1. Removing redundant frames

As described in the Subsection 4.1 of the main paper, the CULane dataset [1] includes a non-negligible amount of redundant frames where the ego-vehicle is stationary and the lane annotations do not change. We remove the frames whose average pixel value difference from the previous frame is below a threshold $T_{\text{diff}}$. Fig. 1 shows the $F_{150}$ val score and the number of train data with different $T_{\text{diff}}$. There is a trade-off between the data cleanliness and data size, thus a peak of $F_{150}$ can be seen at $T_{\text{diff}} \approx 15$. When the redundant data is removed, the spikes in the $\theta$ distribution of ground-truth lanes (Fig. 2) disappear, which results in avoiding overfitting to the spikes. From the experiment above, the optimal threshold ($=15$) is chosen and the remaining 55,698 (62.7%) frames are utilized for training. The $F_{150}$ test score of CLRNet-DLA34 is improved from $80.30 \pm 0.05$ to $80.86 \pm 0.06$ ($N = 5$ each) with the same 15-epoch training.

2. Improved LaneATT with LaneIoU

LaneATT [2] assigns the anchors according to the distance metric $D$. The distance is calculated as the average of the x-coordinate distance at the common horizontal lines between the anchors $X_a$ and ground-truths (GTs) $X_{GT}$. Unlike CLRNet [3], assignment between anchors and GTs does not change throughout the training as the anchors are not learnable. The positive and negative anchors are determined as:

\[
\text{positive: } D(X_a, X_{GT}) < T_{\text{dist}}^{\text{pos}},
\]

\[
\text{negative: } D(X_a, X_{GT}) > T_{\text{dist}}^{\text{neg}},
\]

where $T_{\text{dist}}^{\text{pos}}$ and $T_{\text{dist}}^{\text{neg}}$ are the threshold parameters and set to 15 and 20 respectively in the (640, 360) resolution. For the positive and negative anchors, the learning target of the classification logit is set to positive and negative, otherwise ignored. The regression loss is imposed only on the positive anchors. We keep the above assignment algorithm only for the regression target, and separately assign the positive and negative anchors for the classification target leveraging LaneIoU. We use predictions $X_p$ after refined by the regression-head output for IoU calculation:

\[
\text{positive: } \text{LaneIoU}(X_p, X_{GT}) > T_{\text{iou}}^{\text{pos}},
\]

\[
\text{negative: } \text{LaneIoU}(X_p, X_{GT}) < T_{\text{iou}}^{\text{neg}},
\]

where the thresholds $T_{\text{iou}}^{\text{pos}}$ and $T_{\text{iou}}^{\text{neg}}$ are set to 0.7 and 0.5 respectively. With the LaneIoU-based assignment for classification, the predictions that have high IoU with the corresponding GTs are prioritized and the confidence scores get close to the metric IoU with GTs. The LaneIoU-based assignment improves the $F_{150}$ score from $77.51 \pm 0.10$ to $78.19 \pm 0.06$ as is shown in Table 2 of the main paper.
References

