

# Automatisierte Bilderzeugung und -klassifikation verschiedener Fördererlemente der Intralogistik durch Anwendung maschinellen Lernens

Modelle oder digitale Zwillinge von Anlagen ermöglichen die virtuelle Inbetriebnahme einer Anlage und bieten somit die Möglichkeit eventuell auftretende Probleme, bereits in einem frühen Stadium der Entwicklung, zu identifizieren und gegebenenfalls zu beheben. Bei, in der Intralogistik verwendeten, Förderanlagen ermöglicht eine solche virtuelle Inbetriebnahme beispielsweise die Erprobung von Steuerungscode oder Förderstrategien vor der eigentlichen Fertigstellung der Anlage. Auch bei Nachrüstungen oder Modernisierungen, dem „Retrofitting“ bestehender Förderanlagen, kann durch die Nutzung eines digitalen Anlagenmodells die frühzeitige Identifikation von Problemen ermöglicht und somit eine längere Stilllegung der Anlage vermieden werden. [Mül-2019]

## **Problemstellung und Zielsetzung:**

Der, durch die nachträgliche Erzeugung solcher digitaler Anlagenmodelle, entstehende Aufwand ist insbesondere bei Retrofit Projekten häufig sehr hoch. Dies liegt daran, dass ältere Anlagen oftmals nicht nur über kein aktuelles, digitalisiertes Modell verfügen, sondern auch unvollständige, veraltete oder nicht digitalisierte Dokumentationen vorweisen. Da auch die Kosten zur nachträglichen Erzeugung eines solchen Modells zu hoch sind, wird in vielen Fällen auf deren Verwendung verzichtet. [Mül-2019]

Um die Erzeugung digitaler Modelle zu wirtschaftlich günstigen Bedingungen zu ermöglichen, soll im Rahmen des Forschungsprojektes AGenDa (Automatisierte Generierung digitaler Daten von bestehenden Intralogistikanlagen) ein intelligenter Behälter entwickelt werden, welcher die zur Modellerzeugung benötigten Daten, während einer Behälterfahrt über die Förderanlage, automatisiert extrahiert und in digitaler Form abspeichert. Abschließend sollen die so gewonnenen Daten dann automatisiert ausgewertet und zur (teil)automatischen Modellerzeugung der Anlage genutzt werden. [Mül-2019]

Zur Erzeugung eines digitalen Modells werden unterschiedliche Daten der realen Anlage benötigt. Neben geometrischen Informationen einzelner Fördererlemente, wie bspw. der Länge, Breite oder Steigung werden zum Beispiel auch Informationen über die Art der verwendeten Fördererlemente (Rollen-, Gurt-, Ketten-, Kugelförderer, usw.) erfordert. Zur Extraktion der benötigten Daten soll der, im Projekt AGenDa entwickelte, Behälter über eine Reihe unterschiedlicher Sensoren verfügen, welche eine Ermittlung der benötigten Informationen im laufenden Betrieb der Förderanlage ermöglichen. Anschließend sollen die so gewonnenen Daten dann zu einem digitalen Anlagenmodell zusammengeführt werden.

Die vorliegende Arbeit befasste sich mit der Identifikation der verwendeten Fördererlemente. Das Ziel der Arbeit war somit die Entwicklung einer Möglichkeit zur automatisierten Erkennung unterschiedlicher Fördererlemente anhand von, während der Behälterfahrt über die Förderanlage, aufgezeichneten Daten. In einem ersten Schritt musste hierzu eine Kamera in dem System des intelligenten Behälters integriert werden. Anschließend musste ein Algorithmus zur Bildklassifikation entwickelt werden, welcher eine Zuweisung der aufgezeichneten Bilder in vordefinierte Klassen ermöglicht. Da für die angestrebte Klassifikationsaufgabe keine Echtzeitfähigkeit benötigt wird, wurden die Bilderzeugung und die anschließende Klassifikation getrennt voneinander betrachtet.

## Bilderzeugung

Zur Realisierung der automatisierten Bilderzeugung im laufenden Betrieb der Förderanlage, musste die bereits vorhandene Hardware des intelligenten Behälters genutzt werden. Diese setzte sich aus einem Mikrocontroller, einer Festplatte sowie einer Kamera zusammen. Bei der verwendeten Kamera handelte es sich um eine Stereokamera, welche neben Farbbildern auch Tiefenbilder aufzeichnen kann. Um bei der Behälterfahrt die benötigten Informationen über die Fördererlemente einer Anlage zu erhalten, musste die Kamera so am Behälter angebracht werden, dass sich der jeweils vorausliegende Streckenabschnitt der Förderanlage im Sichtfeld von dieser befindet. Hierzu wurde die Kamera, wie in Abbildung 1 zu erkennen, mithilfe einer Befestigungsvorrichtung in ausreichender Höhe, mittig am Behälter angebracht.

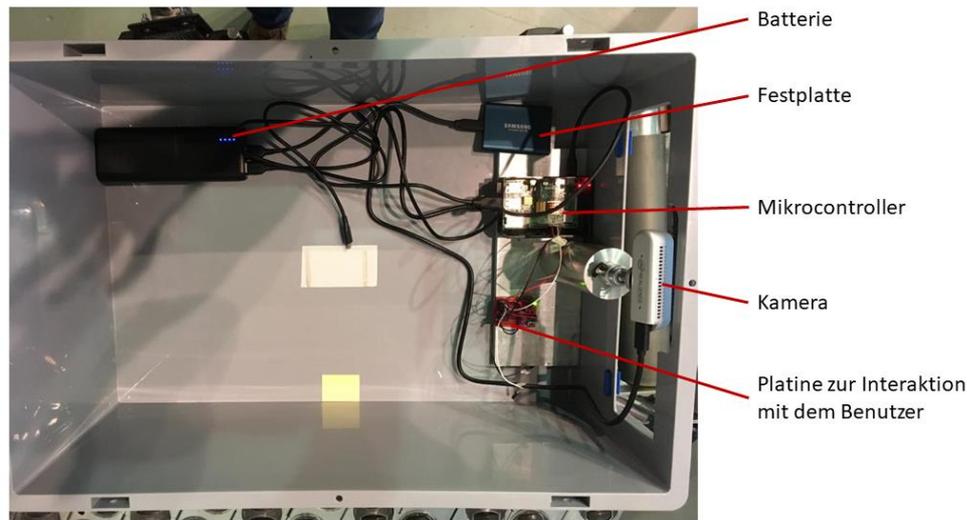


Abbildung 1 - Aufbau des intelligenten Behälters.

Anschließend wurde die Kamera dann auf dem Mikrocontroller in Betrieb genommen. Um den jeweiligen Förderer zu jedem Zeitpunkt der Bildaufzeichnung eindeutig bestimmen zu können, muss eine ausreichend hohe Bildfrequenz garantiert sein. Da der verwendete Mikrocontroller im Vergleich zu leistungsfähigen Computern allerdings eine deutlich geringere Rechenleistung aufweist, ist dieser nicht in der Lage die zur Verfügung stehenden Farb- und Tiefenbilder mit ausreichender Bildfrequenz abzuspeichern.

Zur Steigerung der Bildfrequenz wurde die Bilderzeugung in zwei separate Schritte unterteilt. Im ersten Schritt wurden die Rohdaten der Kamera während der Behälterfahrt über die Förderanlage aufgezeichnet und auf der Festplatte abgespeichert. Das anschließende Auslesen der Rohdaten und somit die Speicherung der eigentlichen Bilder erfolgte dann in einem zweiten Schritt und wurde, zur schnelleren Ausführung, genau wie die spätere Klassifikation auf einem handelsüblichen Computer ausgeführt. Durch diese Aufteilung der Bilderzeugung konnte die erreichte Bildfrequenz von 4 fps (Bilder pro Sekunde) auf etwa 18 fps erhöht werden.

## Bildklassifikation

Die einheitliche Klassenzuweisung kompletter Bilder in vordefinierte Klassen beschreibt das Themengebiet der Bildklassifikation [Süß-2014, S. 589]. Aufgrund verschiedener Anforderungen, Randbedingungen und Beschränkungen unterschiedlicher Anwendungsbereiche, gibt es keine allgemein anwendbare Methode zur Lösung von Bildklassifikationsproblemen [Tre-2010, S. 7, 183].

Probleme der Bildklassifikation lassen sich jedoch grundsätzlich in die vier getrennten Aufgabenblöcke, der Bilderzeugung, Bildverarbeitung, Merkmalsextraktion und der Klassifikation unterteilen. Während sich der Aufgabenblock der Bilderzeugung mit der Erstellung der Bilder,

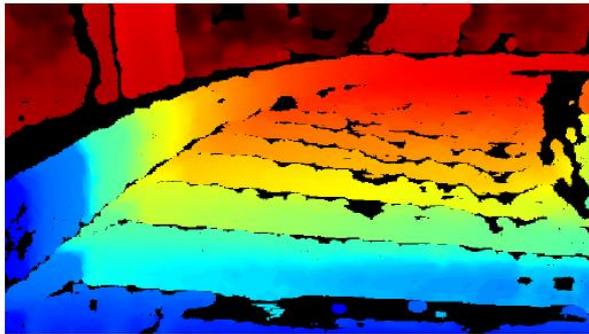
also der numerischen Beschreibung einer realen Szene auseinandersetzt, fokussiert sich die Bildverarbeitung auf die anschließende Bearbeitung der erzeugten Bilder hin zu einer, für die Extraktion von Bildinformationen optimalen Darstellung. Die Merkmalsextraktion setzt sich anschließend mit der numerischen Beschreibung von Bildeigenschaften auseinander, auf der Grundlage welcher im letzten Schritt der Klassifikation dann die Zuweisung zu einer Klasse erfolgt. [Jia-2009]

Neben dieser klassischen Vorgehensweise der Bildklassifikation, finden in den letzten Jahren vermehrt auch Ansätze des „Deep Learnings“ Einzug in das Themengebiet der Bilderkennung. Als „Deep Learning“ wird die Verwendung künstlicher neuronaler Netze bezeichnet. Neuronale Netze sind dabei in der Lage, anhand eines ausreichend großen Datensatzes, eigenständig zur Bildklassifikation geeignete Merkmale zu identifizieren und extrahieren, sowie Bilder auf der Grundlage dieser Merkmale zu klassifizieren [Sch-2012]. Diese Verfahren sind häufig in der Lage höhere Zuweisungsgenauigkeiten als Ansätze der klassischen Vorgehensweise zu erreichen, benötigen hierzu aber einen ausreichend großen und umfangreichen Trainingsdatensatz. [Oma-2020]

Für die angestrebte Bildklassifikation unterschiedlicher Fördererlemente existierte jedoch im Voraus kein geeigneter Datensatz, aus diesem Grund wurde zur Bildklassifikation die klassische Vorgehensweise verwendet. Hierbei sollte eine Möglichkeit entwickelt werden, welche eine Klassifikation von den, in Intralogistikanlagen am häufigsten verwendeten, Förderarten ermöglicht. Bei den so definierten Klassen handelte es sich um Gurt-, Rollen- und Kugelförderer. Um die Bilder eindeutig den vordefinierten Klassen zuweisen zu können, musste in einem ersten Schritt ein geeignetes Verfahren zur mathematischen Bildbeschreibung identifiziert werden. Hierzu wurden eine Reihe unterschiedlicher Merkmalsdeskriptoren betrachtet und es konnte festgestellt werden, dass besonders der HOG-Deskriptor („Histogram of oriented gradients“) für den betrachteten Klassifikationsfall gute Resultate verspricht. Hierbei handelt es sich um ein Verfahren, welches zur Bildbeschreibung die Häufigkeitsverteilung auftretender Gradientenrichtungen in verschiedenen Bildregionen verwendet. Bildgradienten beschreiben dabei die Intensitätsunterschiede benachbarter Bildpunkte und bieten somit die Möglichkeit Bildkanten zu beschreiben.

Beim HOG-Deskriptor handelt es sich um ein globales Bildmerkmal, also um ein Verfahren, welches über das komplette Bild berechnet wird. Es wird folglich nicht zwischen Umgebung und Förderer unterschieden. Um zu vermeiden, dass die Umgebung Einfluss auf die Klassenzuweisung des abgebildeten Förderers hat, muss der Förderer vor der Merkmalsberechnung von der Umgebung hervorgehoben werden. Dieser Schritt wird als Segmentierung bezeichnet.

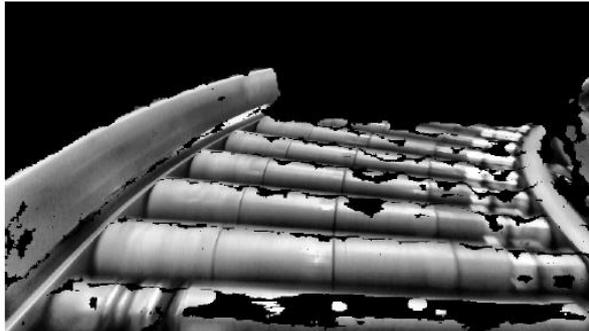
Zur Realisierung der Bildsegmentierung wurden die, von der Kamera ebenfalls aufgezeichneten, Tiefeninformationen verwendet. Anhand einer Schwellwertoperation wurde zunächst eine Bildmaske erzeugt, welche alle Punkte die eine Entfernung größer als 0,8 m zur Kamera aufweisen, ausblendet. Eine anschließende Überlagerung der Bildmaske mit dem jeweiligen Farbbild, führt dann wie in Abbildung 2 zu erkennen, zur Ausblendung der Umgebung. Da Reflexionen und Abschattungen zur fehlerhaften Berechnung von Entfernungen führen können, verfügt die erzeugte Bildmaske über „Flecken“. Durch die nachträgliche Anwendung morphologischer Operationen können diese jedoch beseitigt werden.



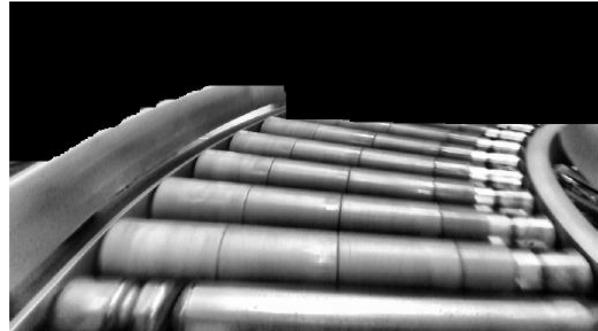
(a) Farbliche Darstellung der Entfernungsinformationen



(b) Erzeugte Bildmaske



(c) Anwendung der Maske auf dem Farbbild



(d) Anwendung der bearbeiteten Maske auf dem Farbbild

Abbildung 2 - Bearbeitungsschritte der Bildsegmentierung.

Die anschließende Berechnung der HOG-Merkmale auf dem segmentierten Bild, führte zu einem 378-dimensionalen Merkmalsvektor. Die Aufgabe des Klassifikators ist es, den so erhaltenen Merkmalsraum unterschiedlichen Klassen zuzuteilen. Bei Problemen der Bilderkennung ist diese Unterteilung des Merkmalsraumes nicht ohne Weiteres zu erkennen, weshalb hier häufig Methoden des maschinellen Lernens zur Klassifikation verwendet werden. Im Rahmen der Arbeit wurde eine Stützvektormaschine zur Klassifikation verwendet. Anhand von Trainingsdaten, werden hierbei zunächst Hyperebenen im Merkmalsraum bestimmt, welche die vorhandenen Klassen bestmöglich gegeneinander abgrenzen. Die auf diese Weise erhaltenen Hyperebenen werden anschließend in der Klassifikationsphase verwendet, um unbekannte Merkmalsvektoren eindeutig einer Klasse zuzuweisen.

Zum Trainieren des Klassifikators wird folglich ein Datensatz aus, bereits einer Klasse zugewiesenen, Bildern benötigt. Da für die angestrebte Klassifikationsaufgabe kein solcher Datensatz existierte, musste zunächst ein eigener Datensatz erstellt werden. Hierzu konnte auf die Testanlage eines Industriepartners des Forschungsprojektes zurückgegriffen werden. Die anhand des intelligenten Behälters gewonnenen Bilder von Rollen-, Gurt- und Kugelförderern wurden anschließend manuell den jeweiligen Klassen zugeteilt. Der auf diese Weise erstellte Datensatz umfasste jeweils 100 Bilder pro Klasse.

### Resultate und Schlussfolgerung

Um nach dem Trainieren des Klassifikators, dessen Klassifikationsgenauigkeit bewerten zu können, muss auch die Klassenzugehörigkeit der hierzu verwendeten Testdaten bereits vor der Ausführung des Klassifikationsalgorithmus bekannt sein. Um solche Daten zu erhalten, wird der Datensatz in der Regel in einen Trainings- und einen Testsatz unterteilt. Die Aufteilung des Datensatzes erfolgte entsprechend in 20% Test- und 80% Trainingsdaten, wobei auf eine gleichmäßige Klassenverteilung geachtet wurde. Nach dem Trainieren des Klassifikators mit den Trainingsdaten, konnte dann eine Klassifikationsgenauigkeit von 100 % festgestellt werden. Anhand einer Hauptkomponentenanalyse des Merkmalsraumes konnte gezeigt werden, dass die guten Klassifikationsresultate durch die paarweise lineare Separierbarkeit der Merkmalsvektoren

der einzelnen Klassen im Datensatz begründet sind.

Schlussendlich wurde der trainierte Klassifikator dann auf den einzelnen Bildern kompletter Kameraaufzeichnungen des intelligenten Behälters angewendet. Die guten Klassifikationsresultate konnten auf diese Weise bestätigt werden. So war der Klassifikator in der Lage Bilder von Rollenförderer mit einer Genauigkeit von 99,83 % zu klassifizieren. Gurtförderer konnten hingegen mit 93,65 % Genauigkeit und Kugelförderer mit einer Genauigkeit von 100 % richtig klassifiziert werden.

Da im Rahmen der Arbeit nur eine Förderanlage zur Verfügung stand, muss bei der Betrachtung der Resultate berücksichtigt werden, dass die klassifizierten Kameraaufzeichnungen auf der gleichen Förderanlage, welche auch zur Erzeugung des Datensatzes verwendet wurde, erstellt wurden. Weiterhin muss auch beachtet werden, dass diese Anlage größtenteils aus Rollenförderern besteht und somit einzelne Fehlklassifikationen von Gurt- oder Kugelförderern, bei der Berechnung der Klassifikationsgenauigkeiten, deutlich stärker ins Gewicht fallen.

Anhand der erhaltenen Resultate konnte gezeigt werden, dass die Bestimmung verschiedener Förderelemente durch die Klassifikation von Kamerabildern gute Resultate verspricht. Weiterhin konnte eine Vorgehensweise zur Bildsegmentierung und Merkmalsextraktion entwickelt werden, welche die Unterscheidung zwischen Rollen-, Gurt und Kugelförderern ermöglicht. Aufgrund der ausschließlichen Verfügbarkeit eines begrenzten Datensatzes, müssen die erhaltenen Ergebnisse jedoch weiter untersucht werden und auf größeren Datensätzen bestätigt werden.

## QUELLEN

- [Jia-2009] Jiang, Xudong: Feature extraction for image recognition and computer vision. In: 2nd IEEE International Conference on Computer Science and Information Technology, 2009.
- [Mül-2019] Müller, Sandra: AGenDa – Automatisierte Generierung digitaler Daten von bestehenden Intralogistikanlagen. Lehrstuhl für Fördertechnik Materialfluss Logistik, 2019.
- [Oma-2020] O’Mahony, Niall u. a.: Deep Learning vs. Traditional Computer Vision. In: Advances in computer vision. Proceedings of the 2019 Computer Vision Conference (CVC), 2020.
- [Sch-2012] Schulz, Hannes; Behnke, Sven: Deep Learning. In: KI - Künstliche Intelligenz, 2012.
- [Süß-2014] Süße, Herbert; Rodner, Erik: Bildverarbeitung und Objekterkennung: Computer Vision in Industrie und Medizin. Computer Vision in Industrie und Medizin. Springer Vieweg, Wiesbaden, 2014.
- [Tre-2010] Treiber, Marco Alexander: An Introduction to Object Recognition: Selected Algorithms for a Wide Variety of Applications. Selected Algorithms for a Wide Variety of Applications. Springer-Verlag London Limited, London, 2010.