

Bis zu 90% weniger Aufwand bei der Lagerinventur

Automatisierte Produkterfassung mit Hilfe von Drohnen und KI



”

In ihrem Whitepaper berichten René Kessler und Ralph Göhrs über die Forschungskooperation mit der Universität Oldenburg in welcher, durch freundliche Unterstützung von „Getränke Essmann“, eine Methode entwickelt wurde, um den Inventurprozess zu optimieren und dabei komplett auf manuelle Zählungen zu verzichten.



AUTOR
RALPH GÖHRS
ralph.goehrs@abat.de

AUTOR
RENÉ KESSLER
rene.kessler@uni-oldenburg.de

Lagerinventur

Die Inventur des Lagers ist in Unternehmen ein manuell geprägter Vorgang, bei dem sehr viele personelle Ressourcen für eine Zählung der physischen Güter gebunden sind und die regulären Abläufe des Tagesgeschäfts zudem gestört werden. Das Problem jedoch: Eine Inventur ist nach §240 HGB gesetzlich verpflichtend und mindestens einmal pro Jahr durchzuführen. Andererseits kann die Inventur aber auch eine Methode sein, um eigene Abläufe zu kontrollieren und für Transparenz im Unternehmen sorgen. So können Fehlbestände festgestellt werden und man kann ihnen auf den Grund gehen, um beispielsweise Diebstahl oder anderweitiges Fehlverhalten auszuschließen.

Gerade bei extremen Warenbewegungen, in kurzer Zeit, kann eine Inventur deshalb von Vorteil sein. Unternehmen befinden sich also in einem Konflikt zwischen Aufwand und Nutzen bei der Durchführung einer Inventur. Das ist ein Grund dafür, dass in der Praxis häufig mit Stichproben oder Teilinventuren gearbeitet wird, bei denen nur ein (bestimmter) Bestand ausgezählt wird. Ein anderer, einfacherer Weg und Fokus des hier vorgestellten Projektes ist die Automatisierung der Inventur, d.h. die Erfassung des Bestandes ohne manuelle Zählung durch einen Menschen.

Bei einem Blick in die Literatur und Projekte der Praxis lassen sich im Wesentlichen zwei unterschiedliche Ansätze erkennen:

1. Lokalisierung und Identifizierung von Produkten mithilfe von Sensoren:

Werden Produkte mit Sensoren ausgestattet, können diese aus dem Umfeld (z. B. Lager) ausgelesen werden. In der Praxis werden hierbei häufig RFID-Tags eingesetzt, um Produkte oder Ladungsträger zu tracken. Doch gerade wenn diese Produkte eher schwierig zu erreichen sind, bspw. wenn sie in Hochregalen lagern, werden in einigen Ansätzen Drohnen eingesetzt, die mit einem RFID-Reader ausgestattet sind, und RFID-Tags von Produkten und Ladungsträgern im Nahfeld auslesen können. Dieser Ansatz ist grundsätzlich funktional und bringt einige Vorteile mit sich, lässt sich aber nicht auf Produkte jeder Art übertragen. In der Regel verfügbare Produkte und Ladungsträger nicht über RFID-Tags und in vielen Fällen würde der Aufwand von Anschaffung und Anbringung der Tags den Nutzen übersteigen.

2. Auslesen von optischen Produktmerkmalen oder -charakteristika:

Ein weiterer Ansatz zur Identifizierung von Produkten und Ladungsträgern ist das Auslesen von Annotationen, wie z.B. Barcodes oder QR-Codes. Grundvoraussetzung ist also stets, dass Annotation vorliegen, die dann über eine Drohne mit Kamera ausgelesen werden können. Dies ist eine Lösung mit viel Potenzial und schnellen Erfolgen. Leider sind aber häufig auch keine Annotationen an Produkten oder Ladungsträgern angebracht.

Auffällig ist insgesamt, dass der KI-basierten Bildverarbeitung zwar gerade in der Kombination mit Drohnen ein riesiges Potenzial in der Logistik nachgesagt wird, aber bisher nur sehr wenige Forschungsansätze zu finden sind. Mit unserem Forschungsprojekt, das wir gemeinsam mit unserem Kunden Getränke Essmann durchgeführt haben, betreten wir somit Neuland und konnten ein verbreitetes Problem prototypisch lösen.



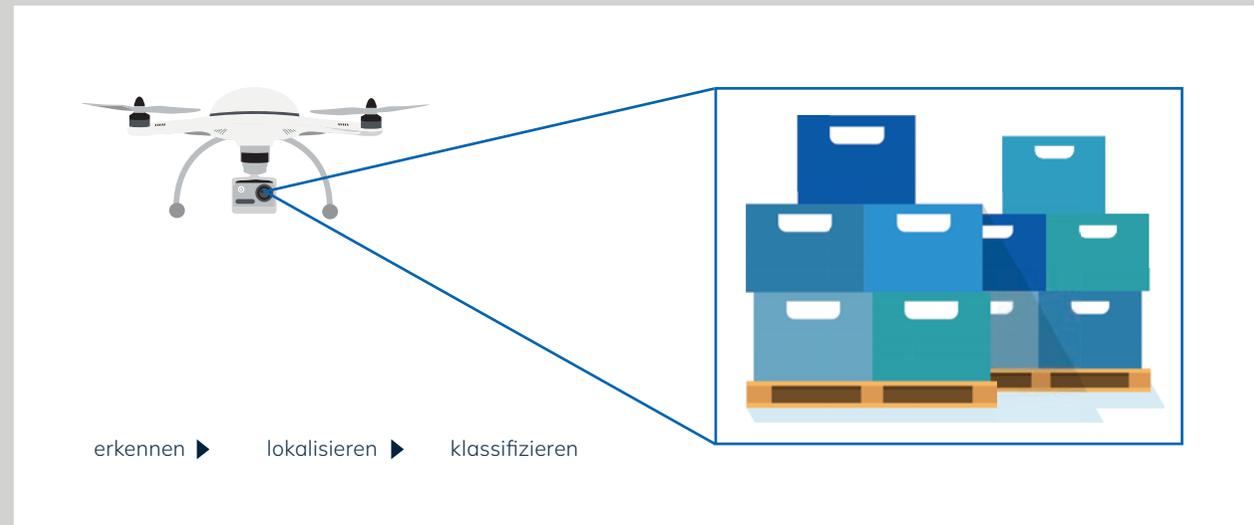
Unser Projekt mit Getränke Essmann

Im Rahmen der Forschungskooperation mit der Universität Oldenburg und gemeinsam mit Getränke Essmann haben wir es uns als Aufgabe gesetzt, den Inventurprozess zu optimieren und komplett auf manuelle Zählungen zu verzichten.

Fokussiert haben wir uns im ersten Schritt auf die Leergutinventur auf dem Außengelände von Getränke Essmann. Bisher wird diese Inventur manuell von mehreren Teams ausgeführt, die jeweils aus zwei bis drei Personen bestehen und mehrere Tage abgestellt werden, um den aktuellen Leergutbestand zu erfassen. Gezählt wird dabei die Anzahl der Paletten. Beachtet werden muss allerdings, dass Paletten, je nach Leergutart einen unterschiedlichen Wert haben können. Dieser Zählvorgang stellte auch die Rahmenbedingungen unseres Vorhabens dar. Ziel war es, eine Lösung zu realisieren, die in menschenähnlicher Präzision den Leergutbestand erfassen kann und dabei ohne manuelle Zählaktivitäten auskommt. Daher haben wir auf eine Kombination von Drohne und KI-Algorithmen zur Objekterkennung gesetzt. Das Vorgehen der prototypischen Umsetzung ist mehrstufig. Zunächst müssen die Objekte (Paletten) auf den Bildern erkannt werden, dann müssen die identifizierten Objekte lokalisiert und schließlich klassifiziert werden, um so den Paletten einen Wert zuzuordnen zu können. Im ersten Schritt haben wir uns hierbei auf die Front-Ansicht der Palettenstapel fokussiert.

VORBEREITENDE SCHRITTE UND AUFBAU DER DATENBASIS

Die Inventur von Paletten mit leeren Getränkekeisten stellt einen sehr spezifischen Vorgang dar, weshalb hierzu keine öffentlichen Datasets existieren. Stattdessen wurde ein eigenes Dataset erstellt. An mehreren Drehtagen wurden auf dem Gelände von Getränke Essmann Videoaufnahmen mit einer handelsüblichen Drohne (Dji Phantom pro 4 v2¹) gemacht. Die Videoaufnahmen wurden im Anschluss manuell weiterverarbeitet. Zunächst wurden die Videos in einzelne Frames geteilt (30 Frames pro Sekunde der Videoaufnahme). Im Anschluss daran wurden die resultierenden Frames manuell sortiert (mit Fokus auf maximale Varianz) und geeignete Daten für die Implementierung ausgewählt. Der finale Datensatz bestand dann aus insgesamt 336 Bildern. Nach Bildauswahl konnte der Labelprozess gestartet werden. Hierzu wurden die Bilder des Datensatzes mithilfe eines Open-Source-Labeltools (labelstud.io²) annotiert. Im ersten Schritt wurden händisch 6732 Bounding Boxes auf den 336 Bildern eingezeichnet, jeweils eine Bounding Box pro Palette (siehe Abb. 1, rechtes Bild).



¹ <https://www.dji.com/de/phantom-4-pro-v2/specs>

² <https://labelstud.io/>



Abbildung 1:
Preprocessing und
Annotation der Datenbasis

Implementierung eines Baseline-Modells

Nach Erstellung und Annotation des Datasets wurde ein erstes Deep Learning-Modell gestaltet, das die Basis für weitere Entwicklungsschritte bildet. Hierzu wurde die etablierte Mask R-CNN-Architektur genutzt und ein Modell mit dem Framework Detectron2³ implementiert. Nach einem Pretraining auf dem MSCOCO17⁴-Dataset wurde das Training auf dem zuvor erstellten Dataset gestartet. In 4.000 Iterationen mit einer Batchsize von 12 Bildern pro Iteration wurde das Modell auf einer Nvidia RTX 2080 Ti⁵ - Grafikkarte trainiert. Die Ergebnisse waren im ersten Schritt nicht zufriedenstellend. Zur Evaluation wurde die mAP-Metrik (Mean Average Precision) genutzt, welche eine gängige Metrik für die Evaluation von Object Detection-Modellen bildet. Das genutzte Modell ist nicht nur in der Lage Bounding Boxes zu bestimmen, sondern liefert auch Masken – bestimmt also die Pixel, die zu einem identifizierten Objekt gehören. Bei der Evaluation konnte lediglich ein Wert von 27.06 für die Bounding Boxes und 27.59 für die Maskierung erreicht werden. Ein ähnliches Bild zeigte sich bei der visuellen Kontrolle, also der Bestimmung von Paletten auf Bildern. Hier hat sich schnell gezeigt, dass die Ergebnisse des Baseline-Modells nicht ausreichend sind.

Optimierung der Vorverarbeitung und Vergleich der Modellarchitekturen

Da das simpel gehaltene Baseline-Modell zu keinen guten Ergebnissen geführt hat, wurden zunächst die Fehlerquellen analysiert. Dabei zeigte sich, dass das Modell vor allem an den seitlichen Bildrändern Schwierigkeiten hatte, Paletten zielsicher zu identifizieren. In einem ersten Optimierungsschritt wurde daher die Bildfläche um 25% auf den beiden Seiten reduziert (siehe Abb. 1, Mitte). Die Reduktion stellte für den vorliegenden Use Case kein Problem dar, da die Drohne ohnehin alle Paletten erfasst – ausgeschnittene Paletten sind also auf anderen Bildaus-schnitten noch vorhanden und werden bei der Inventur erfasst. Nach dem erneuten Training des Modells zeigte sich, dass die Reduktion der Bildfläche einen großen positiven Effekt hatte und die Paletten schon wesentlich besser erkannt wurden (mAP: 47.68 / 46.70).

Eine weitere Verbesserung konnte durch den Verzicht auf unterschiedliche Klassen während der Object Detection erreicht werden. Im Baseline-Modell wurde zwischen „Bier“ und „Andere“ unterschieden, um neben der Object Detection auch direkt das erkannte Objekt – die Palette – zu klassifizieren. Durch den Verzicht der

Klasse und damit der gleichen Behandlung für jede Palette konnte ebenfalls eine Performancesteigerung erreicht werden (mAP: 45.95 / 46.70).

Ein weiterer gängiger Schritt zur Optimierung von Machine Learning-Modellen ist die Anwendung von Data Augmentation-Verfahren, also die Erweiterung der Datenbasis über generierte Bilder.

Auf der Abbildung 2 werden allerdings keine Bilder neu generiert, sondern vielmehr dienen die Bilder des Datensatzes als Basis, um diese leicht zu verändern (bspw. Farbkorrekturen oder andere Transformationen). Es werden also Originalbilder kopiert und transformiert, um die Datenbasis zu erweitern und dem Algorithmus neue Lernreize zu bieten. Da zunächst unklar war, welche Verfahren sich für unseren Ansatz eignen, haben wir die Auswahl der zu testenden Image

³ <https://github.com/facebookresearch/detectron2>

⁴ <https://cocodataset.org/#home>

⁵ <https://www.nvidia.com/de-de/geforce/graphics-cards/rtx-2080-ti>

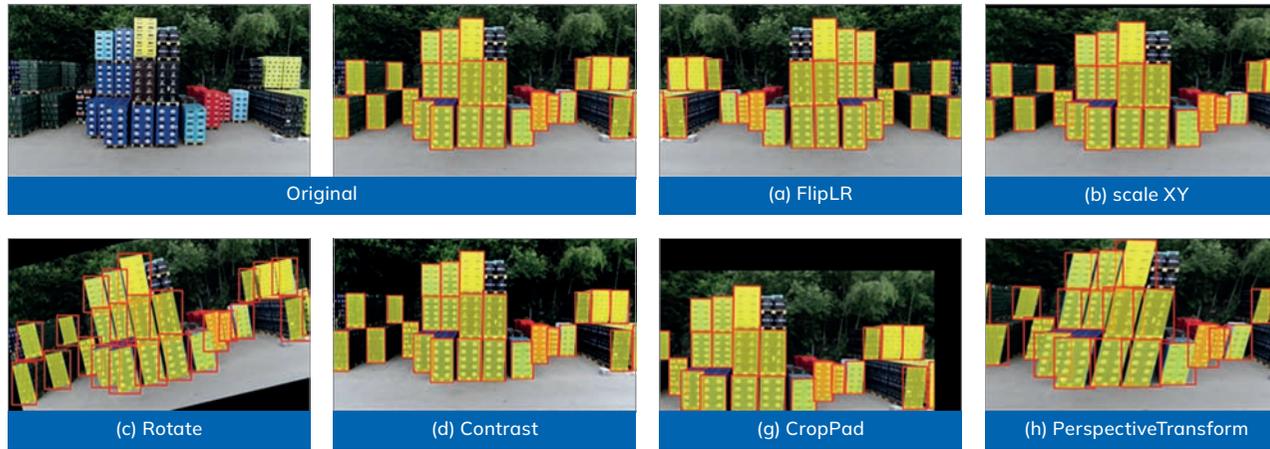


Abbildung 2: Auswahl an Image Augmentation-Verfahren

	Baseline	Keine Klassen	Reduktion der Bildfläche	Keine Klassen und Reduktion der Bildfläche
keine Augmentation	27.06/27.59	45.95/46.70	47.68/48.93	70.39/72.19
Spiegeln (horizontal)	30.03/31.13	46.58/47.48	53.58/55.44	73.94/76.01
Skalierung	49.83/51.00	73.35/75.54	56.01/57.49	78.86/80.93
Rotation	25.48/26.23	41.89/42.81	52.87/53.46	77.07/77.63
Contrast	26.35/26.89	40.22/40.71	48.97/50.05	68.69/70.15
Jpeg Compression	26.12/27.06	42.21/43.29	49.35/50.65	67.83/69.35
Unschärfe	25.27/26.30	43.61/44.76	44.89/46.30	71.19/73.01
Perspektivische Transformation	31.54/32.10	53.13/53.99	54.10/55.42	78.11/79.30
Spiegeln (horizontal) + Skalierung	51.33/52.71	73.84/75.78	57.09/58.85	81.53/83.41

Tabelle 1: Ergebnisse der Vorverarbeitungs- und Image Augmentation-Verfahren

Augmentation-Verfahren, basierend auf gängigen Verfahren aus Forschung und Praxis aufgebaut. Zusätzlich haben wir unsere Beobachtungen aus den bisherigen Experimenten eingebracht und versucht bestimmte Spezifika unseres Versuchsaufbaus einfließen zu lassen. So haben wir gezielt versucht verschiedene Wettereffekte und die Bewegung der Drohne (leichte Neigung der Aufnahme) zu simulieren. Zur besseren Veranschaulichung der Image Augmentation haben wir einige Verfahren und ihr Resultat in Abb. 2 visualisiert.

Alle Augmentation-Verfahren wurden zunächst individuell mit den verschiedenen Modellvarianten (Baseline, Baseline + keine Klassen, Baseline + Reduktion der Bildfläche, Baseline + Reduktion der Bildfläche + keine Klassen) getestet und pro Kombination ein Modell trainiert. Zusätzlich wurden einige Augmentation-Verfahren kombiniert. Insgesamt hat sich gezeigt, dass fast alle Augmentation-Verfahren einen positiven Einfluss auf die Performance haben können. Einzig die Verfahren, die ausgewählt wurden, um Wettereinflüsse zu simulieren, führten zu schlechteren Ergebnissen. Die beste Kombination stellte der Verzicht auf Klassen, die Reduktion der Bildfläche verbunden mit dem horizontalen Spiegeln von Bildern und einer Skalierung beider Achsen dar (mAP: 81.53/83.41). Eine Übersicht zu den Ergebnissen ist in der nachfolgenden Tabelle dargestellt:

Ebenso wichtig wie die Evaluation über Kennzahlen, ist die manuelle Sicht-Evaluation der Ergebnisse. Hierzu haben wir Bilder in das beste Modell gegeben und die vom Modell gesetzten Bounding Boxes auf ihre Güte analysiert. Einige Beispiele für die Vorhersagen des Modells sind in Abb. 3 dargestellt.

Ebenso wichtig wie die Evaluation über Kennzahlen, ist die manuelle Sicht-Evaluation der Ergebnisse. Hierzu haben wir Bilder in das beste Modell gegeben und die vom Modell gesetzten Bounding Boxes auf ihre Güte analysiert. Einige Beispiele für die Vorhersagen des Modells sind in Abb. 3 dargestellt.



Abbildung 3:
Ergebnisse des Baseline-Modells
(Mask R-CNN)



Abbildung 4:
Ergebnisse des finalen
Modells (DetectorS)

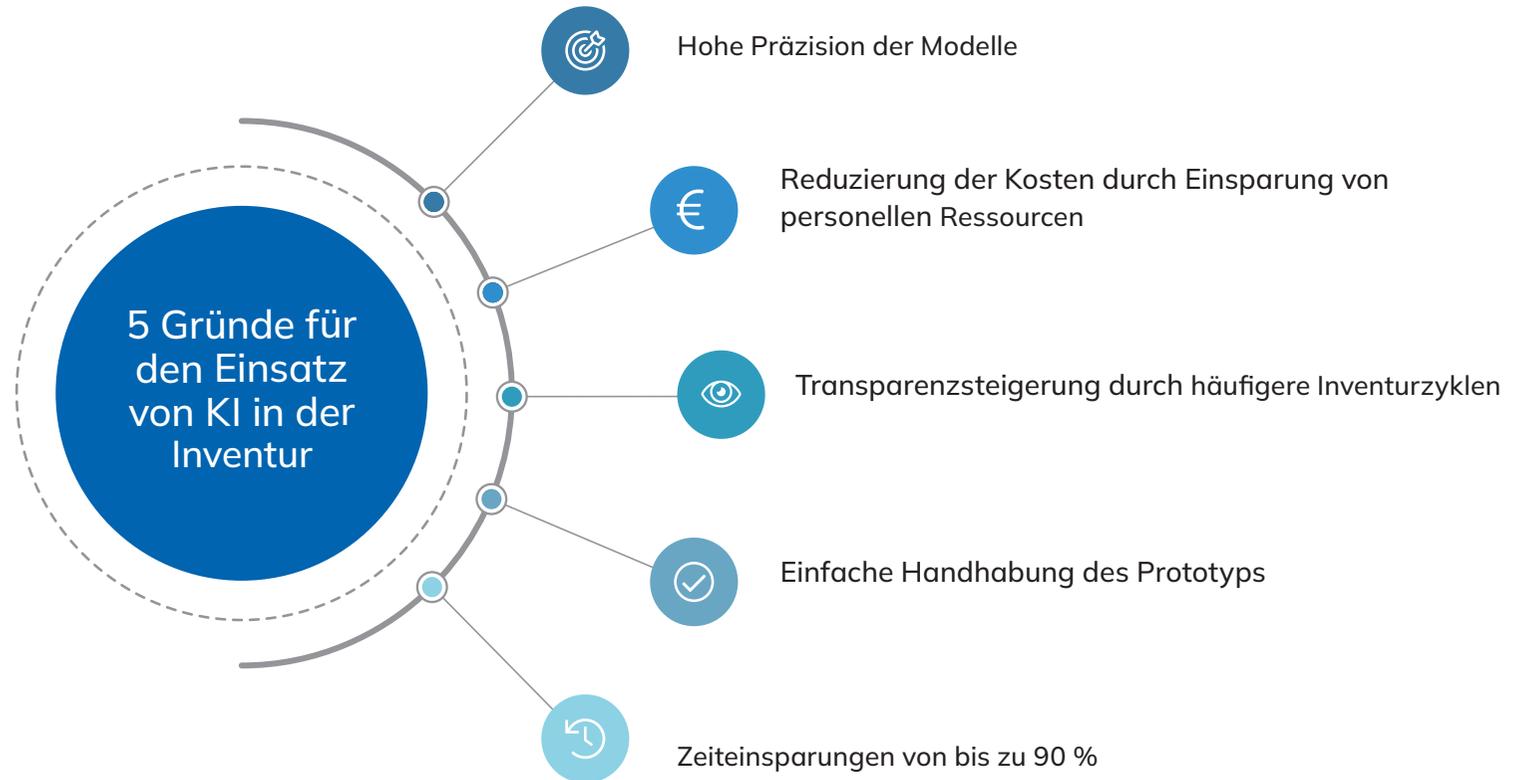
Obwohl das beste Modell schon sehr vielversprechende Ergebnisse lieferte, haben wir weitere Modellarchitekturen herangezogen, da in der Forschung mittlerweile eine Vielzahl von aktuelleren Modellen verfügbar sind, die auf Benchmark-Datensätzen (Laborversuche) auch eine bessere Performance als die von uns verwendete Mask R-CNN-Architektur bieten. Die Wahl fiel dabei auf die Modellarchitekturen DetectoRS (State-of-the-Art und innovative Modelleigenschaften), Yolact (Realtime-

Fähigkeit) und DETR (Transformer-Architektur statt Convolutional Neural Network).

Für das Training der weiteren Modelle wurde die zuvor beste Kombination der Vorverarbeitungsverfahren (Verzicht auf Klassen, Reduktion der Bildfläche und horizontales Spiegeln sowie Skalierung) genutzt. Nach dem Abschluss des Trainings hat sich gezeigt, dass das Modell mit der DetectoRS-Architektur die zufriedenstellenden

Ergebnisse des Baseline-Modells übertreffen konnte. Die Modelle mit der Yolact- bzw. DETR-Architektur konnten das Niveau dagegen nicht erreichen. Die Performance des finalen Modells hat eine Mean Average Precision von 88.3 für die Bounding Boxen und 85.6 für die Maskierung der Paletten erreichen können. Auch in der Sichtprüfung bestand das Modell und konnte auf den Aufnahmen alle Paletten identifizieren, die erkannt werden sollen (Inhalt sowie Holzpalette sichtbar).

Fünf Gründe für den Einsatz von KI in der Inventur.



Potenziale und Anwendung

Zusammenfassend hat sich gezeigt, dass die manuelle Zählung während einer Inventur mit Hilfe von Drohnen und KI-Algorithmen zur Objekterkennung automatisiert werden kann. In unseren Experimenten zeigte sich, dass die resultierenden Modelle Paletten mit einer hohen Zuverlässigkeit erkennen können. Statt mehrere Teams manuell zählen zu lassen, wäre bei der Anwendung dieses Ansatzes nur noch eine Person als Drohnenoperator notwendig, um die Aufnahmen des Lagers zu machen. Die eigentliche Zählung würden die Modelle übernehmen. Vergleicht man den Status quo von manueller mit automatisierter Inventur anhand von messbaren Zahlen, dann werden die Potenziale besonders

sichtbar: Beim Vergleich der Personenaufwände für Vorbereitung, Durchführung und Nachbereitung einer Inventur hat sich herausgestellt, dass, durch den Einsatz einer Drohne und unserer KI-Modelle, bis zu 90% der anfallenden Aufwände eingespart werden können. Besonders großes Einsparungspotenzial liegt dabei in der Vorbereitung der Inventur und in der Durchführung. Für die manuelle Zählung müssen an vielen Stellen Palettenstapel umpositioniert werden, damit diese überhaupt gezählt werden können – mit der Drohne sind dagegen auch Stapel erreichbar, die es ansonsten nicht wären. Den größten Hebel stellt allerdings die eigentliche Zählung dar. Hier hat sich gezeigt, dass eine Drohne nur

ca. 5 % von der für eine manuelle Zählung benötigten Zeit bedarf. Die manuelle Nachbereitung bleibt allerdings bei beiden Ansätzen fast identisch.

Aufgrund der aussichtsreichen Ergebnisse wird dieses Projekt gemeinsam mit unserem Partner Getränke Essmann weiterhin verfolgt, um es zu erweitern und die Algorithmen noch robuster zu gestalten. Neben deren Verfeinerung ist auch noch eine Operationalisierung der Modelle geplant – so sollen sie in Zukunft an die SAP-Landschaft angebunden werden.

Hinweis auf die Publikation

Diese Forschungsprojekt wurde auch wissenschaftlich verwertet, auf dem internationalen Symposium „AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021)“ publiziert und am 23.03.2021 von René Kessler präsentiert. Der Artikel trägt den Titel „Using Camera-Drones and Artificial Intelligence to Automate Warehouse Inventory“ und ist unter folgendem Link einsehbar:

<https://proceedings.aaai-make.info/AAAI-MAKE-PROCEEDINGS-2021/paper38.pdf>

Der Fokus der wissenschaftlichen Publikation lag insbesondere auf dem Vergleich der verschiedenen Image Augmentation-Verfahren und der Modellarchitekturen.

Verweis auf den Artikel:

René Kessler, Christian Melching, Ralph Göhrs und Jorge Marx Gómez (2021): Using Camera-Drones and Artificial Intelligence to Automate Warehouse Inventory. AAAI 2021 Spring Symposium on Combining Machine Learning and Knowledge Engineering (AAAI-MAKE 2021).

RALF GÖHRS

ralf.goehrs@abat.de

RENÉ KESSLER

rene.kessler@uni-oldenburg.de

ÜBER UNS

The logo for 'abat' is written in a clean, white, lowercase, sans-serif font. The letters are bold and have a modern, slightly rounded appearance.

abat ist ein internationaler SAP-Dienstleister und Produkthanbieter, der Unternehmensprozesse optimiert und mit eigenen Lösungen weiterentwickelt. Unsere Leistungen

erbringen wir vorwiegend in den Branchen Automotive, Diskrete Fertigung und Logistik. Auch in den Bereichen Nachhaltigkeitsmanagement sowie Informationssicherheit können wir Sie unterstützen.

Wir sind deutschlandweit mit Standorten in Bremen, München, Oldenburg, St. Ingbert, Walldorf und Wolfsburg vertreten und besitzen Niederlassungen in den USA, Mexiko und Belarus. abat ist SAP Gold Partner sowie Entwicklungspartner im SAP Partner Edge Program for Application Development und besitzt darüber hinaus Recognized Expertise für Automotive, Travel and Transportation, Supply Chain Management, Consumer Products sowie SAP S/4HANA.

Im Bereich Automotive arbeiten wir für viele Hersteller, wie etwa Audi, BMW, Daimler, MAN, Porsche, Volkswagen und Qoros. Mit unserer eigenentwickelten Software PLUS auf Basis von SAP, werden die Produktionsprozesse bei Daimler gesteuert. Für Logistiker und Logistikdienstleister bietet abat Lösungen zur Optimierung der Supply Chain, z.B. in der Intralogistik und im Transportmanagement. Hier arbeiten wir unter anderem für Brose, Bosch, Daimler, DHL, thyssenkrupp und VS HEIBO Logistics. Bei der Sportsoftware SAP Sports One sind wir im Moment der einzige Einführungspartner der SAP. Im Nachhaltigkeitsmanagement bietet abat eine ganzheitliche Beratung an und zusätzlich die Unterstützung bei der Nachhaltigkeitsberichterstattung durch unsere Software ID-Report. Last but not least bieten wir im Bereich Informationssicherheit Beratungsleistungen zu den Themen ISO 27001 sowie TISAX® und helfen so, die Werte unserer Kunden und wiederum deren Kundschaft zu schützen

abat | An der Reeperbahn 10 | 28217 Bremen | +49 421 43 04 60 | www.abat.de