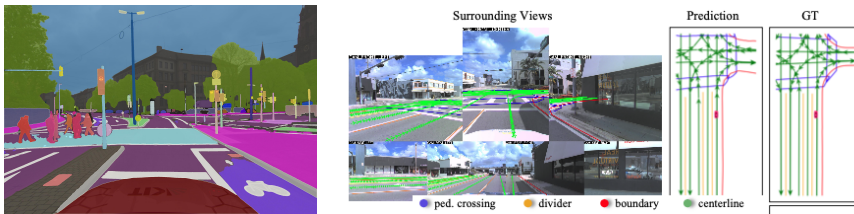


Master's Thesis



Panoptic Segmentation of a Mask2Former² Model at KIT (left), Map Perception results from MapTRv2³ (right)

Deep Learning: Shared ResNet50 Backbone for Mask2Former and MapTRv2 (like Tesla HydraNet)

State-of-the-art computer vision neural networks extract image features using deep backbones, which require a significant amount of computational time. This problem grows with the number of cameras used and the number of different models intended for use in a vehicle.

A sensible integration of various backbones into a shared feature extraction backbone is widely used in industry¹, but is still little explored in research, especially for the joint application of 2D (such as Panoptic Segmentation, left) and 3D models (such as Map Perception, right).

In this thesis, the joint use of once-computed backbone features for various computer vision tasks will be implemented, investigated, and evaluated:

- + Research on Mask2Former² and MapTRv2³
- + Adapting and loading ResNet50 weights from a Mask2Former model trained on the Mapillary Vistas dataset into the ResNet50 backbone of MapTR
- + Evaluating the performance
- + Training/Fine-Tuning MapTRv2 with frozen backbone parameters
- + Implementing a PyTorch nn.Module* that shares the backbone between Mask2Former and MapTRv2
- + Inference and performance evaluation
- + -----Optional-----
- + Automatic generation of semantic segmentation / panoptic segmentation pseudo-labels for Argoverse 2 (AV2) images using a large SwinL model
- + Fine-tuning a Mask2Former model trained on Mapillary Vistas on AV2 images with pseudo-labels
- + Joint training of the entire PyTorch model from*
- + Comparison of performance with the previous approach

Feel free to ask me any questions about the topic, reference literature, or other matters without obligation. Just ask without obligation or apply directly!

Institute for Measurement and Control Systems (MRT)

Prof. Dr.-Ing. Christoph Stiller

Supervisor:

Richard Fehler, M.Sc.

Programming Language(s)¹:

Python advanced
C++ basic

System, Framework(s):

Linux, PyTorch

Additional Requirements:

- Experience with Deep Learning and Neural Networks
- Experience with numerical libraries like NumPy or PyTorch
- Independent and motivated work style

What We Offer:

- Technical supervision by experienced scientists in Deep Learning
- Work with state-of-the-art methods
- Access to large GPU servers and HPC systems

A (partially) successful thesis may lead to a joint **international conference publication** of the scientific work.

Language(s):

German, English

Richard Fehler

Room: 036 → just drop by!

Tel.: +49 721 9654-261

Email: richard.fehler@kit.edu

Or apply directly with a current **transcript of grades** and a brief description of your previous **programming projects**. (preferably with a GitHub link or code as a .zip file)

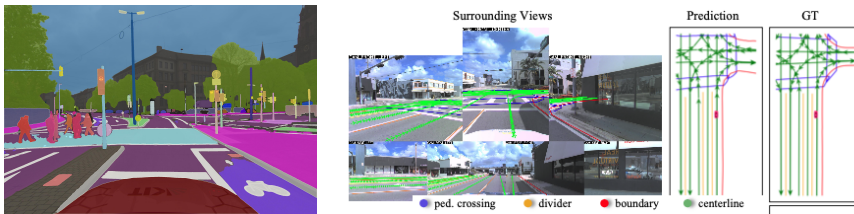
¹Tesla AI Day 2021 from 48min:

<https://www.youtube.com/live/j0z4FweCy4M?si=-b9C6yt0FVXrVxmP&t=2934>

²<https://github.com/facebookresearch/Mask2Former>

³<https://github.com/hustvl/MapTR>

Masterarbeit



Panoptic Segmentation eines Mask2Former² Modells am KIT (links), Map Perception Ergebnisse von MapTRv2³ (rechts)

Deep Learning: Shared ResNet50 Backbone for Mask2Former and MapTRv2 (like Tesla HydraNet)

State-of-the-Art Computer Vision Neural Networks extrahieren Bildmerkmale / Features mithilfe tiefer Backbones, welche einen großen Teil der Rechenzeit benötigen. Das Problem wächst mit der Anzahl an verwendeten Kameras und der Anzahl verschiedener Modelle die in einem Fahrzeug Verwendung finden sollen. Eine sinnvolle Integration der verschiedenen Backbones in ein gemeinsames Feature Extraction Backbone findet in der Industrie¹ breite Verwendung, ist jedoch im Forschungsbereich noch wenig untersucht, insbesondere für die gemeinsame Anwendung von 2D (wie Panoptic Segmentation, links) und 3D Modellen (wie Map Perception, rechts).

In dieser Arbeit soll eben diese gemeinsame Verwendung der einmalig berechneten Backbone Features für verschieden Computer Vision Tasks implementiert, untersucht und evaluiert werden:

- + Recherche zu Mask2Former² und MapTRv2³
- + Anpassen und Laden von ResNet50 Gewichten eines Mask2Former Modells trainiert auf Mapillary Vistas Dataset in das Backbone ResNet50 von MapTR
- + Evaluieren der Performance
- + Training/Fine-Tuning von MapTRv2 mit frozen Backbone Parametern
- + Implementieren eines PyTorch nn.Module * welches das Backbone zwischen Mask2Former und MapTRv2 teilt
- + Inferenz und Evaluation der Performance
- + -----Optional-----
- + Automatische Generierung von Semantic Segmentation / Panoptic Segmentation Pseudo-Labels für Bilder von Argoverse 2(AV2) mit großem SwinL Modell
- + Finetuning eines auf Mapillary Vistas trainierten Mask2Former Modells auf AV2 Bildern mit Pseudo-Labels
- + gemeinsames Trainieren des gesamten PyTorch Modells aus *
- + Vergleich der Performance mit dem vorherigen Ansatz

Gerne beantworte ich dir unverbindlich Fragen zur Thematik, Referenzliteratur oder sonstigen Themen. Frag mich einfach unverbindlich oder bewirb dich direkt!

Institut für Mess- und Regelungstechnik (MRT)
Prof. Dr.-Ing. Christoph Stiller

Betreuer:
Richard Fehler, M.Sc.

Programmiersprache(n)¹:
Python fortgeschritten

System, Framework(s):
Linux, PyTorch

- Weitere Voraussetzungen:**
- Erfahrung mit Deep Learning und Neural networks
 - Erfahrung mit numerischen Bibliotheken wie NumPy oder PyTorch
 - Selbständige und motivierte Arbeitsweise

- Was wir bieten:**
- technische Betreuung durch erfahrene Wissenschaftler im Deep Learning
 - Arbeiten mit State-of-the-Art Methoden
 - Zugriff auf große GPU Server und HPC Systeme

Eine (bereits teilweise) erfolgreiche Arbeit kann zu einer gemeinsamen **internationalen Konferenzveröffentlichung** der wissenschaftlichen Arbeit führen.

Sprache(n):
Deutsch, Englisch

Richard Fehler

Raum: 036 → einfach vorbeikommen!
Tel.: +49 721 9654-261
E-Mail: richard.fehler@kit.edu

Oder bewirb dich direkt mit einem aktuellen **Notenauszug** und einer kurzen Beschreibung deiner vorherigen **Programmierprojekte**. (gerne mit Github Link oder Code als .zip file)

¹Tesla AI Day 2021 ab 48min:

<https://www.youtube.com/live/j0z4FweCy4M?si=-b9C6yt0FVXrVxmP&t=2934>

²<https://github.com/facebookresearch/Mask2Former>

³<https://github.com/hustvl/MapTR>