

Masterarbeit

Evaluation von Vision Foundation Models für Semantic Instance Segmentation in Robotik-Anwendungen

Problemstellung:

Die präzise Objekterkennung (Semantic Instance Segmentation) in hochgradig unübersichtlichen Szenen, wie sie beim Bin Picking auftreten, stellt weiterhin eine große Herausforderung dar. Trotz der beeindruckenden Leistungsfähigkeit von neuartigen Foundation Models bleibt die Anpassung an komplexe und stark variierende Szenarien schwierig. Ziel ist es, die Performance verschiedener Modelle auf einem vorhandenen Datensatz zu evaluieren und zu benchmarken, um deren Potenzial und Limitationen für die Robotik besser zu verstehen.

Aufgabe:

Rahmen der Arbeit sollen folgende Schritte bearbeitet werden:

1. **Literaturrecherche:** Untersuchung aktueller wissenschaftlicher Arbeiten und Identifikation von State-of-the-Art-Modellen.
2. **Modellauswahl:** Auswahl vielversprechender Ansätze auf Basis der Literatur und der Anforderungen an Robotic Bin Picking.
3. **Training:** Anpassung und Training der ausgewählten Modelle auf dem bereits vorhandenen Datensatz, einschließlich möglicher Optimierung von Hyperparametern.
4. **Evaluation:** Durchführung einer systematischen Evaluation und Benchmarking der Modelle anhand geeigneter Metriken.

Abschließend sollen die Ergebnisse analysiert und Empfehlungen für den praktischen Einsatz gegeben werden.

Voraussetzung:

- Hohes Eigeninteresse an Themen wie Computer Vision und maschinellem Lernen.
- Programmiererfahrung in Python, idealerweise auch mit PyTorch und OpenCV.
- Grundkenntnisse in Computer Vision
- Kenntnisse im Bereich Deep Learning sind von Vorteil, aber nicht zwingend erforderlich – eine Einführung in die Grundlagen kann bei Bedarf während der Arbeit erfolgen.

Geboten:

Wir bieten eine spannende Masterarbeit im Bereich Computer Vision mit Zugang zu leistungsstarken Ressourcen für das Training und die Evaluation der Modelle sowie einem modernen Arbeitsplatz am Institut oder zu Hause. Arbeit kann in **Deutsch** oder **Englisch** betreut werden.

Weiterführende Links:

https://medium.com/@tenyks_blogger/the-foundation-models-reshaping-computer-vision-b299a91527fb (Blog-Artikel)

https://openaccess.thecvf.com/content/CVPR2024/papers/Han_Few-Shot_Object_Detection_with_Foundation_Models_CVPR_2024_paper.pdf (Wiss. Paper, CVPR 2024)

Forschungsbereich:

Deep Learning
Machine Vision
Robotics

Beginn: ab sofort

Ausschreibungsdatum:
22.11.2024

Sprache:

German or English

Ansprechpartner:

Maximilian Gilles

maximilian.gilles@kit.edu

Bewerbende fügen ihrer Bewerbung bitte einen **Lebenslauf** und einen aktuellen **Notenspiegel** bei.



Quelle: Grounded-SAM



Quelle: Amazon