

Bachelorarbeit

Thema: Panoptische Indoor-Segmentierung mit Transformer-basierten End-to-End-Ansätzen

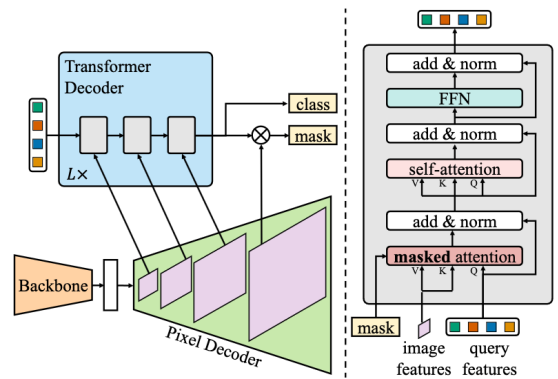
Beschreibung:

Die panoptische Segmentierung spielt eine entscheidende Rolle in der Robotik, da sie detaillierte Informationen zu semantischen Klassen und Instanzen in einer Szene liefert [8]. Dadurch kann ein Roboter ein umfangreiches Wissen über die Szene, in welcher er sich befindet, aufbauen und für seine Aufgaben geeignet berücksichtigen. Bisherige Ansätze erfordern oft eine komplexe Nachverarbeitung, um die einzelnen Teilergebnisse zu der panoptischen Segmentierung zusammenzuführen [1, 2]. Im Gegensatz dazu ermitteln aktuelle Transformer-basierte End-to-End-Ansätze, wie Panoptic SegFormer [3] und Mask2Former [4], panoptischen Masken auf direktem Weg und benötigen dafür eine weniger komplexe Nachverarbeitung. Allerdings wurden diese Verfahren bisher nicht speziell für den Einsatz in der Indoor-Robotik angepasst und evaluiert.

Das Ziel dieser Bachelorarbeit ist es daher, Panoptic SegFormer detailliert aufzuarbeiten und auf gängige Indoor-Datensätze [5-7] anzupassen. Anschließend soll das Verfahren auf diesen Datensätzen trainiert und evaluiert werden. Darüber hinaus sollen der realisierte Ansatz mit anderen Indoor-Verfahren verglichen werden, um die Leistungsfähigkeit zu bewerten.

Detaillierte Aufgabenstellung:

- systematische Aufarbeitung des State of the Art zum betrachteten Themenfeld ausgehend von [1-7]
- Implementierung eines geeigneten Ansatzes in Anlehnung an [3] und [4] mittels PyTorch
- Training und Evaluierung auf geeigneten Indoor-Datensätzen [5-7]
- Ausarbeitung von Präsentationen für den Eröffnungs-, und Abschlussvortrag
- Anfertigen der Masterarbeit entsprechend der Vorgaben des FG NI&KR



Bildquelle: [4]

Notwendige Voraussetzungen:

- Abschluss der Vorlesungen Neuroinformatik
- gute Kenntnisse im Bereich Deep Learning
- gutes mathematisches Verständnis und Erfahrung in der Programmierung mit Python

Literatur:

- [1] Seichter, et al.: [Efficient Multi-Task RGB-D Scene Analysis for Indoor Environments](#), IJCNN, 2022.
 - [2] Fishedick, et al.: [EMSAFormer: Efficient Multi-Task Scene Analysis with RGB-D Transformers](#), IJCNN, 2023.
 - [3] Li, et al.: [Panoptic SegFormer: Delving Deeper into Panoptic Segmentation with Transformers](#), CVPR, 2022.
 - [4] Bown, et al.: [Masked-attention Mask Transformer for Universal Image Segmentation](#), CVPR 2022
 - [5] Silberman, et al.: [Indoor Segmentation and Support Inference from RGBD Images](#), ECCV, 2012.
 - [6] Song, et al.: [SUN RGB-D: A RGB-D scene understanding benchmark suite](#), CVPR, 2015.
 - [7] Dai, et al.: [ScanNet: Richly-annotated 3D Reconstructions of Indoor Scenes](#), CVPR, 2017.
 - [8] Fishedick, et al.: [Bridging Distance with a Collaborative Telepresence Robot for Older Adults – Report on Progress in the CO-HUMANICS Project](#), ISR, 2023.
- Google Scholar scholar.google.com
 - Proceedings der rel. Konferenzen (IROS, ICRA, NIPS, ICML, ICLR, IJCNN, WCCI, ICANN, CVPR, ICCV, ECCV, BMVC, ICPR, ICIP, ...)

Betreuer: Söhnke B. Fishedick, M.Sc. (soehnke.fishedick@tu-ilmeanu.de)
Daniel Seichter, M.Sc. (daniel.seichter@tu-ilmenau.de)
Betr. Hochschullehrer: Prof. Dr. H.M. Groß
Bearbeiter: Leander Felix Heine