






# Automatisierte Erkennung von Störstoffen in Bioabfall mit maschinellem Lernen

## Ansätze und Ergebnisse aus dem Projekt TRACES




Marius Sangel <sup>1</sup>, Emilia Bensch <sup>1</sup>, Hans Brandt-Pook <sup>1</sup>, Timo Röllke <sup>2</sup> und Cedric Markworth <sup>2</sup>



**Abstract:** Die automatisierte Erkennung von Störstoffen in Bioabfall gewinnt zunehmend an Bedeutung, sowohl zur Effizienzsteigerung von Recyclingprozessen und Verbesserung der Kompostqualität als auch durch die Novelle der Bioabfallverordnung. Diese Arbeit entwickelt ein Modell zur Instance-Segmentierung von Störstoffen in einem Müllhaufen, das auf dem YOLACT-Framework basiert, um einzelne Objekte im Abfall präzise zu identifizieren. Dabei wird das Modell auf Bilddaten trainiert, die durch eigene Aufnahmen von Bioabfall erstellt wurden. Ziel ist es, die Erkennung von Störstoffen zu optimieren und einen Beitrag zur Erfüllung der neuen gesetzlichen Anforderungen zu leisten. Durch den Einsatz von Data Augmentation wird das Modell robust gegenüber variierenden Abfallarten, was die Flexibilität und Genauigkeit bei der Klassifikation von Störstoffen im Bioabfall verbessert. Die Ergebnisse zeigen, dass die Anwendung von Instance-Segmentierung eine vielversprechende Methode zur Automatisierung und Verbesserung der Evaluierung des Kontaminierungsgrades sorgt.

**Keywords:** Machine Learning, Computer Vision, Instance Segmentation, CNNs, YOLACT, Data Augmentation, Waste Classification, Trash Detection, Biowaste Analysis

## 1 Einleitung

Verunreinigungen im Biomüll stellen ein zentrales Problem für die biologische Weiterverarbeitung dar. Störstoffe wie Plastik, Glas usw. führen zu Mehraufwand, Kosten und Qualitätseinbußen in Kompostierungsanlagen. Die novellierte Bioabfallverordnung (BioAbfV) schreibt seit 2022 einen maximalen Störstoffanteil vor [Bu25], was automatisierte Verfahren zur frühzeitigen Erkennung notwendig macht. Moderne Verfahren der Bildverarbeitung und KI-gestützten Segmentierung bieten vielversprechende Ansätze zur Detektion solcher Störstoffe. Besonders effizient sind Modelle wie YOLACT (You Only Look At CoefficientS) zur Instanz-Segmentierung [BZ+19] im Vergleich zu klassischen, aber langsameren Ansätzen wie dem Mask R-CNN (Region-Based Convolutional Neural Network)[HZ+16].

<sup>1</sup> HSBI, Fachbereich Wirtschaft, Interaktion 1, 33619 Bielefeld, DE,  
maris.sangel@hsbi.de,  <https://orcid.org/0009-0007-9393-3625>;  
emilia.bensch@hsbi.de,  <https://orcid.org/0009-0009-9962-1918>;  
hans.brandt-pook@hsbi.de,  <https://orcid.org/0009-0002-6668-2684>

<sup>2</sup> c-trace GmbH, Stieghorster Str. 112, 33605 Bielefeld, DE,  
t.roellke@c-trace.de,  <https://orcid.org/0009-0000-1152-2283>;  
c.markworth@c-trace.de,  <https://orcid.org/0009-0001-4476-6361>

Nach der Darstellung der Problemstellung verfolgt dieser Beitrag den Aufbau einer automatisierten Pipeline zur Störstofferkennung: von den Bildaufnahmen über cloud-basierte Datenspeicherung bis hin zur annotierten Datenverarbeitung und Segmentierung mit YOLACT. Schließlich werden erste Ergebnisse vorgestellt und ein Ausblick gegeben. Ziel des Vorhabens ist eine robuste, effiziente und praxistaugliche Lösung zur Analyse der Reinheit von Biomüll.

## 2 Problemstellung

Die Nutzung von Bildanalyseverfahren mit Künstlicher Intelligenz zur Bewertung des Anteils von Störstoffen in Biomüll befindet sich an der Schnittstelle der schnell wachsenden Forschungsfelder KI-gestützte Bilderkennung und Abfallmanagement. Maschinelles Lernen, insbesondere tiefe neuronale Netze, können effektiv in der Analyse und Klassifizierung von Bildmaterialien eingesetzt werden, um verschiedene Materialien und Störstoffe zu identifizieren [AM22; NA23]. Technische Aspekte wie Genauigkeit, Lernfähigkeit und Anpassbarkeit von KI-Modellen sind dabei zentral. Die Fähigkeit, komplexe Muster in Bildern zu erkennen, verbessert sich kontinuierlich, wobei Herausforderungen in der Beschaffung ausreichend annotierter Trainingsdaten und der Übertragbarkeit auf unterschiedliche Kontexte existieren.

Die Detektion von Störstoffen ist entlang der gesamten Prozesskette relevant. Sowohl die Identifikation in der häuslichen Biomülltonne als auch im Müllfahrzeug, beispielsweise durch das System c-detect der c-trace GmbH, bilden die Basis für Maßnahmen, die zu einer Reduktion von Verunreinigungen im Biomüll durch die Bürgerinnen und Bürger beitragen. Im weiteren Verlauf des Prozesses wird der Biomüll von den Müllfahrzeugen in ein Kompostwerk geliefert. Das diesem Beitrag zugrunde liegende Projekt TRACES (<https://www.hsbi.de/wirtschaft/forschung/arbeitsgruppe-angewandte-ki/projekte/traces-trash-recognition-and-ai-controlled-evaluation-of-waste-surfaces>, Aufger. 26.05.2025) bewegt sich in dieser Prozessphase und befasst sich mit der Abschätzung von Störstoffen im Biomüll in einem Kompostwerk. Projektpartnerinnen sind die Hochschule Bielefeld mit dem Fokus auf Datenvorbereitung und Modellierung und die c-trace GmbH mit dem Fokus auf die Implementierung der Hardware. Ablauf und designierte Ergebnisse des Projekts TRACES sind jedoch eng mit Partnerinnen und Partnern der Smart Recycling Factory (<https://smart-recycling-factory.com/>, Aufger. 26.05.2025), einem großen regionalen Innovationsprojekt zur Entwicklung von Recycling und Kreislaufwirtschaft abgestimmt.

Hinsichtlich der Prozessintegration bei der Anlieferung und Bewertung des Biomülls im Kompostwerk spielt die Entwicklung von Echtzeitanalyse-Systemen eine entscheidende Rolle. Denn diese Systeme müssen in bestehende Abfallsortierungs- und -verarbeitungsprozesse integrierbar sein, wobei Studien demonstrieren, dass Automatisierungslösungen das Potenzial haben, Effizienz und Präzision zu steigern [SCH22].

Die BioAbfV sieht verpflichtend eine Sichtkontrolle des beim Abfallentsorgungsunternehmens angelieferten Biomülls vor, um zu entscheiden, ob der darin enthaltene Fremdstoffanteil die Mengenschwelle erreicht, dass er komplett an den Entsorger zurückgewiesen werden kann und/oder zunächst vorzubehandeln ist. Die Vorbehandlung wird durchgeführt, indem zunächst eine Fremdstoffentfernung stattfindet, um den Biomüll dann einer erneuten Beurteilung zuzuführen. Diese Anforderung soll durch die Ergebnisse des TRACES-Projekts systematisch unterstützt werden.

### 3 Lösungsansätze und Herausforderungen

#### 3.1 Bildaufnahmen

Für die Trainingsdaten sollen in einer Kompostierungsanlage regelmäßig Bilder des angelieferten Biomülls aufgenommen werden, um diesen zu bewerten. Das Ziel der Bildaufnahme ist die Darstellung eines  $10m \times 7m$  großen Bereichs, bei dem zu erkennende Objekte in mindestens  $10px \times 10px$  ( $px$  - Pixel) dargestellt sein sollen. Für den Versuchsaufbau wurde eine 5MP-Kamera mit integriertem Rechner und KI-Beschleunigung gewählt. Dies hält Anpassungsmöglichkeiten offen, wie das Integrieren eines KI-Modells bereits auf dem in der Kamera integrierten Rechner. Die Kamera ist an der Hallendecke in ca.  $6,5m$  Höhe installiert und in Vogelperspektive nach unten auf den angelieferten Biomüll ausgerichtet (vgl. Abb. 1). Aufgrund der vorgegebenen Entfernung wurde ein Öffnungswinkel von  $90$  Grad gewählt. Mit der Auflösung von  $5MP$  ( $2560px \times 1920px$ ) entspricht 1 Pixel ungefähr  $4mm$  (Kantenlänge). Somit können Objekte ab einer Größe von  $40mm \times 40mm = 10px \times 10px$  erkannt werden. Um entsprechend der BioAbfV Verunreinigungen ab einer Größe von  $20mm \times 20mm$  zu erkennen, wird eine Auflösung von mindestens  $5000px \times 3500px = 17,5MP$  benötigt und im weiteren Projektverlauf auch geupgraded, da mit dem Bildbereich eine Fläche von  $7m \times 10m$  abgedeckt wird. Wichtig für eine gute Bildqualität ist außerdem die Belichtung des Bildes. Da auch die Lichtquellen  $6,5m$  vom Biomüll entfernt angebracht sind, wird eine relativ lange Belichtungszeit benötigt. Da das Bild allerdings überwiegend unbewegt ist, konnte dies problemlos realisiert werden. Die Kamera nimmt alle 30 Sekunden ein Bild auf und speichert dieses, wenn die Pixeldifferenz zum vorherigen Bild groß genug ist. Über eine Verbindung mit einem LTE-Router werden die relevanten Bilder kontinuierlich in der Cloud gespeichert.



Abb. 1: Kameraanbringung

3.2 Datengewinnung und -vorbereitung

Die Grundlage für die Entwicklung eines leistungsfähigen Modells bildet ein eigenständig aufgebauter Datensatz. Dieser Abschnitt beschreibt die Annotation der Bilddaten sowie die Auswahl eines geeigneten Werkzeugs zur Erstellung der Trainingsdaten.

Die Bilder für das Modelltraining werden durch eigene Aufnahmen von Bioabfallhaufen in einer Kompostierungsanlage erzeugt. Dabei werden die Bioabfallladungen als Haufen aus der Vogelperspektive fotografiert. Ziel des Labelings ist es, nicht nur die Störstoffe als Ganzes zu erkennen, sondern jedes Objekt innerhalb eines Haufens präzise und als individuelle Instanz zu identifizieren. Da sich die Störstoffe in Form, Farbe und Größe stark unterscheiden und häufig teilweise verdeckt oder mit organischem Material vermischt sind, wird eine Bildqualität benötigt, die die Erkennung von kleinen Störstoffen und eine eindeutige Klassifizierung ermöglicht. Wichtig ist auch, dass die Bilder für den Trainingsdatensatz bestimmte Anforderungen erfüllen. So werden ausschließlich Bilder in das Training aufgenommen, in denen der gesamte zu analysierende Bioabfallhaufen abgebildet wurde. Ein weiteres Ausschlusskriterium sind Aufnahmen mit überbelichteten Bildteilen aufgrund eines starken Lichteinfalls von der Sonne in der Aufnahmehalle.

Die Bilder wurden noch vor der 5MP-Kamera in einer Auflösung von  $1920px \times 1080px$  aufgenommen. Diese Auflösung erwies sich in der initialen Phase als praktikabel, stellte jedoch insbesondere bei der Erkennung von kleinen Objekten eine Herausforderung dar. Doch auch größere Störstoffe konnten bei schlechter Beleuchtung oder verschwommenen Aufnahmen nicht immer zuverlässig klassifiziert werden. Zur Annotation der Bilddaten wird das Tool Roboflow verwendet, welches präzise Masken für die Instance-Segmentierung und verschiedene Exportformate (z. B. COCO JSON, YOLO) unterstützt. Für das erste Trainingsset werden 80 Bilder manuell gelabelt und teilweise gespiegelt dupliziert. Insgesamt umfasst der erste Datensatz 150 Bilder. Diese Phase konzentriert sich auf eine grobe Klassifizierung der erkennbaren Störstoffe. Hierfür werden folgende Klassen definiert: Glas, Kunststoff, Metall und Papiertüten. Diese Auswahl basiert auf der Relevanz für Recyclingprozesse und der Sichtbarkeit der Objekte im Bildmaterial. Auffällig ist der überwiegende Anteil von Kunststoff in den annotierten Bildern im Vergleich zu anderen Störstoffklassen.

Störstoff	Störstoffanzahl	Prozentualer Anteil
Glas	1	0,10989 %
Kunststoff	788	86,59341 %
Metall	7	0,76923 %
Papiertüten	114	12,52747 %

Tab. 1: Anteile der Störstoffklassen im ersten Datensatz

Die Annotation erfolgt durch die manuelle Maskierung jeder Fremdstoff-Instanz, wobei insbesondere kleine und halbtransparente Objekte (z. B. Reste von zerrissenen Plastiktüten) eine erhöhte Herausforderung darstellen. Besonders in Aufnahmen mit niedriger Bildqualität



ist es schwierig, feine Details wie zerbrochene Glasstücke oder dünne Folien zu erkennen und korrekt zu markieren. Durch den Wechsel auf eine höhere Auflösung können diese Einschränkungen teilweise kompensiert werden. In der folgenden Abbildung 2 ist anhand der Nahaufnahmen zweier Haufen erkennbar, wie die Verbesserung der Bildqualität eine Sicherheit in der Klassifizierung von Störstoffen gewährt.



Abb. 2: Unterschiede in der Bildqualität bei verbesserter Auflösung

Mit der Einführung von Bildern mit höherer Auflösung ( $2048px \times 1536px$  aus der 5MP-Kamera) können auch kleinere oder schlecht sichtbare Störstoffe besser annotiert werden. Zudem wird erkannt, dass eine feiner abgestufte Klassifikation notwendig ist, um eine höhere Genauigkeit des Modells zu erreichen. Folgend wird die Klasse *Kunststoff* in zwei separate Klassen unterteilt. In die Klassen *Kunststoff* und *Bioabfalltüten*. Letztere sind nun genauer im Bildmaterial zu erkennen und nach § 2 der BioAbfV auch als Fremdstoff-Objekte im Biomüll einzuordnen. Weitere ergänzende Klassen werden eingeführt, um eine realitätsnahe Abbildung der tatsächlichen Müllzusammensetzung zu ermöglichen. So auch die Klasse *Müllhaufen* für die Hintergrundsegmentierung, aber auch die Klassen *Unbekannter Störstoff* und *Unbekanntes Material*. Diese erweiterten Kategorien tragen dazu bei, das Modell auch in Fällen zu stabilisieren, in denen die Objekte nicht eindeutig eingeordnet werden können.

### 3.3 Modelling

Das Projekt setzt zur Segmentierung von Störstoffen in Biomüllhaufen auf das YOLACT-Modell (You Only Look At Coefficients). Diese Wahl begründet sich aus mehreren entscheidenden Faktoren, die sowohl mit der Architektur des Modells als auch mit den spezifischen Anforderungen unseres Anwendungsfalls einher gehen [BZ+19].

YOLACT ist ein frameweises, voll-konvolutionales Modell, das Echtzeit-Segmentierung ermöglicht. Die Architektur basiert auf einem Backbone-Netzwerk (in diesem Fall ResNet-50), welches tiefgehende Merkmalsrepräsentationen extrahiert [HZ+16]. Darauf aufbauend generiert ein Protonet eine Reihe von Prototypen, die mit Detektionsboxen kombiniert werden, um finale Segmentierungsmasken zu erstellen [BZ+19]. Die Schwerpunkte des Modells liegen zum einen in der höheren Effizienz gegenüber zweistufigen Architekturen wie dem Mask R-CNN, da keine Region-Proposal-Phase benötigt wird und zum anderen in der Flexibilität, da durch die Trennung von Prototypen- und Koordinaten-Detektion das

Modell für verschiedene Anwendungsbereiche geeignet ist, ohne dass aufwendige Architekturmodifikationen erforderlich sind [HZ+16]. Darüber hinaus erzielt das YOLACT eine entsprechend solide Segmentierungsleistung, die für unseren Anwendungsfall entscheidend ist. Diese Anforderung ergibt sich aus der späteren Flächenanteilsschätzung der Störstoffe anhand der Segmentierung. Es kommt hinzu, dass eine effiziente und robuste Unterscheidung von Störstoffen in Biomüllhaufen vollzogen werden muss, weshalb schlussendlich das YOLACT einen guten Kompromiss zwischen Rechenaufwand und Genauigkeit darstellt [CT+23].

Die Bilder in unserem Datensatz weisen eine hohe Auflösung auf, wodurch das Modelltraining mit den vollständigen Bildern erheblich erschwert wird. Das Tiling, also das Zuschneiden der Bilder in kleinere Segmente, bringt hierbei mehrere Vorteile. Zum einen reduziert sich die Modellgröße und damit die Rechenzeit, wodurch GPU-Ressourcen gespart und Speicheranforderungen verringert werden können [HZ+17]. Zum anderen ergeben sich bessere Augmentierungsmöglichkeiten, da gängige Methoden nun auf jeden einzelnen Bildausschnitt separat durchgeführt werden können. Augmentation erzeugt Datenvielfalt, sodass das Modell relevante Merkmale auch unter wechselnden Bedingungen sicher erkennt. Hiermit kann eine höhere Robustheit und eine bessere Generalisierungsfähigkeit des Modells sichergestellt werden [SK19]. Ermöglicht wird das Tiling durch die Eigenschaft der Daten, da in der Anwendung der Gesamtzusammenhang des Bildes nicht relevant ist und die Störstoffe somit in isolierten Bildausschnitten erkannt werden können. Dies begünstigt das Tiling-Verfahren (sh. Abb. 3) ohne wesentliche Informationsverluste.

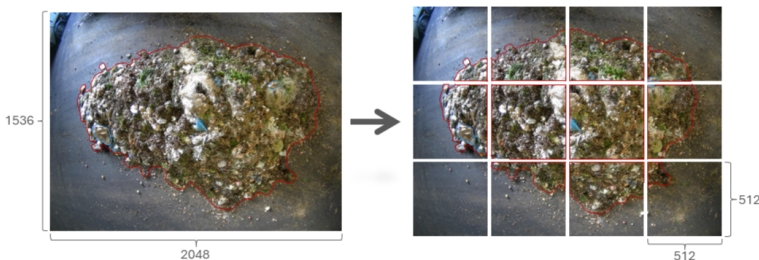


Abb. 3: Aufteilen eines Vollbildes in Einzelsegmente (Tiles)

Da das YOLACT standardmäßig mit einer Eingangsauflösung von  $512px \times 512px$  trainiert wurde, ergibt sich eine natürliche Anpassung an die gewählte Tiling-Strategie. Das Modell wurde mit diesen zerschnittenen Bildausschnitten trainiert, wobei die Annotationen entsprechend angepasst wurden, um die neuen Koordinaten der Objekte korrekt abzubilden. Dies verhindert Verzerrungen und stellt sicher, dass die Netzwerkeingabe mit dem Pretraining des Modells übereinstimmt. Ein Training mit deutlich größeren oder kleineren Eingangsgrößen könnte dagegen zu suboptimalen Ergebnissen führen, da das Modell auf eine spezifische Patch-Größe optimiert ist [BZ+19].

Für die Inferenz in YOLACT wird jedes Bild zunächst in kleinere Tiles (Bereiche) unterteilt, die einzeln analysiert werden. Jedes Tile wird unabhängig voneinander verarbeitet, wobei

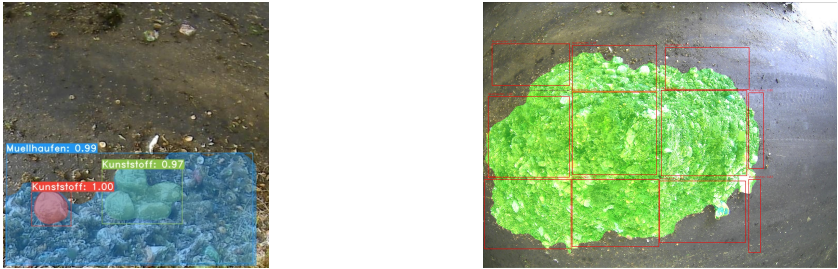


Abb. 4: Inferenz eines Tiles und Zusammenführung der Einzelinferenzen (nur Müllhaufenklasse)

für jedes Tile Instanzmasken und Klassenzuweisungen berechnet werden. Das endgültige Segmentierungsergebnis für das gesamte Bild ergibt sich dann aus der Summe der erkannten Objekte in allen Tiles. Das bedeutet, dass die gesamte Fläche für jede erkannte Klasse im Bild die Summe der Flächen der Objekte in den einzelnen Tiles ist (vgl. Abb. 4). Durch diese Technik kann die Segmentierung schneller und ressourcenschonender durchgeführt werden, da das Bild in handhabbare Teile zerlegt wird. Eine Überlappung der Tiles ist nicht nötig, da die Objektfläche beim Zusammenführen vollständig erfasst wird.

Durch die Kombination von YOLACTs effizienter Architektur mit einem gezielten Tiling-Ansatz konnte ein leistungsfähiges Modell zur Segmentierung von Störstoffen im Biomüll entwickelt werden, welches sich in den folgenden Ergebnissen erkennen lässt.

4 Bisherige Ergebnisse

In den Ergebnissen wird der Fokus auf die Effizienz des Modells sowie die Segmentierungsleistung gesetzt. Zunächst wird die Inferenzdauer des YOLACT betrachtet. Bei einer Gesamtzeit von 18,38ms pro Bild stellt sich unser Modell als besonders schnell heraus (vgl. Tab.: 2). Im Vergleich dazu benötigen andere gängige Modelle für ähnliche Aufgaben deutlich länger, was die Effizienz und Anwendbarkeit von YOLACT in Echtzeitszenarien unterstreicht. Beispielsweise erzielt das Mask R-CNN auf einer vergleichbaren Hardware Inferenzzeiten von über 450ms pro Bild [Jo19]. Dies wurde ermöglicht, indem statt der ursprünglichen Bildauflösung von 2048px × 1536px die Auflösung auf 512px × 512px reduziert und die Bildverarbeitung durch Tiling optimiert wurde. Insgesamt sind in diesem konkreten Fall 12 Durchläufe nötig mit einer sich ergebenden Gesamtzeit von circa 220ms. Die Gesamtzeit dieses Modells liegt dabei sogar unter der des Mask R-CNNs für nur ein Bild. Im Gegensatz dazu bewies sich das YOLO v8 und YOLO v11 mit Zeiten unter 8ms [Jo24].

Task	Backbone	FPN	Proto	Pred Heads	Detect	Other	Total
Time (ms)	6.065	0.452	0.278	1.50	6.986	3.099	18.380

Tab. 2: Benchmark: Inferenz

Ein weiterer wichtiger Aspekt der Ergebnisse ist die Analyse des Verlustverlaufs, die in Abbildung 5 (links) dargestellt ist. Der Verlust für alle Kategorien zeigt einen kontinuierlichen Abfall im Verlauf des Trainings. Dies deutet auf eine effektive Modellanpassung hin, die darauf abzielt, die Segmentierungsergebnisse zu optimieren.

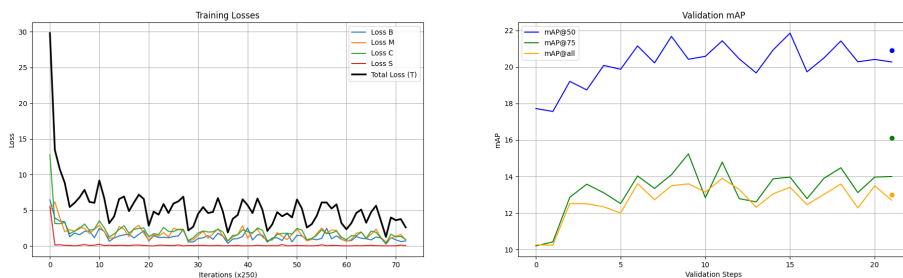


Abb. 5: Verlust und mAP

Während der Validierung wurde der mAP bei Schwellenwerten von 50, 75 und insgesamt ermittelt – ein zentraler Leistungsindikator für Präzision und Vollständigkeit der Segmentierung. Der steigende Verlauf (vgl. Abb. 5) zeigt, dass das Modell zunehmend robuster gegenüber verschiedenen Störstoffkategorien wird. Aktuell liegt der mAP bei 21%, mit dem Ziel, ihn durch weitere Optimierung auf etwa 40% zu steigern. Die in Abbildung 5 gezeigten Punkte beziehen sich auf den Test-Datensatz; weiterführende Ergebnisse folgen im Workshop.

## 5 Fazit und Ausblick

Für das Projekt wurden Bilddaten per installierter Kamera erfasst und über eine Cloudlösung bereitgestellt. Die Datenaufbereitung und das Labeling erfolgten mit Roboflow, ergänzt durch Augmentation und ein eigens entwickeltes Tiling-Verfahren. Anschließend wurde das Modell trainiert und umfassend evaluiert.

Zusammenfassend bestätigen die Ergebnisse, dass das YOLACT-Modell für die Segmentierung von Störstoffen im Biomüll sowohl schnell als auch präzise ist. Die Effizienz des Modells, gepaart mit einem stabilen mAP-Verlauf, belegt die Eignung von YOLACT für Anwendungen im Bereich der Abfallanalyse und -klassifikation.

Dennoch werden Präzision und Generalisierungsfähigkeit durch die geringe Menge an Daten eingeschränkt. Dies impliziert die Notwendigkeit der Datensatzausdehnung, sowie eine Ausdehnung der Repräsentation einzelner Klassen für weiteres Training. Zudem liefern die Inferenzzeiten des YOLO v8 und v11 einen Denkanstoß für einen Modellvergleich für diesen konkreten Anwendungsfall. Darüber hinaus stellt sich die Frage nach der Erkennungsgenauigkeit von kleinen Störstoffen und der Notwendigkeit einer höheren Auflösung, die im Projektverlauf bereits von  $1920px \times 1080px$  auf  $2048px \times 1536px$  erhöht wurde.

Gefördert durch: BMBF im Projekt DATipilot (Förderkennzeichen: 03DPS1130A).

## Literaturverzeichnis

- [AM22] Abdu, H.; Mohd Noor, M. H.: A Survey on Waste Detection and Classification Using Deep Learning. *IEEE Access* 10, S. 128151–128165, 2022.
- [Bu25] Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz: Was ist das Ziel der letzten Änderung der Bioabfallverordnung im Jahr 2022? BMUV, 2025, <https://www.bmuv.de/faq/was-ist-das-ziel-der-aenderung-der-bioabfallverordnung>, Stand: 07. 04. 2025.
- [BZ+19] Bolya, D.; Zhou, C. et al.: YOLACT: Real-Time Instance Segmentation. *IEEE Xplore*, 2019, <https://ieeexplore.ieee.org/document/9010373>, Stand: 02. 04. 2025.
- [CT+23] Corrigan, B. C.; Tay, Z. Y. et al.: Real-Time Instance Segmentation for Detection of Underwater Litter as a Plastic Source. *Journal of Marine Science and Engineering*, 2023, <https://techniumscience.com/index.php/technium/article/view/8345>.
- [HZ+16] He, K.; Zhang, X. et al.: Deep Residual Learning for Image Recognition. *IEEE Xplore*, 2016, <https://ieeexplore.ieee.org/document/7780459>, Stand: 02. 04. 2025.
- [HZ+17] Howard, A. G.; Zhu, M. et al.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv*, 2017, <https://arxiv.org/abs/1704.04861>, Stand: 02. 04. 2025.
- [Jo19] Joya, C.: Mask R-CNN Benchmark. *GitHub*, 2019, [https://github.com/facebookresearch/maskrcnn-benchmark/blob/main/MODEL\\_ZOO.md](https://github.com/facebookresearch/maskrcnn-benchmark/blob/main/MODEL_ZOO.md), Stand: 04. 04. 2025.
- [Jo24] Jocher, G.: Instanz-Segmentierung. *Ultralytics*, 2024, <https://docs.ultralytics.com/de/tasks/segment/#how-do-i-load-and-validate-a-pretrained-yolo-segmentation-model>, Stand: 04. 04. 2025.
- [NA23] Nasir, I.; Aziz Al-Talib, G. A.: Waste Classification Using Artificial Intelligence Techniques: Literature Review. *Technium: Romanian Journal of Applied Sciences and Technology* 5, S. 49–59, 2023, <https://techniumscience.com/index.php/technium/article/view/8345>.
- [SCH22] Sinthiya, N. J.; Chowdhury, T. A.; Haque, A. K. M. B.: Artificial Intelligence Based Smart Waste Management—A Systematic Review. In (Lahby, M.; Al-Fuqaha, A.; Maleh, Y., Hrsg.): *Computational Intelligence Techniques for Green Smart Cities*. Springer International Publishing, Cham, S. 67–92, 2022, [https://doi.org/10.1007/978-3-030-96429-0\\_3](https://doi.org/10.1007/978-3-030-96429-0_3).
- [SK19] Shorten, C.; Khoshgoftaar, T. M.: A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data (SpringerOpen)*, 2019, <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0>, Stand: 02. 04. 2025.