

Introduction and Motivation

- Pre-training commonly used technique in computer vision; only done with ImageNet
- Costly and time-intensive annotation for semantic segmentation limits dataset size and applicability
- Low utilization of large amounts of relatively cheap unlabeled image data
- Standard pre-training has to overcome a task and domain gap harming the efficiency of pre-training

Approach

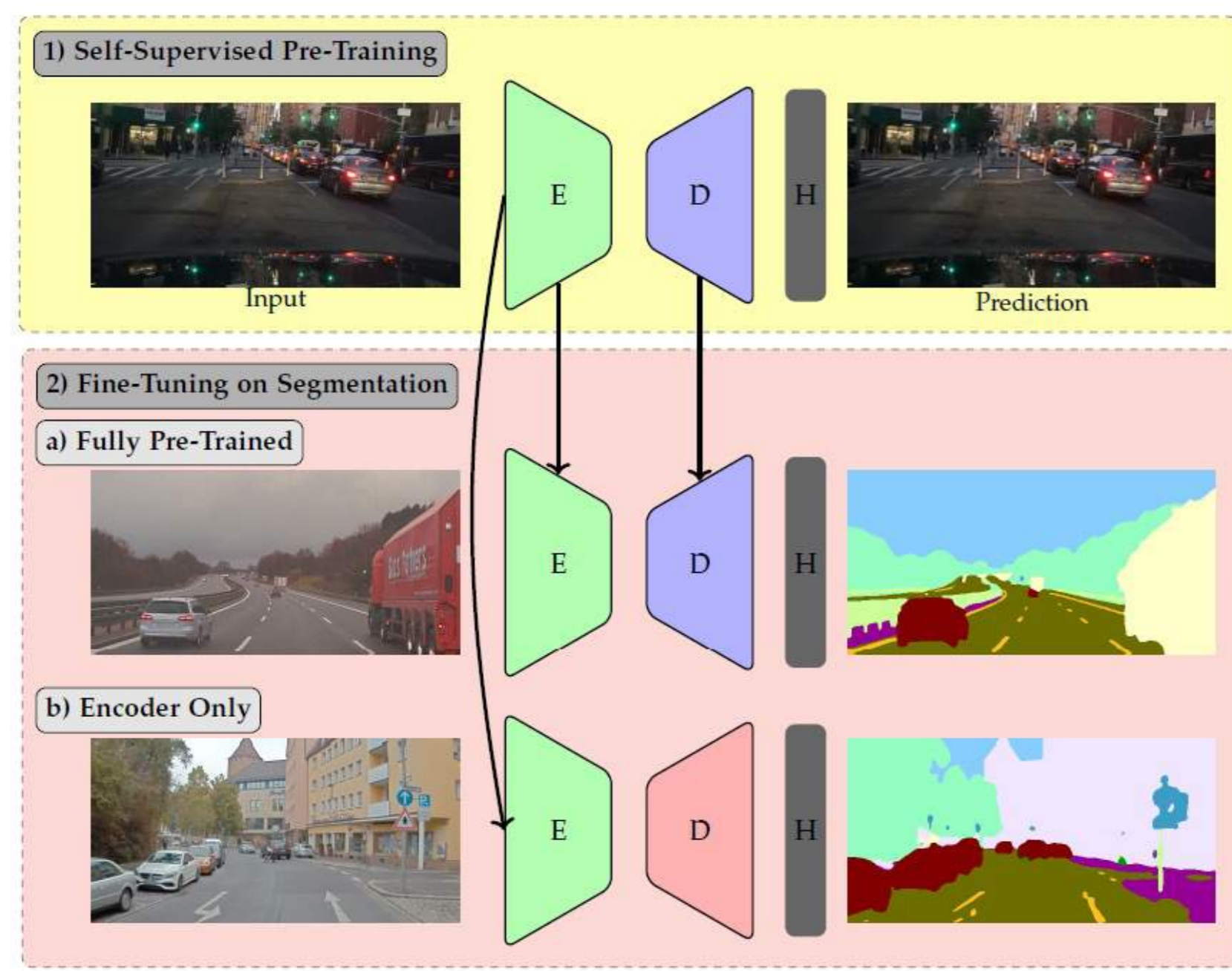


Figure 1: Self-Supervised Learning Approach

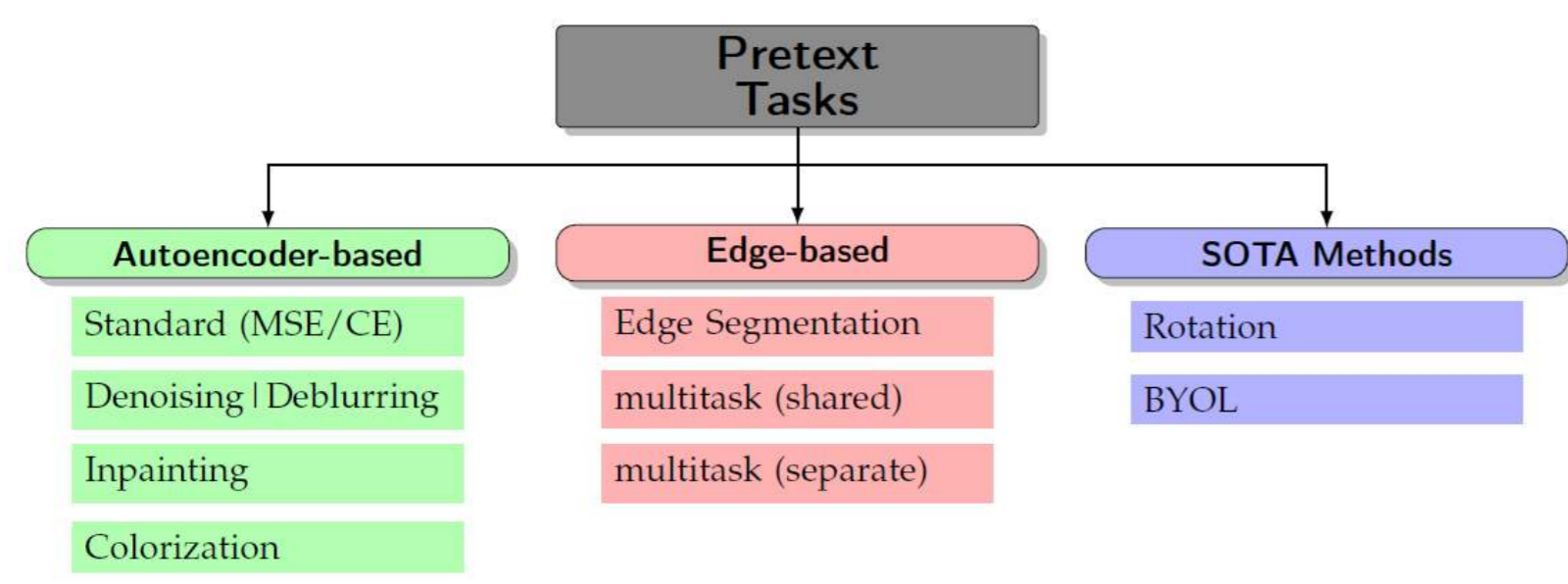


Figure 2: Pretext Tasks (© CARIAD SE)

Idea: self-supervised pre-training on automotive datasets like A2D2 or BDD100k

Pretext Tasks

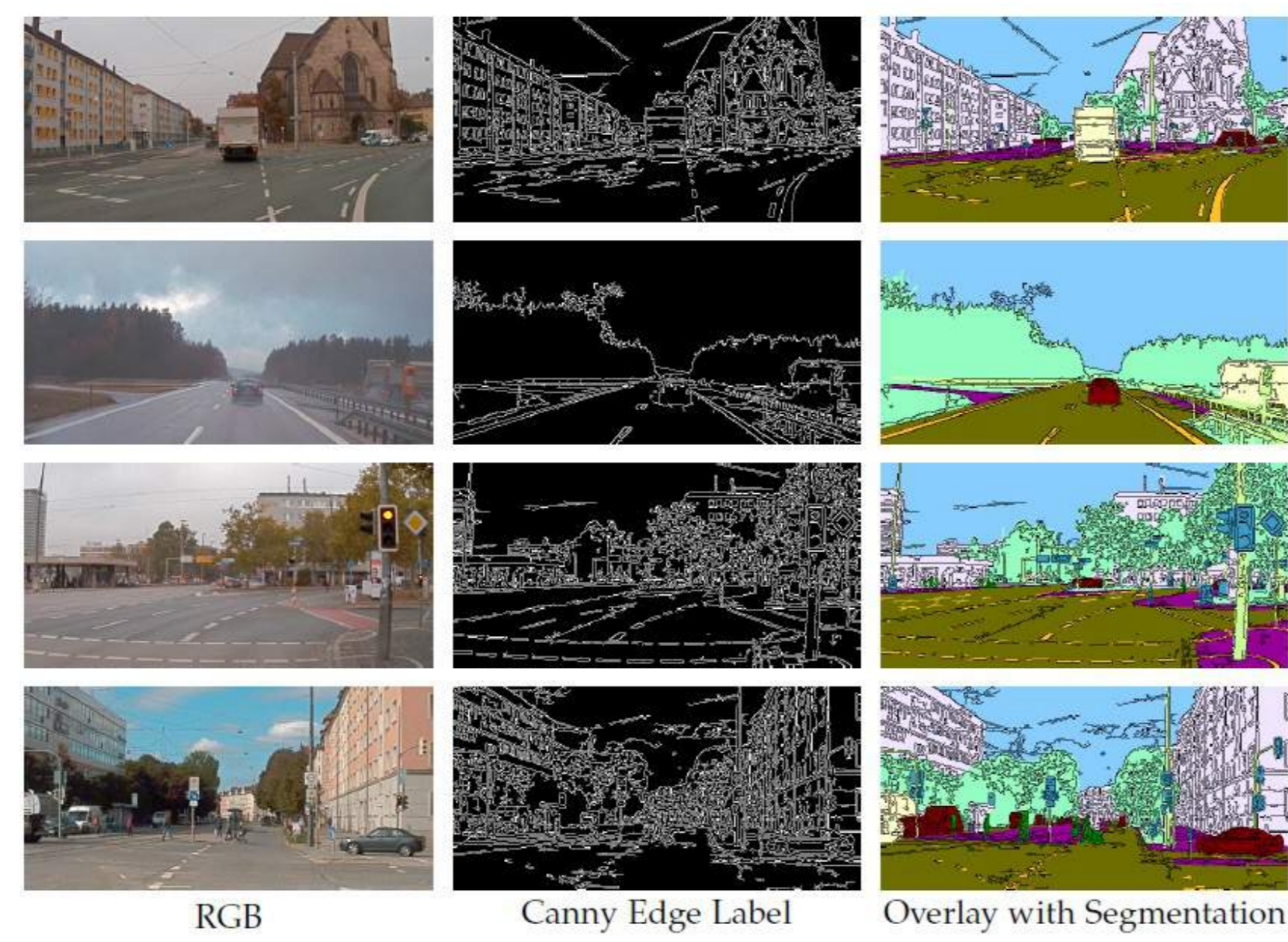


Figure 3: Edge Segmentation

- Semantic correspondence between canny edge detection and semantic segmentation
- Idea: Predict edges as pre-training → useful bias for segmentation
- Also multitask pretext task (autoencoder + edge segmentation)

Results I

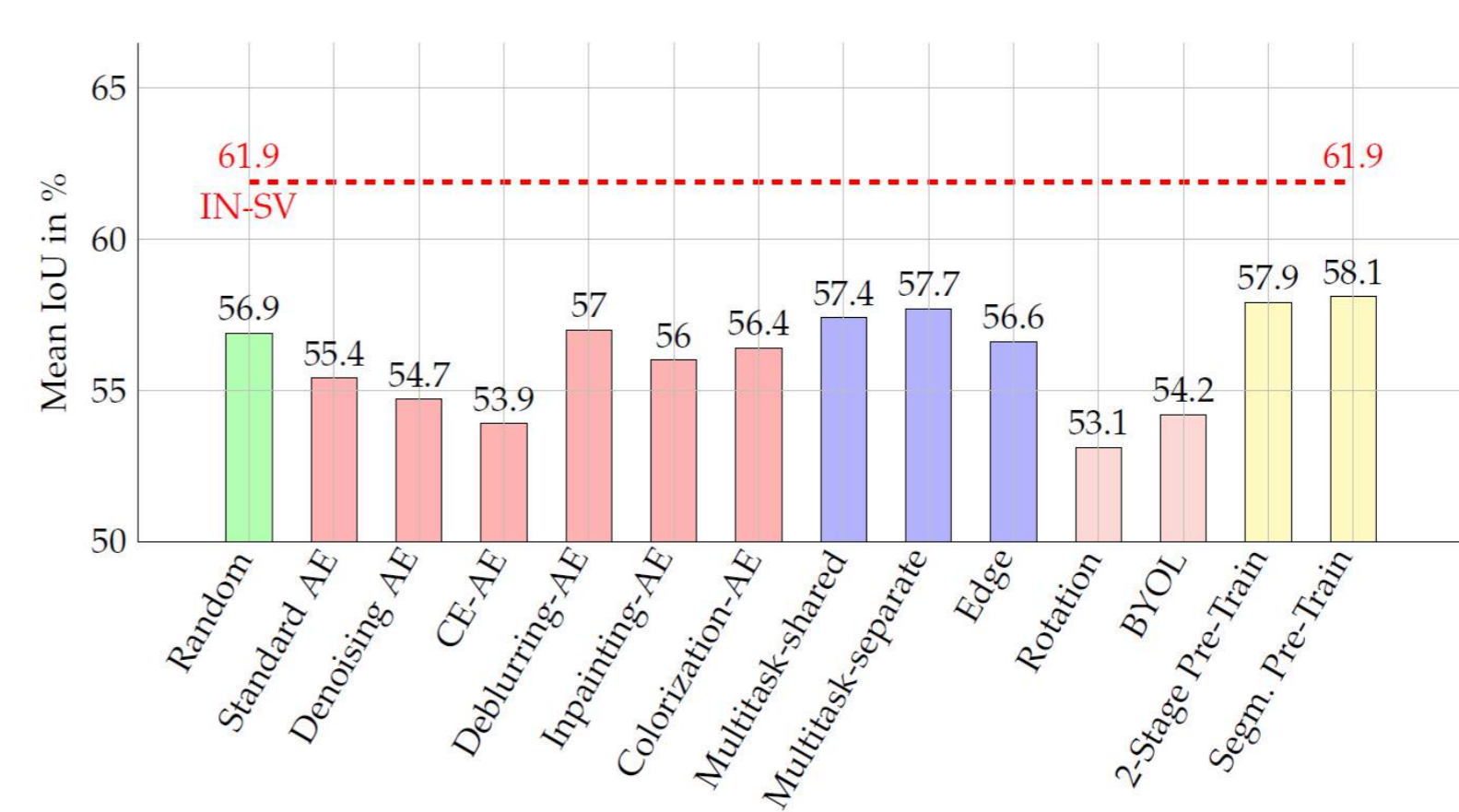


Figure 4: mIoU for fine-tuning on Segmentation (© CARIAD SE)

- None of the evaluated pretext tasks performed competitive with ImageNet pre-training
- Only slight improvements over random initialization → inefficient pre-training

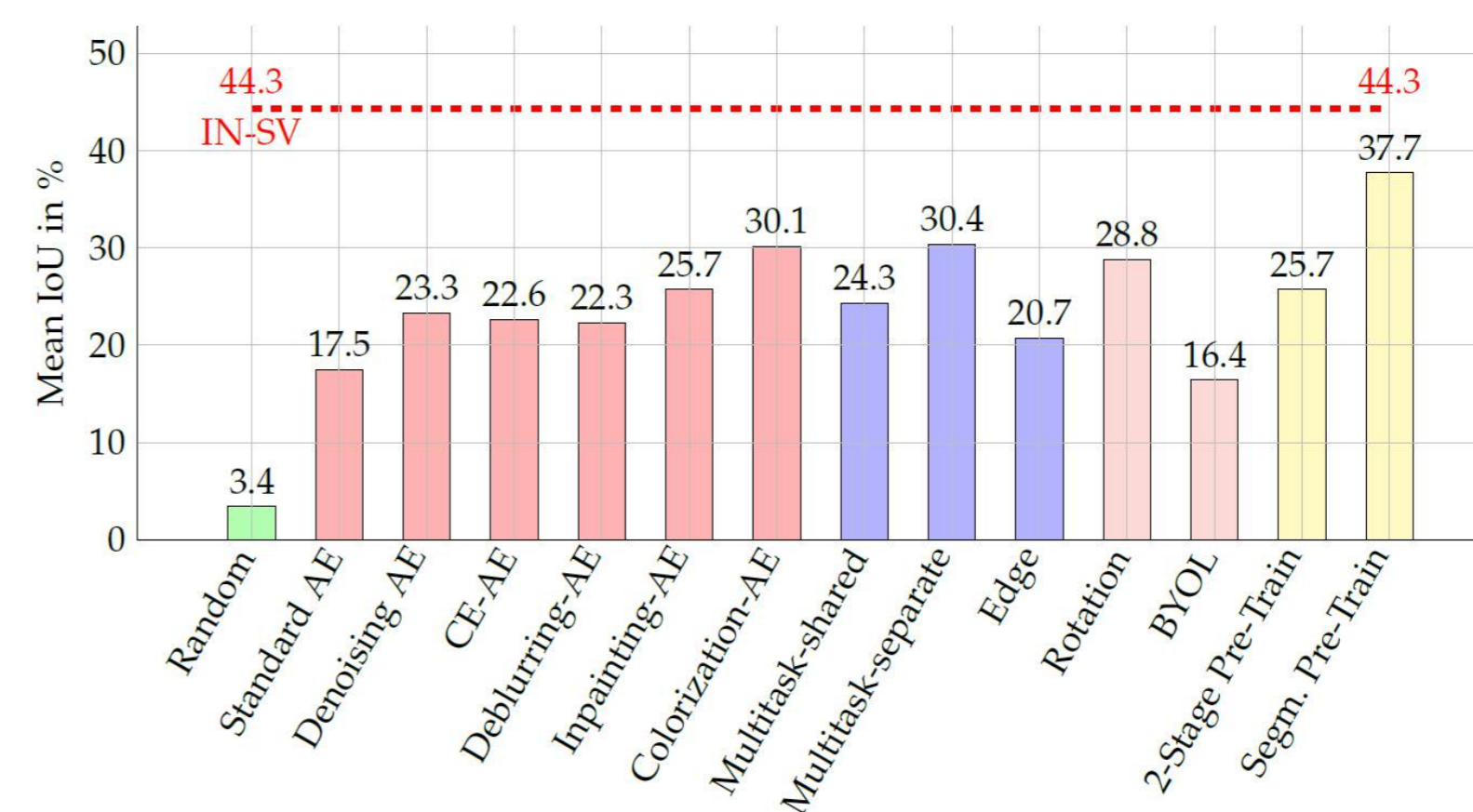


Figure 5: mIoU for frozen encoder (© CARIAD SE)

- Experiments with a frozen encoder revealed that the representations are semantically useful but not competitive with ImageNet pre-training
- Significantly better performance than random init; multitask performs best

Results II

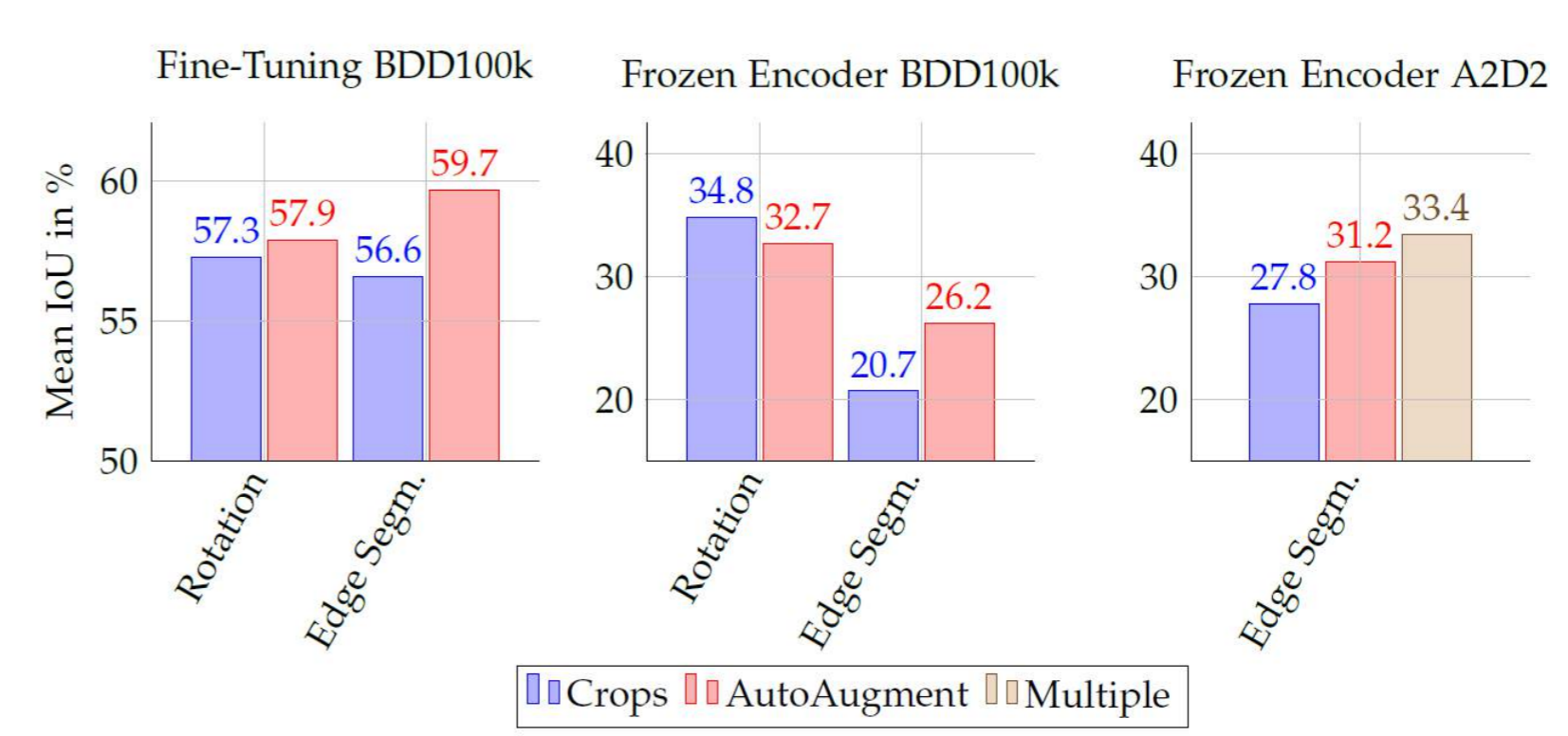


Figure 6: Pre-training under augmentations (@ CARIAD SE)

- When using augmentations during pre-training significant performance gains could be achieved across datasets and pretext tasks

Conclusions

- Self-supervised pre-training on automotive datasets cannot compete with ImageNet pre-training
- Lots of different impacting factors
- Pre-training on one million images provides no gain; augmentations are promising

Partners



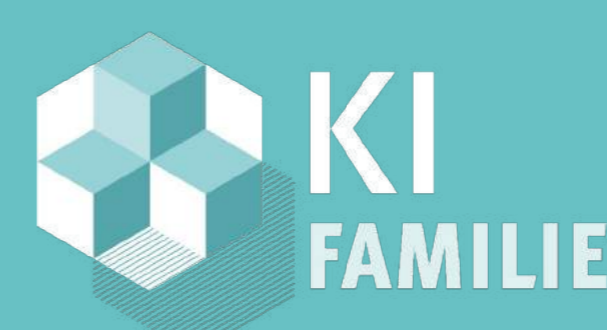
External partners



For more information contact:

nico.schmidt@cariad.technology
manuel.schwonberg@cariad.technology

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 Climate Action.



Supported by:

