe und Position von Objekten, Objektformen und Hintergründe.

3 Verfahren

Das am Fraunhofer IPA entwickelte Verfahren zur sensorrealistischen Bildsimulation folgt der traditionellen Verarbeitungspipeline für Renderings aus 3D-Daten (**Bild 1**) und nutzt als Basis die Open-Source-Software Blender [11].

Zunächst wird ein Modell für das gewünschte Objekt erstellt. Alternativ können vorhandene CAD-Daten verwendet werden. Zusätzlich wird eine UV-Karte für das Objekt generiert, die im Shading-Prozess verwendet werden kann. Beim Shading werden alle Parameter und Texturen für das Material des Objekts definiert. Anschließend wird eine 3D-Szene basierend auf dem realen Messaufbau mit der Simulation von Licht und Sensoren erstellt. In dieser Szene werden zusätzliche Verarbeitungsschritte wie die Simulation von Verformungen, Defekten oder zufälligen Anordnungen von Objekten durchgeführt. Schließlich wird der Renderer für das physikalisch basierte Rendering der Szene angewendet. Der Shading-Prozess und die Verarbeitungsschritte werden in den folgenden Abschnitten erläutert.

Wenn die grundlegende 3D-Szene vorbereitet ist, ermöglicht das Verfahren zur Bildsimulation dem Benutzer, Parameter zu definieren und Schwellwerte für alle weiteren Verarbeitungsschritte und die Bildgenerierung festzulegen. Der Benutzer kann Parameter für alle Teile der 3D-Szene, wie Szenenanordnung, Beleuchtung, Schattierung, Verformung, Defekte und Rendering, angeben. Anhand der Benutzereinstellungen werden Szenenvariation, Rendering und Verarbeitung vollständig automatisiert, um mit geringem Aufwand große Datensätze zu erzeugen. Außerdem werden ausgewählte Annotationen während des Rendering-Prozesses automatisch erzeugt.

3.1 Shading

Objekte in der realen Welt weisen häufig eine große Vielfalt in ihrem Erscheinungsbild auf. Dies erschwert die Oberflächenmodellierung, da ein Shader nicht alle möglichen Variationen abdecken kann. Dies ist auch bei den zur Untersuchung der Defektentstehung verwendeten Unterlegscheiben zu beobachten. Daher wurden für die Unterlegscheiben vier Basis-Shader definiert, um extreme Fälle der Vielfalt abzubilden. **Bild 2** zeigt ausgewählte Kamera-Referenzbilder in der oberen Reihe und Beispiele von simulierten Bildern für die vier Shader in der unteren Reihe. Es ist offensichtlich, dass die Unterschiede zwischen den realen und synthetischen Daten minimal sind. Darüber hinaus wird die in den realen Daten vorhandene Varianz auch in den synthetischen Daten abgebildet.

Der Großteil des Erscheinungsbildes wird durch Oberflächenstrukturen bestimmt, die durch Hinzufügen einer Normalenkarte (Normal Map) zum Shader reproduziert werden können. Für die Unterlegscheiben wird der Shader mit verschiedenen Strukturen und Kratzmustern erweitert. Die Grundstruktur wird durch eine

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

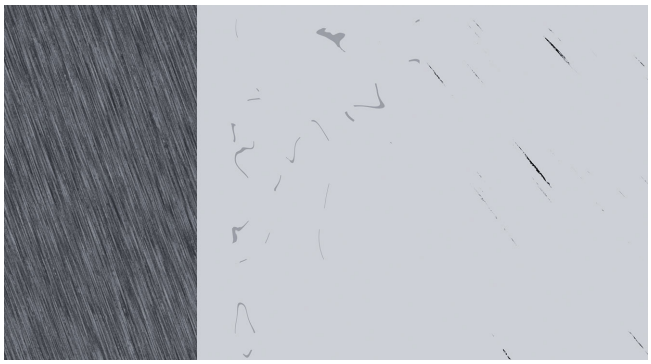

Bild 3. Grundlegende Struktur für die Modellierung der Unterlegscheiben (links) und verschiedene Kratzmuster (Mitte und rechts). *Grafik: Fraunhofer IPA*

Rauschtextur vorgegeben. Verschiedene Kratzmuster werden durch die Kombination von Rausch-, Wellen- und Voronoi-Texturen erzeugt, wie in **Bild 3** dargestellt.

Die Kombination dieser Strukturen in Farb- und Normalenkarten ermöglicht die Modellierung komplexer Oberflächenerscheinungsformen. Da alle Texturen prozedural generiert werden, haben alle generierten Farb- und Normalenkarten ein zufälliges und einzigartiges Erscheinungsbild, was stark variierende synthetische Daten ermöglicht.

3.2 Erzeugung von Defekten

Für die automatische Sichtprüfung wurde am Fraunhofer IPA ein Defektgenerator implementiert, der durch Änderung der Objektgeometrie automatisch Defekte auf 3D-Modellen erzeugen kann. Dieses Verfahren wurde für die Defektklassen Kratzer, Dellen, Beulen und Kerben entwickelt und wird in Abschnitt 4.1 am Anwendungsfall der Defekterkennung und -klassifikation für Unterlegscheiben demonstriert. Jeder dieser Defekte wird dabei an einer zufälligen Stelle und mit einem zufälligen aber repräsentativen Erscheinungsbild erzeugt. In den folgenden Abschnitten werden die unterschiedlichen Fehlerarten und deren synthetische Erzeugung beschrieben.

3.2.1 Kratzer und Beulen

Displacement Maps ermöglichen die Simulation von Beulen und Kratzern auf einer Oberfläche, indem sie eine Verschiebung von Punkten des Oberflächennetzes entlang des Normalenvektors ermöglichen. Ein Grauwertbild codiert dabei den Verschiebungsgrad für jeden Punkt durch seinen Intensitätswert. Um diese Karten auf die Oberfläche eines 3D-Objekts anzuwenden, ist es notwendig, ein UV-Layout zu erstellen, das die dreidimensionale Oberfläche des Objekts in eine zweidimensionale Darstellung übersetzt (siehe Bild 1).

Um Beulen zu simulieren, werden Displacement Maps mit einem glockenförmigen Intensitätsprofil genutzt, um das Erscheinungsbild realer Beulen nachzubilden. Displacement Maps für Kratzer werden mit prozeduralen Techniken erstellt, welche eine erhebliche Varianz in Form, Größe und Tiefe ermöglichen. Für das Einbringen der feinen Details von Kratzern in das virtuelle 3D-Modell ist ein hochauflösendes Oberflächennetz des Bauteils nötig, das sicherstellt, dass die Nuancen jedes Kratzers realistisch dargestellt werden können.

3.2.2 Dellen und Kerben

Für die synthetische Erzeugung der Defekte Dellen und Kerben wird ein Hilfsobjekt erstellt, das als Negativform für die Verformung von Dellen und Kerben dient. Dieses Hilfsobjekt wird zufälligen Verformungs-, Translations- und Skalierungsprozessen unterzogen, um die Vielfalt der resultierenden Bilder zu erhöhen. Anschließend wird ein Boolescher Operator auf das Original- und das Hilfsobjekt angewendet, der eine subtraktive Schnittmenge erzeugt. Diese Schnittmenge entfernt die Form des Hilfsobjekts aus dem Originalobjekt, wodurch das verformte Originalobjekt entsteht. **Bild 4** zeigt eine vereinfachte Darstellung dieser Methode zur Erzeugung von Kerben.

3.3 Zufällige Objektanordnungen für die Objektlageerkennung

Für die Erzeugung zufälliger Objektanordnungen wird eine Physics Engine, zum Beispiel aus Blender [11], verwendet. Bei diesem Verfahren werden mehrere Instanzen eines unverformbaren Objekts erstellt und dann eine Starrkörpersimulation angewendet, um ihr Fallverhalten zu modellieren.

Durch die Simulation können zufällige Anordnungen und Schüttungen von Objekten automatisiert erzeugt werden. Vor allem für das Objekthandling ist es wertvoll zu wissen, welche Objekte gegriffen werden können. Daher wird in der virtuellen 3D-Szene überprüft, welche Objekte nicht von anderen Objekten verdeckt werden. Diese Objekte können dann automatisch annotiert werden. Im Anschluss können die annotierten Datensätze für das Training eines KI-Modells zur Erkennung greifbarer Objekte verwendet werden.

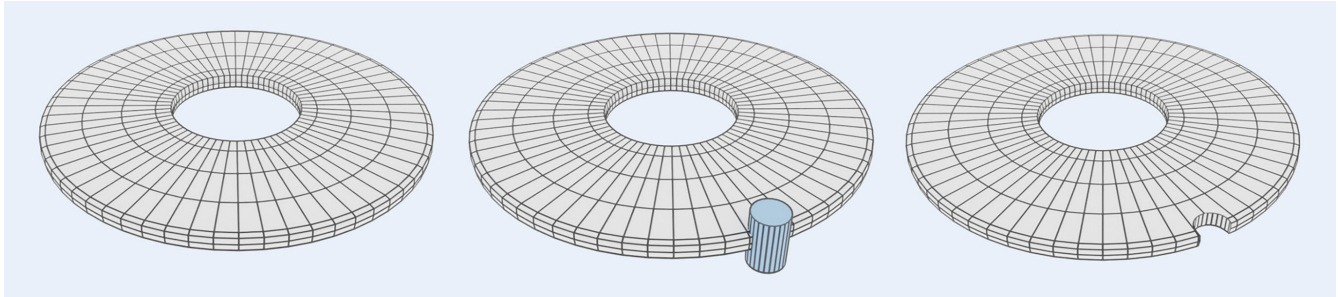

Bild 4. Vereinfachte Darstellung der Erzeugung von Kerben mithilfe eines Booleschen Operators. *Grafik: Fraunhofer IPA*

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

4 Anwendungstests mit erzielten Ergebnissen

Der Nutzen synthetischer Daten in realen Bildverarbeitungsanwendungen wird anhand zwei verschiedener Anwendungsfälle unter Anwendung der vorgestellten Methoden und Prozessschritte aufgezeigt. Die beiden Anwendungsfälle stammen aus dem Bereich der optischen Inspektion und Objekterkennung.

4.1 KI-basierte Sichtprüfung an Metallscheiben

Um das vorgestellte Verfahren zur sensorrealistischen Bildsimulation im Bereich der automatischen KI-basierten Sichtprüfung zu demonstrieren, wurden defekte Metallscheiben mit Beulen, Dellen, Kratzern und Kerben betrachtet. Für den Testdatensatz wurden 40 Bilder für jede Defektklasse aufgenommen. Für den Trainingsdatensatz wurden 500 synthetische Bilder für jede Defektklasse mit dem Bildsimulationsverfahren erzeugt. Diese Bilder wurden verwendet, um das neuronale Netz „EfficientDet“ [12] als modernen Objektdetektor zu trainieren. 80 % der Bilder wurden für das Training genutzt und 20 % für die Validierung des Modells. Das trainierte KI-Modell wurde anschließend mit den realen Kamerabildern der Metallscheiben getestet. Neben dem rein auf synthetischen Daten trainierten KI-Modell wurden zum Leistungsvergleich zwei zusätzliche KI-Modelle mit einem realen Datensatz (35 Bilder pro Defektklasse) beziehungsweise mit einem gemischten Datensatz aus realen und synthetischen Bildern trainiert (500 synthetische Bilder und 35 reale Bilder je Defektklasse).

Für den Vergleich wurde jeweils die Precision, also wie viel Prozent der von der KI erkannten Defekte sind tatsächlich Defekte, und der Recall, das heißt wie viel Prozent der vorhandenen Defekte wurden von der KI korrekt erkannt, ausgewertet. Mit einer erreichten Precision von 99 % und einem Recall von 100 % schnitt das Modell auf Basis gemischter Daten am besten ab, da die Kombination aus synthetischen und realen Daten potenziell zu robusteren Merkmalen der Klassen führt. Die Performance des Modells auf Basis rein synthetischer Daten (Precision: 96 %, Recall: 89 %) unterscheidet sich kaum von dem Ergebnis mit realen Daten (Precision: 99 %, Recall: 86 %). Die Modelle können als gleichwertig angesehen werden, wobei die Kratzererkennung auf Basis der synthetischen Daten deutlich zuverlässiger funktionierte (Recall von knapp 89 %). Bei der visuellen Inspektion können somit kleinere reale Datensätze mit synthetischen Daten für eine robuste Fehlererkennung angereichert werden. Die **Tabelle** bietet eine Zusammenfassung der Ergebnisse unter Verwendung realer, synthetischer und gemischter Trainingsdaten.

4.2 KI-basierte Objekterkennung für Kunststoffspritzen

Die Anwendung der Objekterkennung wird anhand eines Griff-in-die-Kiste-Szenarios demonstriert. Ziel war es Plastikspritzen zu erkennen, die für einen Roboter greifbar sind, also nicht von anderen Spritzen verdeckt werden. Mithilfe der in Abschnitt 3.3 vorgestellten Methode zur Generierung zufälliger Anordnungen wurden automatisch 2000 synthetische Bilder von Spritzen in einer Kiste erstellt. Der Fallprozess der leicht transparenten Spritzenmodelle wurde mithilfe der Physics Engine simuliert. Ein Beispiel für eine synthetische Szene zeigt **Bild 5**.

Tabelle. Ergebnisse der KI-basierten Sichtprüfung im Überblick.

		Precision	Recall
Reale Daten	*Bump*	1.0	0.971
	Dent	1.0	0.971
	Notch	1.0	0.971
	Scratch	0.950	0.543
	Total	0.992	0.864
Synthetische Daten	*Bump*	1.0	0.914
	Dent	0.909	0.857
	Notch	1.0	0.914
	Scratch	0.939	0.886
	Total	0.962	0.893
Gemischte Daten	*Bump*	0.972	1.0
	Dent	1.0	1.0
	Notch	1.0	1.0
	Scratch	0.972	1.0
	Total	0.986	1.0

Als Objektdetektor kommt das neuronale Netz „EfficientDet“ zum Einsatz, das problemspezifisch angepasst wurde. Das KI-Modell wurde rein mit synthetischen Daten trainiert, um die greifbaren Spritzen zu identifizieren. Anschließend wurden Tests an realen Kamerabildern von den ungeordneten Spritzen in der Kiste durchgeführt. Die Tests erreichten eine Precision von 96 % und einen Recall von 93 %. Die erfolgreichen Erkennungsraten demonstrieren die Leistungsfähigkeit des ML-Modells bei der Erkennung greifbarer Spritzen. Die Visualisierung der Annotationen und eines Testbilds mit den Erkennungen des KI-Modells ist in **Bild 6** dargestellt.

5 Zusammenfassung und Ausblick

In diesem Beitrag wurde eine Methode zur Generierung realistischer synthetischer Datensätze auf Basis von sensorrealistischer Bildsimulation beschrieben. Der Schwerpunkt lag auf dem Realismus der Daten durch physikalisch korrektes Rendering, einerseits zur Simulation von Defekten für die Sichtprüfung und andererseits zur Simulation von Szenen mit ungeordneten Objekten für die Bauteilzuführung oder -handhabung. Die Varianz der Daten wurde durch flexible, an die Zielanwendung angepasste Parameterdefinitionen erreicht. Die Datengenerierung wurde automatisiert durchgeführt anhand der vorgestellten Prozesskette zur Bildsimulation mit Komponenten, vor allem für die Defektgenerierung, die zufällige Objektanordnung in einer Szene und das Rendering unter Berücksichtigung von Kamera, Beleuchtung und Materialeigenschaften der zu untersuchenden Bauteile.

In zwei praktischen Anwendungsfällen wurde die Leistungsfähigkeit und Anwendbarkeit der Bildsimulationsmethode des Fraunhofer IPA demonstriert. Synthetische Daten, die sowohl für die Objekterkennung als auch für die visuelle Inspektion erzeugt wurden, konnten erfolgreich zum Training neuronaler Netze ver-

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - CC BY

Bild 5. Simulation der Szene: synthetische Spritzen fallen in Kiste (links), simulierte Szene mit ungeordneten Spritzen (rechts). *Foto: Fraunhofer IPA*

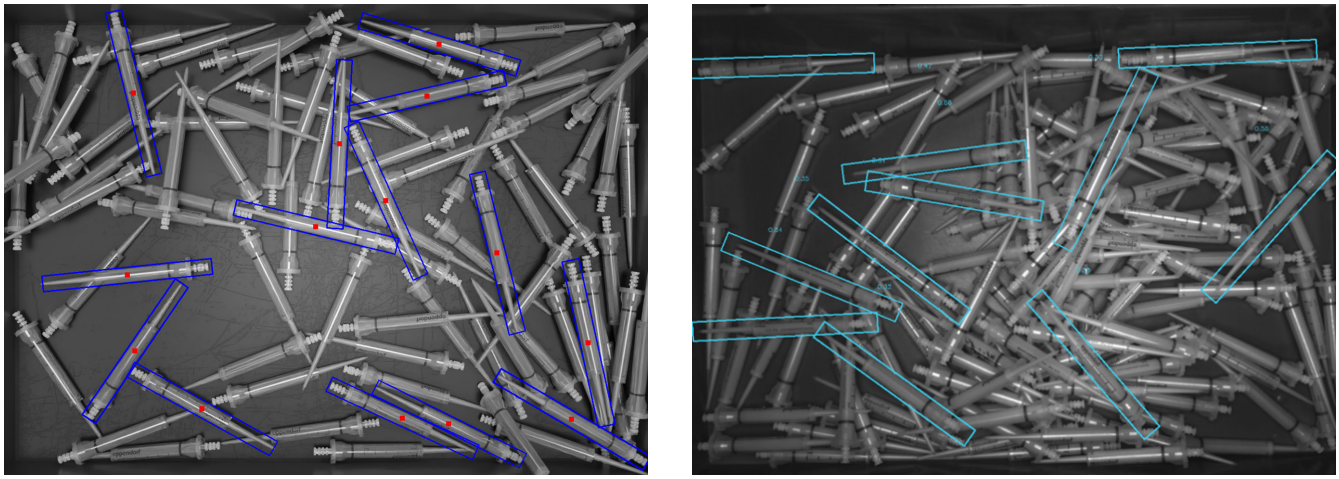

Bild 6. Annotationen greifbarer Spritzen, das heißt ohne Verdeckung, in der synthetischen Szene (links); von der KI erkannte greifbare Spritzen im realen Kamerabild der Szene (rechts). *Foto: Fraunhofer IPA*

wendet werden. Die trainierten ML-Modelle waren in der Lage, bei der Anwendung auf realen Kamerabildern genaue und zuverlässige Ergebnisse zu erzielen. Somit können synthetische Daten als Alternative zu realen Daten verwendet werden, wenn ein Mangel an Daten besteht. Zudem sind sie von großem Vorteil, wenn die Erfassung realer Daten zu zeitaufwendig ist oder reale Datensätze angereichert werden müssen, um robustere KI-Modelle zu erhalten.

Zukünftige Forschungsarbeiten zielen darauf ab, KI-basierte Bildverarbeitungslösungen vollständig virtuell auslegen und testen zu können, ohne dass ein Hardwareaufbau des optischen Prüf- oder Erkennungssystems oder die zu untersuchenden oder zu handhabenden Bauteile verfügbar sein müssen.

Literatur

[1] Denninger, M.; Winkelbauer, D.; Sundermeyer, M. et al.: BlenderProc2: A Procedural Pipeline for Photorealistic Rendering. Journal of Open Source Software 82 (2023) 8, #4901, doi.org/10.21105/joss.04901

[2] Greff, K.; Belletti, F.; Beyer, L. et al.: Kubric: A scalable dataset generator. ArXiv 2022, doi.org/10.48550/arXiv.2203.03570

[3] Unity Technologies: Unity Perception Package. Stand: 2020. Internet: github.com/Unity-Technologies/com.unity.perception. Zugriff am 14.08.2024

[4] A. Raistrick, Lipson, L.; Ma, Z. et al.: Infinite Photorealistic Worlds using Procedural Generation. ArXiv 2023, doi.org/10.48550/arXiv.2306.09310

[5] Napier, C. C., Cook, D.; Armstrong, L. et al.: A Synthetic Wheat L-System to Accurately Detect and Visualise Wheat Head Anomalies. Proceedings of the 3rd International Conference on Smart and Innova-

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

tive Agriculture, ICoSIA 2022, pp. 379–391. DOI: doi.org/10.2991/978–94–6463–122–7_36
[6] Schmedemann, O. et al.: Procedural synthetic training data generation for AI-based defect detection in industrial surface inspection. Procedia CIRP 107 (2022), pp. 1101–1106
[7] Nikolenko, S. I.: Synthetic Data for Deep Learning. arXiv 2019. DOI: doi.org/10.48550/arXiv.1909.11512
[8] Gopalan, R.; Li, R.; Chellappa, R.: Domain adaptation for object recognition: An unsupervised approach. International Conference on Computer Vision, Barcelona, Spain, 2011, pp. 999–1006, doi. org/10.1109/ICCV.2011.6126344
[9] Pharr, M.; Jakob, W.; Humphreys, G.: Physically Based Rendering: From Theory To Implementation. Burlington: Morgan Kaufmann 2016
[10] Tobin, J.: Fong, R.; Ray, A. et al.: Domain Randomization for Transferring Deep Neural Networks from Simulation to the Real World. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 2017, pp. 23–30, doi: 10.1109/IROS.2017.8202133
[11] Blender Foundation: Homepage. Internet: www.blender.org. Zugriff am 14.08.2024
[12] Tan, M.; Pang, R.; Le, Q. V.: EfficientDet: Scalable and Efficient Object Detection. ArXive 2019, arxiv.org/abs/1911.09070

Dr.-Ing. Ira Effenberger
ira.effenberger@ipa.fraunhofer.de
Tel. +49 711 / 970-1853
Foto: Fraunhofer IPA

Frederik Seiler, M.Sc.

Verena Eichinger, M.Sc.

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA
Nobelstr. 12, 70569 Stuttgart
www.ipa.fraunhofer.de

LIZENZ

Dieser Fachaufsatz steht unter der Lizenz Creative Commons Namensnennung 4.0 International (CC BY 4.0)

https://doi.org/10.37544/1436-4980-2024-09-34 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY