Supplemental Materials

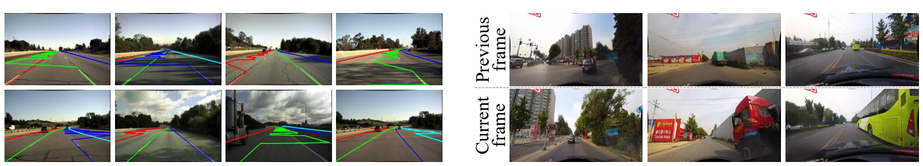
OMR: Occlusion-Aware Memory-Based Refinement for Video Lane Detection

Dongkwon Jin[1,2] and Chang-Su Kim[1]

[1] School of Electrical Engineering, Korea University, Seoul, Korea
[2] Samsung Advanced Institute of Technology
dongkwonjin@mcl.korea.ac.kr, changsukim@korea.ac.kr

A Other Datasets

Besides VIL-100 and OpenLane-V, there are some lane detection datasets, including TuSimple [1], LLAMAS [2], and CULane [3]. However, these datasets are not suitable for video lane detection. TuSimple consists of multiple video clips, each containing 20 consecutive frames. But, only the last frame in each video is annotated, so the video metrics cannot be used. In LLAMAS, many annotation errors were found, as in Fig. 1(a). Also, CULane exhibits low inter-frame correlation because the frame interval is too long, as in Fig. 1(b). Thus, it is regarded as an image-level dataset.



(a) LLAMAS (b) CULane

Fig. 1: (a) There are many annotation errors in LLAMAS. (b) Two adjacent frames do not exhibit high temporal correlation in CULane.

B Data Augmentation

We introduce a data augmentation scheme for training the OMR module. We employ the KINS dataset. It is an amodal instance segmentation dataset, in which a fully-shaped mask per object is given including its occluding parts. Given an original video sequence, we randomly select an object from the KINS dataset and then overlay its full shape onto the video frames. We also vary the size and position of the object linearly over the frames. Fig. 2 shows some examples of the synthetic data.

Fig. 2: Some examples of the synthesized frames.

C Detection Results

Fig. 3 and Fig. 4 compare the lane detection results of the proposed algorithm with those of RVLD for three consecutive video frames in VIL-100 and OpenLane-V, respectively. We see that the proposed algorithm maintains the temporal stability of lanes whereas RVLD does not.

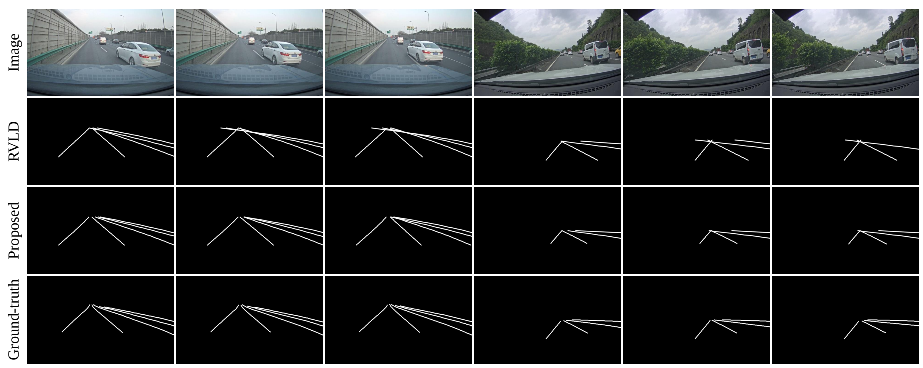


Fig. 3: Comparison of lane detection results for three consecutive video frames in the VIL-100 dataset.

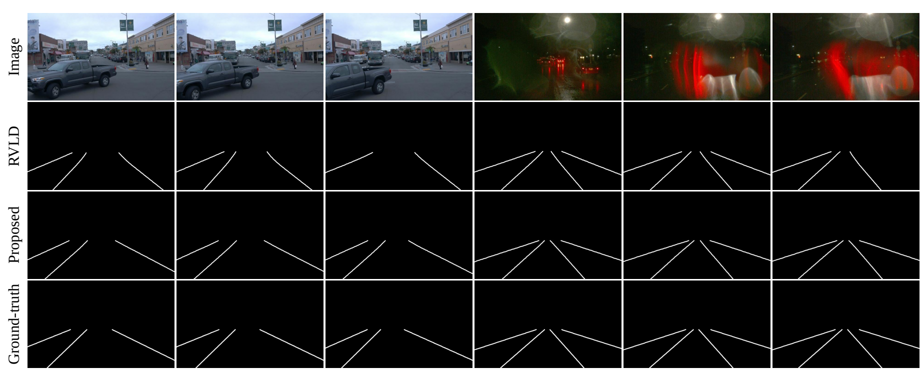


Fig. 4: Comparison of lane detection results for three consecutive video frames in the OpenLane-V dataset.

D Visualization

Fig. 5 visualizes the obstacle mask $\tilde{O}_t$, the feature map $\tilde{F}_t$, the lane probability map $\tilde{P}_t$ of a current frame I_t, and the refined ones F_t and P_t. Some lane parts in I_t are occluded by nearby obstacles, so their features and probabilities are defective. However, those features and probabilities are enhanced reliably through the proposed OMR module.

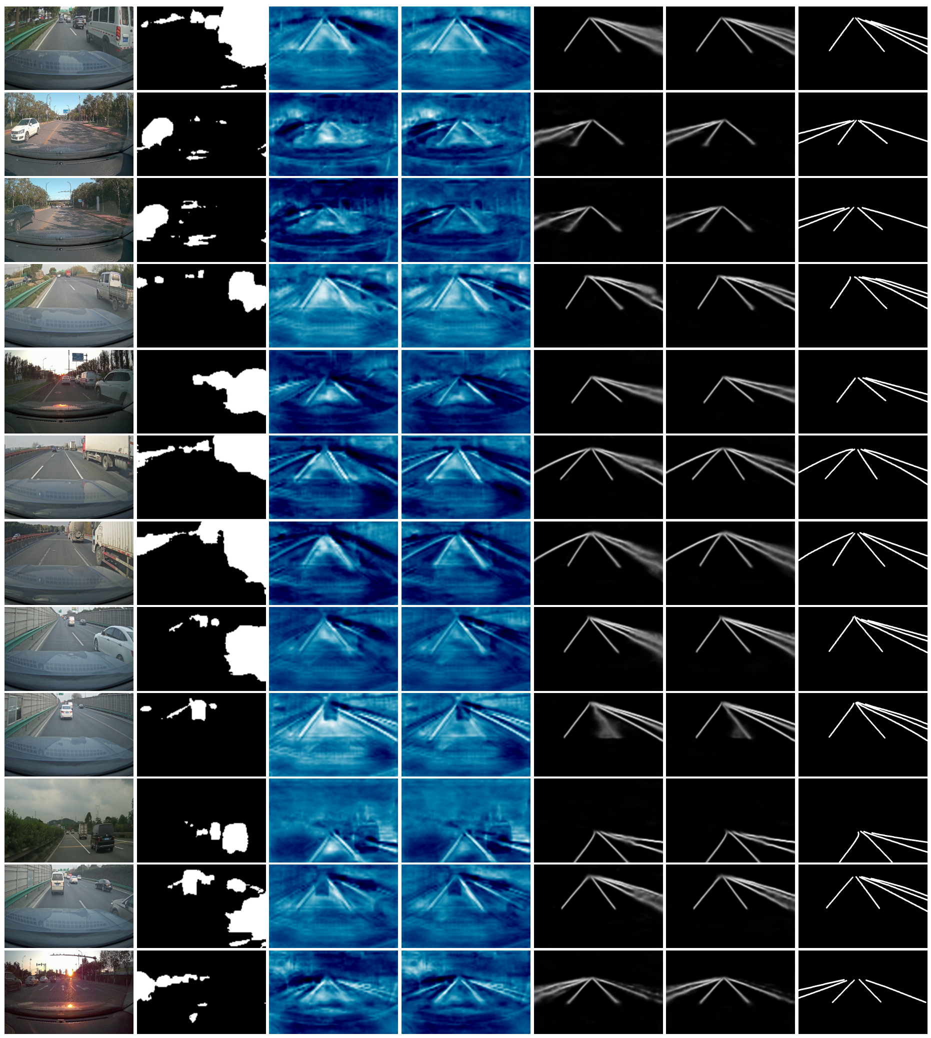


Fig. 5: Visualization of the obstacle mask $\tilde{O}_t$, the feature map $\tilde{F}_t$, the probability map $\tilde{P}_t$, and the enhanced F_t and P_t.

References

1. TuSimple benchmark. [Online]. Available: `https://github.com/TuSimple/tusimple-benchmark`
2. Behrendt, K., Soussan, R.: Unsupervised labeled lane markers using maps. In: Proc. IEEE ICCV Workshops (2019)
3. Pan, X., Shi, J., Luo, P., Wang, X., Tang, X.: Spatial as deep: Spatial CNN for traffic scene understanding. In: Proc. AAAI (2018)