

KMU-innovativ - Verbundprojekt

IMPRO - Image-supported Product Data Creation Processor

**- Automatisierte bildgestützte Produktdatenanlage in Informationssystemen
des Handels -
(Förderkennzeichen: 01IS20085)**

Schlussbericht

Kooperationspartner im Verbundprojekt

Unternehmen (Projektkoordinator)	Forschungspartner
retailsolutions GmbH Otto-Kaiser-Straße 4 66386 St. Ingbert	Hochschule Trier, Standort Birkenfeld Institut für Softwaresysteme Postfach 1380, 55761 Birkenfeld

Projektstart: 01.01.2021

Projektende: 31.05.2023

Autoren: Andreas Mohr (retailsolutions GmbH)
Email: andreas.mohr@retailsolutions.de

Prof. Dr. Rolf Krieger (Hochschule Trier, Standort Birkenfeld)
Email: r.krieger@umwelt-campus.de

Datum: 24. November 2023

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS20085 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Inhalt

1.	Einleitung.....	1
2.	Problemstellung	1
3.	Projektpartner	3
4.	Darlegung der Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten	4
4.1	Arbeitspaket 1: Projektmanagement	5
4.2	Arbeitspaket 2: Extraktion produktbezogener Information aus Bildern	6
4.3	Arbeitspaket 3: Gewinnung von Produktinformationen aus Webquellen	14
4.4	Arbeitspaket 4: Intelligente Produktbild-basierte Stammdatenanlage.....	17
4.5	Arbeitspaket 5: Produktbild-basierte Prüfung von Stammdaten	21
4.6	Arbeitspaket 6: Evaluation und Verwertung	26
5.	Die wichtigsten Positionen des zahlenmäßigen Nachweises	28
6.	Fortschritt anderer Stellen während des Vorhabens.....	28
7.	Voraussichtlicher Nutzen, insbesondere Verwertbarkeit des Ergebnisses	29
8.	Erfolgte und geplante Veröffentlichungen der Ergebnisse.....	31
9.	Fazit.....	31

1. Einleitung

Handelsunternehmen verfolgen heute vielfach eine Omni-Channel-Retailing-Strategie, die alle Absatzkanäle wie Ladengeschäfte, Online-Shops und Kataloge gleichermaßen unterstützt. Sie erfordert, dass die Absatzkanäle parallel bedient, gesteuert, ausgewertet und optimiert werden müssen. Die zu verwaltende Informationsmenge je Produkt hat sich dadurch in den letzten Jahren erheblich erhöht. Dies stellt hohe Anforderungen sowohl an das Produktinformationsmanagement (PIM) als auch an das Stammdatenmanagement im zentralen ERP-System (Enterprise Resource Planning). Ein ERP-System ist eine integrierte Softwarelösung, die Daten und Funktionen zur Unterstützung der Geschäftsprozesse der zentralen Funktionsbereichen eines Unternehmens wie z.B. Einkauf, Bestandsmanagement, Logistik, Verkauf, Rechnungswesen etc. bereitstellt.

Durch den E-Commerce haben gerade Produktbilder zunehmend an Bedeutung gewonnen. Für die Unternehmen bedeutet dies, dass die Produktdaten sehr umfangreich und heterogen sind. Sie bestehen aus strukturierten und unstrukturierten Daten, wozu auch Bild- und Videodaten gehören. Vielfach müssen die Daten noch sehr kostenintensiv und fehleranfällig manuell erfasst werden. Für die Unternehmen ist es daher schwierig sicherzustellen, dass die Produktdaten und die von ihnen repräsentierten Informationen, die aus unterschiedlichen Quellen stammen, korrekt sind. Der hierfür erforderliche Abgleich der auf dem Produktbild enthaltenen Informationen mit den bereitgestellten Produktdatensätzen erfolgt, wenn überhaupt, nur manuell und ist entsprechend aufwändig.

Ziel des Forschungsprojekts war es, den Erstellungsprozess von Produktdaten durch die Einbeziehung von Produktbildern und weiterer Informationen aus Webquellen oder internen Datenquellen mittels eines Image supported Product Data Creation Processors (IMPRO) durch den Einsatz von Verfahren des maschinellen Lernens zu automatisieren. Damit sollen die Kosten insbesondere für die Erfassung von Produktdaten reduziert und die Effizienz und Qualität der erforderlichen Pflegeprozesse erhöht werden.

Der vorliegende Bericht informiert über den Projektverlauf und fasst die Projektergebnisse zusammen. Dazu wird die zugrundeliegende Problemstellung zunächst weiter konkretisiert. Danach werden die geleisteten Projektarbeiten, die wichtigsten Positionen des zahlenmäßigen Nachweises, der voraussichtliche Nutzen und die Verwertbarkeit der Ergebnisse erläutert. Der Bericht schließt mit einer Auflistung der bereits erfolgten bzw. geplanten Veröffentlichungen und einem Fazit.

2. Problemstellung

In großen Handelsunternehmen müssen täglich die Datensätze für Tausende von Produkten neu angelegt, geändert und aufbereitet werden. Dabei muss sichergestellt sein, dass die Produktdaten, die immer wieder Änderungen unterliegen, aktuell, vollständig und korrekt sind, um Kosten zu vermeiden, die durch fehlerhafte Daten, z.B. fehlerhafte Logistikdaten, falsche Preise etc. entstehen. Die produktbezogenen Daten werden entweder direkt von Herstellern übermittelt oder über

Datenpools (z.B. atrify, www.atrify.com, markant, www.markant.com etc.) bezogen. Die Daten sind dabei sehr heterogen. Oftmals liegen zu einem Produkt strukturierte und unstrukturierte produktbezogene Daten vor, die zum Teil unvollständig sind. Im Fall von Produktbildern werden häufig auch mehrere Ansichten gespeichert. Diese Daten müssen zur Erstellung eines vollständigen Produktdatensatzes in einem ERP- oder einem PIM-System (Product Information Management) zusammengeführt und abgeglichen werden. Das PIM-System ist eine Softwarelösung, die es Unternehmen ermöglicht, alle relevanten Daten, Informationen und Marketingmaterialien ihrer Produkte zentral zu verwalten und zu pflegen, um die Produktkommunikation über verschiedene Vertriebskanäle hinweg konsistent und effizient zu gestalten. Deshalb müssen Inkonsistenzen aufgedeckt und behoben werden. Dabei ist zu berücksichtigen, dass ein Produktbild eine Vielzahl von produktbezogenen Daten, wie in Abbildung 1 dargestellt, beinhaltet.



Abbildung 1 Beispiel für eine Verpackung mit darauf dargestellten Produktinformationen

Wesentliches Ziel des Forschungsprojekts war es, Verfahren zu erforschen, um die Produktinformationen aus den Bilddaten zu extrahieren und mit weiteren bereitgestellten Produktdaten, Bild- und Textdaten von Webseiten und ähnlichen Produkten aus einem existierenden ERP-System zu fusionieren. Abschließend sollte die Generierung eines vollständigen Produktdatensatzes erfolgen.

Zur Evaluation sollten die Verfahren prototypisch implementiert und zu einem Image-supported Product Data Creation Processor zusammengeführt werden. Abbildung 2 erläutert die Arbeitsweise des Image-supported Product Data Creation Processors, der in einer Verwertungsphase weiterentwickelt und in einer Systemlandschaft mit SAP Retail (bzw. SAP S/4HANA Retail for Merchandise Management) eingesetzt werden soll.

Aus Abbildung 2 lassen sich vier Funktionsbereiche ableiten, die im Forschungsprojekt genauer betrachtet wurden. Die Funktionsbereiche 1 - 3 beziehen sich auf die Automatisierung des Prozesses der Produktdatenanlage. Funktionsbereich 4 bezieht sich auf Verfahren zur Prüfung der Fehlerfreiheit und der Vollständigkeit von bereits angelegten Produktdatensätzen unter Verwendung der aus Produktbildern extrahierten Informationen.

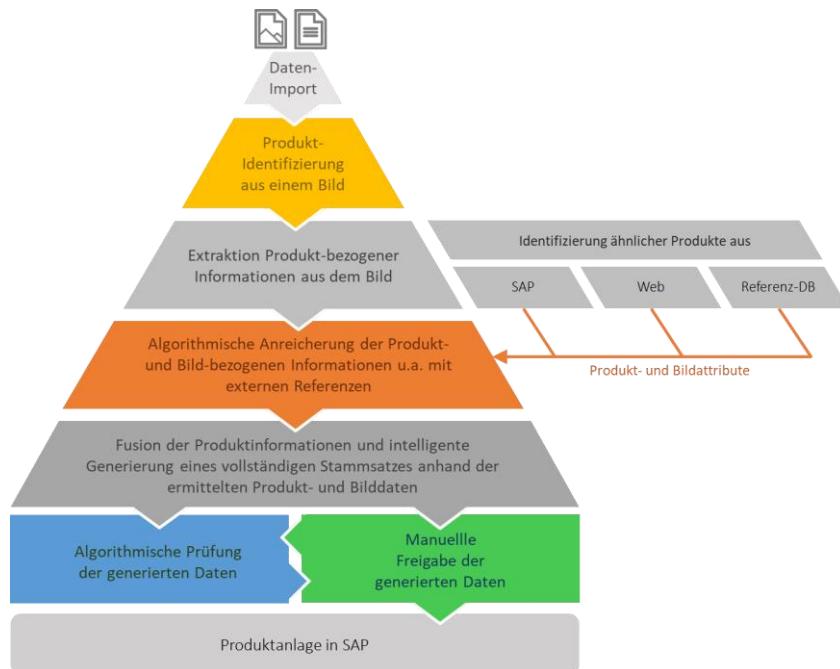


Abbildung 2 Konzept des Image-supported Product Data Creation Processors mit Darstellung der zentralen Komponenten

Im Vorhaben sollte zudem berücksichtigt werden, dass durch IMPRO insbesondere auch Endbenutzer bei der interaktiven Stammdatenpflege unterstützt werden sollen. Es sollten auch Erkenntnisse über das Potenzial des maschinellen Lernens zur Automatisierung von operativen Geschäftsprozessen im Stammdatenmanagement eines Handelsunternehmens gewonnen werden.

3. Projektpartner

Bei dem Forschungsprojekt handelte es sich um ein Verbundprojekt. An dem Projekt waren folgende Partner beteiligt:

1. retailolutions GmbH (Projektkoordinator)

Die retailolutions GmbH mit Standorten in St. Ingbert, Essen, Zug (Schweiz), Salzburg (Österreich), Stockholm (Schweden), Zaragoza (Spanien) und Watford (England) bietet Beratung und IT-Lösungen für den Handel und die Konsumgüterindustrie an. Schwerpunkt bildet die Implementierung und Erweiterung der Branchenlösung SAP for Retail. Das Unternehmen verfügt über tiefgehendes Prozesswissen und bietet SAP-spezifische Add-ons zur Qualitätssicherung von Produktdaten an. Es verfügt über eine ausgewiesene Expertise über den Einsatz von KI-Methoden im Handel.

2. Hochschule Trier, Institut für Softwaresysteme (Forschungspartner)

Das Institut für Softwaresysteme in Wirtschaft, Umwelt und Verwaltung der Hochschule Trier hat seinen Arbeitsschwerpunkt in der Angewandten Informatik. In verschiedenen Forschungsprojekten hat sich das Institut bereits mit Problemen der Datenqualitätssicherung und dem Einsatz und der Entwicklung von KI-Anwendungen im Stammdatenmanagement beschäftigt.

Zusätzlich waren in das Projekt drei Handelsunternehmen als assoziierte Projektpartner, Coop Genossenschaft (Basel), GLOBUS Markthallen Holding GmbH & Co. KG (St. Wendel) und Transgourmet Deutschland (Riedstadt), eingebunden. Die Unternehmen haben das Forschungsprojekt wie geplant durch die Bereitstellung von Daten und bei der Evaluation der entwickelten Verfahren unterstützt.

4. Darlegung der Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Über das Forschungsprojekt wurden verschiedene mit dem Förderprogramm KMU-innovativ verbundene Förderziele erreicht. So hat das Forschungsprojekt als Verbundprojekt zur Intensivierung von Forschungskooperationen zwischen kleinen und mittleren Unternehmen (KMU) und Partnern aus Wissenschaft und Wirtschaft beigetragen. Durch die enge Zusammenarbeit der beiden Projektpartner Hochschule Trier und retailsolutions GmbH wurde innovatives Know-how und Anwendungswissen ausgetauscht und geteilt. Bei beiden Partnern hat das Projekt zu einer Steigerung der Forschungs- und Entwicklungsaktivitäten geführt. Durch gezielte Fortbildungen und den Einsatz neuer Technologien konnten insbesondere die Fähigkeiten der Mitarbeiter:innen bei der retailsolutions GmbH signifikant gesteigert werden, was zu einer höheren Produktivität und Wettbewerbsfähigkeit beiträgt. Ein weiterer Mehrwert, der mit dem Forschungsprojekt in Zusammenhang steht, ist die gestiegene Anzahl an öffentlichkeitswirksamen Präsentationen, wie beispielsweise bei Veranstaltungen wie der Fachtagung Innovative Technologien für den Handel des DFKI und dem retailforum in Frankfurt im Jahr 2022. Diese Veranstaltungen wurden als Plattform genutzt, um die erzielten Fortschritte und Innovationen einem breiteren Publikum vorzustellen, wodurch das Interesse an KI in der Anwendungsdomäne Retail und damit auch das Interesse an der retailsolutions GmbH gesteigert und neue Geschäftsmöglichkeiten geschaffen wurden.

In den nachfolgenden Abschnitten werden die im Forschungsprojekt geleisteten Arbeiten und erzielten Ergebnisse dargestellt und die Erreichung der obengenannten Förderziele belegt. Der bei der Planung des Forschungsvorhabens ausgearbeitete Projektstrukturplan bestand aus sechs Arbeitspaketen. Die nachfolgende Tabelle gibt die Arbeitspakete und die mit ihrer Bearbeitung verbundenen Aufwände an.

Arbeitspaket	Aufwand [PM] Gesamt	Aufwand [PM] retailsolutions	Aufwand [PM] HS Trier
AP1: Projektmanagement	2	1	1
AP2: Extraktion produktbezogener Information aus Bildern	19	8	11

AP3: Gewinnung von Produktinformationen aus Webquellen	10	6	4
AP4: Intelligente Produktbild-basierte Stammdatenanlage	14	11	3
AP5: Produktbild-basierte Prüfung von Stammdaten	11	6	5
AP6: Evaluation und Verwertung	7	4	3

Tabelle 1 Arbeitspakete und Aufwände in Personenmonate

Jedem Arbeitspaket wurden Aktivitäten zugeordnet, um die jeweiligen mit dem Arbeitspaket verbundenen Ziele bzw. Meilensteine zu erreichen. Der Gesamtaufwand betrug 63 PM, wobei 36 PM von der retailolutions GmbH und 27 PM von der Hochschule Trier erbracht wurden. Die Arbeiten begannen bei der retailolutions GmbH wie geplant in 01/2021, an der Hochschule Trier aufgrund von Problemen bei der Stellenbesetzung verspätet in 03/2021. Aufgrund des verspäteten Projekteintritts der Hochschule wurde einem Antrag der Projektpartner auf kostenneutrale Verlängerung zugestimmt. Der geplante Gesamtaufwand von 60 PM wurde mit Genehmigung des Projektträgers um drei Monate überschritten. Das Projekt endete in 05/2023. Die nachfolgende Beschreibung der Arbeiten orientiert sich am Projektstrukturplan, der für das Projekt erstellt wurde.

4.1 Arbeitspaket 1: Projektmanagement

retailolutions GmbH: Die Projektkoordination lag im Verantwortungsbereich der retailolutions GmbH. In der Detailplanung wurden die den Arbeitspaketen zugeordneten Aktivitäten auf die Projektpartner unter Berücksichtigung ihrer Kernkompetenzen verteilt. Für die organisatorische und fachliche Abstimmung wurden wöchentliche Telefonkonferenzen und mehrere Arbeitstreffen am Standort St. Ingbert der retailolutions GmbH durchgeführt. Zudem wurden zwei Status-Meetings unter Beteiligung des Projektträgers durchgeführt. Diese Meetings fanden am 28.01.2022 und am 10.05.2023 statt.

Der Projektfortschritt wurde unter anderem anhand von sechs Meilensteinen überprüft:

1. Meilenstein (Monat 1): Projektplan mit allen Aktivitäten und Zuordnung von Mitarbeitern ist erstellt.
2. Meilenstein (Monat 8): Gelabelte Trainingsdaten und Konzept zur Informationsextraktion aus Produktbildern liegen vor.
3. Meilenstein (Monat 12): Verfahren zur Informationsextraktion aus Produktbildern sind prototypisch implementiert.
4. Meilenstein (Monat 18): Prototyp zur Produktbild-basierten Stammdatenanlage ist erstellt.
5. Meilenstein (Monat 24): Verfahren für eine Produktbild-basierte Stammdatenprüfung, Ergebnisbericht zur Evaluation des Prototyps, aktualisierter Verwertungsplan und Abschlussbericht liegen vor.

Die mit der Projektplanung festgelegten Meilensteine wurden unter Berücksichtigung des verspäteten Projekteintritts der Hochschule Trier und der genehmigten kostenneutralen Verlängerung um zwei Monate alle termingerecht erreicht.

Hochschule Trier: Die Hochschule Trier hat die retailsolutions GmbH bei der Projektkoordination unterstützt. Sie hat die für die Planung und Durchführung der Meetings notwendigen Informationen bereitgestellt und die Ausführung der ihr zugeordneten Aktivitäten geplant und überwacht.

4.2 Arbeitspaket 2: Extraktion produktbezogener Information aus Bildern

Es wurde von beiden Projektpartnern eine umfangreiche Literaturrecherche zum Einsatz von maschinellen Lernverfahren im Produktdatenmanagement insbesondere zur Extraktion von Informationen aus Produktbildern durchgeführt. Die Ergebnisse wurden zusammengeführt. Darüber hinaus wurde die Entwicklungsumgebung, Programmiersprachen und die Verwendung von Softwarebibliotheken abgestimmt. Die Entwicklungsumgebung basiert auf den ML-Bibliotheken Tensorflow, PyTorch und scikit-learn sowie der Programmiersprache Python.

Um die diesem Arbeitspaket zugeordneten Extraktionsverfahren (Aktivitäten 4.2.2 und 4.2.3) voneinander abzugrenzen und in einem späteren Schritt aufeinander abzustimmen, wurde der gesamte Extraktionsprozess modelliert und graphisch dargestellt.

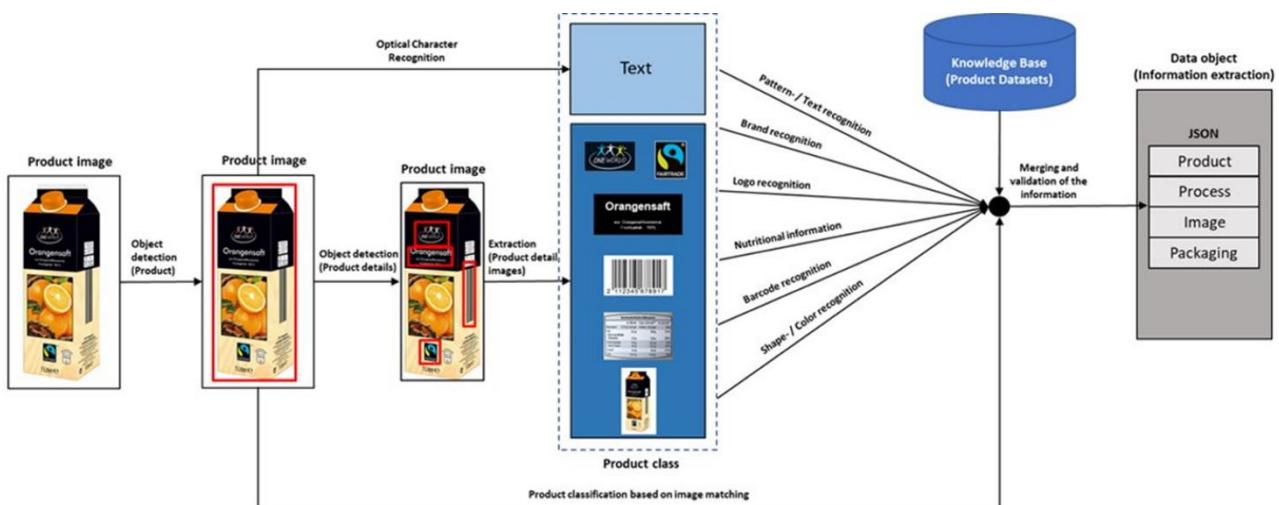


Abbildung 3 Prozess zur Extraktion von Produktinformationen aus einem Produktbild

Die Abbildung zeigt die wesentlichen Schritte ausgehend von einem Produktbild bis zum strukturierten Datenobjekt, das für die Produktdatenanlage in einem ERP- bzw. PIM-System verwendet werden kann.

4.2.1 Definition einer Trainingsmenge für die ausgewählten Sortimentsbereiche

Über eine Literaturrecherche wurden einige Datenmengen mit gelabelten Produktbildern für das Training von maschinellen Lernverfahren identifiziert. Diese Daten wurden hinsichtlich ihrer Eignung für den Einsatz im Forschungsprojekt geprüft. Dabei stellte sich heraus, dass ein Großteil der

öffentliche zugänglichen Daten nicht geeignet war. In den meisten Datenmengen waren Lebensmittelprodukte nur mit einem geringen Anteil enthalten. Außerdem waren oftmals auch Produkte enthalten, die nicht für den deutschen bzw. europäischen Markt bestimmt waren. Die entsprechenden Produktdaten konnten daher nicht verwendet werden, da die auf den Produktverpackungen dargestellten Informationen landespezifisch sind und im Forschungsprojekt ausschließlich Produkte für den deutschsprachigen Markt betrachtet werden sollten. Im Forschungsprojekt wurde daher auf Daten zurückgegriffen, die von den assoziierten Projektpartnern zur Verfügung gestellt wurden. Die Produktdatensätze mit dazugehörigen Bilddaten aus dem Produktsegment Lebensmittel / Getränke / Tabakwaren wurden arbeitsteilig analysiert, dokumentiert, aufbereitet und bereinigt. Die Datensätze hatten folgenden Umfang: (Partner 1 - Globus) ca. 72.000 Bilder zu 37.000 Produkten, (Partner 2 - Transgourmet) ca. 41.000 Bilder zu 20.000 Produkten und (Partner 3 - Coop) ca. 88.000 Bilder zu 9.200 Produkten.

retailsolutions GmbH: Für die nachfolgenden Arbeiten zur Informationsextraktionen wurde entschieden, die Produktdaten von Partner 3 (Coop) zu verwenden. Grund war, dass auf den entsprechenden Produktverpackungen mehrsprachige Beschreibungstexte aufgedruckt waren (deutsch, französisch, italienisch). Die Produktdaten bzw. Bilder wurden gemäß Klassifikationssystem GS1 Global Product Classification (GPC) manuell klassifiziert. Weitergehende Produktinformationen wurden vom Partner zur Verfügung gestellt. Zusätzlich wurden ca. 6.000 Bilder aus der Datenmenge von Partner 1 für die Objekterkennung mit Label versehen. Mit dieser Datenmenge sollten Modelle für die Vorhersage der Verpackungsart und des Materials sowie der Erkennung des Produktnamens auf einem Bild erstellt bzw. trainiert werden.

Hochschule Trier: Für das Training von Verfahren zur Objekterkennung waren die von den assoziierten Projektpartnern bereitgestellten Daten unzureichend gelabelt. Zudem gab es zu einem Produkt oftmals nur ein Frontbild, welches i.d.R. eine niedrige Auflösung aufweist (1000x1000 Pixel). Es wurde daher in Abstimmung mit dem Projektpartner entschieden, eine eigene Bilddatenmenge für das Training von Modellen zur Objekterkennung zu erstellen. Hierfür wurde ein Drehteller entworfen, der in eine Fotobox für die Erstellung der Produktbilder gestellt wird. Auf dem Drehteller wird das Produkt positioniert. Über ein Programm, das den Drehteller und eine Kamera ansteuert, können dann Bilder erstellt werden, die die verschiedenen Produktseiten zeigen. Die so erstellten Produktbilder wurden anschließend für die Objekterkennung gelabelt. Dabei wurden 30 verschiedene Label berücksichtigt, z.B. Produktnamen, Logos, Nährwerttabelle, Füllmenge, Recyclingcode, Barcode, Mindesthaltbarkeitsdatum etc. Die resultierende Produktdatenmenge umfasst 250 Produkte und 1.034 Produktbilder.

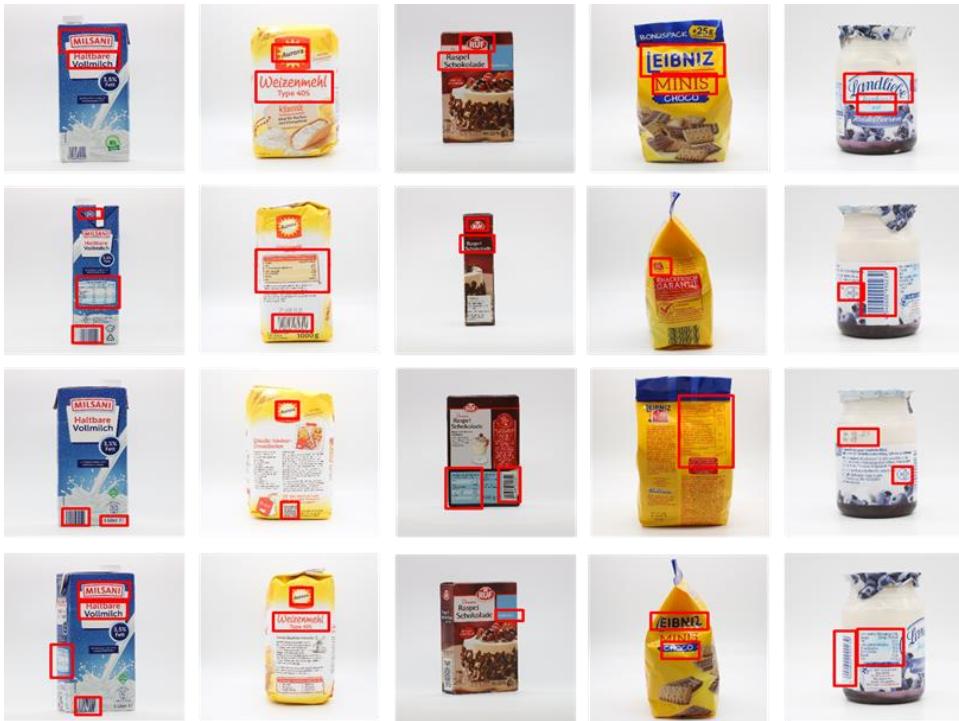


Abbildung 4 Produktbilder mit Label für das Training von Maschinellen Lernverfahren zur Objekterkennung

Je Produkt wurden durchschnittlich 4,14 Bilder erstellt. Je Bild wurden 5 Klassifikationslabel (z.B. Verpackung, Bildtyp etc.) und durchschnittlich 8,14 Label für die Objekterkennung vergeben. Abbildung 4 zeigt Beispiele der erstellten Produktbilder mit den entsprechenden Labeln zur Objekterkennung. Für die Zuordnung der Label wurde anfangs die Software LabelImg und später die Open Source Software Label Studio verwendet. Abbildung 5 zeigt, wie häufig die einzelnen Label vorkommen.

Da die Erstellung und das Labeln von Produktbildern mit erheblichem Aufwand verbunden sind, wurde noch an Verfahren zur Erzeugung von synthetischen Produktbildern gearbeitet. Die Verfahren wurden prototypisch implementiert. Mit dem Prototyp ist es möglich, vollständig gelabelte Produktbilder, zu Produkten, die in Tetra-Paks angeboten werden, zu generieren. Bei der Generierung werden zufällige Informations- und Bildobjekte aus einer Datenbank ausgewählt. Im Ergebnis entstehen dadurch eine Vielzahl von unterschiedlichen Produktbildern. Experimentelle Analysen haben jedoch gezeigt, dass ML-Modelle, die mit den synthetischen Bildern trainiert wurden, wesentlich schlechter waren als die Modelle, die mit Realdaten trainiert wurden. Wir vermuten jedoch, dass das von uns umgesetzte Generierungsverfahren verbessert werden kann. Hierzu sind jedoch weitere Forschungsarbeiten notwendig.

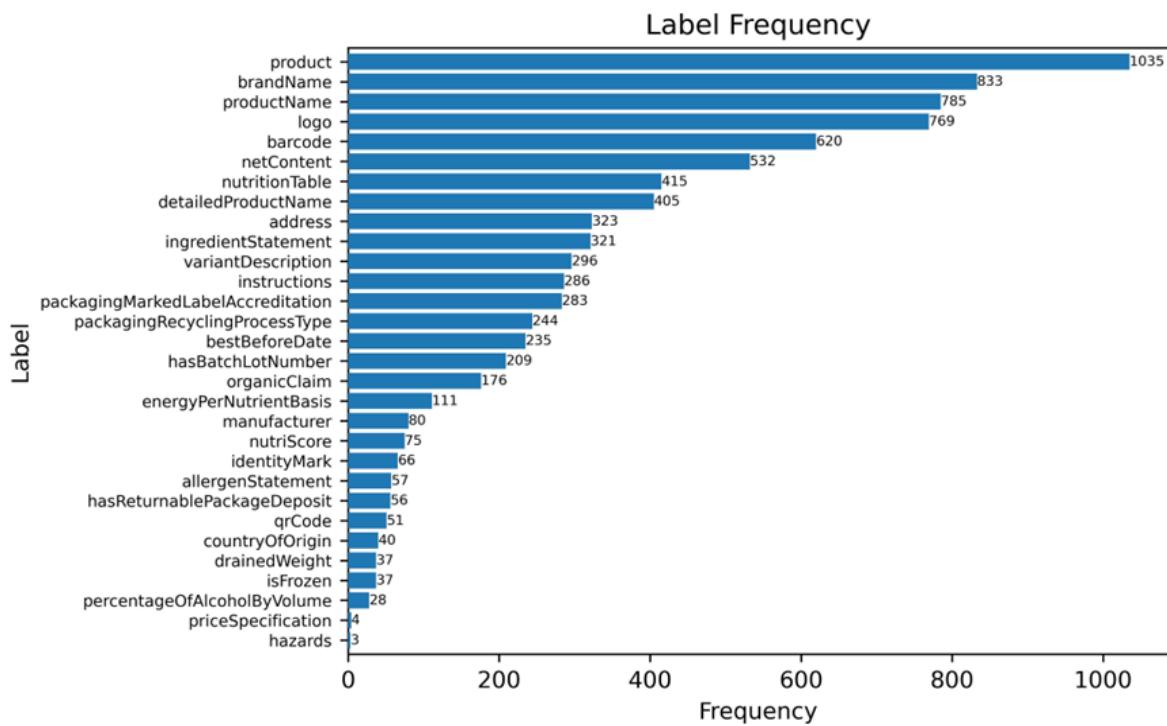


Abbildung 5 Häufigkeit der Label, die den Bildern der Trainingsmenge zugeordnet worden sind.
(n=1.034)

4.2.2 Extraktion von textbasierter Information

Zur Extraktion von textbasierter Information aus Produktbildern wurden Verfahren des Optical Character Recognition (OCR) eingesetzt. Hierfür wurden verschiedene OCR-Bibliotheken geprüft. Zusätzlich wurden Modelle mithilfe der extrahierten Texte trainiert.

retailsolutions GmbH: Für die Extraktion von einzelnen Textbausteinen aus den Produktbildern wurden neben einer eigenen OCR-Implementierung auch die Google Vision API sowie die Amazon Recognition API evaluiert. Aufgrund der Qualität der Ergebnisse wurde im weiteren Verlauf des Vorhabens meistens die Google Vision API für die Textextraktion verwendet. So wurden die Texte aus den von den assoziierten Projektpartner Globus und Coop bereitgestellten Produktbildern mithilfe der Google Vision API ausgelesen und aufbereitet (Umwandlung von Großbuchstaben in Kleinbuchstaben, Eliminierung von Verbindungswörtern, Zahlen und Stopwörtern). Die Texte wurden von beiden Projektpartnern verwendet, um verschiedene Verfahren zur Extraktion textbasierter Informationen zu implementieren und zu testen. In einem Experiment wurden die mehrsprachigen Texte zur Vereinheitlichung mit dem maschinellen Übersetzungstool DeepL, das über ein API angebunden wurde, ins Englische übersetzt. Um die extrahierten unstrukturierten Daten für ML-Verfahren nutzbar zu machen, wurden sie mithilfe der statistischen Verfahren Frequency-Inverse Document Frequency (TF-IDF) und Weighted Log Odds Ratio in Vektoren umgewandelt, die u.a. die Häufigkeiten und Wichtigkeiten einzelner Wörter beinhalten. Diese Vektoren wurden im weiteren Prozess für das Training von eigenen ML-Modellen zur textbasierten Klassifikation verwendet.

Hochschule Trier: Anstatt mit den vollständig extrahierten unstrukturierten Texten zu arbeiten, wurde entschieden die textbasierten Informationsobjekte über den in Abbildung 3 dargestellten Extraktionsprozess auszulesen und zu verarbeiten. In diesem Prozess werden in einem ersten Schritt die interessanten Bereiche auf dem Produktbild mithilfe von Objekterkennung erkannt und klassifiziert (siehe Abschnitt 4.2.3). Entsprechend der Klasse des erkannten Bereichs wird dann entweder ein spezielles Verfahren, wie der NutritionTableReader für Nährwerttabellen, oder eine OCR-Methode, realisiert unter Verwendung der Bibliotheken Tesseract oder EasyOCR, aufgerufen. Bereiche, in denen nur der Text ausgelesen werden musste, sind beispielsweise der Produktnname, oder der Markenname. Bei der Entwicklung des NutritionTableReader wurden auch OCR-Bibliotheken Tesseract bzw. EasyOCR eingesetzt. Für die Evaluation wurden Label priorisiert betrachtet, die möglichst unabhängig von der Produktklasse sind. Beispiele hierfür sind Produktnname, Markenname und Nährwerttabellen.

Ergänzend zu den Arbeiten der retailSolutions GmbH wurde zur Bewertung der verschiedenen OCR-Bibliotheken die OCR-Software Tesseract genauer betrachtet. Diese ermöglicht das Erweitern der vortrainierten Modelle und stellt hierfür auch dokumentierte Schnittstellen bereit. Während der Forschungsarbeit ist jedoch klar geworden, dass der notwendige Aufwand für das manuelle Aufbereiten der Bilder für das erforderliche Training zu groß ist. Im Projekt wurde sich deshalb auf die vortrainierten Modelle von Tesseract und EasyOCR beschränkt.

Anhand der extrahierten vollständigen Texte der Frontbilder wurde ein Markenerkennungs-Dienst entwickelt. Dieser gleicht die Wörter im extrahierten Text mit einer Wissensbasis ab und liefert den Markennamen als Ergebnis. Die Wissensbasis entspricht einem Markenverzeichnis, das 6.399 Marken wie "Leibniz", "Dr. Oetker", "Haribo", "0,5 Original" etc. enthält. Mit diesem Verfahren erzielten wir in Experimenten eine Genauigkeit von 65% (Top-1) bzw. 73% (Top-3). Die Genauigkeit gibt den Prozentsatz der Fälle an, in denen das Modell die korrekte Kategorie oder in diesem Fall die korrekte Marke vorhersagt. Der Dienst ist damit eigenständig bzw. isoliert aktuell nur begrenzt einsetzbar. Er kann jedoch mit anderen Verfahren kombiniert werden, z.B. kann er zur Unterstützung der Markenerkennung mithilfe der Objekterkennung eingesetzt werden.

Mithilfe der extrahierten Texte wurde ein Textklassifikationsmodell basierend auf der Transformer-Architektur mithilfe des Python-Packages HuggingFace trainiert. Das Modell unterscheidet zwischen 312 verschiedenen GPC-Brick Codes im Lebensmittel-Segment. Insgesamt erreicht das Modell in unseren Experimenten einen F1-Score von 76,29%. Zusätzlich wurde analog ein Modell für die Artikelkurzbeschreibungen trainiert, welche wir von den assoziierten Projektpartnern erhalten haben. Dieses weist in unseren Experimenten einen F1-Score von 86,70% auf.

Alle entwickelten Modelle wurden mit den Kooperationspartnern geteilt und sind entsprechend ihrer Performance in die Demonstratoren integriert.

4.2.3 Extraktion von bild-basierter Information (Verpackung, Form etc.)

retailsolutions GmbH: Für die Extraktion von bild-basierter Information wurde auf den von der Hochschule Trier entwickelten Verfahren zur Objekterkennung aufgebaut. Mit diesem Verfahren können wie nachfolgend dargelegt Informationsobjekte auf Produktbildern erkannt und extrahiert werden. Für die Extraktion von Texten wurden wieder OCR-Verfahren eingesetzt. Darüber hinaus wurde auch an der Erkennung von Bildobjekten gearbeitet. Beispielsweise wurde hierzu ein Verfahren entwickelt, mit dem Energieeffizienz-Label erkannt und ausgelesen werden können. Das Verfahren wurde dann in Arbeitspaket 4.5.1 eingesetzt.

Hochschule Trier: Es wurden mehrere Open Source ML-Modelle zur Objekterkennung (u.a. YOLOv5), zur Klassifikation von Produktbildern (u.a. das Residual Neural Network ResNet50) und zur Extraktion der Informationen über OCR-Verfahren (EasyOCR und Tesseract) getestet und bewertet. Zur Veranschaulichung der entwickelten Verfahren wurde ein erster Demonstrator umgesetzt. Für die Objekterkennung wurde das Modell YOLOv5 mithilfe der in Arbeitspaket 2.1 (siehe Abschnitt 4.2.1) erarbeiteten Daten trainiert. Der Trainingsdatensatz hat dabei 1.034 Bilder umfasst, welche mit den für uns interessanten Bereichen annotiert waren. Die Ergebnisse des finalen Trainings sind Abbildung 6 zu entnehmen.

class	labels	P	R	mAP@.5	mAP@.5:95
all	1,665	0.698	0.496	0.521	0.394
address	69	0.450	0.246	0.299	0.189
percentageOfAlcoholByVolume	1	1.000	0.000	0.000	0.000
barcode	133	0.917	0.970	0.982	0.826
bestBeforeDate	40	0.748	0.700	0.712	0.450
brandName	166	0.538	0.536	0.541	0.352
energyPerNutrientBasis	19	0.945	0.737	0.771	0.630
countryOfOrigin	13	1.000	0.000	0.002	0.000
drainedWeight	11	0.884	0.545	0.642	0.536
variantDescription	44	0.385	0.318	0.255	0.132
isFrozen	7	0.282	0.121	0.049	0.040
hazards	1	1.000	0.000	0.000	0.000
identityMark	19	0.886	0.737	0.840	0.747
ingredientStatement	62	0.512	0.387	0.371	0.225
instructions	54	0.198	0.167	0.129	0.084
logo	151	0.694	0.649	0.692	0.525
manufacturer	19	0.392	0.263	0.319	0.248
productName	140	0.572	0.621	0.608	0.344
detailedProductName	85	0.455	0.259	0.259	0.155
nutriScore	22	0.899	0.909	0.912	0.670
netContent	90	0.784	0.789	0.787	0.614
nutritionTable	87	0.821	0.828	0.831	0.702
organicClaim	21	0.640	0.667	0.704	0.551
priceSpecification	1	1.000	0.000	0.497	0.348
product	213	0.877	0.995	0.995	0.954
qrCode	19	0.961	0.737	0.747	0.556
packagingMarkedLabelAccreditation	69	0.668	0.493	0.519	0.403
packagingRecyclingProcessType	55	0.830	0.764	0.798	0.656
hasReturnablePackageDeposit	9	0.878	1.000	0.995	0.657
hasBatchLotNumber	37	0.389	0.324	0.315	0.195
allergenStatement	8	0.323	0.125	0.046	0.029

Abbildung 6 Performance der Modelle zur Objekterkennung abhängig vom betrachteten Label (P: Precision, R: Recall, mAP: mean Average Precision).

Bereits früh im Entwicklungsprozess ist uns aufgefallen, dass unterschiedliche Klassen nach dem Training unterschiedlich erkannt werden können. Dies ergibt sich aus den Spalten P (Precision) und R (Recall). Die unterschiedliche Performance erklärt sich wahrscheinlich zum einen aus der Häufigkeit der Label (siehe Abbildung 5) und zum andern aus der Komplexität der jeweiligen Klasse. Die Komplexität wird bestimmt durch die Ähnlichkeit zu anderen Objekten innerhalb- und außerhalb einer Objektklasse. So besitzt beispielsweise der Barcode eine sehr hohe Ähnlichkeit zwischen allen Objekten innerhalb seiner Klasse und eine sehr niedrige Ähnlichkeit zu Objekten außerhalb seiner Klasse. Dies ist eine ideale Kombination für das Objekterkennungsmodell. Der Produktnamen liegt im Vergleich im Mittelfeld. Bei ihm haben wir eine niedrige Ähnlichkeit zwischen Objekten innerhalb seiner Klasse und eine niedrige Ähnlichkeit zu Objekten anderer Klassen, ausgenommen von der Klasse "Markenname", festgestellt. Die niedrige Ähnlichkeit innerhalb der Klasse erschwert es dem Objekterkennungsmodell zu verallgemeinern. Zusätzlich führt die visuelle Ähnlichkeit zur Klasse der Markennamen häufig zu Verwechslungen der beiden Klassen in den Vorhersagen. Klassen, die anhand der berechneten Metriken schlecht abschneiden, sind die Klassen Manufacturer, Address und DetailedProductName. Sie haben gemeinsam, dass es sich um reine Textklassen handelt. Sie besitzen eine hohe Ähnlichkeit innerhalb der Klasse, aber auch eine hohe Ähnlichkeit zu anderen (Text-) Klassen. Sie besitzen allerdings als reine Textklassen auch keine signifikanten visuellen Merkmale, die sie voneinander oder von anderen Bereichen auf dem Produktbild, welche für uns nicht interessant sind, unterscheiden. Unsere Ergebnisse lassen vermuten, dass diese Textklassen mit Objekterkennungsmodellen nicht mit zufriedenstellender Genauigkeit gefunden werden. In Zukunft könnten hierfür Verfahren der Named-Entity-Recognition (NER) untersucht werden.

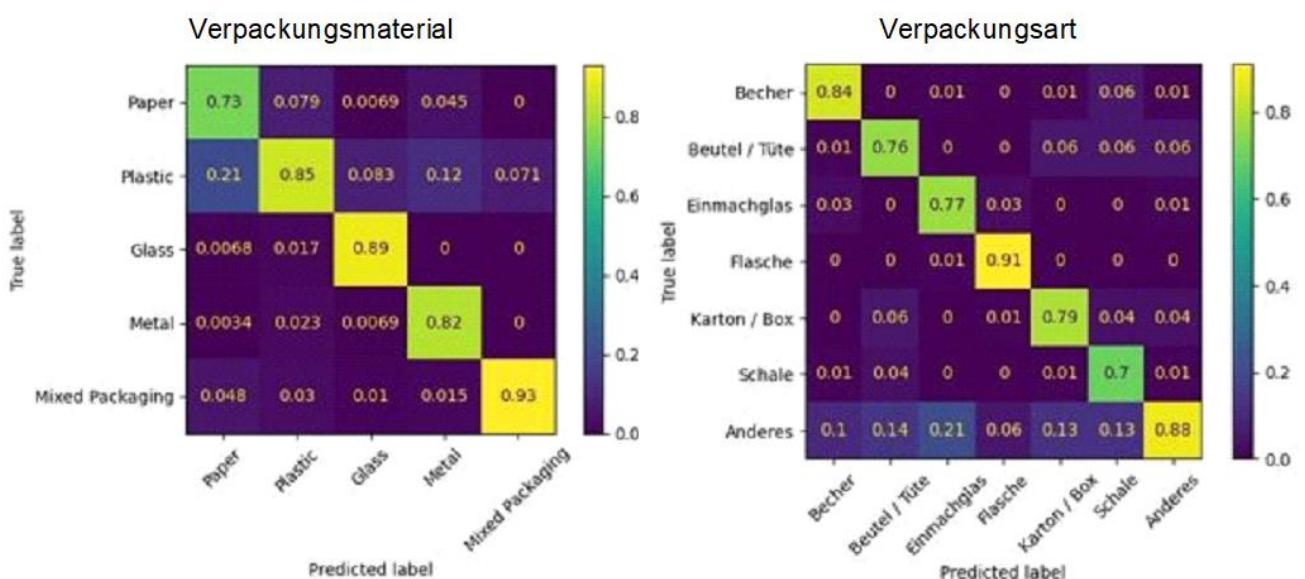


Abbildung 7 Performance der Modelle zur Klassifikation des Verpackungsmaterials und der Verpackungsart.

retailsolutions GmbH und Hochschule Trier: Neben den Modellen zur Objekterkennung wurden auch mehrere Klassifikationsmodelle entwickelt. Um das Verpackungsmaterial und die Verpackungsart zu bestimmen, wurde anhand der Produktbilder aus unserem Datensatz und weiteren Datensätzen, die von den assoziierten Projektpartnern stammen, zwei Modelle trainiert. Die Modelle basieren auf der ResNet-Architektur und wurden jeweils mit ca. 7.000 Bildern trainiert. Die in Abbildung 7 dargestellten Ergebnisse sind vielversprechend. Wir gehen davon aus, dass über eine Erweiterung der Trainingsmenge die Performance der Modelle gesteigert werden kann.

Hochschule Trier: Um die Vertrauenswürdigkeit der konzipierten ML-Verfahren zu bewerten, wurde der Einfluss von verschiedenen Bildeigenschaften auf das Ergebnis der OCR-Verfahren betrachtet. So wurden mehrere Experimente zur Bestimmung des Einflusses von Bildqualität und Bildauflösung auf die Extraktion von Bildinhalten durchgeführt. Dazu wurden unter anderem verschiedene Modelle mit und ohne Skalierungsverfahren zum verlustfreien Vergrößern der Bildauflösung von Produktbildern getestet. Das Ziel hierbei war es herauszufinden, wie groß Bilder sein müssen, um eine gute OCR-Qualität zu erreichen und welche Rolle die Bildschärfe hierin spielt. Die Experimente wurden mit Bildern von Nährwerttabellen, die aus den Produktbildern gewonnen wurden, durchgeführt. Für die Informationsextraktion wurde der von uns entwickelte NutritionTableReader, ein Programm zur Extraktion von Nährwerttabellen, eingesetzt. Abbildung 8 illustriert ein wesentliches Ergebnis.

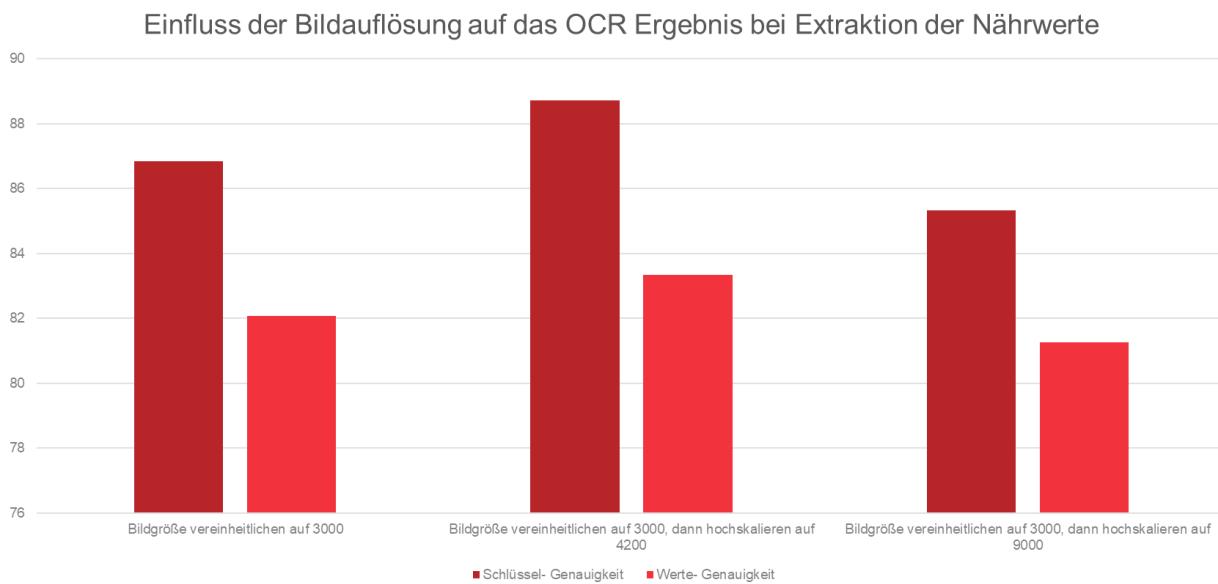


Abbildung 8 Einfluss der Bildauflösung auf das OCR-Ergebnis bei Extraktion der Nährwerte

Aus den Experimenten lassen sich einige Hinweise für die Optimierung der Verfahren zur Informationsextraktion ableiten. So ist ein Hochskalieren von Bildern, die in geringer Auflösung vorliegen, nur bis zu einem gewissen Grad sinnvoll, da bei zu hoher Auflösung die Genauigkeit wieder abnimmt. Dies bedeutet, dass in unseren Experimenten bei Extraktion der Nährwertangaben

mit steigender Auflösung die Fehlerrate zunimmt. Die besten Ergebnisse wurden erzielt, nachdem die Bilder auf 4.200 Pixel hochskaliert wurden.

Ähnliche Untersuchungen wurden im Bereich der Produktklassifikation basierend auf Image Matching durchgeführt. Hier wurde untersucht, welchen Einfluss eine Standardisierung von Seitenverhältnis und Auflösung der Produktbilder auf das Klassifikationsergebnis hat. In der Regel ist es so, dass Produktbilder diesbezüglich Unterschiede aufweisen. Die Experimente haben gezeigt, dass die Standardisierung nur zu einer geringfügigen Verbesserung des Klassifikationsergebnisses führt. Des Weiteren wurde der Einfluss von Farben auf das Ergebnis der Produktklassifikation untersucht.

4.3 Arbeitspaket 3: Gewinnung von Produktinformationen aus Webquellen

4.3.1 Auswahl und Analyse von Webseiten für verschiedene Sortimentsbereiche

Für dieses Arbeitspaket wurden verschiedene Webquellen ausgewählt. Bei den Webquellen handelt es sich um öffentlich zugängliche Produktdatenbanken und Online-Shops von deutschsprachigen Einzelhandelsunternehmen. Aktuelle Ansätze für das Web Scraping wurden recherchiert und bewertet. Ziel war es möglichst generische Verfahren für die Gewinnung von Produktinformationen aus Produktseiten zu entwickeln. Für das Crawling wurden die einzelnen Webquellen detailliert analysiert. Bei Online-Shops wurden folgende Informationen, die für das Crawlen relevant sind, ausgewertet: robots.txt, sitemap.xml, Menge und Umfang der bereitgestellten Produktinformationen etc. Insgesamt wurden während der Projektlaufzeit sechs Online-Shops genauer analysiert. Ebenso wurden 15 Produktdatenbanken vergleichend gegenübergestellt, wobei auf Ergebnissen aus einem vorangegangenen Forschungsprojekt aufgebaut werden konnte. Für den Vergleich wurden die Schnittstellentechnologie, Datenformate, Attributgruppen, Lizenzvereinbarungen etc. betrachtet. Die Ergebnisse wurden in zwei Tabellen zusammengefasst.

retailsolutions GmbH: Neben den ausgewählten Online-Shops wurden zusätzlich offene Datenbanken mit Produktinformationen betrachtet und untersucht, wie diese über Schnittstellen an den Prototypen angebunden werden können. Dies wurde exemplarisch für die OpenFoodFacts-Datenbank (URL de.openfoodfacts.org) umgesetzt. Bei der Datenbank handelt es sich um eine weltweit offen zugängliche Datenbank, in die jeder Produkte mit ihren Merkmalen eintragen kann. Die Daten werden in der dokumentenbasierten Datenbank MongoDB abgelegt. Die Produktinformationsextraktion wurde mit der Hochschule Trier abgestimmt, um die Daten besser zu kombinieren, die aus dem entwickelten Web Scraping-Verfahren gewonnen werden. Zur Evaluation der Daten aus den Webquellen wurden Daten aus der OpenFoodFacts-Datenbank und Daten der Projektpartner verwendet. Der Zugriff auf die OpenFoodFacts-Datenbank erfolgte über das bereitgestellte API.

Hochschule Trier: Für die Entwicklung der generischen Web Crawling und Scraping-Verfahren wurden strukturierte Datenformate untersucht, die für die Einbettung von Produktinformationen in Webseiten eingesetzt werden. Im Wesentlichen sind dies: JSON-LD, Microdata, RDFa, Open Graph Protocol und CSS-Klassen. Für die Extraktion der Produktinformationen wurden zahlreiche darauf angepasste Algorithmen und Heuristiken entwickelt und untersucht. Zusätzlich wurde erforscht, wie die Produktkategorien in den Online-Shops identifiziert werden können. Damit sollte ermöglicht werden, gezielt Informationen über Lebensmittel zu gewinnen, die bei der Entwicklung des IMPRO-Demonstrators fokussiert betrachtet werden. Das resultierende Verfahren wurde anschließend basierend auf dem Framework Scrapy in der Programmiersprache Python umgesetzt und anhand von sechs ausgewählten Webquellen bewertet.

4.3.2 Erarbeitung eines Konzeptes zur Knowledge Representation

retailsolutions GmbH: Das Datenmodell und das API der OpenFoodFacts-Datenbank wurden analysiert. Darauf aufbauend wurde ein Konzept erarbeitet, um die OpenFoodFacts-Datensätze mit dem darin enthaltenen Wissen in einem Datenobjekt abzulegen, das eine Stammdatenanlage in einem ERP-System ermöglicht (siehe Abschnitt 4.4). Hierzu wurde das SAP ERP-Datenmodell mit dem OpenFoodFacts-Datenmodell abgeglichen und ein Mapping erarbeitet, das die Verwendung der OpenFoodFacts-Objekte in SAP ERP ermöglicht.

Hochschule Trier: Für die Speicherung wurde in Abstimmung mit dem Projektpartner die dokumentenorientierte Datenbank (MongoDB) verwendet. Bei der Entwicklung des Datenbankschemas wurde sich am schema.org-Vokabular orientiert. In der Datenbank sind sowohl Informationen zur Steuerung des Crawl-Prozesses als auch die extrahierten Produktinformationen je Webquelle gespeichert. Beispiele hierfür sind: Produktnname, Beschreibung, Preis, Marke, GTIN, Hersteller, Nährwertangaben etc. Zusätzlich sind Informationen hinterlegt, um die Qualität der gespeicherten Daten beurteilen zu können. Z.B. wird zu jedem Datenobjekt ein Vollständigkeitswert angegeben. Außerdem werden bereits bei der Extraktion von Produktinformationen, unzureichend beschriebene Produkte verworfen und nicht in die MongoDB eingefügt. Zu einem Datenobjekt können gemäß dem entwickelten Schema ca. 20 Attributwerte abgespeichert werden.

4.3.3 Entwicklung von Web Crawling & Scraping-Verfahren für die Informationsextraktion

Hochschule Trier: Für die Extraktion von Produktinformationen aus Online-Shops wurde mithilfe der Open-Source Bibliothek Scrapy ein Webcrawler entwickelt, der ausgehend von der Shop-Startseite den gesamten Produktkatalog automatisch aufruft, die geforderten Informationen aus den Webseiten extrahiert und final in der dokumentenorientierten Datenbank MongoDB speichert.

Es wurden sechs Online-Shops angelegt, die in einem Crawl-Prozess berücksichtigt werden können. Das Hinzufügen neuer Webshops oder anderer Webquellen ist nach einem manuellen Ergänzen von spezifischen Informationen zur Steuerung des Crawl-Prozesses möglich.

Über eine schlüsselwortbasierte Analyse der Produktkategorien wird sichergestellt, dass nur Lebensmittelprodukte in der Datenbank gespeichert werden. Dazu wurden verschiedene Heuristiken betrachtet und experimentell analysiert. Weiterhin ist es möglich, einige Attribute als Schlüsselattribute zu kennzeichnen, die zwingend im Produktdatenobjekt vorhanden sein müssen. Damit wird sichergestellt, dass nur vollständige Datenobjekte in der Mongo-DB gespeichert werden.

4.3.4 Evaluation der extrahierten Informationen

retailsolutions GmbH: Die Datenverwaltung der Open Food Facts-Datenbank (de.openfoodfacts.org) wurde untersucht, um Verlässlichkeit und Vertrauenswürdigkeit der Datenbank bewerten zu können. Da es sich um eine Nutzer-geführte Datenbank handelt, wurden Stichproben von Produkten mit unseren Klassifikationsmodellen abgeglichen. Ergebnis dieser Untersuchung war, dass die Datenbank nicht als alleinige Informationsquelle genutzt werden sollte, aber sich durchaus in Kombination mit anderen Modellen zur Informationsanreicherung eignet. Über die Stichproben konnte ein Konzept für ein systematisches Profiling erarbeitet werden, das gemeinsam mit dem Projektpartner umgesetzt wurde.

Hochschule Trier: Um die Vertrauenswürdigkeit und Qualität der Webquellen zu bewerten, wurden verschiedene Analysen durchgeführt. Abbildung 9 zeigt eine Auswertung der durchschnittlichen Vollständigkeit der extrahierten Produktinformationen bzgl. vorgegebener Attribute für ausgewählte Shops. Mit dem Web Crawler können gezielt HTML-Strukturen angefragt werden. Die Inhalte dieser füllen dann die Attribute der Datenstruktur. Die unterschiedlichen Vollständigkeitswerte ergeben sich daher in erster Linie aus der Verfügbarkeit der Informationen auf Seite der Online-Shops.

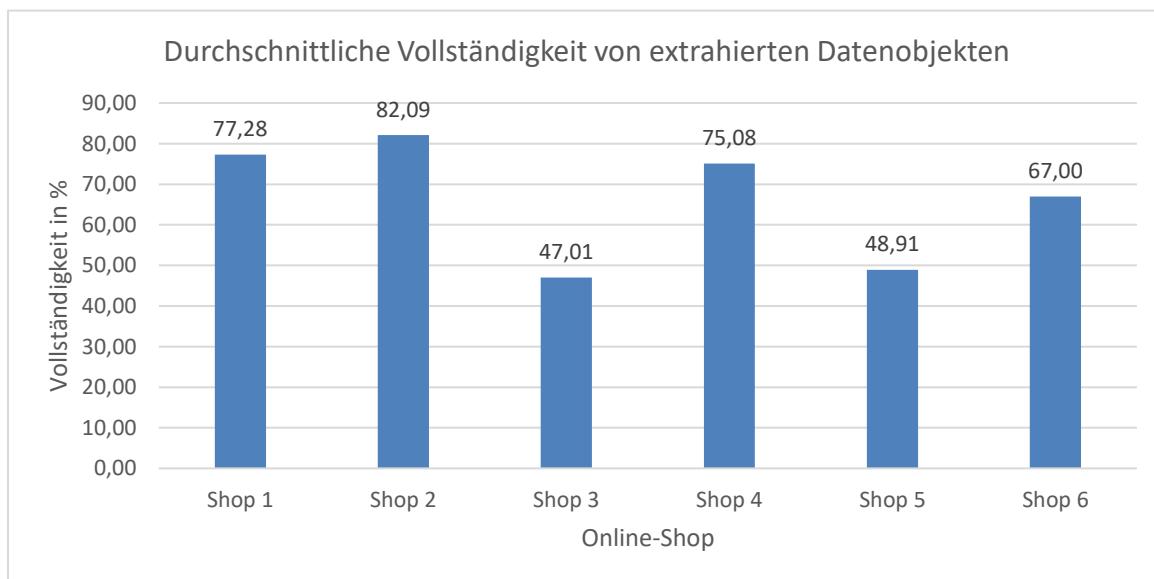


Abbildung 9 Vollständigkeit von extrahierten Datenobjekten für verschiedene Online-Shops.

Darüber hinaus wurde, um die Korrektheit der extrahierten Produktdaten zu gewährleisten, ein statischer Referenzdatensatz erstellt, mit dessen Hilfe nach Code-Änderungen an der Crawler-Komponente die Ergebnisse gegengeprüft werden können.

Zusätzlich wurde in Abstimmung mit dem Projektpartner die Open Food Facts-Datenbank (de.openfoodfacts.org) einem Data Profiling unterzogen, um die Qualität der dort bereitgestellten Daten abschätzen zu können. Über das Data Profiling wurden systematisch Informationen über die Vollständigkeit und die Aktualität der in der Datenbank enthaltenen Produktdaten bestimmt. Die Datenbank enthielt zum Auswertungszeitpunkt im Mai 2022 ca. 300.000 Produkte, die der DACH-Region zugeordnet waren. Dabei wurde beispielsweise festgestellt, dass etwa 45% dieser Produkte vor dem 01.01.2021 angelegt worden sind. Die Datenbank enthält damit wahrscheinlich auch viele Produktinformationen, die nicht mehr aktuell sind bzw. sich auf Produkte beziehen, die nicht mehr hergestellt werden. Außerdem konnten nur ca. 55.000 Produkte dem deutschen Markt zugeordnet werden. Die Qualität der Daten ist für ihre Anwendung im IMPRO-Demonstrator nach diesen Ergebnissen wie bereits dargelegt kritisch zu bewerten.

4.4 Arbeitspaket 4: Intelligente Produktbild-basierte Stammdatenanlage

4.4.1 Produktklassifikation auf Basis der Produktbilder durch Machine Learning

Ein wesentlicher Schritt bei der Erstanlage von Produktdaten ist die Produktklassifikation. Dabei wird ein Produkt einer Produktkategorie zugeordnet. In diesem Arbeitspaket wurden Produktkategorien gemäß dem Global Product Classification Standard (GPC) betrachtet. Der GPC-Standard ist eine international anerkannte vierstufige Hierarchie, die im Lebensmittelbereich in der im Projekt betrachteten Version auf der untersten Ebene 884 verschiedene Kategorien (Bricks) unterscheidet.

retailsolutions GmbH: Aufbauend auf der in Abschnitt 4.2.2 beschriebenen Extraktion von Texten und Umwandlung in Feature-Vektoren wurde das Klassifikationsverfahren Logistic Regression für die verschiedenen GPC-Hierarchien trainiert. Dabei konnte auf einer früheren Arbeit zur Produktklassifikation aufgebaut werden, die von Allweyer et al. im Jahr 2020 auf der 9th International Conference on Data Science, Technology and Applications (DATA 2020) veröffentlicht wurde. Zusätzlich wurden verschiedene Heuristiken zur Bereinigung, Standardisierung und Segmentierung geprüft. Es hat sich gezeigt, dass das Auslassen stark unterrepräsentierter Klassen (< 10 Datenpunkte) und das Auslassen zu kurzer Texte (< 3 Wörter) das Modell signifikant verbessern konnte. Außerdem konnte eine Differenzanalyse der verschiedenen Hierarchiestufen für einen verlässlicheren Konfidenzwert sorgen.

Hochschule Trier: Es wurde ein Maschinelles Lernverfahren erarbeitet, um Produkte anhand von gegebenen Produktbildern den entsprechenden GPC-Produktkategorien zuzuordnen. Das Verfahren basiert auf einem Image Matching-Ansatz, den die nachfolgende Abbildung skizziert.

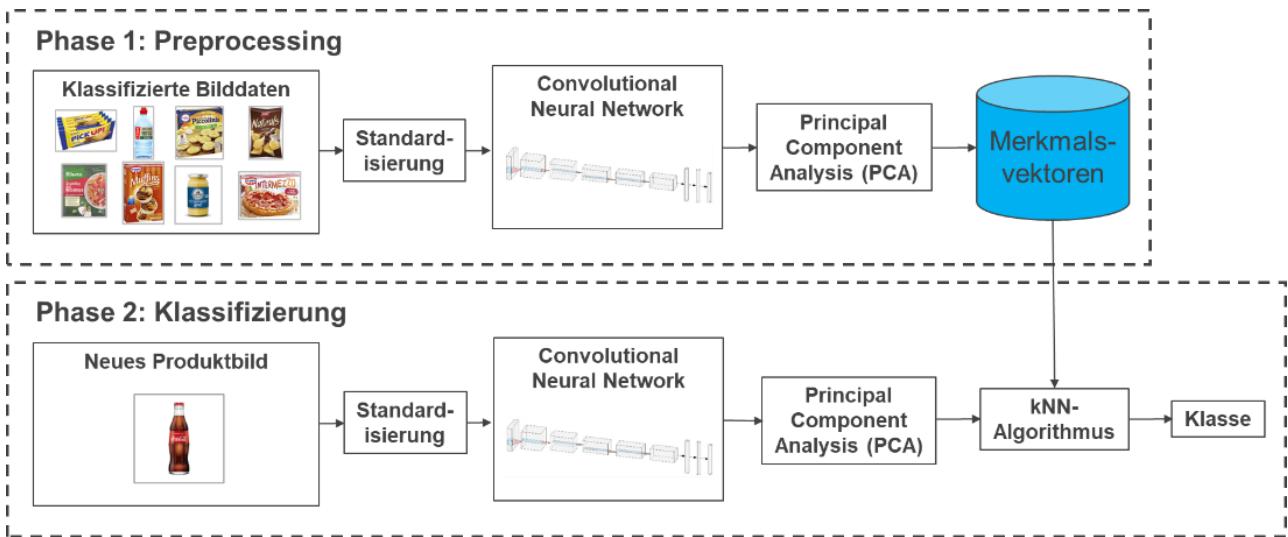


Abbildung 10 Verfahren zur Produktklassifikation gemäß dem Global Product Classification Standard basierend auf Image Matching.

In einer Preprocessing-Phase werden bereits klassifizierte Produktbilder standardisiert. Danach wird zu jedem Bild über ein Convolutional Neural Network (ResNet50) mit anschließender Hauptkomponentenanalyse ein Merkmalsvektor erzeugt und in einer Vektordatenbank abgespeichert. Soll nun ein Produkt anhand eines Bildes klassifiziert werden, wird in analoger Weise ein Merkmalsvektor erzeugt. Zu diesem Vektor werden dann ähnliche Vektoren ermittelt, über die dann die Produktkategorie bestimmt werden kann. Zur Bewertung des Verfahrens wurde eine Vielzahl von Experimenten durchgeführt. Hierfür wurden die von den assoziierten Projektpartnern bereitgestellten Produktbilder verwendet. Die betrachtete Bildmenge umfasste ca. 70.000 Bilder von ca. 37.000 Produkten. In den Experimenten wurden verschiedene Netzwerkarchitekturen, Standardisierungsverfahren und Parameter der Lernverfahren hinsichtlich ihres Einflusses auf das Klassifikationsergebnisse untersucht. Aus den Experimenten wurde die beste Konfiguration ausgewählt (F1-Score 86,6% auf GPC-Brick-Ebene). Die Ergebnisse wurden in einem wissenschaftlichen Paper zusammengefasst und auf einer Konferenz vorgestellt.

retailsolutions GmbH & Hochschule Trier: Es wurde an einem Verfahren gearbeitet, um das Ergebnis der bildbasierten Produktklassifikation mit dem Ergebnis der textbasierten Klassifikation heuristisch zusammenzuführen. Hierbei hat sich gezeigt, dass die Kombination aus text- und bildbasierter Klassifikation zu besseren Klassifikationsergebnissen führt. So konnte beispielsweise in unseren Experimenten der F1-Score um 1 % verbessert werden. In Tests konnte so auf der

feinsten Klassifikationsebene, der Brick-Ebene, ein F1-Score von 88,2 % erreicht werden. Auf der darüberliegenden Class-Ebene wurde ein F1-Score von 95,3 % erreicht.

4.4.2 Entwicklung von Verfahren für das Product Matching; Ermitteln von zusätzlichen ERP-Produktdaten über eine Ähnlichkeitsanalyse

retailsolutions GmbH: Es wurde ein Framework zur Bestimmung ähnlicher Produkte in einer Mehrsystemlandschaft erarbeitet. Bei den Systemen kann es sich um interne als auch externe Datenquellen handeln, über die Produktdaten bereitgestellt werden. Dabei wurde auf dem von der Hochschule Trier bereitgestellten Ergebnissen, die nachfolgend beschrieben sind, aufgebaut.

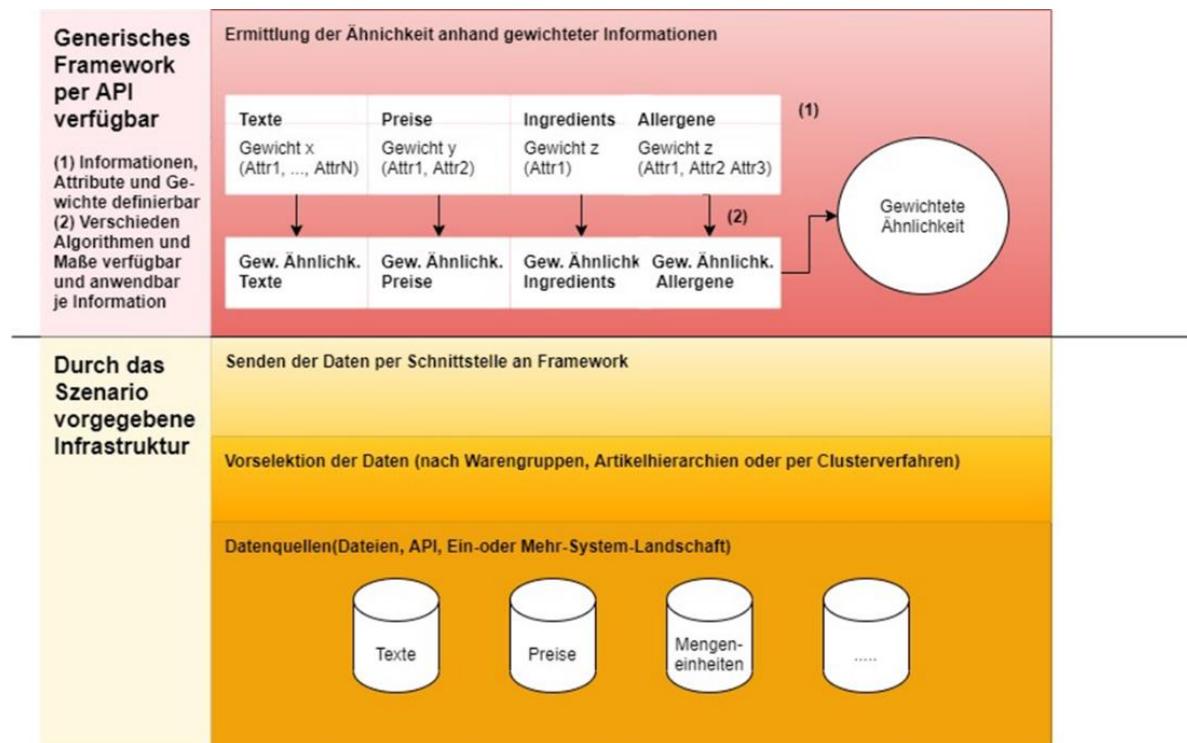


Abbildung 11 Framework zur Bestimmung ähnlicher Produkte in einer Mehrsystemlandschaft.

Der Zugriff auf die einzelnen Datenquellen erfolgt über von den jeweiligen Betreibern öffentlich bereitgestellte APIs. Eine zentrale Datenquelle bildet in diesem Framework das SAP ERP-System mit seinem komplexen Artikelstamm. Eine Vorselektion der Daten anhand von Warengruppen und Artikelhierarchien garantiert eine ausreichende Performance auch in großen Datenbeständen.

Hochschule Trier: Für das Product Matching wurde ein Verfahren entwickelt und untersucht, das auf einer Ähnlichkeitsanalyse vektorieller Produktdaten basiert. Das Verfahren basiert auf einem grundlegenden Ansatz für das Data Matching, der von Peter Christen im Jahr 2012 beschrieben wurde. Die Ähnlichkeit von Produkten wird dabei über die gewichtete Ähnlichkeit von ausgewählten Produktattributen bestimmt, wobei numerische, alphabetische und alphanumerische Werte möglich sind. Bei dem von uns umgesetzten Ansatz können die Gewichte der Attribute abhängig vom

jeweiligen Anwendungsszenario frei gewählt werden. Um das Verfahren zu evaluieren, wurden verschiedene Experimente ausgeführt. Dabei wurde unter anderem das Szenario betrachtet, in dem Datensätze in zwei verschiedenen Datenbeständen zu finden sind, die sich auf das gleiche Produkt beziehen. Der eine Datenbestand umfasste 35.138 Produktdatensätze, der andere 4.635, 876 Datensätze beziehen sich auf das gleiche Produkt. Als Attribute wurde der Produktname, die Produktbeschreibung, der Preis, die Nährwerte, enthaltene Allergene und die Inhaltsstoffe der Lebensmittelprodukte betrachtet. Zeichenketten als Attributwerte wurden bereinigt, standardisiert und segmentiert. Die Gewichte der Attribute wurden experimentell bestimmt. In diesem Anwendungsszenario hat sich ergeben, dass Attribute wie Preise und Nährwertangaben stärker zu gewichten sind als beispielsweise die Produktbeschreibungen, da diese oft variieren.

4.4.3 Entwicklung von Verfahren zur Fusion der Produktinformationen unter Berücksichtigung der Vertrauenswürdigkeit der Informationsquellen

retailsolutions GmbH: Die Fusion von Attributen aus Bildern und Webquellen wurde prototypisch für die Attribute Vegan, Vegetarisch, Bio und den GPC-Baustein in den in Abschnitt 4.5.3 beschriebenen Fiori Wizard integriert. Hierbei wurden die Bild-Informationen mit einer Mischung aus Objekterkennung und Klassifikation gewonnen und die Web-Informationen mithilfe der OpenFoodFacts API. Da weder die Bildverfahren noch die OpenFoodFacts API perfekte Ergebnisse liefern, wurde anstatt einer Gewichtung eine Differenzanalyse durchgeführt, sodass der Nutzer bei widersprüchlichen Informationen eingreifen und den Konflikt auflösen kann.

Hochschule Trier: Es wurden an Verfahren gearbeitet, um die Ergebnisse der textbasierten Klassifikation mit den Ergebnissen der bildbasierten Klassifikation basierend auf Image Matching zusammenzuführen. Dazu wurden mehrere Experimente unter Verwendung der Produktbilder eines assoziierten Partners (siehe Abschnitt 4.4.1) durchgeführt. Unter Verwendung der Ergebnisse wurde ein heuristisches Verfahren nach dem Mehrheitsprinzip (Majority Rule) umgesetzt, um die Produktklasse zu bestimmen. Dadurch konnte die Performance (F1-Score) bezogen auf den zugrundeliegenden Datensatz von 86,6% (nur bildbasiert) auf 88,2% (bild- und textbasiert) gesteigert werden. Mit dem Projektpartner wurde die Schnittstelle abgestimmt, um das Ergebnis in den Prototyp zur Produktdatenanlage zu integrieren.

4.4.4 Entwicklung von Verfahren zur automatischen ERP-Produktanlage unter Berücksichtigung von standort- und funktionsbereichsabhängigen Attributen und User Feedback

retailsolutions GmbH: Für die finale Anlage des Produktdatensatzes im ERP-System wurde ein Wizard konzipiert, der den Benutzer durch den Erstellungsprozess führt und ihm anzeigt, aus welchen Quellen die Vorschlagsdaten stammen. Der Wizard wurde als SAP Fiori Custom Control entwickelt, sodass dieser leicht in bestehende SAP-Applikationen eingebunden werden kann. Da

der Wizard auf eine eigene entwickelte Schnittstelle zugreift, kann IMPRO in jede beliebige App integriert werden. Aktuell wurde der Wizard an ein SAP-Testsystem angebunden.

Um die automatisch erzeugten Werte nachvollziehbar zu machen, wurde eine Informationsbox implementiert, die z.B. Auskunft über erkannte Siegel und abgeleitete Attribute gibt. Außerdem ist es Nutzern möglich, Feedback zu geben. Das Feedback kann verwendet werden, um die Qualität der erzeugten Werte zu überwachen (siehe Abschnitt 4.5.3) und um neue Daten für ein Training zu gewinnen, um die Modelle zu verbessern (siehe Abschnitt 4.5.2).

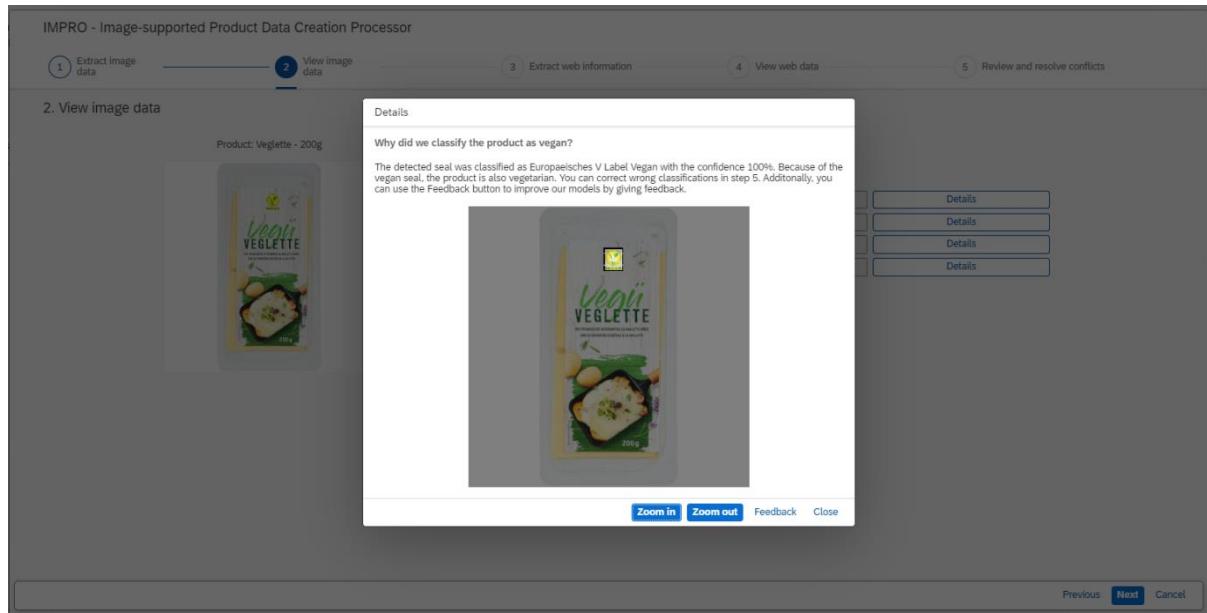


Abbildung 12 Detailansicht Vegan. Der Nutzer kann nachvollziehen, warum das Produkt als vegan identifiziert wurde.

4.5 Arbeitspaket 5: Produktbild-basierte Prüfung von Stammdaten

4.5.1 Differenzanalyse zwischen vorhandenen Produktdaten und den aus dem Produktbild extrahierten Informationen

retailsolutions GmbH: In den in Abschnitt 4.4.4 beschriebenen Fiori Wizard wurde für die in Abschnitt 4.4.3 beschriebene Fusion von Bild- und Webinformationen eine Differenzanalyse integriert. Hierbei wurden die Werte unserer Modelle mit den Werten der Webquelle OpenFoodFacts abgeglichen, sodass Nutzer bei Konflikten informiert werden und gezwungen sind den Konflikt aufzulösen. Zusätzlich wurde eine Problemstellung aus einem anderen Sortimentsbereich betrachtet (siehe Abschnitt 4.2.3). Auslöser war eine neue EU-Verordnung zur Kennzeichnung von Elektrogeräten, die zum 1. März 2021 in Kraft getreten ist. Im Zuge der Verordnung wurden neue Energieeffizienzlabel mit neuer Klasseneinteilung definiert. In Unternehmen ist dadurch oftmals das Problem entstanden, dass die in den PIM-Systemen gepflegten Angaben zur Energieeffizienz nicht mit den Angaben im Online-Shop oder auf der Verpackung übereinstimmten.

Hierzu wurde ein ML-Modell entwickelt, um die Differenzen zu identifizieren. Dabei kamen sowohl Verfahren der Bildklassifikation als auch Verfahren zur Informationsextraktion zum Einsatz. Das resultierende Verfahren kann auch in ähnlicher Weise zur Klassifikation von Bildobjekten im Lebensmittelbereich (z.B. Nutri-Score) eingesetzt werden, um Inkonsistenzen zwischen Bilddaten und tabellarischen Daten aufzudecken.

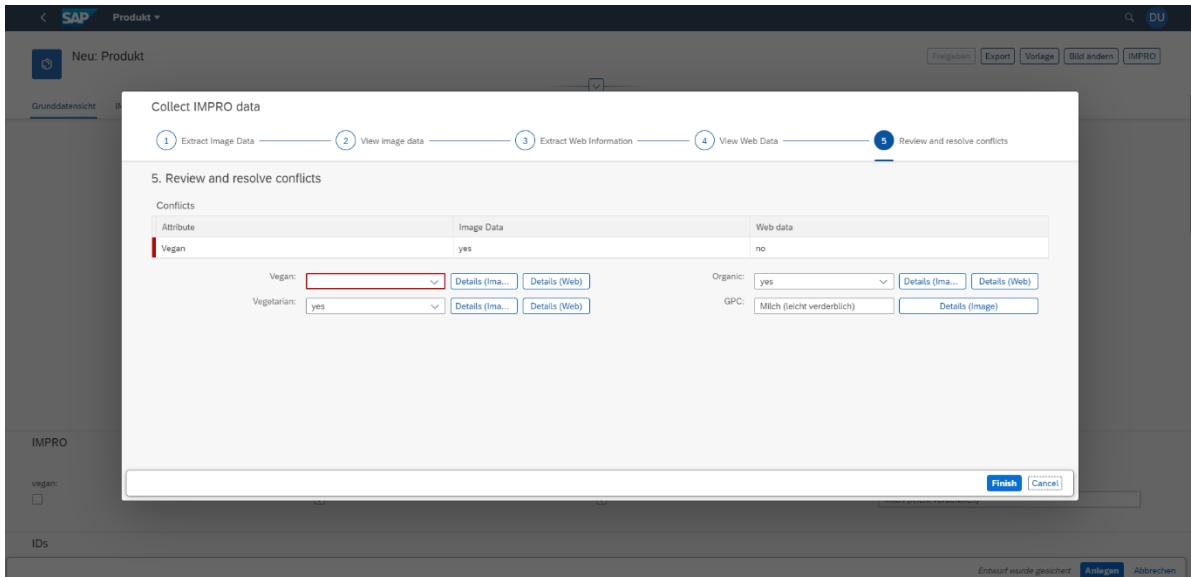


Abbildung 13 Differenzanalyse: Da Bild- und Webquelle widersprüchliche Informationen ergeben, muss der Nutzer den Konflikt auflösen.

Hochschule Trier: Es wurde an verschiedenen Verfahren gearbeitet, um Fehler in tabellarischen Daten durch ihren Vergleich mit aus Bildern extrahierten Daten zu erkennen. Dabei konnte auf den in Arbeitspaket 2 (siehe Abschnitt 4.2) entwickelten Modellen aufgebaut werden. So können beispielsweise über die bildbasierte als auch über die textbasierte Klassifikation vorhandene Produktkategorien überprüft und fehlerhafte Kategorien identifiziert werden. Die Ergebnisse wurden den assoziierten Projektpartnern zur Bewertung zur Verfügung gestellt. Als weitere Problemstellung wurde unter anderem die Überprüfung der Nährwertangaben betrachtet. Hierfür wurde das Verfahren für die Extraktion der Nährwertangaben optimiert und implementiert (NutritionTable-Reader). Anschließend wurde in verschiedenen Experimenten die Qualität des Extraktionsverfahrens evaluiert, in dem die extrahierten Daten mit vorliegenden tabellarischen Daten abgeglichen wurden.

4.5.2 Optimierung durch Active Learning Algorithmen und User Feedback

retailsolutions GmbH: In den in Abschnitt 4.4.4 beschriebenen Fiori Wizard wurden zwei Feedbackverfahren integriert, welche die Nutzung von Active Learning ermöglichen (siehe Abbildung 14 und Abbildung 15).

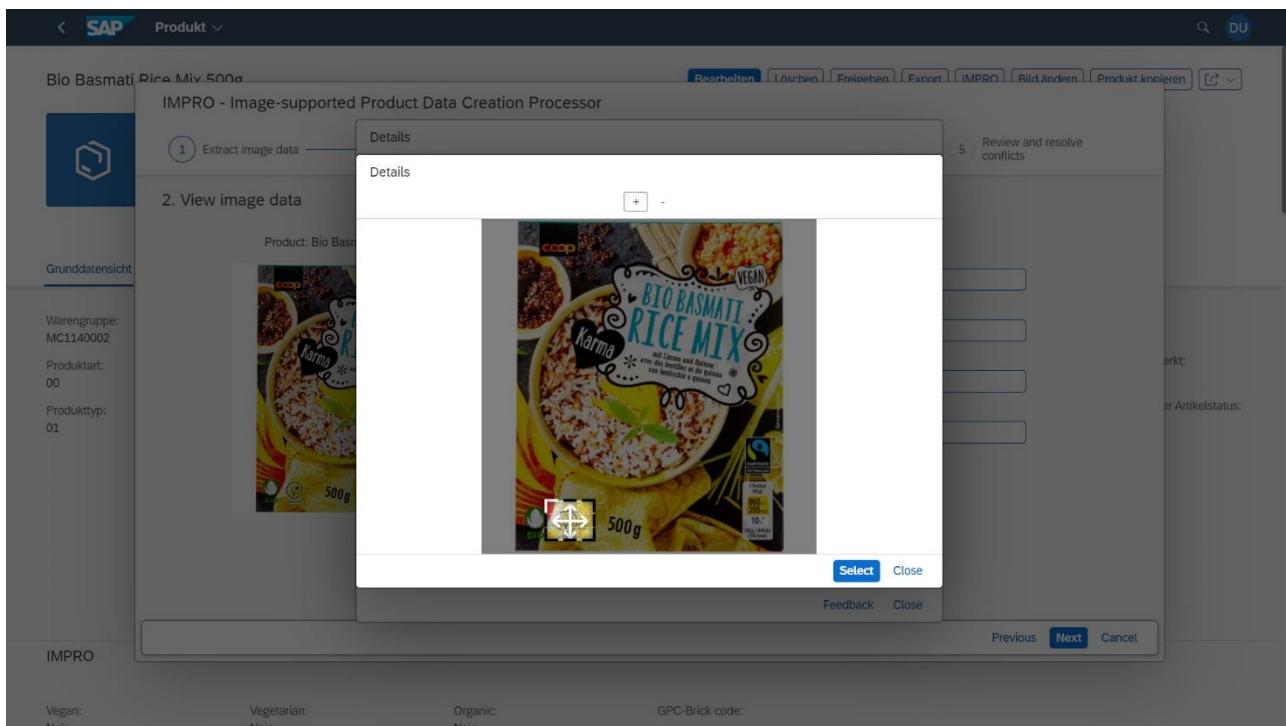


Abbildung 14 Feedback für Objekterkennung: Der Nutzer wählt in einem ersten Schritt die Position des Siegels.

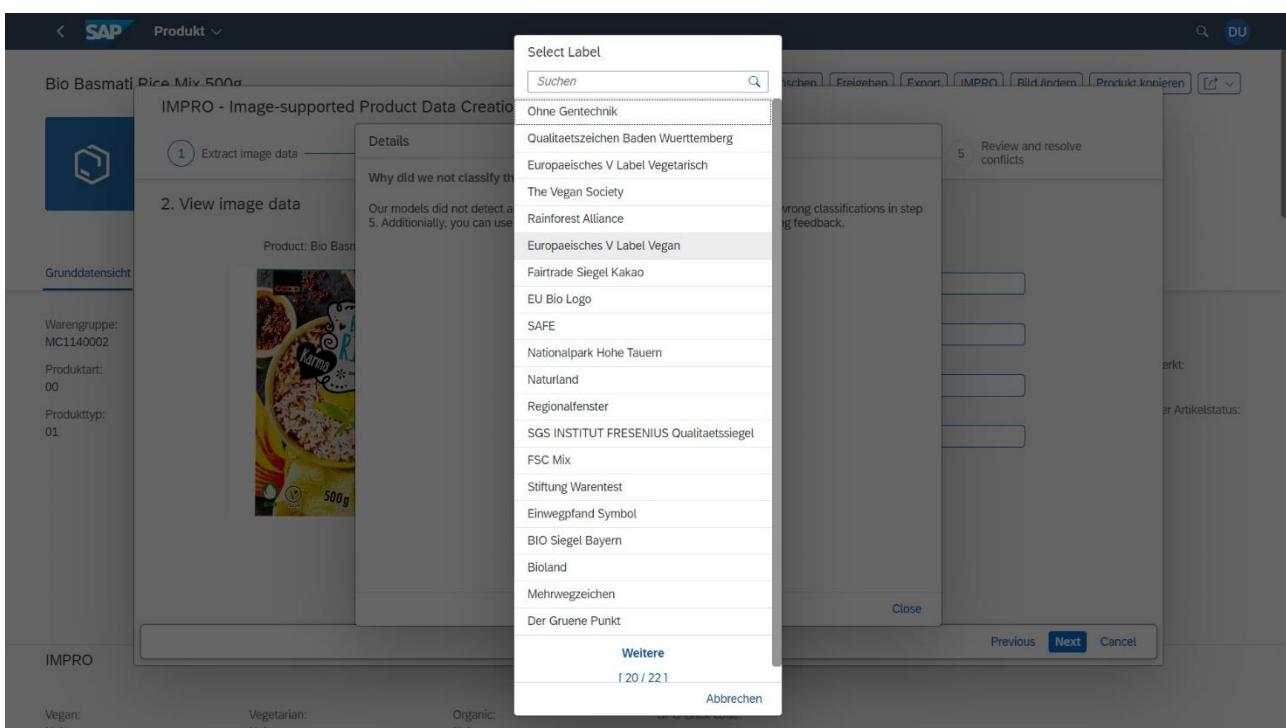


Abbildung 15 Feedback für Objekterkennung: Der Nutzer wählt nach der Position in einem zweiten Schritt die richtige Klasse aus.

Nutzer können bei einem falsch klassifizierten Attribut die richtige Klasse auswählen. Diese Änderung wird automatisch getrackt und in einer Datenbank gespeichert, sodass unser Datensatz um ein neues Produkt mit neuem Bild und mit der durch den Nutzer richtig bestimmten Klasse erweitert wird. Auf den so gewonnenen Daten können die verwendeten Modelle neu trainiert werden, um die Performance zu verbessern und die Modelle aktuell zu halten. Da Objekterkennung solch ein implizites Feedback nicht zulässt (die Verbesserung des Attributes vegan gibt z.B. keinen Aufschluss über die Position des veganen Siegels) wurde für diesen Fall zusätzlich explizites Feedback implementiert, bei dem die Nutzer mithilfe einer beweglichen Bounding Box nicht erkannte Siegel auf dem hochgeladenen Bild markieren und das entsprechende Siegel auswählen können. Somit können auch die Modelle zur Objekterkennung auf dem Feedback trainiert werden.

Hochschule Trier: Der Einsatz von Active Learning wurde im Problemfeld der bildbasierten Produktklassifikation untersucht. Hierzu wurde ein Szenario mit einem Machine Learning Modell betrachtet, welches dem Benutzer (engl. human annotator) ein entsprechendes Label (GPC-Klasse) vorschlägt. Dieses Label wird in diesem Beispiel in Form eines Brick-, Class- oder Family-Codes nach GPC ausgegeben. Diese Klassifizierung wird durch das ML-Modell anhand bereits klassifizierter Daten vorgeschlagen. Der Benutzer kann die vorgeschlagene Klasse korrigieren oder ohne Änderungen bestätigen. Anschließend werden die Daten zu den bereits klassifizierten Produktdaten hinzugefügt und zur Klassifizierung von neuen Daten verwendet. Das Konzept wurde mit einem Datensatz aus zu klassifizierenden Produktbildern überprüft. Die Bilder der Siegel wurden aus den von den assoziierten Partnern bereitgestellten Bildern extrahiert.

4.5.3 Entwicklung eines Monitoring-Cockpits mit User Feedback Integration

retailsolutions GmbH: Mit dem in Abschnitt 4.5.2 beschriebenen User-Feedback, das in unsere Datenbank abgespeichert wurde, können nicht nur unsere Modelle verbessert sondern auch deren Performance überwacht werden. Jedes Attribut, dass vom Nutzer korrigiert wird, entspricht einer falschen Vorhersage, jedes Attribut, das nicht korrigiert wird, einer richtigen Vorhersage. Zudem gibt es den Fall, dass gar keine Vorhersage gemacht werden kann, da beispielsweise kein Siegel erkannt wird oder die verwendeten Webquellen keine Informationen enthalten. Zur Visualisierung dieser Performance-Daten wurde prototypisch ein Monitoring Cockpit entwickelt, in welchem das Feedback und die Qualität der Modelle angezeigt werden. Hierbei ist zum einen die gesamte Performance, aber auch eine Aufschlüsselung von jedem verwendeten Modell und jeder Webquelle angezeigt, sodass Schwachpunkte gegebenenfalls identifiziert werden können (siehe Abbildung 16 und Abbildung 17).

Hochschule Trier: Es wurden Schnittstellen für die von uns entwickelten Verfahren und Modelle mit dem Projektpartner abgestimmt. Der Projektpartner wurde bei der Integration der Verfahren in das Monitoring Cockpit und dem anschließenden Integrationstest unterstützt.

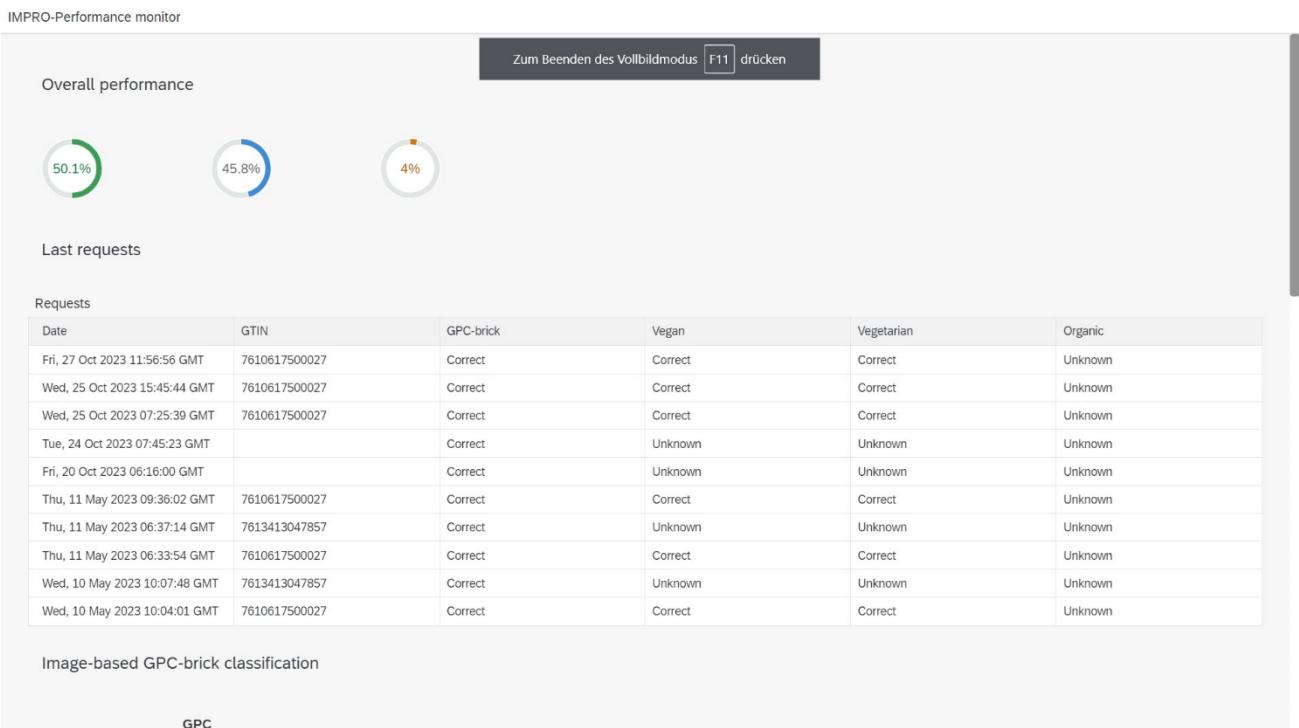


Abbildung 16 Darstellung der Gesamtperformance und Tabelle mit den letzten Anfragen im Monitoring-Cockpit.



Abbildung 17 Darstellung der Aufgeschlüsselte Performance einzelner Modelle und Webquellen im Monitoring Cockpit

4.6 Arbeitspaket 6: Evaluation und Verwertung

4.6.1 Evaluation ausgewählter Funktionen mit Projektpartnern

Für die Evaluation wurden Workshops mit den drei eingebundenen assoziierten Projektpartnern durchgeführt. Die Grundlage bildete der in den vorangegangenen Arbeitspaketen entwickelte Prototyp des Image-supported Data Creation Processors. Ziel war es, gemeinsam mit den assoziierten Projektpartnern im Hinblick auf einen Einsatz in einem produktiven Umfeld die verschiedenen Funktionsbereiche des Prototyps zu analysieren und zu diskutieren.

retailsolutions GmbH: Mit den assoziierten Projektpartnern wurden Termine für die Workshops festgelegt und der jeweilige Teilnehmerkreis abgestimmt. An jedem Workshop haben mehrere Mitarbeiter:innen aus der für das Stammdatenmanagement verantwortlichen Fachabteilung und der IT-Abteilung teilgenommen. Die Workshops waren wie folgt aufgebaut: (1) Kurzvorstellung, (2) Vorführung des IMPRO-Demonstrators, (3) Vorführung des Integrationsszenarios IMPRO/SAP ERP, (4) Mögliche Strategien zur Einbettung von KI-basierten Verfahren in Produktivumgebungen und (5) Vorstellung verschiedener Use Cases (z.B. bildbasierte Klassifikation, Extraktion von Informationsobjekten wie Nährwerte, Abtropfgewicht etc.). Abschließend erfolgte in jedem Workshop eine Bewertung der IMPRO-Funktionalität und eine Diskussion über mögliche weitergehende Anwendungsszenarien. Die Vorführung des Demonstrators wurde auf den jeweiligen Projektpartner abgestimmt. So wurden beispielsweise bei der Objekterkennung oder Produktklassifikation Produktbilder des jeweiligen Partners verwendet. Die Mitarbeiter:innen der Fachabteilungen konnten sich so die Einbettung des Demonstrators in ihre Produktivumgebung besser vorstellen und fundiertes Feedback bei der Evaluation geben.

Hochschule Trier: Die entwickelten Modelle wurden für die einzelnen Workshops optimiert und getestet. Für die Demonstration wurden Beispieldaten der assoziierten Projektpartner ausgewählt. Die Daten wurden so ausgewählt, dass die verschiedenen Verarbeitungsprozesse des IMPRO-Demonstrators vorgeführt werden konnten. Z.B. wurden einerseits Produktbilder verwendet, bei denen die Daten fehlerfrei extrahiert werden konnten. Andererseits wurden Produktbilder verwendet, bei denen die Extraktion fehlerhaft verlief und die Benutzer über die Feedbackfunktionalität zur kontinuierlichen Verbesserung der Modelle beitragen können (Active Learning). Zusätzlich konnte so die Erklärbarkeit der Ergebnisse besser beurteilt werden. In den Workshops wurden die von der Hochschule entwickelten Modelle den Projektpartnern kurz erläutert.

Das Feedback der assoziierten Projektpartner wurde systematisch ausgewertet. Von einem der Projektpartner wurden weitere Daten bereitgestellt, um weitere Anwendungsszenarien zu

untersuchen. Die dabei gewonnenen Ergebnisse wurden in mehreren nachfolgenden Workshops diskutiert.

4.6.2 Konzeption von Exponaten für Messen, Präsentation der Ergebnisse und Vorführung des Prototyps auf Anwenderkonferenzen und Fachtagungen

Im Laufe des Projektes wurden verschiedene Exponate für Messen erstellt:

retailsolutions GmbH: Der in den vorangegangenen Arbeitspaketen entwickelte IMPRO-Prototyp wurde zu einem Forschungsdemonstrator weiterentwickelt, der auf Fachtagungen einem größeren Publikum vorgestellt werden kann. Er entspricht einer browserbasierten Anwendung, die es ermöglicht, ausgewählte Funktionsbereiche von IMPRO mit beliebigen Produktbildern zu präsentieren. Der Schwerpunkt liegt auf der Extraktion der auf den Produktbildern enthaltenen Informationen und der automatischen Produktklassifikation. Über die Integration mit dem in der Wirtschaft weitverbreiteten SAP ERP können realitätsnahe Szenarien dargestellt werden.

Hochschule Trier: Es wurden Jupyter Notebooks zur Evaluation der verschiedenen ML-Modelle, z.B. für die Produktklassifikation, für die Objektdetektierung (z.B. Produktnamen, Nährwerttabelle etc.) etc. entwickelt. Ausgewählte Modelle wurden in den IMPRO-Demonstrator der retailsolutions GmbH integriert. Darüber hinaus wurde ein Live-Demonstrator entwickelt, der an eine Kamera angebunden ist. Damit ist es möglich, ausgewählte IMPRO-Funktionen an realen Produkten, die auf einem Drehteller positioniert werden, vorzuführen. Der Live-Demonstrator wurde am Tag der offenen Tür der Hochschule Trier eingesetzt. Zusätzlich wurden ein wissenschaftliches Poster und ein Rollup Display erstellt, um die wissenschaftlichen Ziele und Ergebnisse des Forschungsprojektes zu beschreiben.

Ausgewählte Forschungsergebnisse wurden auf Anwenderkonferenzen und Fachtagungen vorgestellt. Die Hochschule Trier hat auf folgenden Tagungen vorgetragen:

1. Workshop DTIXperience for researchers and academics, (Veranstalter SAP AG, März 2022)
2. DATA 2022 - 11th International Conference on Data Science, Technology and Applications in Lissabon (Juli 2022)
3. DATA 2023 - 12th International Conference on Data Science, Technology and Applications in Rom (Juli 2023)

Die Projektpartner, die retailsolutions GmbH und die Hochschule Trier, haben zudem gemeinsam die Ergebnisse auf der Fachtagung Innovative Technologien für den Handel (ITH 2022) des Deutschen Forschungszentrums für künstliche Intelligenz im Oktober 2022 vorgestellt. Auf der Fachtagung wurde zusätzlich im Rahmen der Begleitausstellung ein erster Prototyp des Demonstrators vorgestellt. Der Stand wurde gemeinsam von beiden Projektpartnern betreut. Dabei wurde auch das Rollup Display eingesetzt.

4.6.3 Überprüfung und Fortschreibung der Verwertungsmöglichkeiten

retailsolutions GmbH: Zur Überprüfung der wirtschaftlichen Verwertung wurden erste Anwendungsfälle identifiziert, die in den oben genannten Workshops vertieft diskutiert werden sollen. Die Anwendungsfälle beziehen sich auf Klassifikationsprobleme (z.B. Klassifikation nach Warengruppe) und auf das Problem, Unterschiede zwischen Bilddaten und bereits vorhandenen gepflegten Attributen im Informationssystem zu identifizieren (z.B. Energieeffizienzlabels, Gefahrgutkennzeichnung, Inhaltsangaben).

Hochschule Trier: Die Hochschule Trier plant keine wirtschaftliche Verwertung. Eine Verwertung in Lehre und Forschung wurde projektbegleitend geprüft.

4.6.4 Wissenschaftliche Publikationen und Abschlussbericht

Ausgewählte Forschungsergebnisse wurden von der Hochschule Trier bereits projektbegleitend auf wissenschaftlichen Konferenzen vorgestellt und als wissenschaftliche Aufsätze in den Konferenz-Proceedings publiziert (siehe Abschnitt 8).

5. Die wichtigsten Positionen des zahlenmäßigen Nachweises

Der weitaus größte Teil der Zuwendung wurde wie geplant für Personal eingesetzt (über 95%). An der Hochschule Trier wurden zwei wissenschaftliche Mitarbeiter in der Entgeltgruppe E12 bzw. E13 mit Expertise im Bereich Statistik und Maschinelles Lernen im Umfang von 50 % der regelmäßigen wöchentlichen Arbeitszeit und ein studentischer Hilfswissenschaftler über die gesamte Projektlaufzeit beschäftigt. Bei der retailsolutions GmbH waren 0,5 Senior Berater und 1 Application Developer in das Projekt eingebunden. Zudem sind Reisekosten für Arbeitstreffen mit dem Projektpartner sowie für Konferenzbesuche wie folgt entstanden:

- 11.-13. Juli 2022, Teilnahmegebühr und Reisekosten für Konferenz DATA 2022, Lissabon
- 06.-07. Oktober 2022, Teilnahmegebühr und Reisekosten für die DFKI-Fachtagung, Innovative Technologien für den Handel, St. Wendel
- 11.-13. Juli 2023, Teilnahmegebühr und Reisekosten für die Konferenz DATA 2023, Rom
- Regelmäßige Arbeitstreffen der Projektpartner in den Jahren 2021, 2022 und 2023 am Standort der retailsolutions GmbH in St. Ingbert.

Des Weiteren sind bei der retailsolutions GmbH Ausgaben für Server (Clouddienste), Datenbanken (mongoDB) und das Trainieren der Modelle entstanden. An der Hochschule Trier sind Sachausgaben für die Beschaffung von Literatur und von technischen Komponenten (Kamera, Fotobox mit Drehteller etc.) für die Erstellung von Produktbildern entstanden.

6. Fortschritt anderer Stellen während des Vorhabens

Begleitend zur Bearbeitung der Arbeitspakete wurden von beiden Projektpartnern kontinuierlich Literaturrecherchen durchgeführt. Dabei wurden insbesondere die Literaturdatenbanken

scholar.google.com, ieeexplore.ieee.org, dl.acm.org etc. berücksichtigt. Bezuglich Arbeitspaket 2 (siehe Abschnitt 4.2) ist der Übersichtsartikel „Deep Learning for Retail Product Recognition: Challenges and Techniques“ von Wei et al., der im Jahr 2020 veröffentlicht wurde, erwähnenswert. Er gibt einen ausgezeichneten Überblick über den aktuellen Stand der Technik. Zusätzlich wurde über eine Internetrecherche und über eine Auswertung von Studien von Marktforschungsunternehmen (z.B. Gartner Inc.) nach Aktivitäten von Unternehmen im Forschungsgebiet recherchiert. Es wurde festgestellt, dass sich eine Vielzahl von Handelsunternehmen mit Verfahren zur Automatisierung von Filialprozessen beschäftigen (z.B. die Just walk out-Technologie von Amazon, das Intelligent Retail Lab von Walmart und der Checkout-Free Store von Aldi). Dabei kommen auch Verfahren zum Einsatz, um Produkte auf Bildaufnahmen zu identifizieren. Das amerikanische Unternehmen Walmart beschäftigt sich in einer Arbeit „Detecting and Recognising Highly Arbitrary Shaped Texts from Product Images“ von Rajesh Shreedhar Bhat (Walmart Global Tech India, 2020) mit dem Thema der Extraktion von Informationen aus Produktbildern. Die Ergebnisse können helfen, einzelne Komponenten des IMPRO-Demonstrators zu verbessern. Außerdem wurden mehrere Datenmengen mit Produktbildern identifiziert, die für das Training von Maschinellen Lernverfahren eingesetzt werden können. Die Datenmengen haben jedoch den Nachteil, dass sie nur wenige Lebensmittelprodukte enthalten und die Produkte oftmals nicht dem deutschen Markt entstammen.

Zusätzlich beschäftigen sich mehrere amerikanische Tech-Konzerne mit Themen, die einen Bezug zum Forschungsprojekt aufweisen. Erwähnenswert sind die Google Vision API und der Chatbot ChatGPT. Es ist davon auszugehen, dass auf Grund der rasanten Entwicklung im Bereich der KI beide Services zu einer Steigerung der Leistungsfähigkeit vom IMPRO beitragen können. Aufgrund der länderspezifischen Besonderheiten von Produktbeschreibungen sind hier jedoch zur Bewertung der Einsatzmöglichkeiten der Technologie in IMPRO weitere Forschungsarbeiten notwendig.

Nach unserer Recherche bleibt festzustellen, dass das Gesamtkonzept von IMPRO auf dem aktuellen Stand der Technik ist und mit der Spezialisierung auf die komplexen Artikeldaten in ERP-Systemen nicht in anderen am Markt erhältlichen Lösungen in ähnlicher Weise umgesetzt wurde. Es ist jedoch davon auszugehen, dass über die umfangreichen Forschungsaktivitäten im KI/ML-Bereich Erkenntnisse gewonnen werden, die eine Umsetzung und Einführung des IMPRO-Konzeptes zukünftig vereinfachen bzw. unterstützen werden.

7. Voraussichtlicher Nutzen, insbesondere Verwertbarkeit des Ergebnisses

retailsolutions GmbH: Die im Rahmen des Vorhabens entwickelte Funktionalität von IMPRO ist in vielen Anwendungsbereichen von Bedeutung und Nutzen. Die zentrale Komponente soll langfristig nach Projektende in ein marktfähiges Produkt überführt werden, wobei wie geplant zunächst SAP-Systemlandschaften fokussiert werden. Hierbei muss der Prototyp weiterentwickelt werden.

Begleitend wird geprüft, wie sie auch in nicht SAP-geprägten Systemlandschaften in einer serviceorientierten Architektur eingesetzt werden kann. Des Weiteren müssen für eine erfolgreiche Einführung von IMPRO Schulungen konzipiert werden, um den Anwendern das notwendige theoretische Grundlagenwissen für die Nutzung KI/ML-basierter Technologien praxisnah zu vermitteln. Insgesamt wurden die ursprünglichen Ziele erreicht. Die retailsolutions konnte ihr Wissen im Bereich der künstlichen Intelligenz insbesondere im Bereich des Image Processing erweitern und auf die Aufgabenstellungen im Kontext von SAP übertragen. Basierend auf den gewonnenen Ergebnissen und den Erfahrungen im Forschungsprojekt können sich nach Projektende zudem neue Dienstleistungen entwickeln, die von retailsolutions GmbH zukünftig angeboten werden können. Auch öffentlich wird die retailsolutions durch das KI-Forschungsprojekt mehr wahrgenommen.

Hochschule Trier: Die Hochschule plant keine wirtschaftliche Verwertung der gewonnenen Ergebnisse. Ausgewählte Ergebnisse sind jedoch bereits projektbegleitend in Lehrveranstaltungen eingeflossen. So hatten Studierende über das Forschungsprojekt die Möglichkeit, Projekt- und Abschlussarbeiten in einem innovativen Umfeld zu erstellen und sich für eine anspruchsvolle Tätigkeit im FuE-Bereich eines Unternehmens zu qualifizieren. Es wurden u.a. sieben Projektarbeiten, vier Bachelorarbeiten und eine Masterarbeit erstellt, in denen Themen mit Bezug zum Forschungsprojekt bearbeitet wurden. Über die Lehre und die Absolventen wird so das erworbene Wissen in die Unternehmen transportiert. Für die beteiligten Mitarbeiter besteht zudem die Möglichkeit, das erworbene Wissen in anderen Forschungsprojekten oder in Promotionsvorhaben zu verwerten. Ausgewählte Ergebnisse wurden auf wissenschaftlichen Tagungen vorgestellt. Die Publikationen haben den Bekanntheitsgrad der Hochschule als Forschungspartner im Bereich KI & Retail für andere Hochschulen und für Unternehmen gesteigert. Das gewonnene Wissen und die gewonnenen Erfahrungen in der Umsetzung von Forschungsprojekten im KI/Retail-Bereich kommen damit der Hochschule in vielfältiger Weise zugute.

Aufgrund der in diesem und anderen Projekten aufgebauten Kompetenzen hat das Institut zwischenzeitlich weitere Forschungsvorhaben mit KI-Bezug erfolgreich beantragt. Das Projekt IMPRO hat das Potenzial des Maschinellen Lernens zur Unterstützung der Produktdatenanlage in Informationssystemen des Handels deutlich gemacht. In einem weiteren Projekt, das in das Förderprogramm KI4KMU aufgenommen wurde, wollen die beiden Projektpartner gemeinsam untersuchen, inwieweit die Erklärbarkeit und Leistungsfähigkeit von maschinellen Lernverfahren verbessert werden kann, in dem sie mit regelbasierten Verfahren der KI zu sogenannten hybriden Verfahren kombiniert werden. Wesentlich sind dabei Wissensgraphen bzw. Produktgraphen, in denen Faktenwissen über Produkte verwaltet wird. In diesem Forschungsprojekt soll eine KI-basierte Plattform zur automatischen Extraktion von Produktinformationen und zur Erstellung und Exploitation von Produktgraphen für das Stammdatenmanagement im Einzelhandel entstehen.

8. Erfolgte und geplante Veröffentlichungen der Ergebnisse

Ein Teil der im Forschungsprojekt gewonnenen Ergebnisse wurde projektbegleitend publiziert und auf Konferenzen vorgestellt:

Schriftliche Publikationen bzw. Konferenzbeiträge

[Bast 2022] Bast, S., Brosch, C., Krieger, R.: "A Hybrid Approach for Product Classification based on Image and Text Matching", Proceedings of 11th International on Data Science, Technology and Applications (DATA 2022), Lissabon, Portugal, Juli 2022

[Bast 2023] Bast, S., Brosch, C., Krieger, R.: "Combining Image and Text Matching for Product Classification in Retail", Data Management Technologies and Applications, Springer book of DATA 2022 Revised Selected Papers, 2023, pp. 134-153

[Brosch 2023] Brosch, C., Bouwens, A., Bast, S., Haab, S., Krieger, R.: Creation and Evaluation of a Food Product Image Dataset for Product Property Extraction, Proceedings of 12th International on Data Science, Technology and Applications (DATA 2023), Rome, Juli 2023

Vorträge auf Fachtagungen und Messen

1. Krieger, Rolf: Image-supported Product data creation processor, Workshop DTIXperience for researchers and academics, SAP AG, St. Ingbert, März 2022
2. Mitarbeiter:innen der Hochschule Trier: Vorstellung des Demonstrators für Schüler- und Besuchergruppen bei Workshops der MINT-Nationalparkregion Hunsrück-Hochwald plus, Mai 2022
3. Mitarbeiter:innen der Hochschule Trier: Vorstellung des Demonstrators am Tag der offenen Tür der Hochschule Trier am Umwelt-Campus Birkenfeld, Juni 2022
4. Mohr, Andreas und Krieger, Rolf: Automatisierte bildgestützte Produktdatenanlage in Informationssystemen des Handels, Tagung „Innovative Technologien für den Handel“, Veranstalter Deutsches Forschungszentrum für Künstliche Intelligenz, St. Wendel, Oktober 2022
5. Mitarbeiter der retailsolutions GmbH: Vorstellung des Projekts auf dem International Retail Forum 2022, Frankfurt, November 2022

9. Fazit

Die Zusammenarbeit der Projektpartner über die gesamte Projektlaufzeit war durchweg positiv. Die regelmäßigen Projekttreffen förderten den Austausch über Arbeitsergebnisse und praktische Erfahrungen. Das Management durch die retailsolutions GmbH hat die Zusammenarbeit unterstützt. Die Arbeitspakete konnten wie geplant bearbeitet werden. Das Vorhaben hat das Potenzial von Verfahren der Künstlichen Intelligenz zur Unterstützung der Prozesse im Stammdatenmanagement von Handelsunternehmen deutlich gemacht. Die notwendige Erstellung und Aufbereitung von

qualitativ hochwertigen Daten für das Training der maschinellen Lernverfahren zur Informationsextraktion aus Produktbildern war mit einem erheblichen Aufwand verbunden. Die Trainingsdaten wurden zur Unterstützung anderer Forschungsarbeiten veröffentlicht. Ausgewählte Projektergebnisse konnten bereits während der Projektlaufzeit auf wissenschaftlichen Tagungen vorgestellt werden. Die Forschungsarbeiten haben gezeigt, dass die Entwicklung insbesondere von vertrauenswürdigen KI/ML-Anwendungen und ihre Einbettung in betriebliche Abläufe sehr komplex ist und weiterer Forschungsbedarf besteht. Dabei kann auf den gewonnenen Ergebnissen aufgebaut werden. Insgesamt bleibt festzuhalten, dass KI-basierte Verfahren erheblich zur Effizienz- und Qualitätssteigerung im Stammdatenmanagement beitragen können. Ausgewählte Funktionen des IMPRO-Prototyps (z.B. zur bild- und textbasierte Produktklassifikation) sollen in der anstehenden Verwertungsphase ausgebaut werden, um die speziellen Anforderungen eines Produktbetriebs zu erfüllen.

Kurzbericht

Fördervorhaben: KMU-innovativ - Verbundprojekt: IMPRO - Automatisierte bildgestützte Produktdatenanlage in Informationssystemen des Handels

Geplante Projektlaufzeit:	01.01.2021- 31.12.2022
Projektlaufzeit mit genehmigter kostenneutraler Verlängerung:	01.01.2021- 31.05.2023
Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS20085 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren Andreas Mohr, retailsolutions GmbH und Prof. Dr. Rolf Krieger, Hochschule Trier.	

Projektpartner

1. retailsolutions GmbH, St. Ingbert (Verbundkoordinator)
Kontakt: Andreas Mohr, andreas.mohr@retailsolutions.de
2. Hochschule Trier, Standort Birkenfeld, Institut für Softwaresysteme
Kontakt: Prof. Dr. Rolf Krieger, r.krieger@umwelt-campus.de

Aufgabenstellung

Handelsunternehmen verfolgen heute vielfach eine Omni-Channel-Retailing-Strategie, die alle Absatzkanäle wie Ladengeschäfte, Online-Shops oder Kataloge gleichermaßen unterstützt. Sie erfordert, dass alle Absatzkanäle parallel bedient, gesteuert, ausgewertet und optimiert werden müssen. Die zu verwaltende Informationsmenge je Produkt hat sich dadurch in den letzten Jahren erheblich erhöht. Dies stellt hohe Anforderungen sowohl an das Produktinformationsmanagement als auch an das Stammdatenmanagement im zentralen ERP-System. Durch den E-Commerce haben gerade Produktbilder zunehmend an Bedeutung gewonnen. Für die Unternehmen bedeutet dies, dass die Produktdaten sehr umfangreich und heterogen sind. Vielfach müssen die Daten noch sehr kostenintensiv und fehleranfällig manuell erfasst werden. Für die Unternehmen ist es daher schwierig sicherzustellen, dass die Produktdaten und die von ihnen repräsentierten Informationen, die aus unterschiedlichen Quellen stammen, korrekt sind. Der hierfür erforderliche Abgleich der auf dem Produktbild enthaltenen Informationen mit den bereitgestellten Produktdatensätzen erfolgt, wenn überhaupt, manuell und ist entsprechend aufwändig.

Ziel des Forschungsvorhaben war es, den Erstellungsprozess von Produktdaten durch die Einbeziehung von Produktbildern und weiterer Informationen aus Webquellen oder internen Datenquellen mittels eines Image supported Product Data Creation Processors (IMPRO) durch den Einsatz von Verfahren des maschinellen Lernens zu automatisieren. Mit dem Prozessor sollen die Kosten insbesondere für die Erfassung von Produktdaten reduziert und die Effizienz und Qualität der erforderlichen Pflegeprozesse erhöht werden.

Ablauf des Vorhabens

Der für das Forschungsvorhaben ausgearbeitete Projektstrukturplan bestand aus sechs Arbeitspaketen: AP1-Projektmanagement, AP2-Extraktion von produktbezogener Information aus Bildern, AP3-Gewinnung von Produktinformationen aus Webquellen, AP4-Intelligente Produktbild-basierte Stammdatenanlage, AP5-Produktbild-basierte Prüfung von Stammdaten und AP6-Evaluation und Verwertung. Die Projektkoordination lag im Verantwortungsbereich der retailsolutions GmbH. In der Detailplanung wurden die den Arbeitspaketen zugeordneten Aktivitäten auf die Projektpartner unter Berücksichtigung ihrer Kernkompetenzen verteilt. Die für die Entwicklung der maschinellen

Lernverfahren notwendigen tabellarischen Daten und Bilddaten wurden von Handelsunternehmen, die als assoziierte Projektpartner (Globus SB-Warenhaus Holding GmbH & Co. KG, Coop Genossenschaft, Transgourmet Deutschland GmbH & Co. OHG) eingebunden waren, bereitgestellt. Die Daten wurden analysiert und für die Modellentwicklung aufbereitet. Für die Erforschung der Verfahren zur Informationsextraktion aus Produktbildern wurden zudem eigene Produktbilder erstellt und gelabelt. Darüber hinaus wurde für weitere Untersuchungen auf Produktdaten aus verschiedenen Webquellen zurückgegriffen. Die entwickelten Modelle und Verfahren zur Extraktion, Fusion und Bewertung von Produktinformationen wurden detailliert untersucht, verglichen und optimiert. Für die Evaluation wurde zudem ein Prototyp entwickelt.

In allen Arbeitspaketen wurde an den aktuellen wissenschaftlich-technischen Stand angeknüpft. So wurden beispielsweise bei der Entwicklung der maschinellen Lernverfahren auf aktuellen Arbeiten aufgebaut. Es wurden zur Objekterkennung YOLOv5, zur Bildklassifikation RestNet50 und zur Textklassifikation das Open Source Framework BERT eingesetzt. Für die Implementierung der OCR-Verfahren wurden aktuelle Funktionen von Google Vision AI und easyOCR verwendet. Zusätzlich wurden dem Stand der Technik entsprechende Standards wie der Global Product Classification Standard (GPC) und der GS1 Product Image Specification Standard berücksichtigt.

Wesentliche Ergebnisse

Basierend auf tabellarischen Produktdaten und Bilddaten wurden eine Vielzahl von maschinellen Lernfahrungen entwickelt, um Produktinformationen aus verschiedenen Datenobjekten zu extrahieren und in einem zentralen Datensatz für die Stammdatenanlage in einem ERP-System bereitzustellen. Für die Extraktion von Informationen wie z.B. Produktnamen, Hersteller, Marke, Nährwerttabelle, Logos und Inhaltsstoffe etc. aus Produktbildern wurden Verfahren zur Objekterkennung erforscht. Notwendige Produktbilder wurden erstellt und gelabelt. Verschiedene Experimente haben die Anwendbarkeit des auf YOLOv5 basierenden Verfahrens zur Objekterkennung bestätigt. Insbesondere bei Informationsobjekten, die immer ähnlich strukturiert sind, waren die Ergebnisse sehr vielversprechend. Für die Extraktion von Informationen aus den erkannten Objekten wurden bei textuellen Daten Object Character Recognition-Verfahren eingesetzt. Darüber hinaus wurden weitere ML-Modelle für die Erkennung und Klassifizierung von bildbasierten Informationsobjekten (z.B. Bio-Logos, Qualitätssiegel etc.) untersucht und experimentell basierend auf den von den assoziierten Projektpartnern bereitgestellten Daten evaluiert. Das entwickelte Verfahren für die Stammdatenanlage fusioniert die verschiedenen Informationen zu einem Datensatz, informiert den Benutzer über Konflikte und ermöglicht dem Benutzer über eine Feedback-Funktion die Modelle zu optimieren. Bei der Fusion werden auch Daten aus Webquellen berücksichtigt, die ebenfalls automatisch extrahiert werden. Die Ergebnisse und der entwickelte Prototyp wurde den assoziierten Projektpartnern zur Evaluation in Workshops mit Endanwendern vorgestellt. Die Ergebnisse wurden unter Berücksichtigung der besonderen Anforderungen von produktiven Systemumgebungen diskutiert. Insgesamt wurden die Ergebnisse sehr positiv bewertet. Es wurden zusätzliche Anwendungsmöglichkeiten identifiziert, die in einer Verwertungsphase geprüft werden sollen. Ausgewählte Ergebnisse wurden bereits auf wissenschaftlichen Tagungen vorgestellt. Die Trainingsdaten für die Objekterkennung und dazugehörende Modelle wurden ausführlich dokumentiert und wurden anderen Forschern zum Download zur Verfügung gestellt.