Active Learning for Lane Detection: A Knowledge Distillation Approach Supplementary Materials

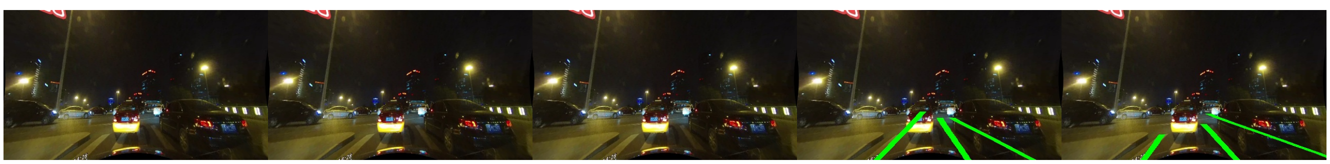
Fengchao Peng, Chao Wang, Jianzhuang Liu, Zhen Yang
Noah's Ark Lab, Huawei Technologies
{pengfengchao, wangchao165, liu.jianzhuang, yang.zhen}@huawei.com

1. Examples of Uncertain Cases

In Section 3.2 of the paper, we divide the uncertain samples of lane data into four cases, based on D_{SS} and D_{ST}. In the paper we show an example of case 4, with large D_{SS} and large D_{ST}. Here we show an example of each of the other three cases in Figs. 1–3.

(a) Input **(b)** GT **(c)** M_T **(d)** M_{S-KD} **(e)** M_S

Figure 1: An example of case 1. (a) The original image. (b) The ground truth. (c) The prediction of the teacher. (d) The prediction of the distilled student. (e) The prediction of the student without distillation. The difference D_{ST} between (c) and (d) is small, and D_{SS} between (d) and (e) is also small. The three models make very similar predictions on this image (in fact, these predictions match well with the ground truth). This means that this image is easy to detect and there is no need to annotate it.



(a) Input **(b)** GT **(c)** M_T **(d)** M_{S-KD} **(e)** M_S

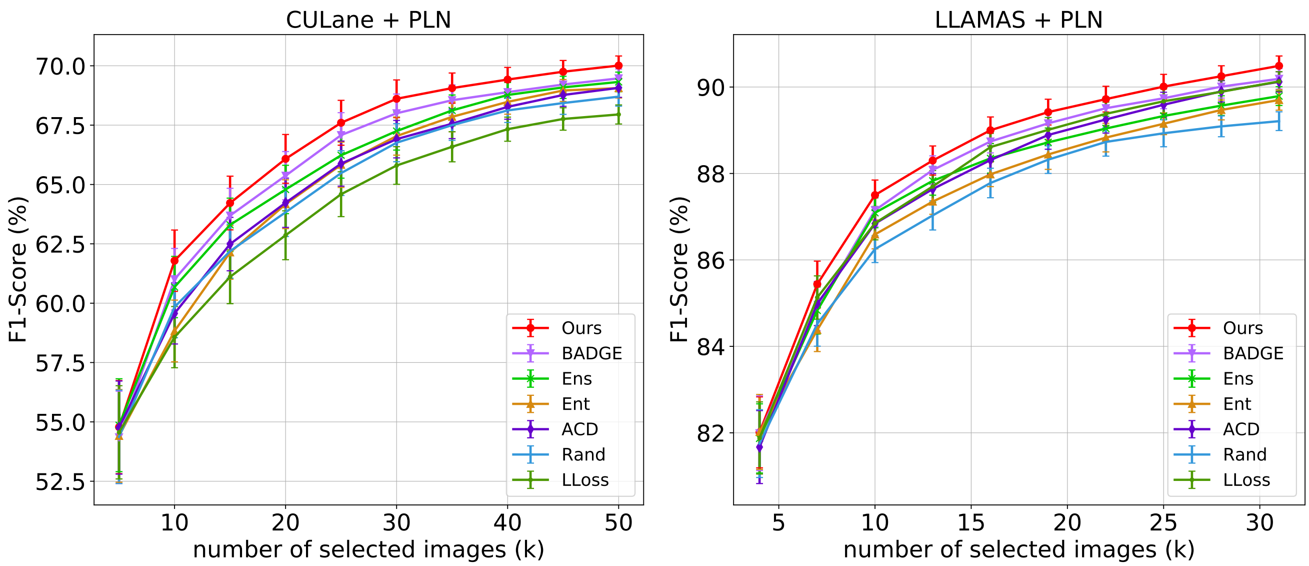
Figure 2: An example of case 2. (a) The original image. (b) The ground truth. (c) The prediction of the teacher. (d) The prediction of the distilled student. (e) The prediction of the student without distillation. The difference D_{ST} between (c) and (d) is large, and D_{SS} between (d) and (e) is small. There is no lane in this image and the teacher model M_T yields the correct prediction. But the student model M_S mistakes the zebra crossing segments as lanes. The student M_{S-KD}, though learning from the teacher, makes the same mistake. This means that this image is currently too difficult for the student model to learn, and it is valuable to annotate it.

(a) Input **(b)** GT **(c)** M_T **(d)** M_{S-KD} **(e)** M_S

Figure 3: An example of case 3. (a) The original image. (b) The ground truth. (c) The prediction of the teacher. (d) The prediction of the distilled student. (e) The prediction of the student without distillation. The difference D_{ST} between (c) and (d) is small, and D_{SS} between (d) and (e) is large. The lane missed by M_S is difficult to detect because it is occluded and its visible part is very short. Though the teacher model M_T manages to detect it and passes the knowledge to the student M_{S-KD}, it is still desirable to use human annotations to validate the correctness of the passed knowledge.

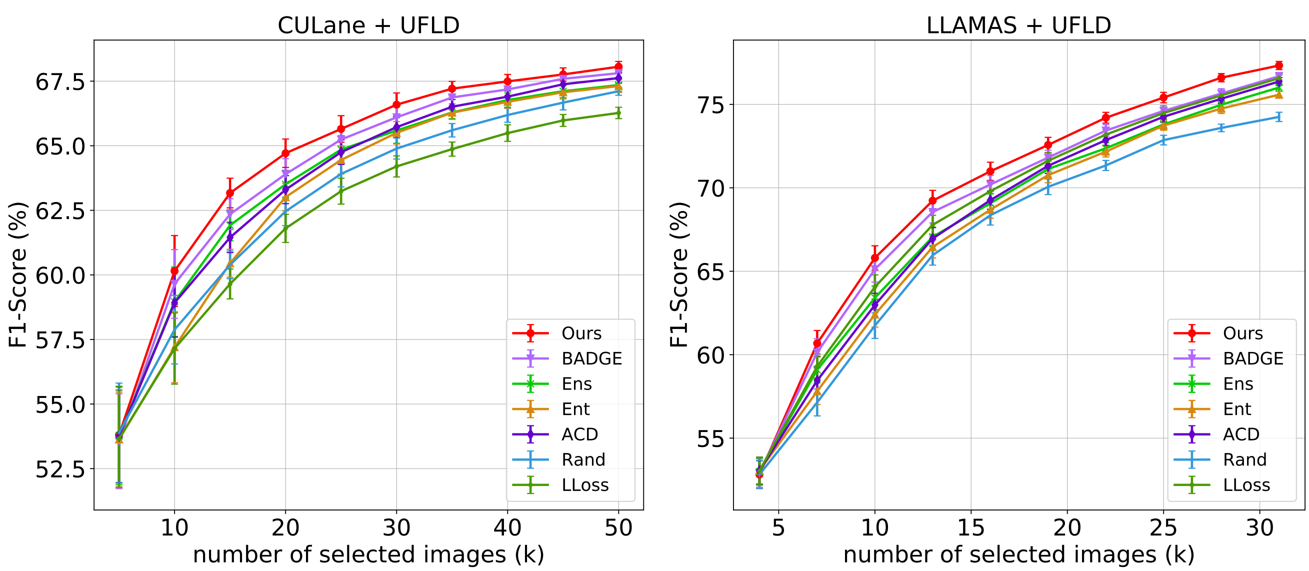
2. Results with Standard Deviations

In the paper, we only show the curves of average F1-Score. We do not show the standard deviations because otherwise there will be too many curves and markers in the small figures, making the curves messy. In Figs. 4–6, we show the results with standard deviations in larger figures to provide a more complete comparison among all the methods. We can see that all the methods have similar deviations.



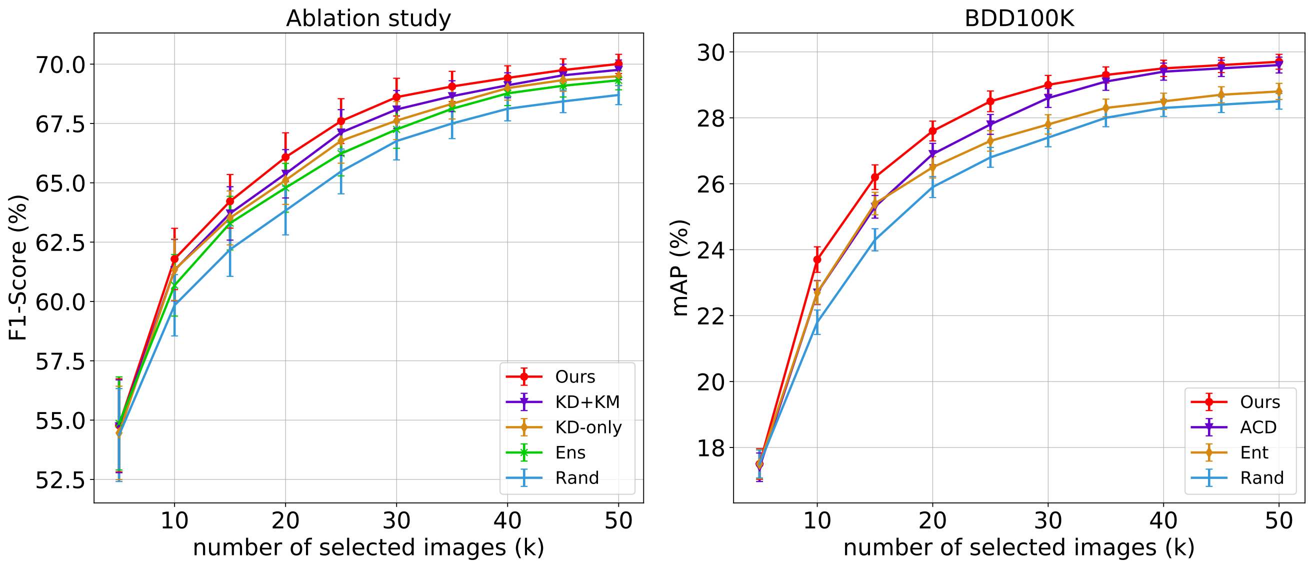
(a) **(b)**

Figure 4: Results using PointLaneNet(PLN).



(a) **(b)**

Figure 5: Results using the UFLD model.



(a) **(b)**

Figure 6: (a) Results of the ablation study. (b) Results of 2D object detection.

3. Discussion of the Tusimple Dataset

Tusimple [1] is another widely used lane detection benchmark, but we do not discuss it in the paper because none of the other methods obtains obvious improvement over the random selection in our experiments. The reason is that this dataset is too small. It contains only 3626 training images and 2782 test images. There is very limited space for active learning methods to further reduce the annotation cost. In addition, we find that this dataset is too easy for lane detection models. An example is shown in Fig. 7. We start from using 700 images to train the UFLD model, and iteratively add 200 images to the training set and re-train the model. Each training lasts for 90 epochs. We find that even if only using 700 images, UFLD already achieves an accuracy of 91.48%, which is only 4% lower than that using the entire training set. It is not quite meaningful to perform active learning on such a small and simple dataset. Therefore, we do not discuss Tusimple in the paper.

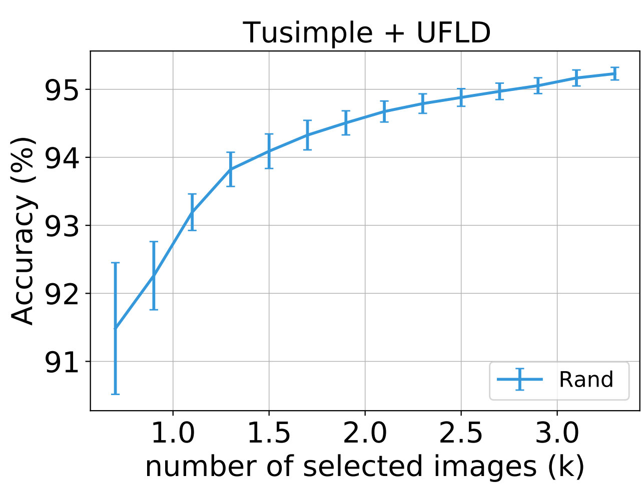


Figure 7: Results of the Tusimple dataset.

References

[1] Tusimple benchmark, 2017. Accessed 20 Dec 2020. https://github.com/TuSimple/tusimple-benchmark. 3