

# Masterarbeit

**Thema:** Erlernen von Instanz-Embeddings zur panoptischen Segmentierung

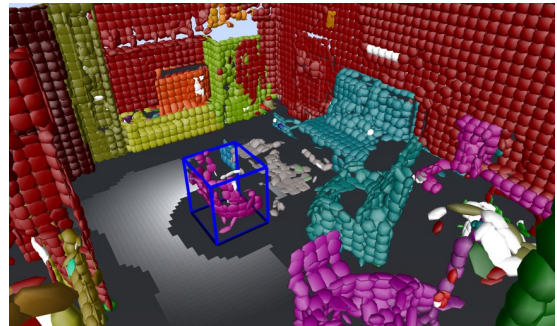
## Beschreibung:

Die panoptische Segmentierung ermöglicht ein tiefgehendes Verständnis von Szenen, indem sie detaillierte Informationen über semantische Klassen und Instanzen einer Szene bereitstellt und somit wichtige Informationen über die Umwelt, in welcher sich ein mobiler Roboter befindet [5], liefert. Speziell für diesen Einsatz wurden in [1-3] echtzeitfähige Verfahren zur panoptische RGB-D Segmentierung vorgestellt, die auf der Grundidee aus [4] aufbauen. Diese Ansätze führen einerseits eine pixelgenaue semantische Segmentierung durch und identifizieren andererseits Instanzen anhand ihres Massenschwerpunkts. In [1-2, 4] werden Pixel durch zusätzlich erlernte Verschiebungsvektoren zu einem Zentrum zugeordnet. Eine alternative Methode wurde in [3] vorgestellt, bei der anstelle von Verschiebungsvektoren Embeddings der Instanzen erlernt werden, um Pixel mit einem ähnlichen Embedding einer Instanz zuzuordnen.

Das Ziel dieser Masterarbeit besteht darin, das Verfahren aus [2] so zu modifizieren, dass ähnlich wie in [3] Embeddings anstelle von Verschiebungsvektoren erlernt werden. Zum Beurteilen der Leistungsfähigkeit, sollen hierfür umfangreiche Experimente auf den Indoor-Datensätzen aus [1, 2] verschiedenen Datensätzen durchgeführt werden. Dabei sollen auch mögliche sinnvolle Modifikationen an der Gesamtarchitektur erarbeitet und deren Einfluss untersucht werden. Darüber hinaus soll die Laufzeit des entwickelten Verfahrens betrachtet werden.

## Detaillierte Aufgabenstellung:

- systematische Aufarbeitung des State of the Art zum betrachteten Themenfeld ausgehend von [1-4]
- Implementierung eines geeigneten Ansatzes in Anlehnung an [2] und [3] mittels PyTorch
- Training und Evaluierung auf geeigneten Indoor-Datensätzen
- Überführung und Optimierung [6] für den Einsatz auf mobilen Robotern des FG NI&KR (NVIDIA Jetson AGX Orin 64GB)
- Ausarbeitung von Präsentationen für den Eröffnungs-, und Abschlussvortrag
- Anfertigen der Masterarbeit entsprechend der Vorgaben des FG NI&KR



Bildquelle: [5]

## 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] Sodano, et al.: [Robust Double-Encoder Network for RGB-D Panoptic Segmentation](#), ICRA, 2023.
  - [4] Bown, et al.: [Panoptic-DeepLab: A Simple, Strong, and Fast Baseline for Bottom-Up Panoptic Segmentation](#), CVPR 2020
  - [5] Fishedick, et al.: [Bridging Distance with a Collaborative Telepresence Robot for Older Adults – Report on Progress in the CO-HUMANICS Project](#), ISR, 2023
  - [6] NVIDIA TensorRT: <https://developer.nvidia.com/tensorrt> (online: 28.11.2023)
- Google Scholar [scholar.google.com](https://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](mailto:soehnke.fishedick@tu-ilmeanu.de))

Daniel Seichter, M.Sc. ([daniel.seichter@tu-ilmenau.de](mailto:daniel.seichter@tu-ilmenau.de))

**Betr. Hochschullehrer:**

Prof. Dr. H.M. Groß

**Bearbeiter:**

Xinxin Zhou