TRA
TRANSPORT
RESEARCH
ARENA
BUDAPEST
18-21/05/26

Lane Segmentation from Street-Level Imagery via Noisy Label Generation and Contrastive Self-Supervision

JULIA ALVES PORTO, APOSTOLOS ZIAKOPOULOS, DANIEL FELIPE LOPEZ, GEORGE YANNIS

21/05/2026

INTRODUCTION

BACKGROUND AND MOTIVATION

BACKGROUND

- Lane detection has an important role in **automated navigation tools** and **behaviour monitoring**
- When well maintained, road lane markings have very strong **shape and colour cues**
- In those cases, lane detection can be done with **geometry and colour transformation tools**
- For more complex scenarios, there is availability of **benchmark datasets**
- **AI models** with different configurations have been published for lane detection or segmentation

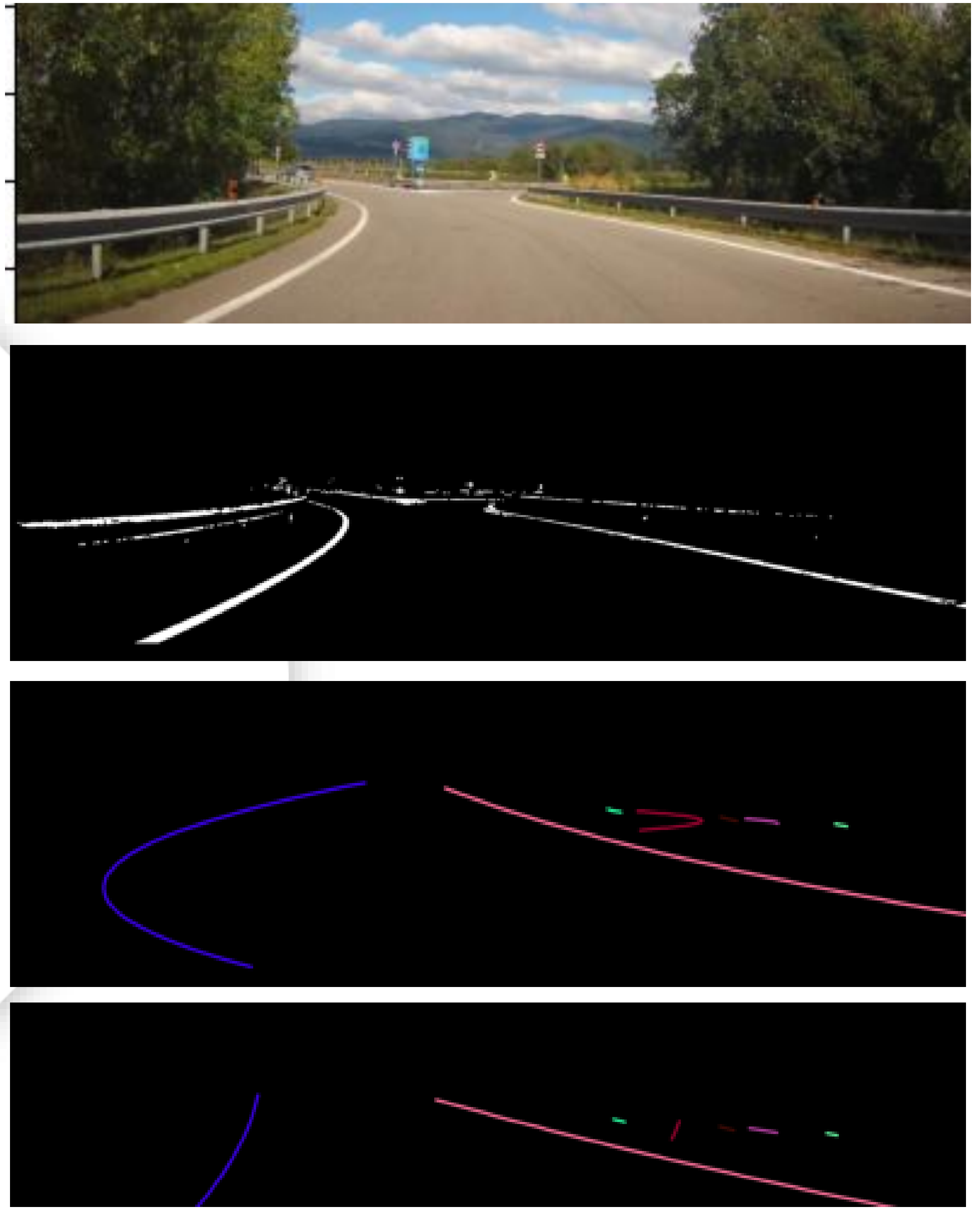
MOTIVATION

- Label generation using geometric cues
  - Mohammadi, H.: Road Lane Line TuSimple Dataset Preparation. Kaggle.
  - Soumya044: Lane Line Detection. Kaggle.
- Contrastive learning for lane segmentation improvement
  - Zoljodi et al., 2024: Contrastive Learning for Lane Detection via cross-similarity
  - Nie et al., 2025: LaneCorrect: Self-Supervised Lane Detection
  - Khan et al., 2025: Adapt, But Don’t Forget: Fine-Tuning and Contrastive Routing for Lane Detection under Distribution Shift

METHODS
AND DATASETS

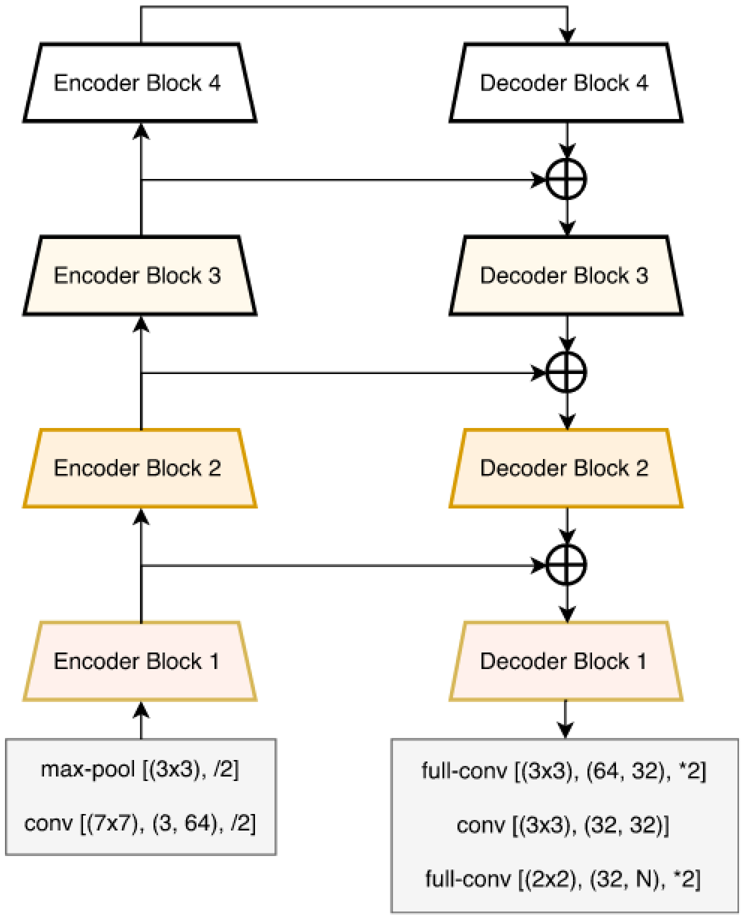
NOISY LABEL GENERATION

- **Color thresholding**
  - RGB pixels >= [150, 150, 150], or
  - RGB pixels >= [160, 160, :] and <= [:, :, 100]
- **Edge detection**
- **Countour detection**
- **DBSCAN clustering**
- **Polyline fitting**
  - L2 regularization
- **Merging similar lanes**
  - Angle difference
  - Distance difference
  - Define endpoints for continuity detection
- **Fitting to linear function**



FULLY SUPERVISED SEGMENTATION LEARNING

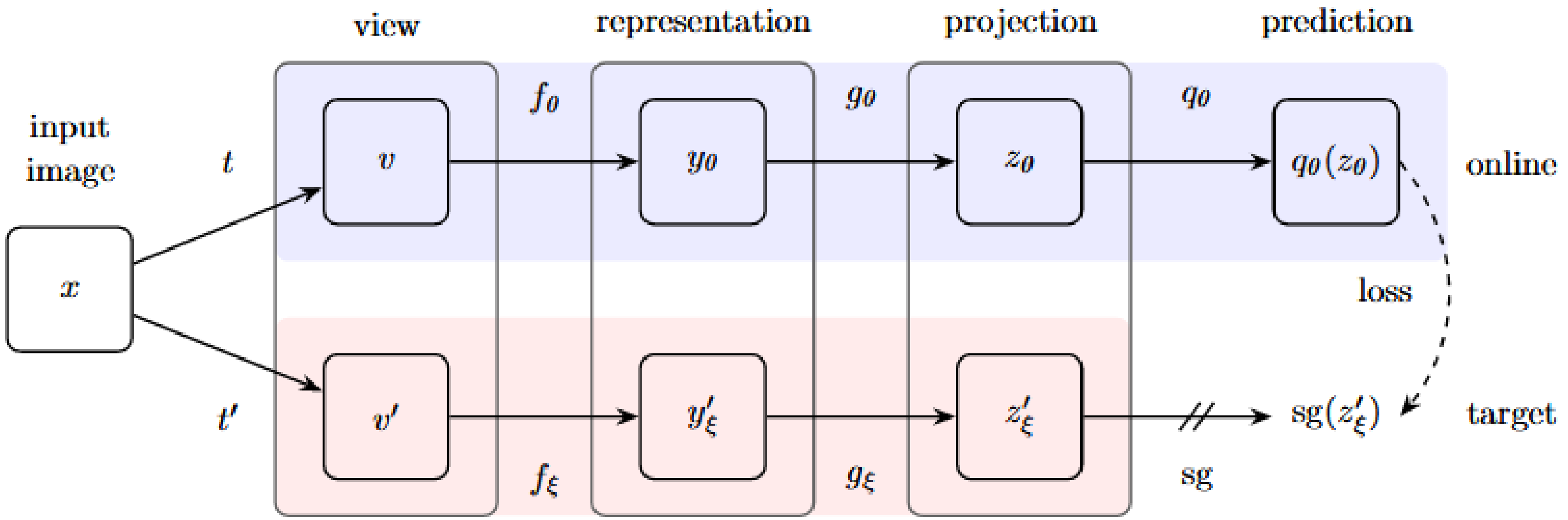
- **LinkNet Model** (Chaurasia and Culurciello, 2017)
  - Hyperparameters:
  - 4 layers on the decoder
  - ResNet backbone (He et al., 2016)
  - No pretrained weights
  - Activation: ‘sigmoid’ or None
  - Segmentation Models Pytorch library (Iakubovskii, 2019)
- Loss Function:
  - Binary Cross Entropy Loss
  - Binary Cross Entropy With Logits Loss
- Optimization:
  - Adam
  - Learning rate: 0.001



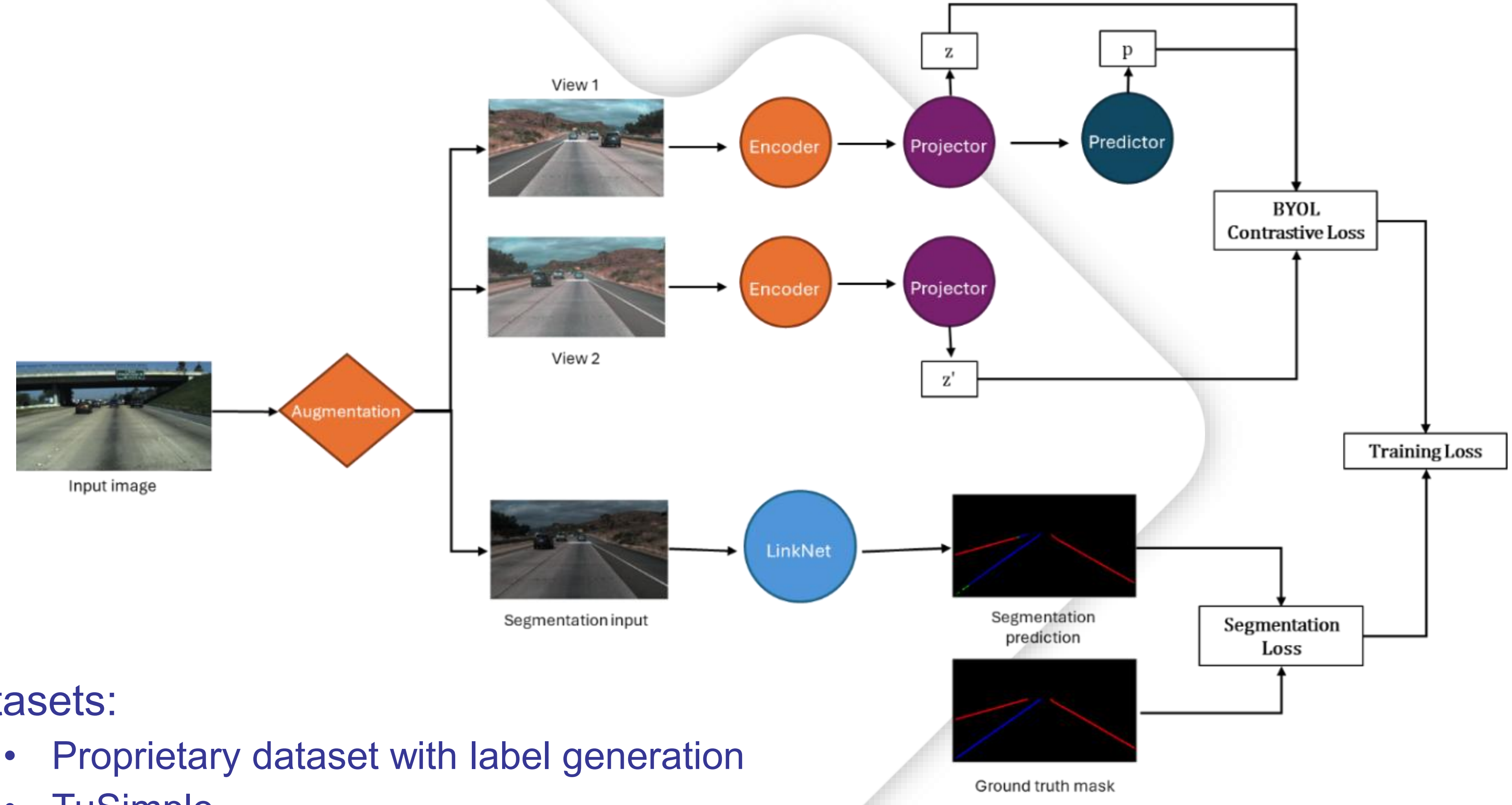
SELF SUPERVISED CONTRASTIVE LEARNING

- **Bootstrap Your Own Latent – BYOL** (Grill et al., 2020)
  - Does not require negative examples
  - Uses asymmetrical loss to avoid collapsing
  - Two differently augmented views of the same image
  - Loss: 2 – 2 * cosine similarity between projection and prediction features

$$\mathcal{L}_{\theta,\xi} \triangleq \left\|\overline{q_\theta}(z_\theta) - \overline{z}'_\xi\right\|_2^2 = 2 - 2 \cdot \frac{\langle q_\theta(z_\theta), z'_\xi \rangle}{\left\|q_\theta(z_\theta)\right\|_2 \cdot \left\|z'_\xi\right\|_2}.$$



WEAK SUPERVISION APPROACH



- Datasets:
  - Proprietary dataset with label generation
  - TuSimple

RESULTS

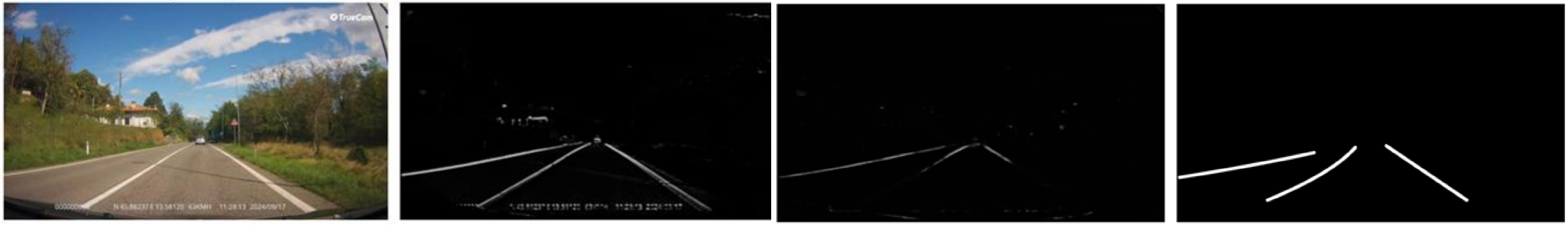
PROPRIETARY DATASET

- While **Weak Supervision** was still subject to **noise, Full Supervision** presented **high precision but low recall** (i.e., smaller confidence when defining lane pixels)

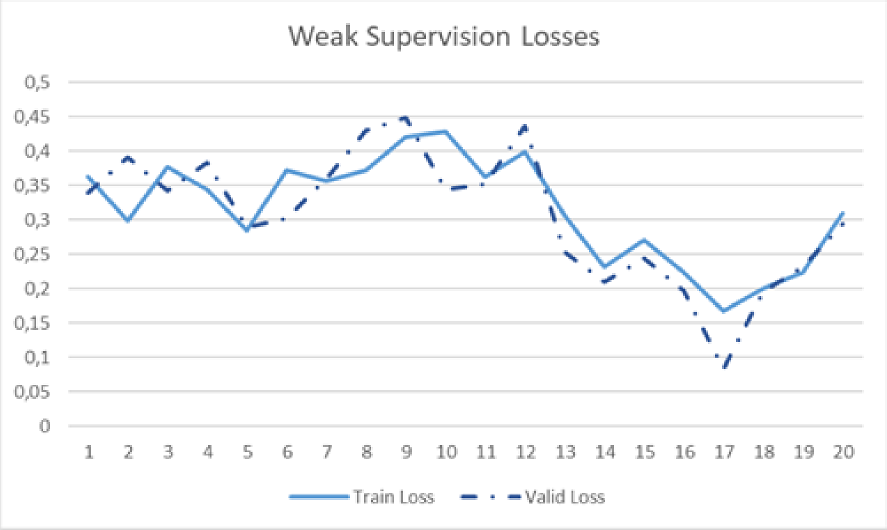
- **Weak Supervision learning was mainly driven by the BYOL Loss** and converged to a higher loss value in the segmentation branch compared to Full Supervision

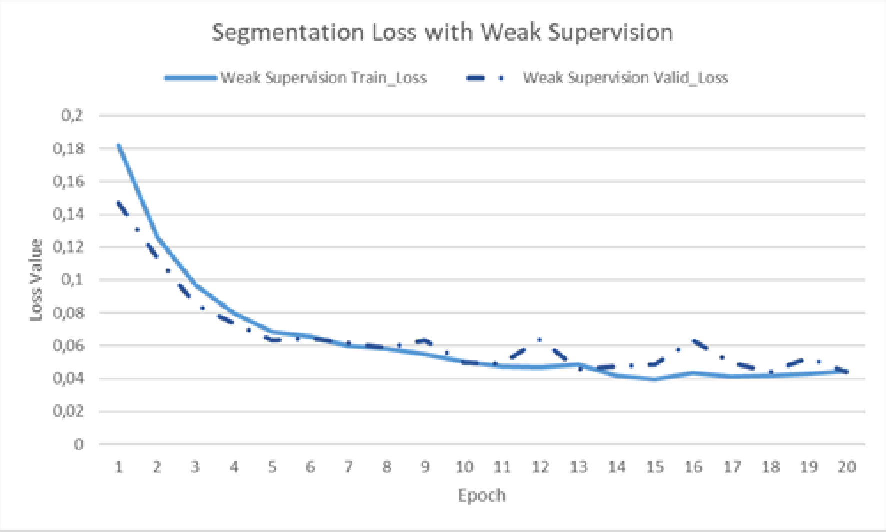
- Is the model **learning different representation** in its encoder or has it **not fully converged?**

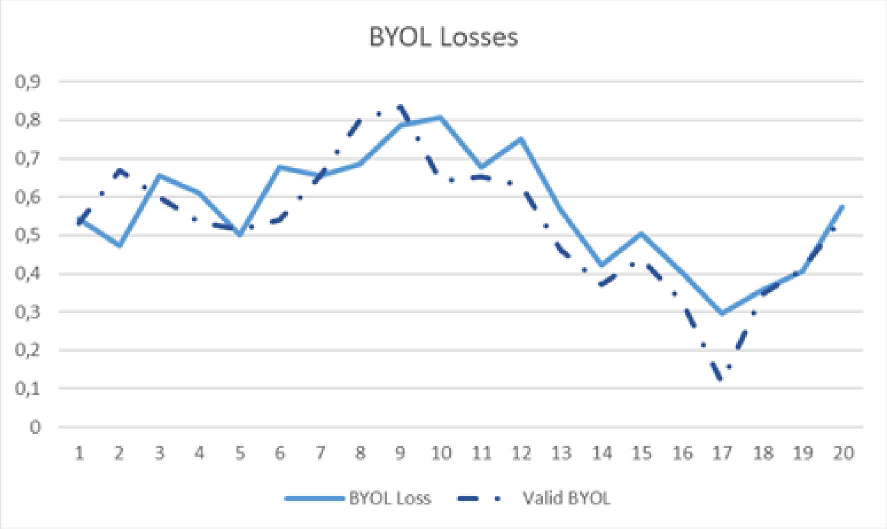
Original Image | Predicted Mask | Predicted Mask | Ground Truth Mask

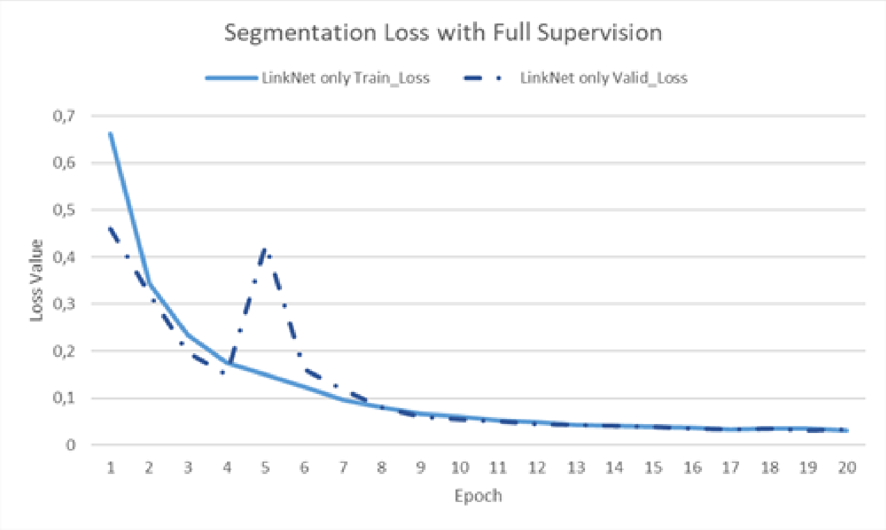


PROPRIETARY DATASET (LOSS)

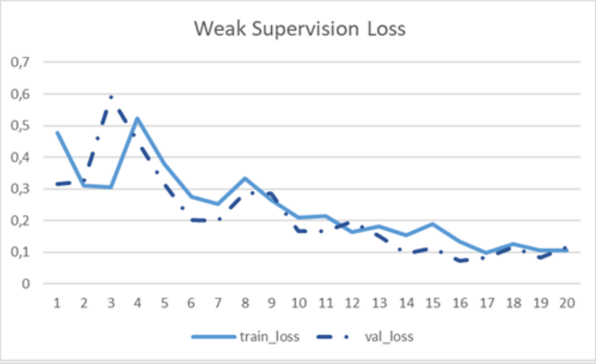


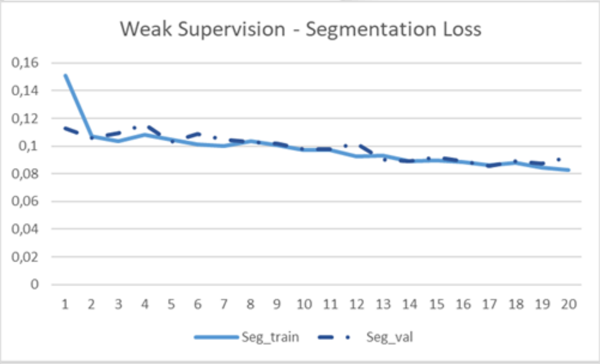


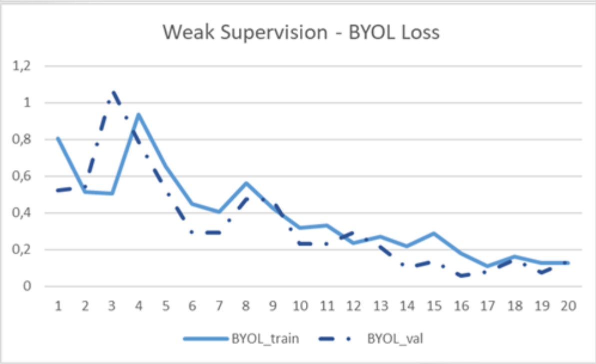


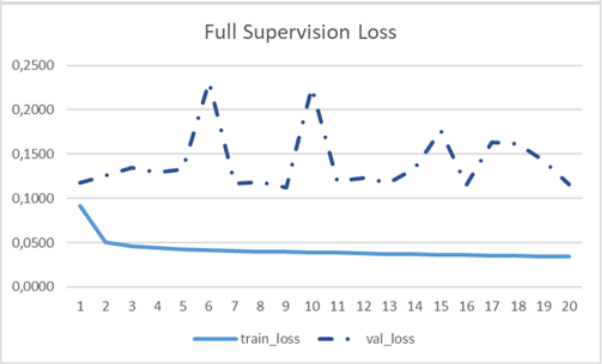


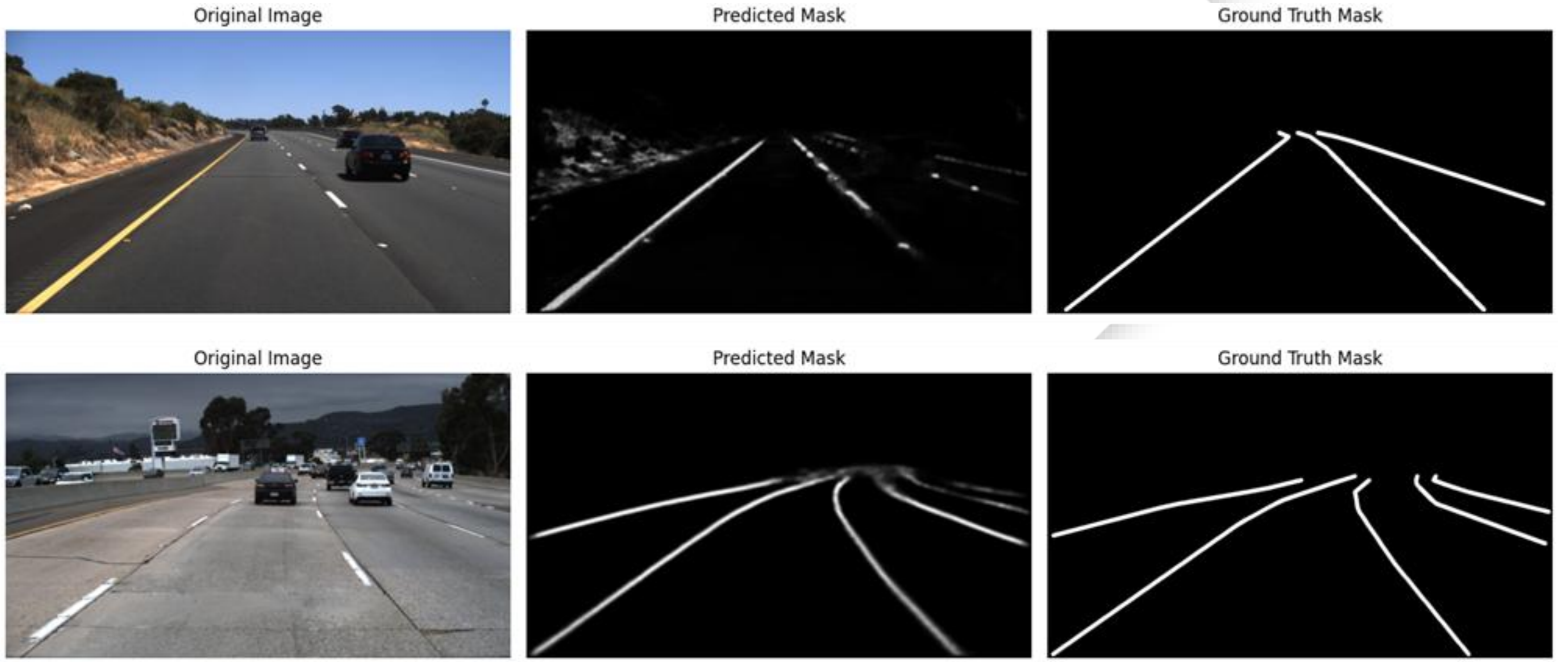
TUSIMPLE DATASET







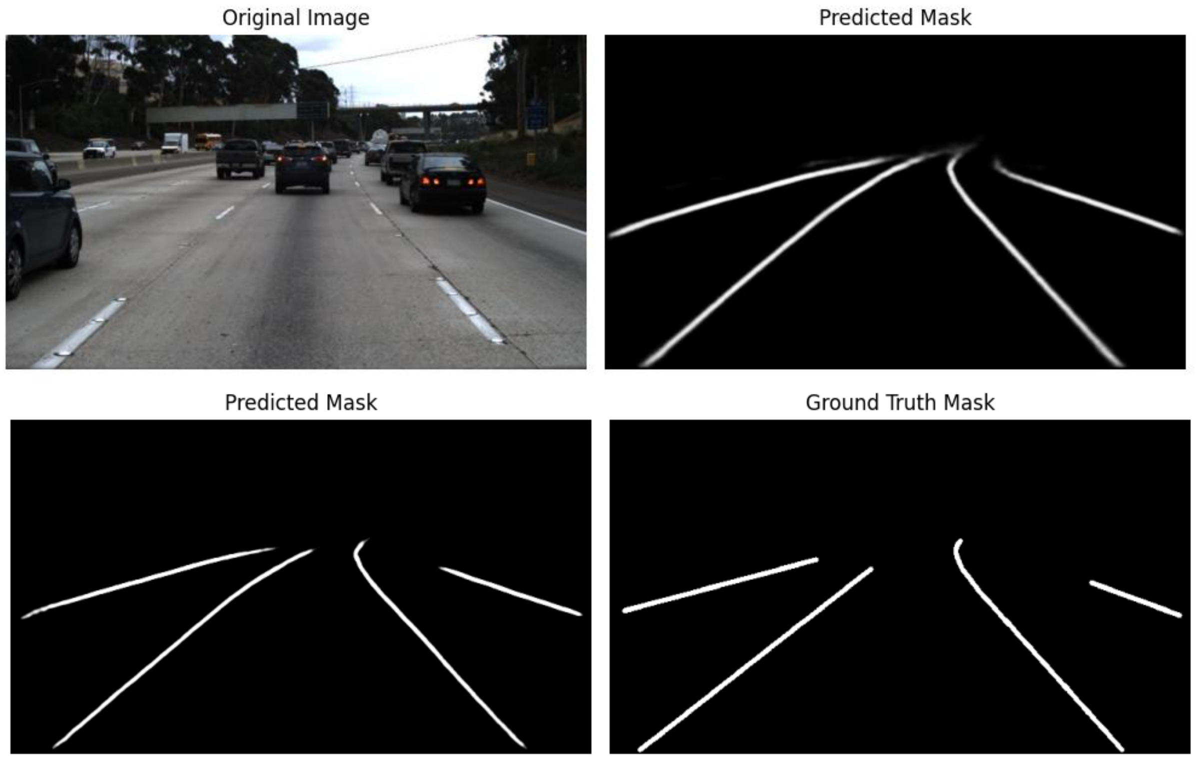


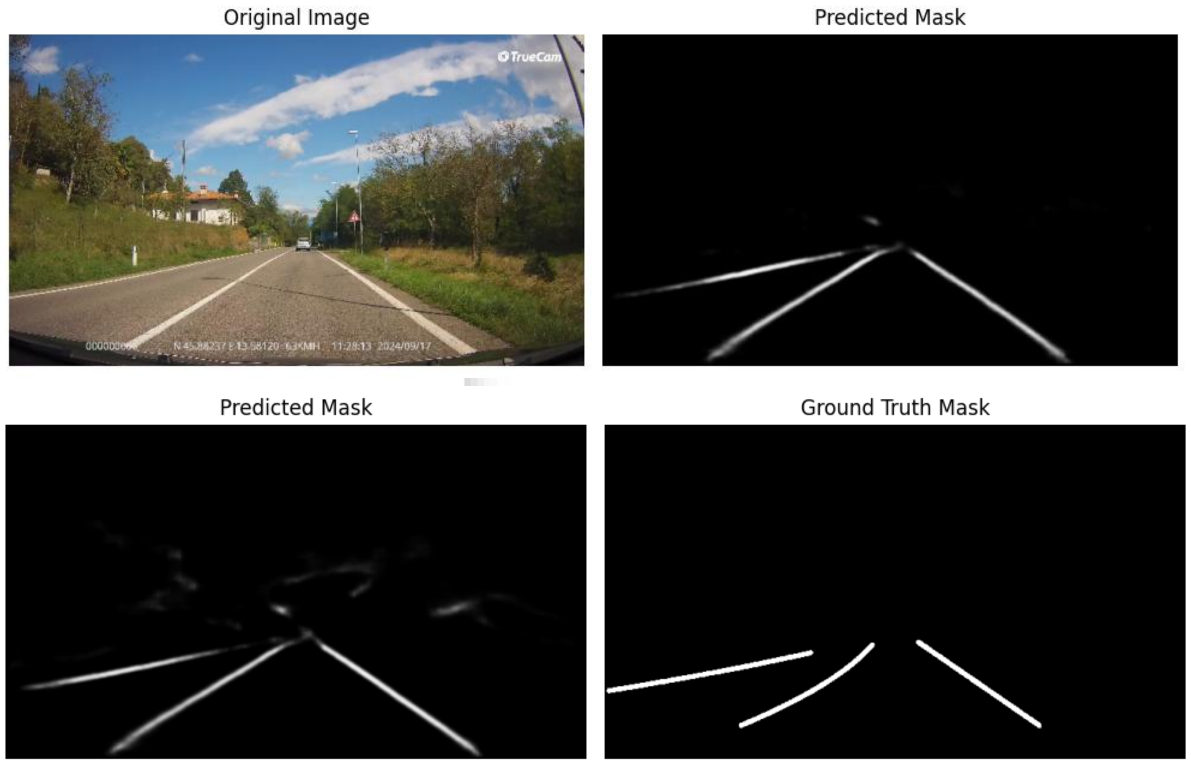


ADRESSING REVIEWERS COMMENTS

- Dataset was reduced from 720x1080 to **368x640** pixels
- Training configurations:
  - **Early stop** criteria: stop training after validation loss doesn't show improvement for 20 consecutive epochs
  - **Fixed number of epochs**, but saving only the best model at **20, 50, 100 and 300** epochs
  - **Transfer learning**: metrics using Proprietary dataset for training and TuSimple dataset for testing, and vice-versa
  - **Fine Tuning**: training the encoder only with self-supervision, or training half the epochs with full supervision and half with weak supervision
- Overall, the same behavior was noted: **lower but steady convergency** with the Weak Supervision model and output possibly subject to **noise**

EXAMPLE: 50 EPOCHS





PERFORMANCE UNDER DIFFERENT THRESHOLDS

Full Supervision training with TuSimple dataset for 50 epochs

Data	Threshold	F1-Score	IoU	Precision	Recall
FRED	0.1	0.316806	0.189816	0.223056	0.562959
FRED	**0.3**	**0.380016**	**0.236343**	**0.368388**	**0.401518**
FRED	0.5	0.345407	0.210838	0.480225	0.27471
FRED	0.7	0.233763	0.134163	0.589172	0.148708
FRED	0.9	0.045175	0.023388	0.658618	0.023747
TuSimple	0.1	0.668512	0.503822	0.532467	0.901657
TuSimple	**0.3**	**0.723288**	**0.569852**	**0.662734**	**0.798701**
TuSimple	0.5	0.713328	0.558639	0.752275	0.68058
TuSimple	0.7	0.640472	0.475006	0.831489	0.523292
TuSimple	0.9	0.374568	0.232736	0.922115	0.237108

PERFORMANCE UNDER DIFFERENT THRESHOLDS

Fine Tuning with WS with TuSimple dataset for 50 epochs

Data	Threshold	F1-Score	IoU	Precision	Recall
FRED	0.1	0.408361	0.257876	0.301272	0.644645
FRED	**0.3**	**0.448255**	**0.291532**	**0.411743**	**0.500581**
FRED	0.5	0.424694	0.273274	0.494907	0.380394
FRED	0.7	0.355547	0.205625	0.577257	0.24345
FRED	0.9	0.115339	0.063013	0.699203	0.06942
TuSimple	0.1	0.679183	0.515951	0.543464	0.908371
TuSimple	**0.3**	**0.729922**	**0.577879**	**0.665256**	**0.810804**
TuSimple	0.5	0.723566	0.57087	0.750613	0.700456
TuSimple	0.7	0.661701	0.498113	0.826659	0.553771
TuSimple	0.9	0.427825	0.274492	0.914126	0.281357

CONCLUSION

CONCLUSION

- The **Label Generator** was successful in providing training labels for the model, albeit limited by the dataset size
- Best IoU scores when binary threshold is defined at **0.3** in most cases
- The Weak Supervision approach seams to **prevent overfitting** and, therefore, allow for better **generalization** of the model
- Future work should consider the application of weak supervision for **multiclass lane segmentation**

Lane Segmentation from Street-Level Imagery via Noisy Label Generation and Contrastive Self-Supervision

TRA TRANSPORT RESEARCH ARENA

BUDAPEST 18-21/05/26

Re-Generation in transport