Lane Detection

Zhanwei Zhang, Jianan Xie

Content

- Background, Motivation, and Related Work
- Innovation
- Methodology
- Experiment
- Conclusion
- Future Work

Background


High-Precision Lane Detection - NVIDIA DRIVE Labs Ep. 7
On the left we see frame-by-frame pixel-level
lane detections from High-Precision LaneNet
0:17 / 1:29
Scroll for details

Motivation

- Visual Interferences
- Environmental Variablility
- Variety of Lane Markings
- Missing Lane Lines
- Multiple Lane Tracking
- Road Changes

Related Work

- Vpgnet[12] uses vanishing point estimation to improve robustness under various conditions.
- A reliable CNN based regression approach[11] is used to detect thin and elongated lane boundaries.
- A newly published work[3] based on transformer can learn BEV representations to eliminate the changes in road height.

Issues and limitations

- model including post-processing stage can take high computational complexity[12].
- many models only take ego lanes into consideration[8] or can only detect fixed numberof predefined lanes[15][11].
- undulating terrain sets up big difficulties for lane detection.
- Intuitively, learnable perspective transformation will overperform fixed one.

LaneNet Architecture

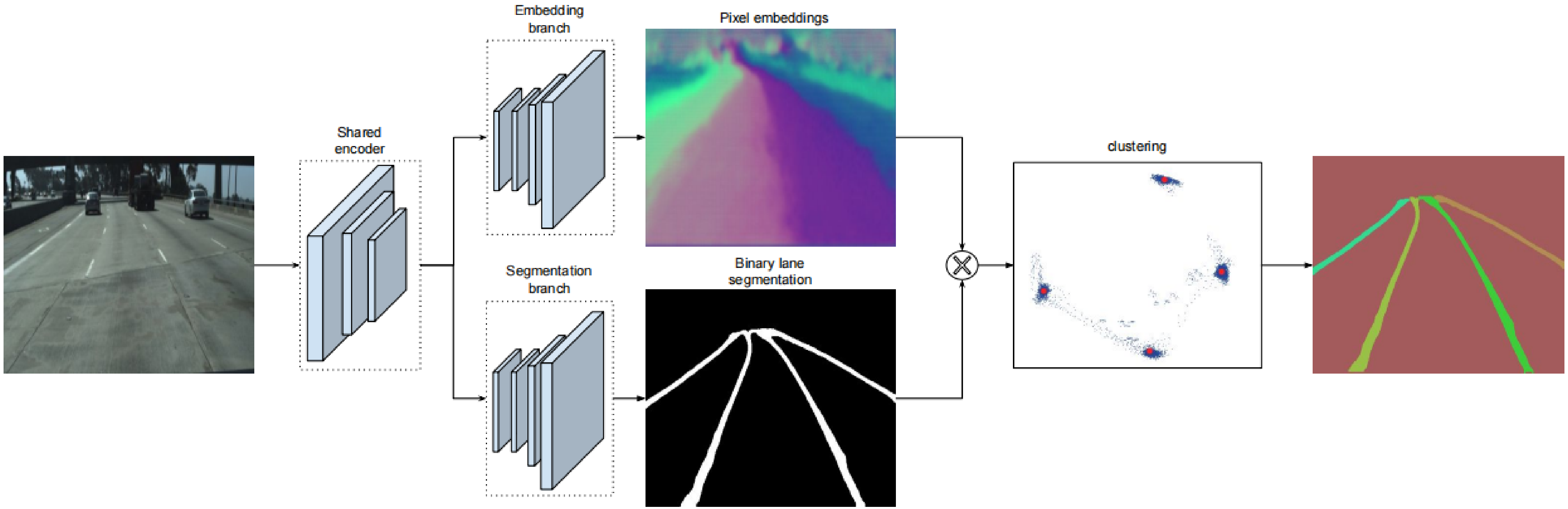


Fig. 2. LaneNet architecture. It consists of two branches. The segmentation branch (bottom) is trained to produce a binary lane mask. The embedding branch (top) generates an N-dimensional embedding per lane pixel, so that embeddings from the same lane are close together and those from different lanes are far in the manifold. For simplicity we show a 2-dimensional embedding per pixel, which is visualized both as a color map (all pixels) and as points (only lane pixels) in a xy grid. After masking out the background pixels using the binary segmentation map from the segmentation branch, the lane embeddings (blue dots) are clustered together and assigned to their cluster centers (red dots).

Initial Result

- We observe the outcomes of applying a pre-trained model from a lane detection study to the SUSCape dataset as shown.

Initial Results

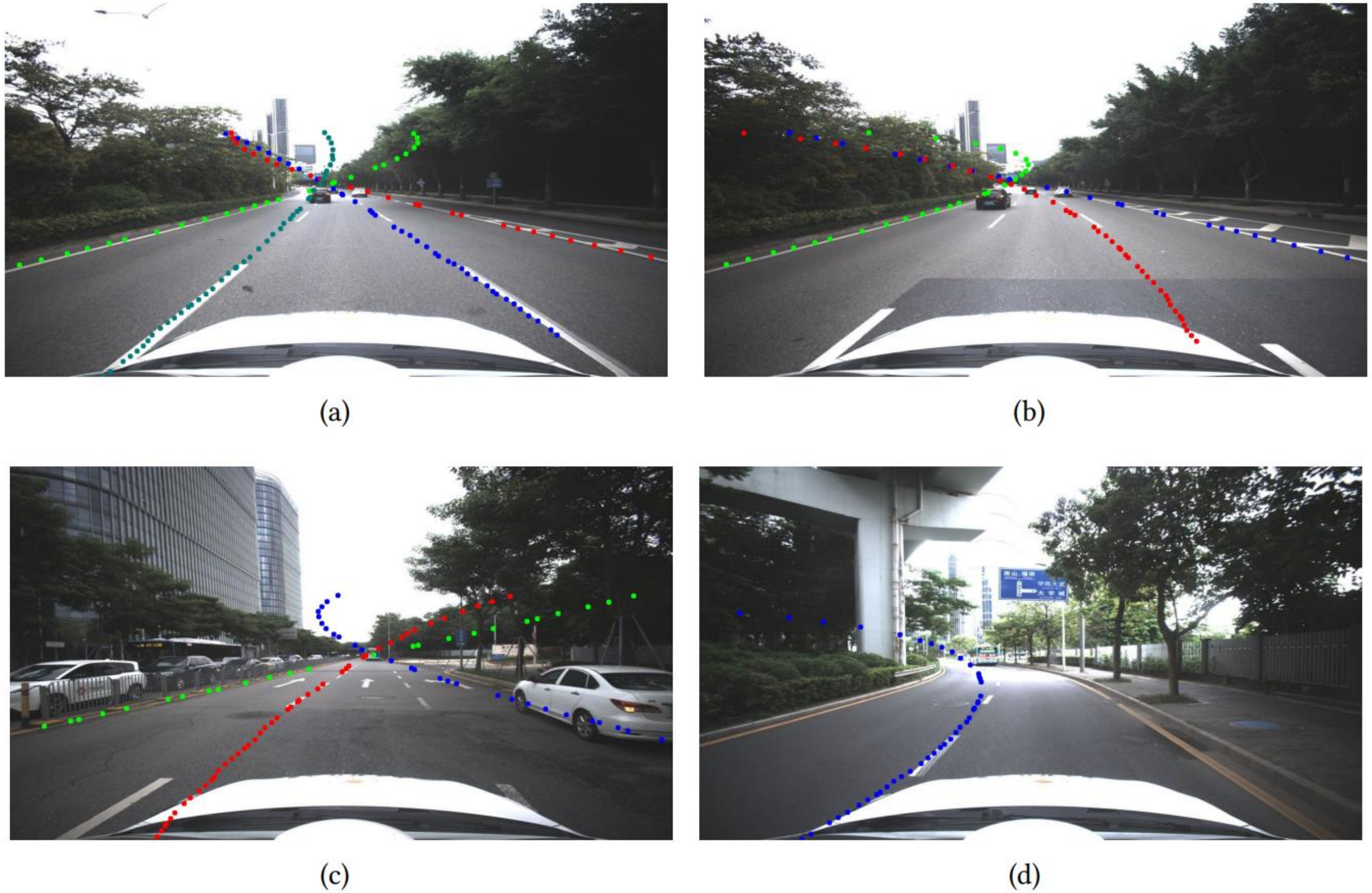


Fig. 1. Pre-trained model in SUSCape dataset

Initial Results

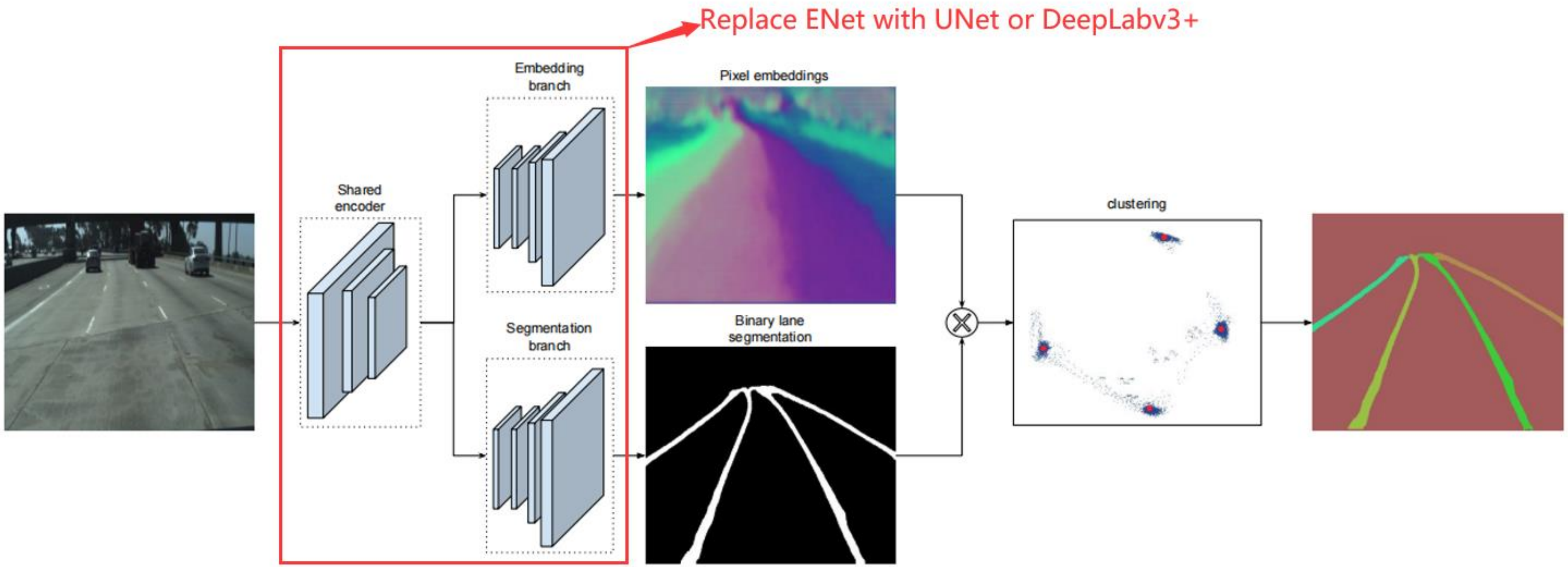
Primary issues were identified in the results:

- Incomplete Lane Detection
- Excessive Lane Fitting Extension

Innovation

- Backbone Improvements
  - ENet
  - UNet
  - DeepLabv3+(Resnet101)
- IPM(inverse perspective mapping) Improvements
  - H-net
  - Transformer(considering)

Methodology



UNet

U-Net: Convolutional Networks for Biomedical Image Segmentation
arXiv:1505.04597 [cs.CV]

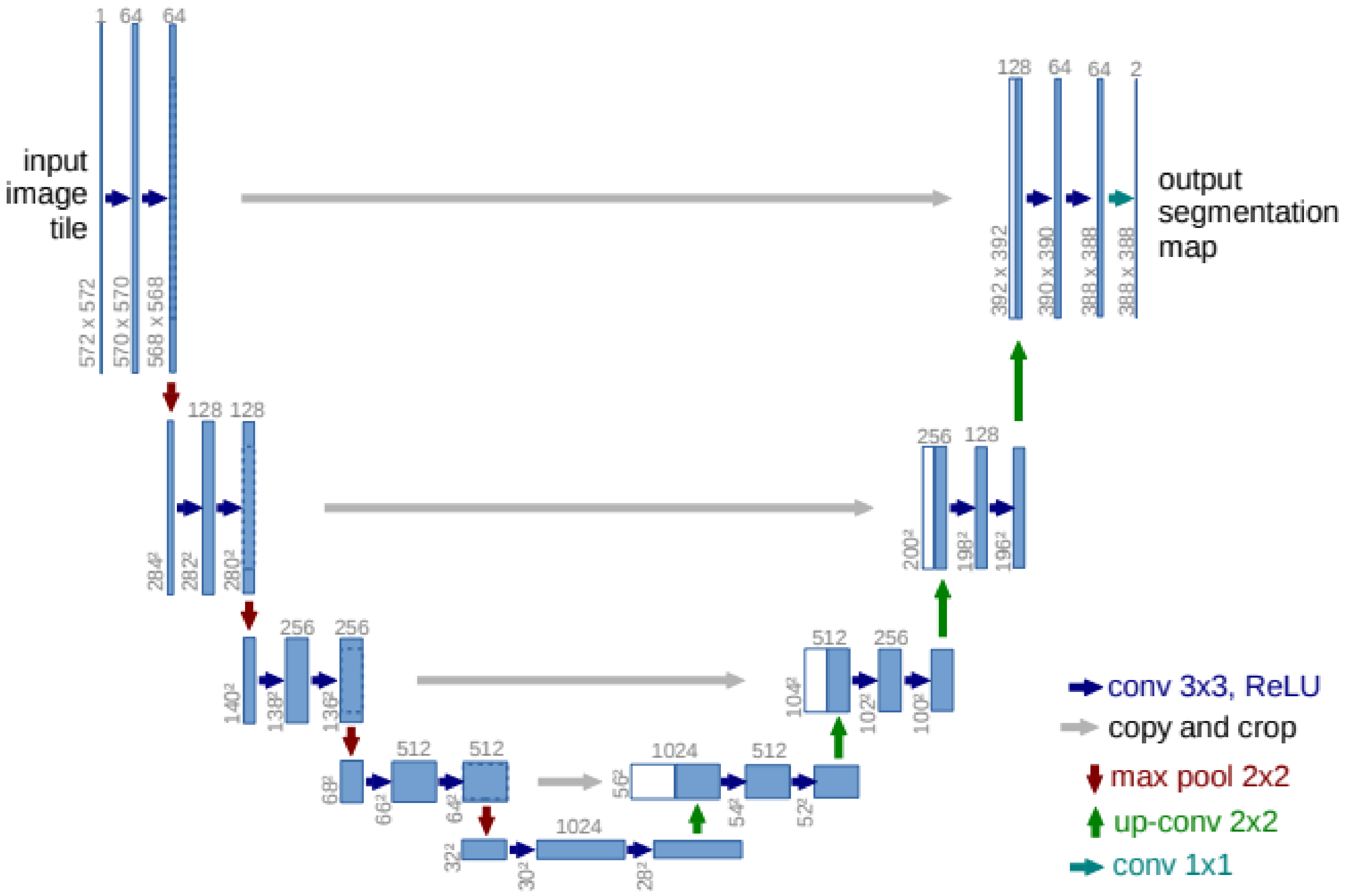
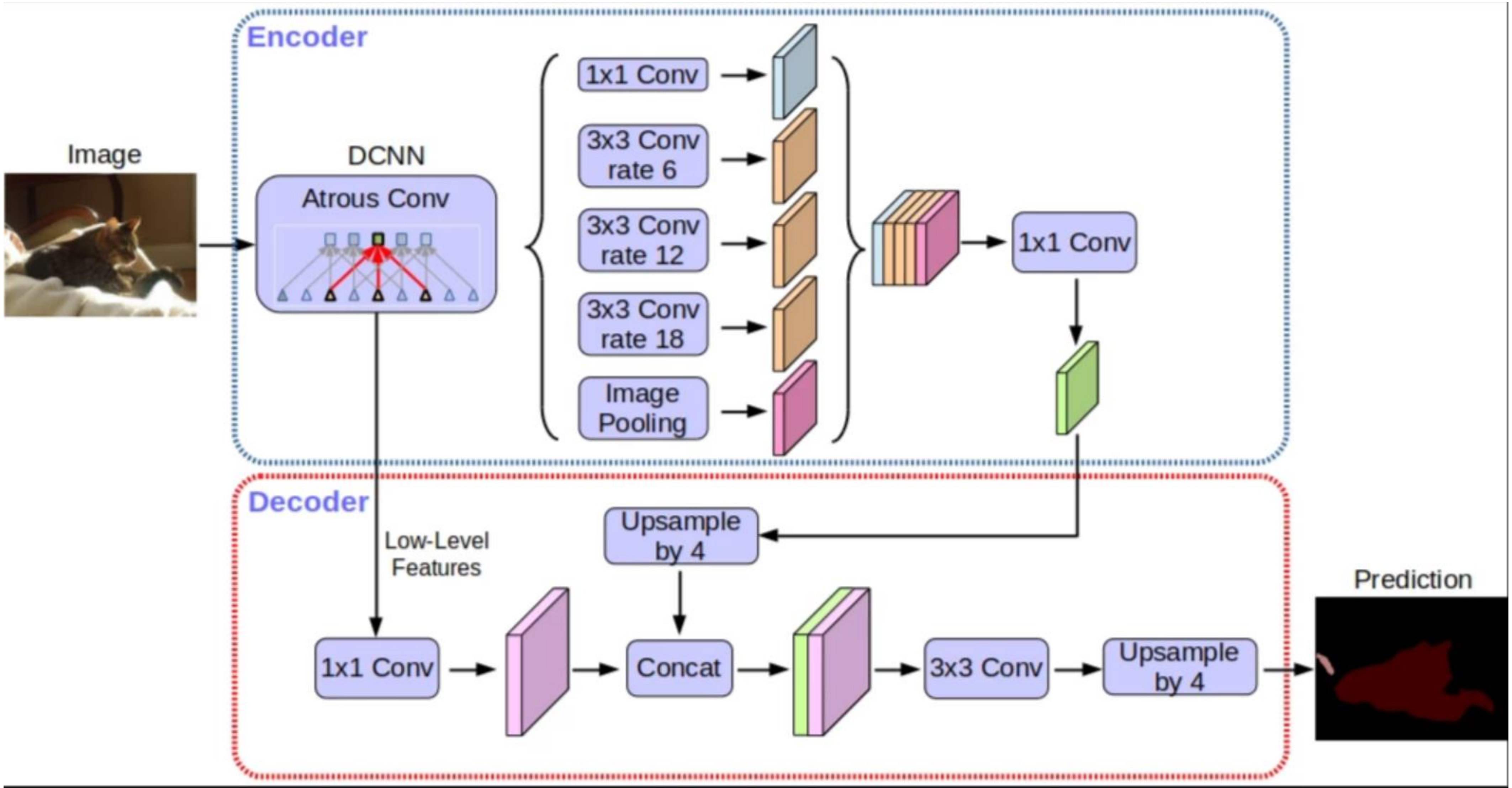


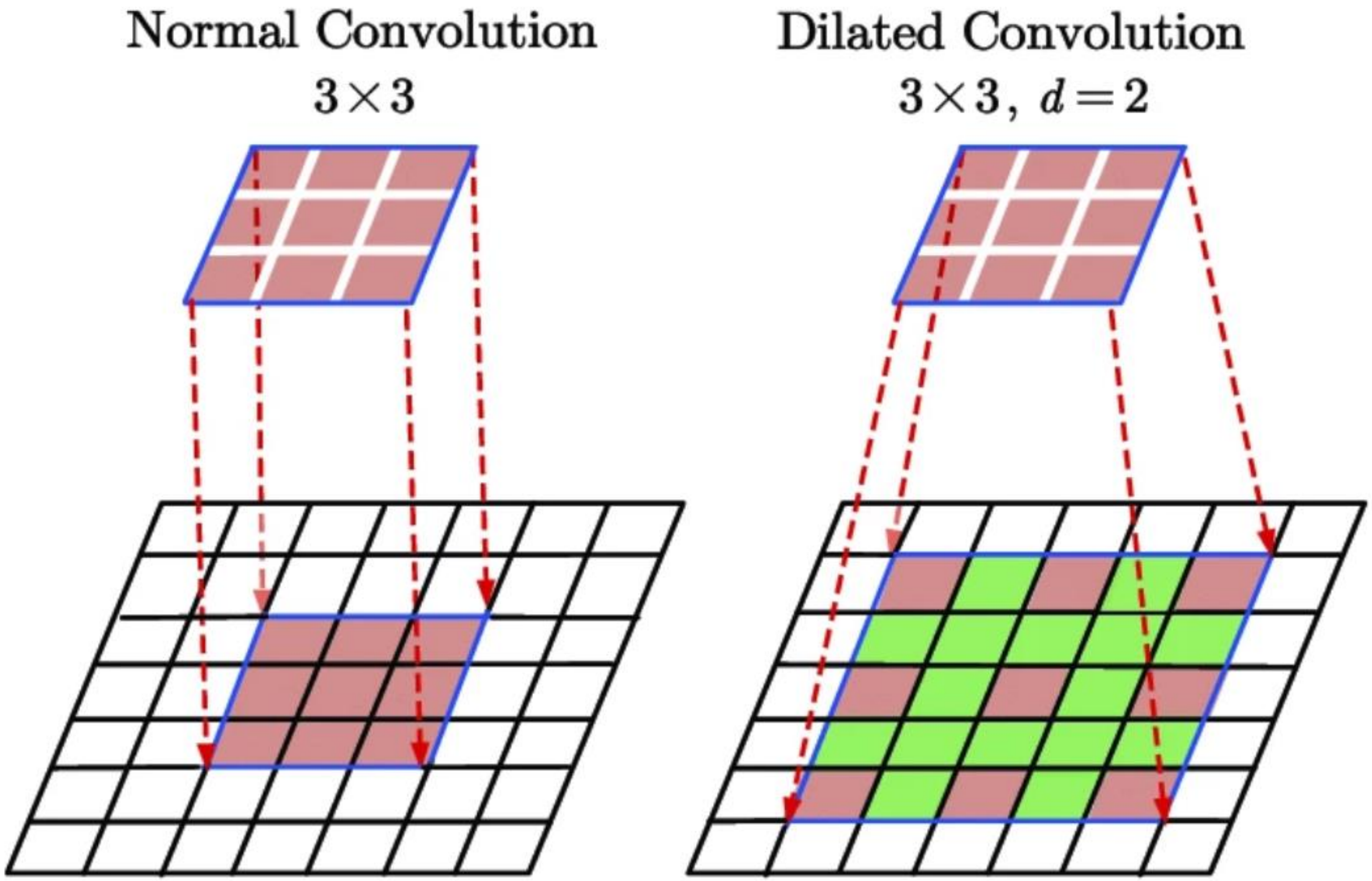
Fig. 1. U-net architecture (example for 32x32 pixels in the lowest resolution). Each blue box corresponds to a multi-channel feature map. The number of channels is denoted on top of the box. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

DeepLabv3+

[2] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation



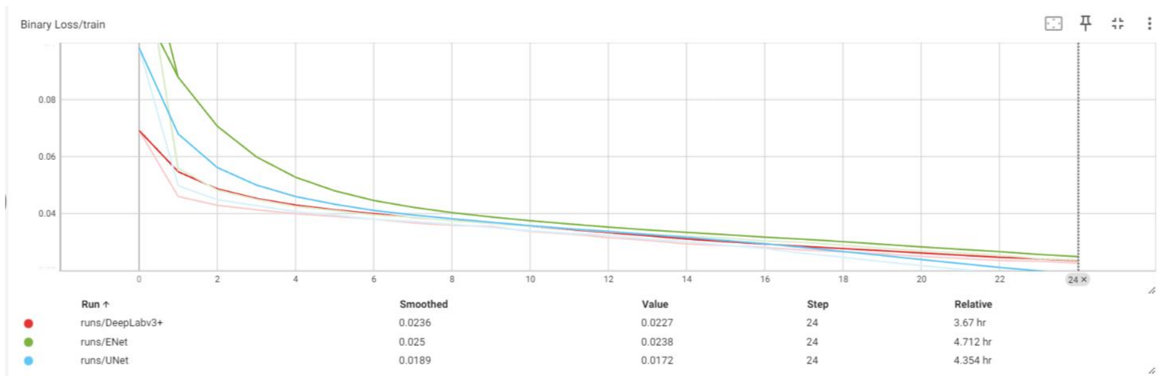
DeepLabv3+: ASPP

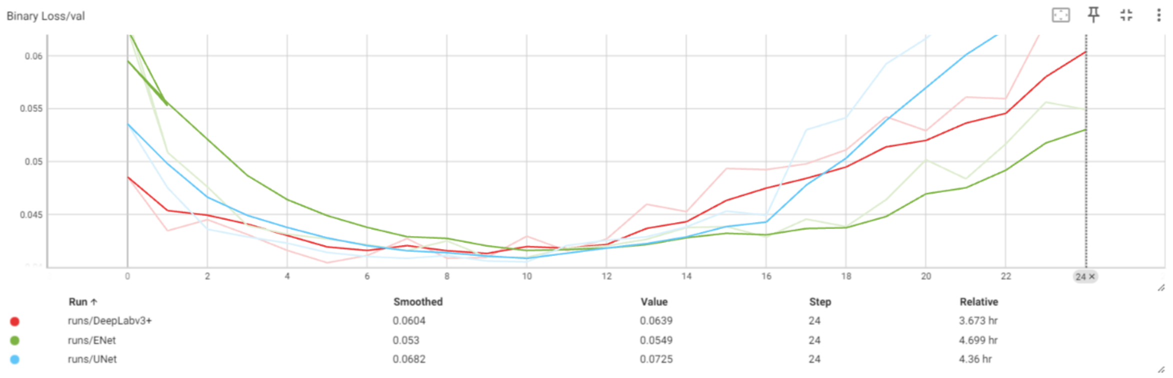


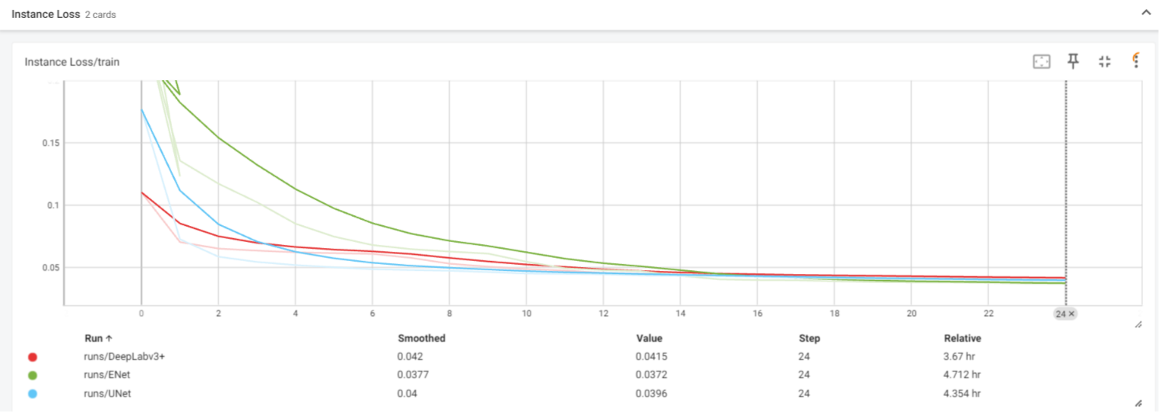
$$FL(p_t) = -\alpha_t(1-p_t)^\gamma \log(p_t)$$

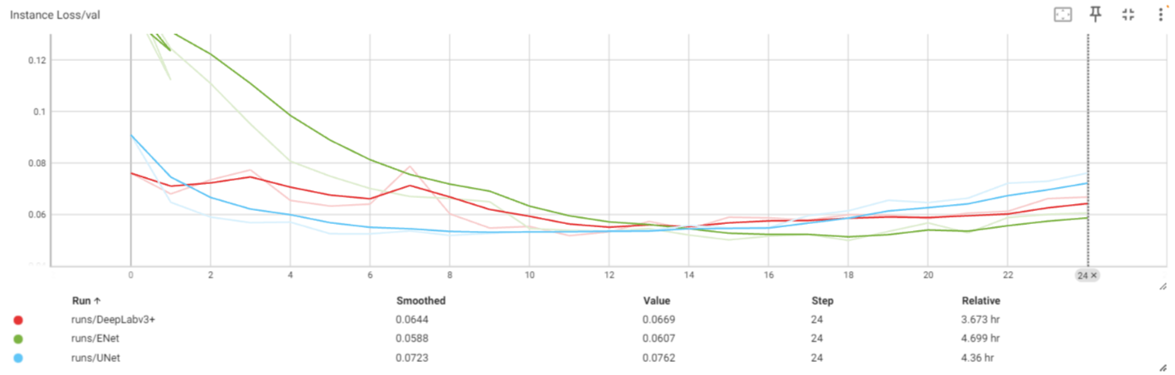
Experiment

- We trained our LaneNet using UNet and DeepLabv3+ as backbone on TuSimple dataset respectively.(Epoch=25, Loss function=Focal Loss)

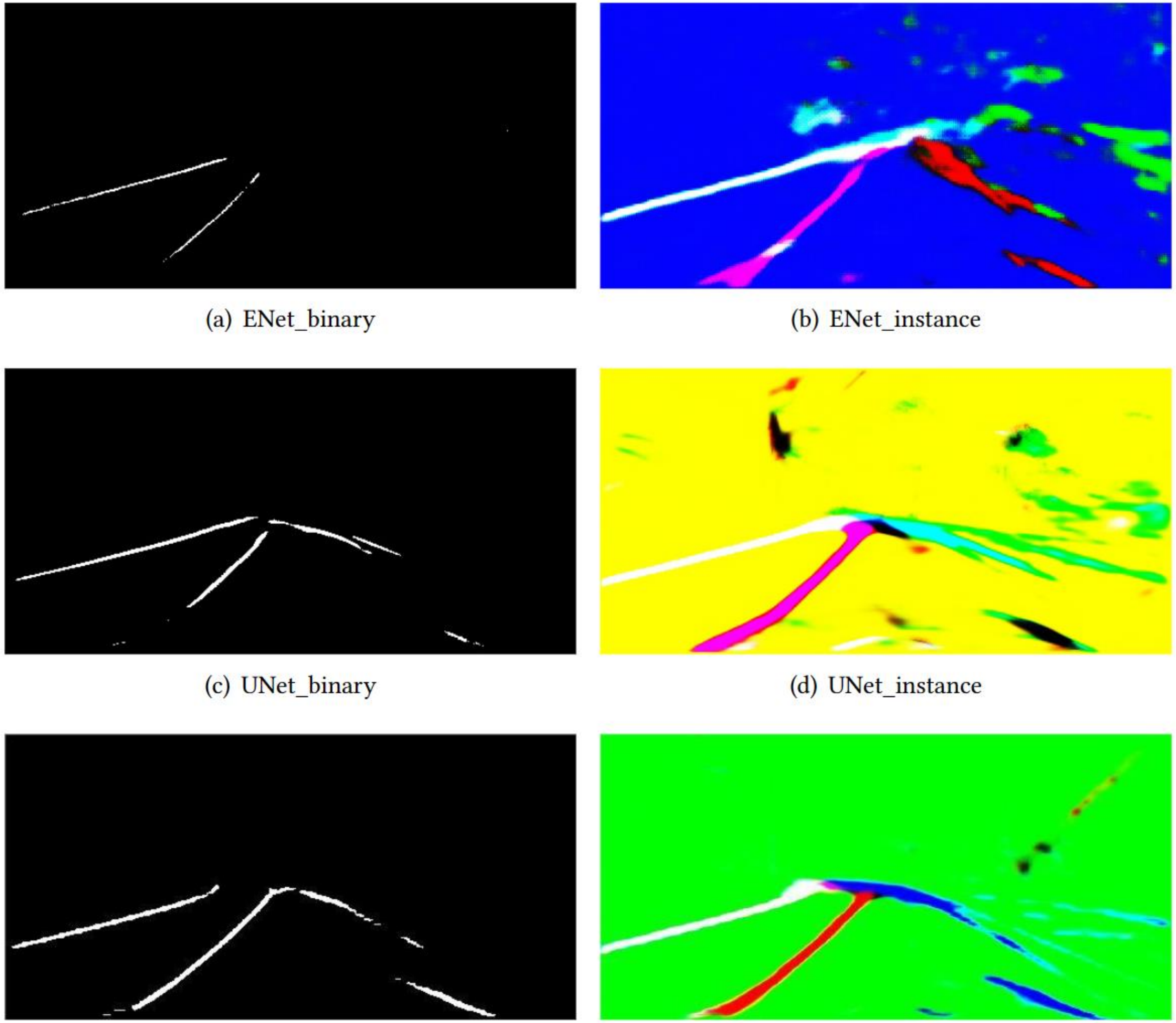








Result



(a) ENet_binary

(b) ENet_instance

(c) UNet_binary

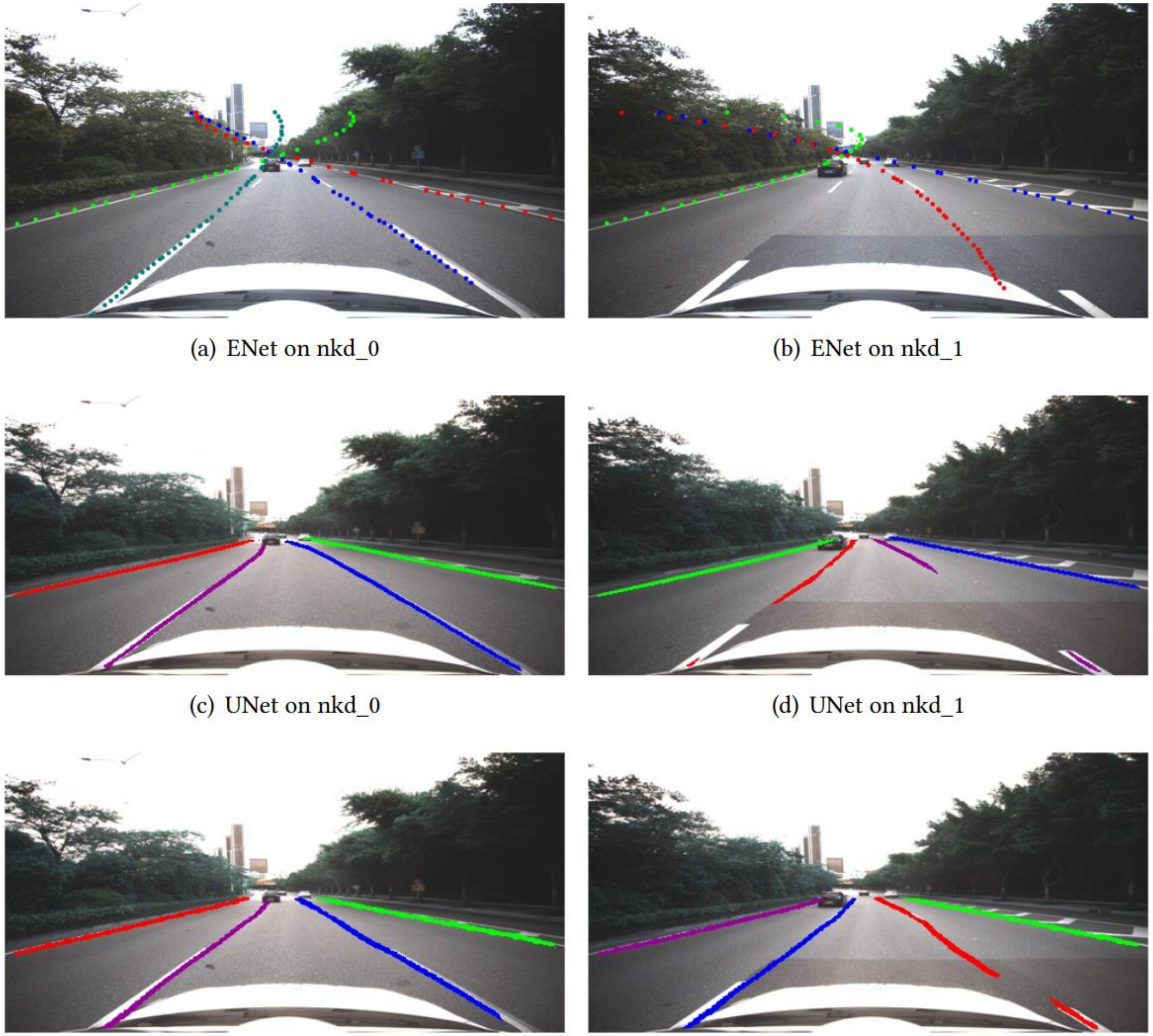
(d) UNet_instance

(e) DeepLabv3+_binary

(f) DeepLabv3+_instance

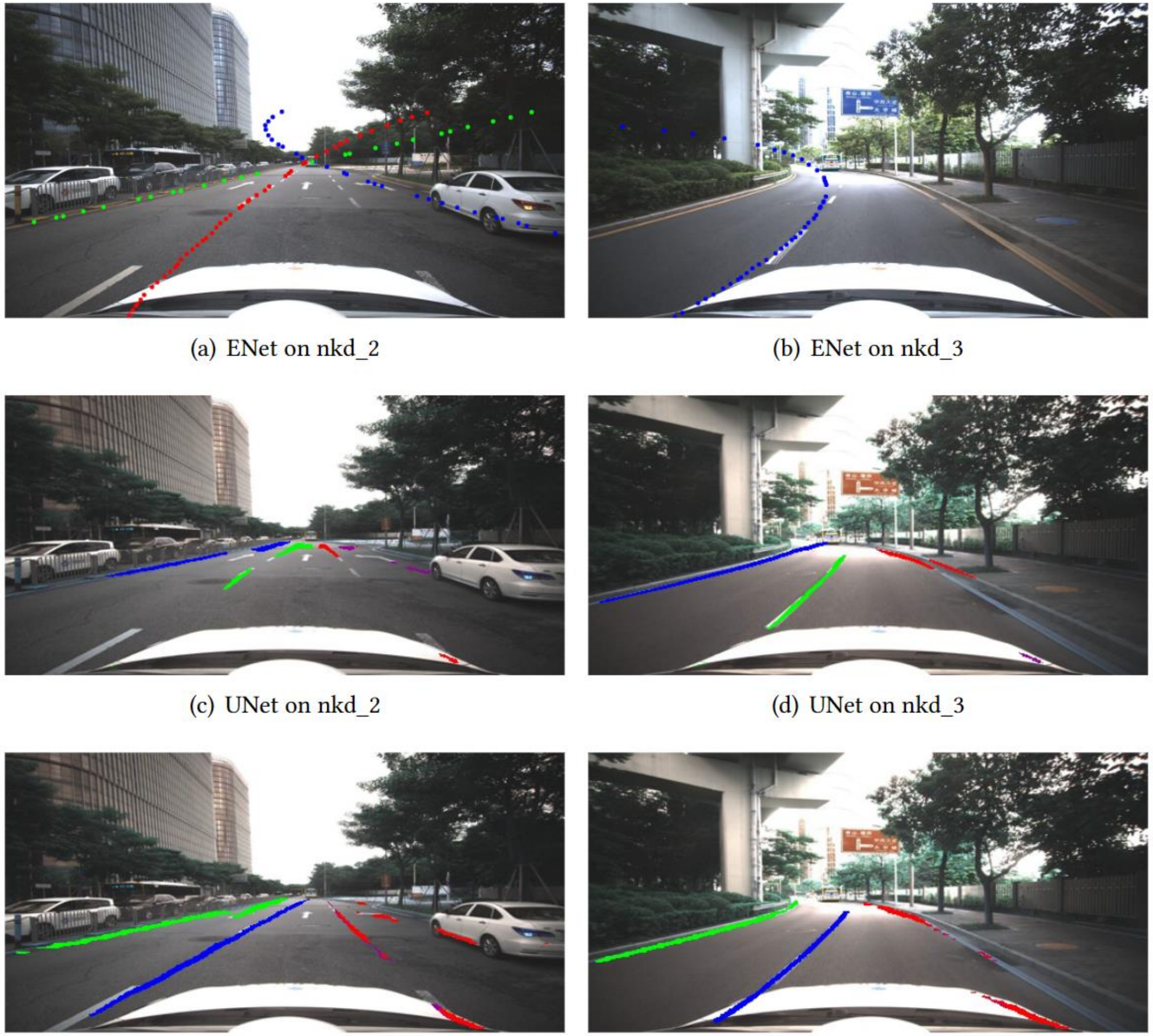
Fig. 2. Binary and Instance segmentation Comparison for nkd_3

Result



(a) ENet on nkd_0
(b) ENet on nkd_1
(c) UNet on nkd_0
(d) UNet on nkd_1
(e) DeepLabv3+ on nkd_0
(f) DeepLabv3+ on nkd_1

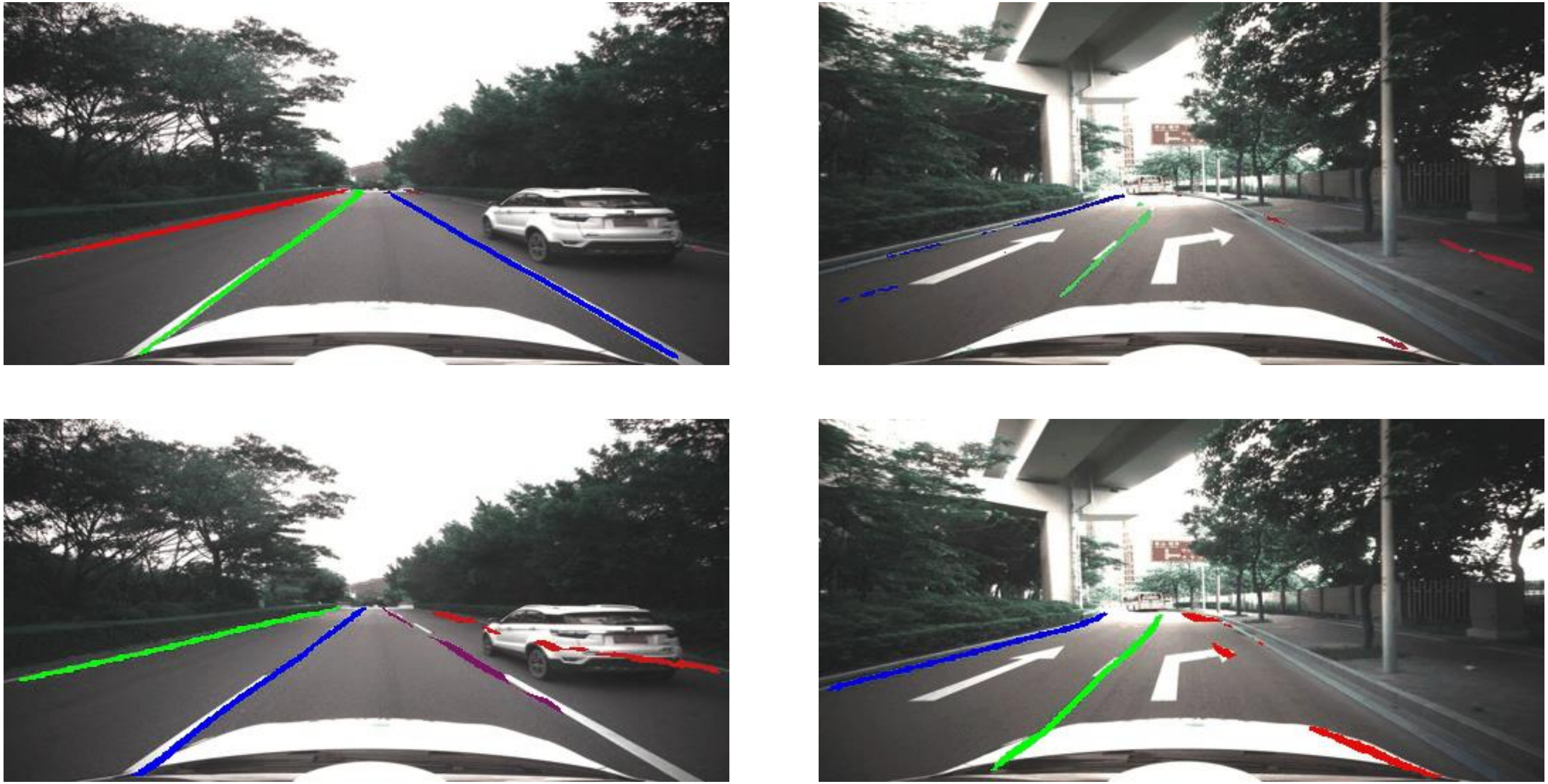
Fig. 2. Mask image comparison(1)



(a) ENet on nkd_2
(b) ENet on nkd_3
(c) UNet on nkd_2
(d) UNet on nkd_3
(e) DeepLabv3+ on nkd_2
(f) DeepLabv3+ on nkd_3

Fig. 3. Mask image comparison(2)

Result



Conclusion

- Our models overperform on real-word scenes than original Enet-based LaneNet.
- DeepLabv3+ based model has more powerful ability to learn the structure of the lanes, and recognizes lanes well within 4~5.

Future Work

- Cope with the color mistakes
- Try to use more powerful encoder-decoder model to learn the lane structure
- H-Net

References

[1] M. T. Smith A. Borkar, M. Hayes. 2012. A Novel Lane Detection System With Efficient Ground Truth Generation. IEEE Trans. Intelligent Transportation Systems (2012).

[2] S. Kim E. Culurciello A. Paszke, A. Chaurasia. 2016. ENet: A deep neural network architecture for real-time semantic segmentation. CoRRabs/1606.02147 (2016).

[3] Ziye Chen, Kate Smith-Miles, Bo Du, Guoqi Qian, and Mingming Gong. 2023. An Efficient Transformer for Simultaneous Learning of BEV and Lane Representations in 3D Lane Detection. arXiv:2306.04927 [cs.CV]

[4] S. Reuter M. Szczot M. Konrad K. Dietmayer H. Deusch, J. Wiest. 2012. A random finite set approach to multiple lane detection. ITSC (2012).

[5] Z. Zhang et al. H. Zhang, C. Wu. 2020. Resnest: split-attention networks. arXiv:2004.08955

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. Deep Residual Learning for Image Recognition. arXiv:1512.03385 [cs.CV]

[7] Ma Z. Liu C. Loy C. C Hou, Y. 2019. Learning lightweight lane detection cnns by self attention distillation. In Proceedings of the IEEE/CVF international conference on computer vision (2019).

References

[8] C. Park J. Kim. 2017. End-to-end ego lane estimation based on sequential transfer learning for self-driving cars. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (2017).

[9] M. Xu P. Lv X. Zhao J. Niu, J. Lu. 2016. Robust lane detection using two-stage feature extraction with curve fitting. Pattern Recognit (2016).

[10] De Brabandere B. Georgoulis S. Proesmans M. Van Gool Neven, D. 2018. Towards end-to-end lane detection: an instance segmentation approach. IEEE intelligent vehicles symposium (IV) (2018).

[11] A. Ismail G. Adam V. Narayan M. Schulze S. Chougule, N. Koznek. 2018. Reliable multilane detection and classification by utilizing CNN as a regression network. Proceedings of the European Conference on Computer Vision(ECCV) (2018).

[12] J. Shin Yoon S. Shin O. Bailo N. Kim T.-.H. Lee H. Seok Hong S.-.H. Han I. So Kweon S. Lee, J. Kim. 2017. Vpgnet: vanishing point guided network for lane and road marking detection and recognition. Proceedings of the IEEE International Conference on Computer Vision (2017).

[13] Berriel R. Paixao T. M. Badue C. De Souza A. F. Oliveira-Santos T Tabelini, L. 2021. Keep your eyes on the lane: Real-time attention guided lane detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (2021).

[14] Mingxing Tan and Quoc V. Le. 2020. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv:1905.11946 [cs.LG]

[15] D. Neven M. Proesmans L. Van Gool W. Van Gansbeke, B. De Brabandere. 2019. End-to-end lane detection through differentiable least-squares fitting. Proceedings of the IEEE International Conference on Computer Vision Workshops (2019).