KI DELTA LEARNING
Scalable AI for Automated Driving

Self-Supervised Deep Visual Representation Learning as Pre-Training

Manuel Schwonberg, Dr. Nico Schmidt

WP 3.1

Networks

Use-Case in Delta Learning

Our goal is to develop evaluate self-supervised deep representation learning methods as an alternative to the standard supervised ImageNet-based pre-training and to reduce the labeling effort. This can be achieved by using a pretext task which does not need any human-based labels.

Technical Challenge

The initialization of the weights of a Deep Neural Network (DNN) has a strong impact on the achievable performance. To ease training even with a small number of samples supervised ImageNet pre-training is applied to provide a good initialization of the network. That is called Transfer Learning.
There are a couple of challenges coming with Transfer Learning; namely a task and feature domain gap between pre-training and target training that may harm the performance and a potentially intense labeling effort if ImageNet is not used.
Additionally unlabeled automotive data is available at less cost and strong self-supervised learning methods can help to utilize this hidden potential.

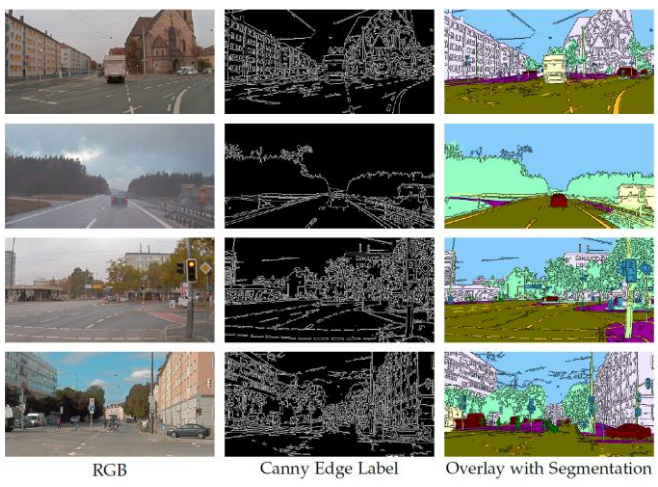


Figure I: Edge Images

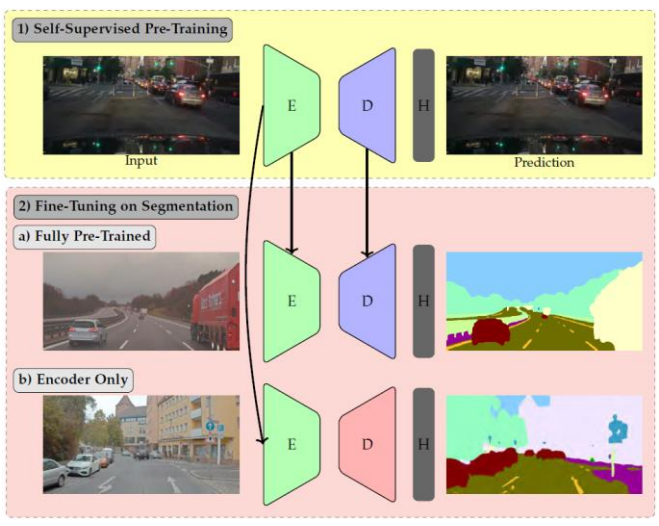


Figure II: 2-Stage Training

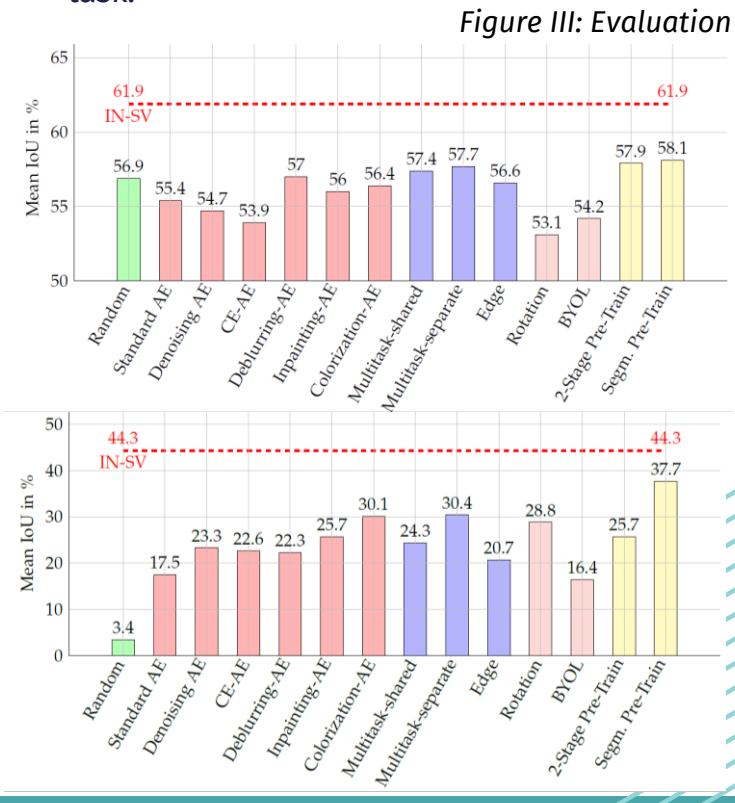
Technical Solution

The technical solution are self-supervised learning methods which can use unlabeled data to learn deep visual representations. We applied a 2-stage method (Fig. 2) where both self-supervised pre-training and the supervised target-training on segmentation data are conducted sequentially. The self-supervised representations serve as the initialization for the supervised target training. The crucial point is the pretext task that makes use of the unlabeled data. We tested a wide range of different learning methods starting from a convolutional Autoencoder to simple pretext tasks like rotation prediction or patch-wise reconstruction. We developed a new pretext task, namely edge prediction (Fig. 1), and a new multi-task pretext learning scheme where the network both learns to predict the edges and the image itself. Also a new contrastive approach called *Bootstrap Your Own Latent* was tested.
Another novelty of our approach is that pre-training was conducted on an automotive dataset which should reduce the domain gap to the target training.

Evaluation

As it can be seen from the graphs below (Fig. 3) none of the self-supervised methods was capable to outperform supervised ImageNet pre-training but that all methods learned useful visual representations.
There are different explanations such as overfitting to the pretext-task or a misalignment between the pretext and target task.

Figure III: Evaluation



CARIAD
A VOLKSWAGEN GROUP COMPANY

For more information contact:
manuel.schwonberg@cariad.technology or
nico.schmidt@cariad.technology

Partners

External partners

KI Delta Learning is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Energy.

KI FAMILIE

Supported by:
Federal Ministry for Economic Affairs and Energy
on the basis of a decision by the German Bundestag

www.ki-deltalearning.de @KI_Familie KI Familie