

Opponent Review

for the master thesis

"Enhancing Vehicle Interior Action Recognition using Contrastive Self-Supervised Learning with 3D Human Skeleton Representations"

Reviewer: Prof. Dr. Phillipp Torkler
Student: Yasser El Bachiri
Matriculation No: 12102782

Summary

The thesis aims to recognize human actions utilizing deep learning self-supervised learning techniques on 3D human skeleton data to enhance driver safety in cars. Various driver actions, such as hand gestures, head movements, and body postures, shall be classified in real-time to detect potential distractions and unsafe behaviors for timely intervention and accident prevention. Two approaches are proposed and evaluated in the thesis. First, the DINO approach, originally developed for image data, is adapted to work with 3D skeleton data and uses the STTFormer transformer as backbone. Second, partial spatio-temporal learning (PSTL) framework is combined with a spatial temporal graph convolutional network (ST-GCN) backbone. The NTU RGB+D dataset and the Drive&Act dataset are used for evaluation.

Review

The thesis is structured in a coherent and logical fashion, adheres to common scientific standards, and is written very well. Both, the extensive literature review and comprehensive method description demonstrate a solid understanding of the topic and the research objectives are stated precisely.

The explanation of the used datasets and performance evaluations could have been more detailed. The following points could have been more pronounced:

- A motivation is missing why the NTU dataset was used at all, since it is not connected to driving.
- Splits for training, validation and testing of the datasets are lacking, and cross-validation and hyperparameter optimization is not mentioned.
- Strong performance differences between the NTU and Drive&Act datasets raise many questions.
- Evaluation only reports overall accuracies, but does not focus on the motivation of the thesis - the detection of *potential dangerous driver actions*. The Drive&Act dataset for sure contains many *normal* actions. An emphasis in the evaluation on *potential dangerous driver actions* would have been advantageous.

In conclusion, I recommend rating the thesis in the range between *very good* and *good*.

Deggendorf, 29.08.2023

Prof. Dr. Phillipp Torkler