

Supplementary Material - PolyTransform: Deep Polygon Transformer for Instance Segmentation

Justin Liang¹ Namdar Homayounfar^{1,2}
Wei-Chiu Ma^{1,3} Yuwen Xiong^{1,2} Rui Hu¹ Raquel Urtasun^{1,2}
¹Uber Advanced Technologies Group ²University of Toronto ³MIT
{justin.liang,namdar,weichiu,yuwen,rui.hu,urtasun}@uber.com

In this supplementary material, we report additional qualitative results. In particular:

- In Table 1 and 2 we show the boundary metric performance at 1 and 2 pixels respectively of Deep Level Sets [1] and our method. The results are for the interactive annotation setting using the Cityscapes Stretch test dataset following [1]. We see that our model significantly out performs [1] in almost every class for both boundary metrics at a threshold of 1 and 2 pixels. This suggests that our method performs much better at the boundaries.
- In Figures 1 and 2, we show more instance segmentation results of our model on the validation set of Cityscapes on the fully automatic setting. In these images, we see that our model is able to handle occlusion; Occluded objects with multiple components are represented by the same instance.
- In Figures 3 and 4, we show more instance segmentation results of our model on the validation set of our new dataset. Like the results in Cityscapes, we see that our model is able to handle occlusion well, and produces accurate boundaries.
- In Figures 5 and 6, we show results of our model run in the interactive annotation setting where the ground truth bounding box is given. As we can see when given the ground truth boxes, our predictions are very strong for each object.
- Finally, we run our model on the video sequences from cityscapes. These sequences are made up of frames captured at 17 Hz. They can be downloaded [here](#).

	$F_{1px,mean}$	bicycle	bus	person	train	truck	mcycle	car	rider
Deep Level Sets [1]	60.29	55.67	60.75	69.70	45.24	60.65	56.10	72.25	62.02
Ours	62.32	57.28	63.26	71.62	44.20	64.31	58.25	74.92	64.77

Table 1. **Interactive Annotation (Cityscapes Stretch)**: This table shows our boundary performance at a threshold of 1px on the test set. This is in the setting of interactive annotation where we are given the ground truth boxes and we stretch the images following [1].

	$F_{2px,mean}$	bicycle	bus	person	train	truck	mcycle	car	rider
Deep Level Sets [1]	74.40	70.39	74.93	84.02	57.94	75.35	70.16	85.51	76.86
Ours	76.55	72.01	77.89	85.54	58.17	78.66	72.55	87.67	79.87

Table 2. **Interactive Annotation (Cityscapes Stretch)**: This table shows our boundary performance at a threshold of 2px on the test set. This is in the setting of interactive annotation where we are given the ground truth boxes and we stretch the images following [1].



Input Image

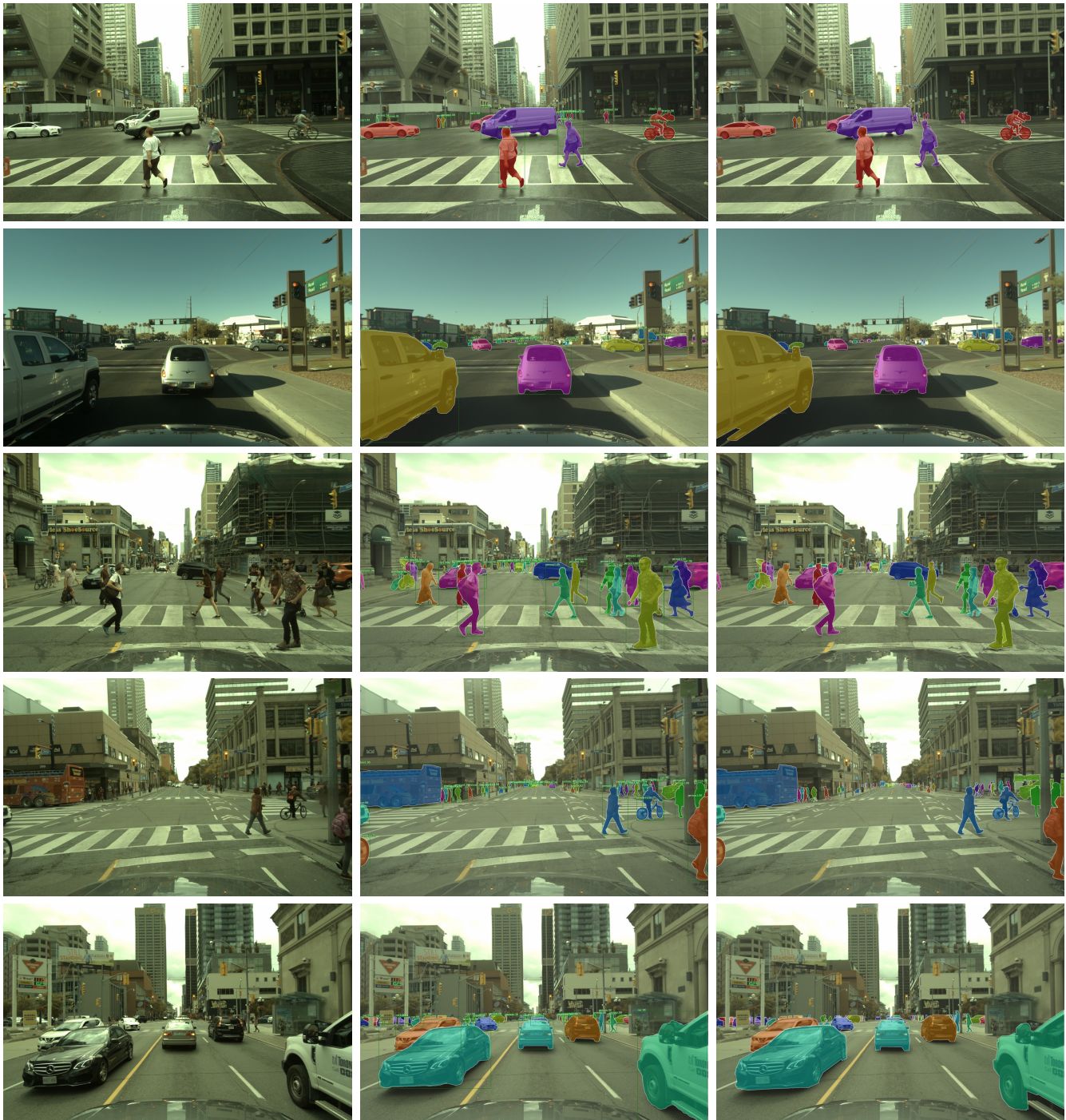
Our Instance Segmentation

GT Instance Segmentation

Figure 1. We showcase qualitative results of our instance segmentation model on the validation set.



Input Image Our Instance Segmentation GT Instance Segmentation
 Figure 2. We showcase qualitative results of our instance segmentation model on the Cityscapes validation set.



Input Image

Our Instance Segmentation

GT Instance Segmentation

Figure 3. We showcase qualitative results of our instance segmentation model on our new dataset's validation set.



Input Image Our Instance Segmentation GT Instance Segmentation

Figure 4. We showcase qualitative results of our instance segmentation model on our new dataset’s validation set.

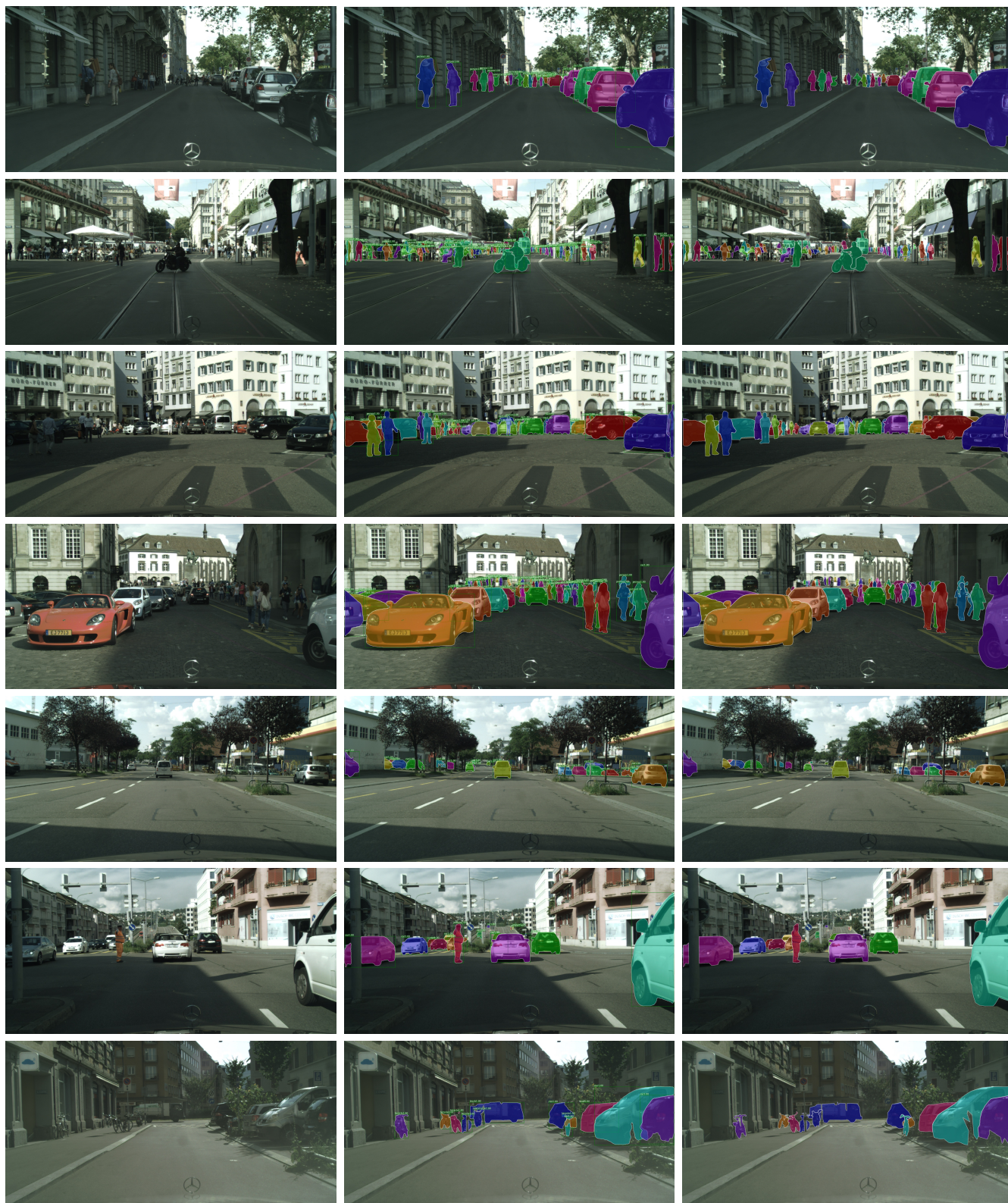


Input Image

Our Instance Segmentation

GT Instance Segmentation

Figure 5. We showcase qualitative results of our model in the setting of interactive annotation on the Cityscapes Stretch validation set.



Input Image

Our Instance Segmentation

GT Instance Segmentation

Figure 6. We showcase qualitative results of our model in the setting of interactive annotation on the Cityscapes Stretch validation set.

References

- [1] Zian Wang, Huan Ling, David Acuna, Amlan Kar, and Sanja Fidler. Object instance annotation with deep extreme level set evolution. In *CVPR*, 2019.