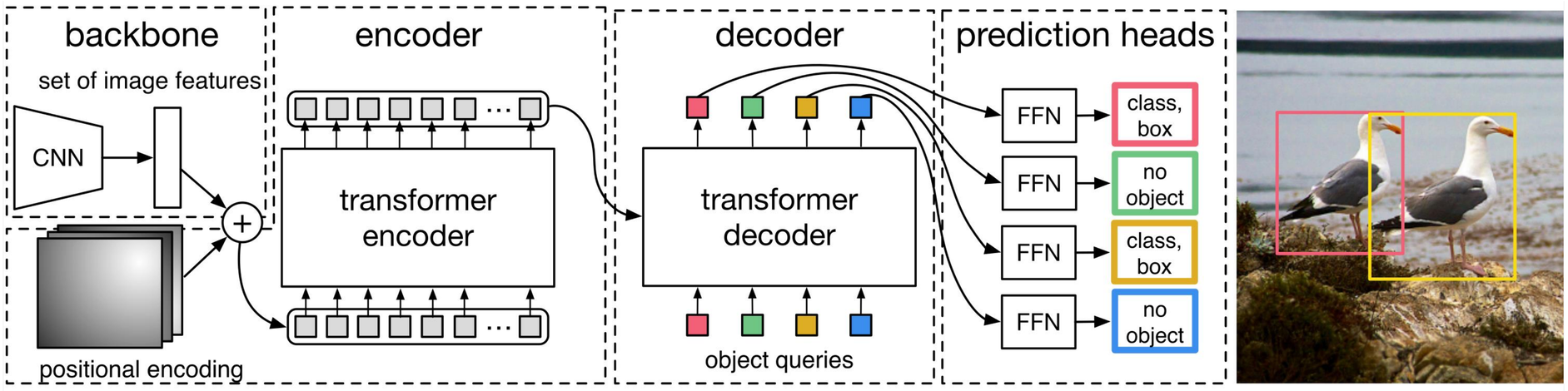
HSLU Hochschule Luzern

Technik & Architektur
Master of Science in Engineering
Specialization in Data Science

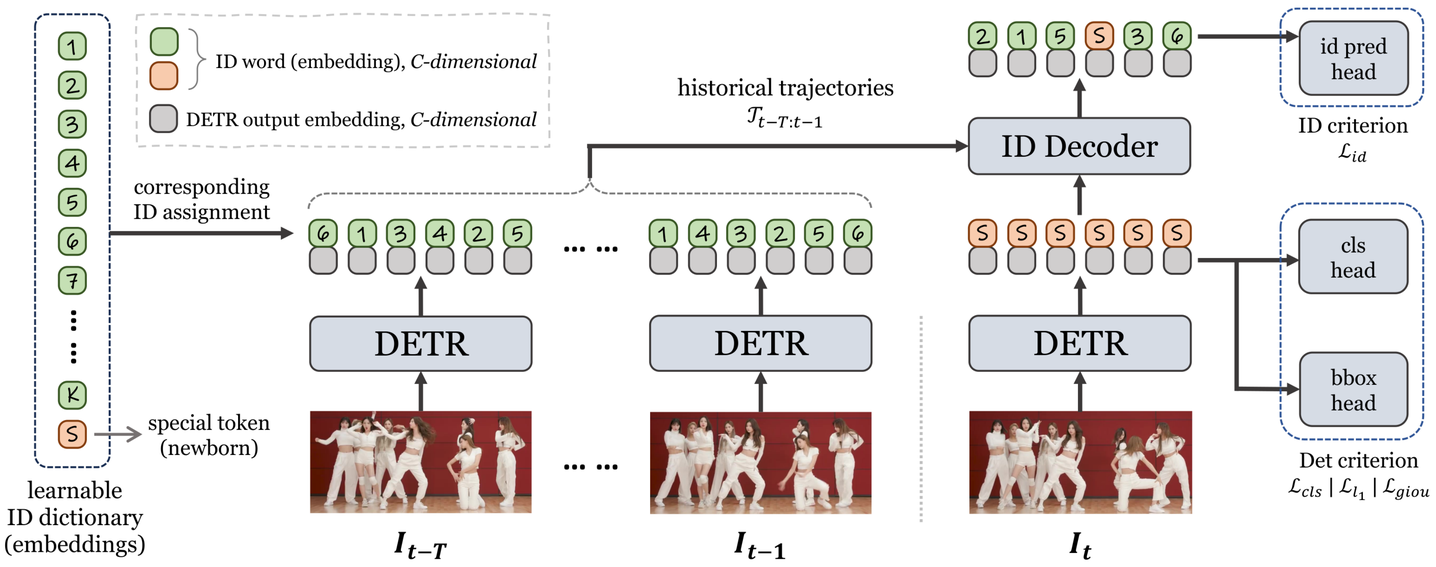
Master-Thesis Master of Science in Engineering

End-to-End Model Evaluation for Multiobject Tracking Tasks

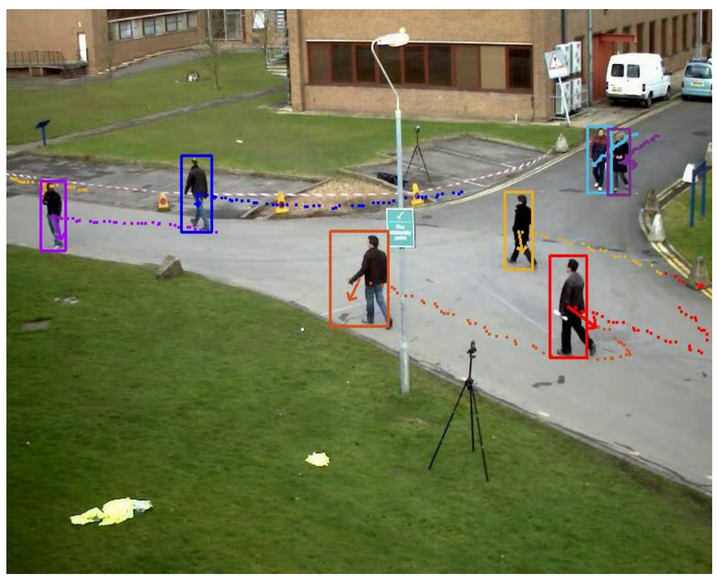
Evaluation of Methods for Multiobject Tracking and Trajectory Prediction



DETR model architecture



End-to-End model architecture of MOTIP



MOTIP output visualization

Problem Description

Multiobject tracking and trajectory prediction for pedestrians is a complex task, which typically is processed using a pipeline of individual components. A common approach is the use of a detection model, which provides high-quality detections, combined with a tracking algorithm for association. Trajectory prediction then uses the track history to estimate the future position.

The novel transformer architecture has revolutionized sequential modelling and driven progress in applications such as natural language processing. In recent years, the transformer has also found increasing attention in computer vision and enabled new models to integrate a joint detection and tracking approach. This work evaluates, how well these models can integrate end-to-end capabilities for multiobject tracking and trajectory prediction in the context of pedestrian tracking.

Solution

This work evaluates various methods in a utility analysis, where MOTIP is identified as the most promising candidate. The model is trained on a custom pedestrian dataset to evaluate application specific capabilities. Further, the model is deployed on dedicated hardware to assess real-time capabilities and the impact of optimization measures on tracking performance.

Results

The training of MOTIP on application specific data has shown improved performance when compared with a tracking-by-detection method such as YOLOX combined with ByteTrack. A significant improvement is achieved on association capabilities of the ID Decoder, which is responsible for the mapping of detection outputs with identities in the temporal domain. Real-time capabilities are evaluated on the Nvidia Jetson Nano Orin, where real-time performance is achieved on reduced resolution. But the lower resolution also affects tracking performance, so a trade-off between resolution and frame rate is required, for which various ideas are proposed. Attempts to integrate trajectory prediction with MOTIP have been investigated. However, performance evaluation did not produce any successful results.

Maurus Fritsche

Advisor:
Prof. Dr. Klaus Zahn

External Expert:
Marco Hauri

Industrial Partner:
agtatec ag

FH Zentralschweiz

MSE | MASTER OF SCIENCE IN ENGINEERING