Coarse-to-Fine Domain Adaptive Semantic Segmentation with Photometric Alignment and Category-Center Regularization
Supplementary Material

Haoyu Ma[1*] Xiangru Lin[1*] Zifeng Wu[2] Yizhou Yu[1†]
[1]The University of Hong Kong [2]Deepwise AI Lab
mahaoyu@connect.hku.hk,xrlin2@cs.hku.hk, wuzifeng@deepwise.com, yizhouy@acm.org

1. Training Settings and Hyper-prameters

Datasets The **Cityscapes** dataset is the target domain dataset with $2,957$ 2048×1024 training images and 500 validation images of the same resolution. **Cityscapes** has 19 categories in total. **GTA5** and **Synthia** are two source domain datasets of computer generated synthetic images, which contain $24,966$ 1914×1052 training images and 9400 1280×760 training images respectively.

Category Centers The category center f_c for each category is updated using source features at the beginning of each learning step defined in Section 3.1 (Coarse-to-Fine Pipeline). They are introduced to make the calculation of category-oriented triplet loss practical. If we use the original triplet loss without category centers, we need to store pairwise distances among all pixels with a tremendous GPU memory overload (HW x HW x c).

Hyper-parameters The values of P_h and p are chosen according to the study in [2]. Experiments suggest mIoU is not sensitive to the values of hyper-parameters. For example, when P_h varies between 0.7 and 0.9 and p varies between 5 and 20, the final performance ranges from 55.3% to 56.1%. The degree of performance improvement varies from 4.3% in the first learning step to 0.1% in the last step.

2. Cross-Category Margin for the Target Domain

Two category-level feature distribution regularization methods are proposed in our work. In the source domain, a category-oriented triplet loss is introduced to enlarge category-level margins, In the target domain, a target domain consistency regularization is used to minimize intra-domain variances. The category-oriented triplet loss is applied to the source domain samples only because it relies heavily on labeled hard samples which are only available in the source domain. Although the category-oriented triplet loss is only applied to the source domain, it cooperates with the consistency regularization to enlarge category-level margins in the target domain as well. We propose a metric called cross-category margin (CCM) for the target domain to verify our claim. Let $x_{i,j}$ denote the pixel feature at location (i,j) in a target domain image, the corresponding ground-truth label is denoted as $y_{i,j}$. The CCM is defined as

$$\begin{aligned} CCM(c) =& \frac{1}{S_c} \sum_{S_c} \min_{k, k \neq c} (\delta_{i,j,k}) \\ \delta_{i,j,k} =& \frac{\|x_{i,j} - f_{c=k}\| - \|x_{i,j} - f_{c=y_{i,j}}\|}{\|f_{c=y_{i,j}} - f_{c=k}\|}, \end{aligned} \tag{1}$$

where $CCM(c)$ is the cross-category margin for category c, S_c is the set of pixels in category c across all images in the target domain, f_c is the category center in the source domain for category c, and $\delta_{i,j,k}$ is the category-level margin at location (i,j) between groundtruth category $y_{i,j}$ and another category k. This margin is a normalized one with respect to the category center distance between categories $y_{i,j}$ and k. We choose the minimum margin, which corresponds to the most confusing class, to be the category-level margin at a certain pixel. Then such margins for all pixels in category c are averaged to define the margin for category c. Note that the target domain groundtruth labels, $y_{i,j}$'s, are not used during training, and are only used to evaluate cross-category margins in the target domain. This metric can be considered as an indicator of category-level feature distributions in the feature space.

We compare cross-category margins computed for the models trained in the ablation study. Compared to the model with coarse alignment only, the model trained with additional target domain consistency regularization achieves improved cross-category margins, and the further integration of the category-oriented triplet loss improves cross-category margins further as shown in Figure 1. By

*These authors have equal contribution.
†Corresponding author.

		road	sidewalk	building	wall	fence	pole	light	sign	vege	terrace	sky	person	rider	car	truck	bus	train	motor	bike	mIoU
DeeplabV2	BDL [2]	91.0	44.7	84.2	34.6	27.6	30.2	36.0	36.0	85.0	**43.6**	83.0	58.6	31.6	83.3	35.3	49.7	3.3	28.8	35.6	48.5
	IDA [3]	90.6	36.1	82.6	29.5	21.3	27.6	31.4	23.1	85.2	39.3	80.2	59.3	29.4	86.4	33.6	53.9	0.0	32.7	37.6	46.3
	DTST [7]	90.6	44.7	84.8	34.3	28.7	31.6	35.0	37.6	84.7	43.3	85.3	57.0	31.5	83.8	**42.6**	48.5	1.9	30.4	39.0	49.2
	FGGAN [6]	91.0	50.6	86.0	**43.4**	**29.8**	36.8	43.4	25.0	**86.8**	38.3	87.4	64.0	**38.0**	85.2	31.6	46.1	6.5	25.4	37.1	50.1
	FDA [8]	**92.5**	**53.3**	82.3	26.5	27.6	36.4	40.5	**38.8**	82.2	39.8	78.0	62.6	34.4	84.9	34.1	53.1	**16.8**	27.7	46.4	50.4
	coarse align. (ours)	84.6	37.4	81.0	25.6	12.9	35.7	33.8	16.5	83.5	31.2	82.7	64.8	35.7	85.3	30.0	31.9	8.0	25.7	32.2	44.1
	coarse-to-fine (ours)	89.8	46.0	**85.8**	32.5	22.3	**41.0**	**43.9**	28.9	86.4	31.0	**89.4**	**65.6**	36.9	**87.9**	42.4	**54.4**	6.5	**38.9**	**56.2**	**51.9**
DeeplabV3+	CAG [9]	90.4	51.6	83.8	**34.2**	27.8	**38.4**	25.3	**48.4**	**85.4**	38.2	78.1	58.6	34.6	84.7	21.9	42.7	41.1	29.3	37.2	50.2
	coarse align. (ours)	83.9	37.5	82.7	28.7	18.9	35.3	41.3	31.1	85.2	29.5	86.6	62.8	30.9	82.4	23.0	39.3	33.0	26.0	39.7	47.3
	coarse-to-fine (ours)	**92.5**	**58.3**	**86.5**	27.4	**28.8**	38.1	**46.7**	42.5	**85.4**	**38.4**	**91.8**	**66.4**	**37.0**	**87.8**	**40.7**	**52.4**	**44.6**	**41.7**	**59.0**	**56.1**

Table 1: Performance comparison with state-of-the-art methods on the GTA5→Cityscapes task. Results with coarse alignment only and our whole coarse-to-fine pipeline are both presented.

		road	sidewalk	building	wall	fence	pole	light	sign	vege	sky	person	rider	car	bus	motor	bike	mIoU	mIoU*
DeeplabV2	BDL [2]	**86.0**	**46.7**	80.3	-	-	-	14.1	11.6	79.2	81.3	54.1	27.9	73.7	42.2	25.7	45.3	-	51.4
	IDA [3]	84.3	37.7	79.5	5.3	0.4	24.9	9.2	8.4	80.0	**84.1**	57.2	23.0	78.0	38.1	20.3	36.5	41.7	48.9
	DTST [7]	83.0	44.0	80.3	-	-	-	17.1	15.8	80.5	81.8	59.9	33.1	70.2	37.3	28.5	45.8	-	52.1
	FGGAN [6]	84.5	40.1	**83.1**	4.8	0.0	34.3	**20.1**	27.2	**84.8**	84.0	53.5	22.6	**85.4**	**43.7**	26.8	27.8	45.2	52.5
	FDA [8]	79.3	35.0	73.2	-	-	-	19.9	24.0	61.7	82.6	61.4	31.1	83.9	40.8	**38.4**	51.1	-	52.5
	coarse align. (ours)	76.4	28.8	71.6	7.7	0.5	31.0	13.8	27.8	69.3	70.0	59.7	26.4	75.7	29.9	22.1	25.2	39.7	45.9
	coarse-to-fine (ours)	81.9	33.7	78.5	**11.0**	**1.9**	**36.7**	**32.6**	**33.4**	79.6	78.2	**67.3**	**33.6**	84.0	33.5	25.9	**47.6**	**47.5**	**54.6**
DeeplabV3+	CAG (13 classes) [9]	**84.8**	**41.7**	**85.5**	-	-	-	13.7	23.0	**86.5**	78.1	**66.3**	28.1	81.8	21.8	22.9	49.0	-	52.6
	CAG (16 classes) [9]	84.7	40.8	81.7	7.8	0.0	35.1	13.3	22.7	84.5	77.6	64.2	27.8	80.9	19.7	22.7	48.3	44.5	-
	coarse align. (ours)	64.0	25.7	73.9	9.6	0.8	33.3	12.3	25.9	81.6	85.5	62.4	26.2	80.6	30.9	26.8	23.8	41.5	47.7
	coarse-to-fine (ours)	75.7	30.0	81.9	**11.5**	**2.5**	**35.3**	**18.0**	**32.7**	86.2	**90.1**	65.1	**33.2**	**83.3**	**36.5**	**35.3**	**54.3**	**48.2**	**55.5**

Table 2: Performance comparison with state-of-the-art methods on the Synthia→Cityscapes task (mIoU: 16-class; mIoU*: 13-class).

comparing to the experimental results in Table 1 of the main paper, we can draw the conclusion that both our category-level feature distribution regularization methods improve cross-category margins in the target domain.

3. Experiments with DeeplabV2

In addition to the experiments discussed in the main paper, here we discuss experiments using another commonly adopted segmentation backbone DeeplabV2 [6, 7, 2]. As shown in Table 1, although the performance improvement achieved by our proposed framework using this backbone is not as significant as DeeplabV3+, our proposed method still outperforms all previous methods, achieving a new state-of-the-art mIoU (51.9%), which is 1.5% higher than the previous best result using the same backbone on the GTA5→Cityscapes task. On the Synthia→Cityscapes task, our proposed method also achieves a new state-of-the-art mIoU (54.6%), which is 2.1% higher than the previous best result using the same backbone (Table 2). Although DeeplabV3+ is more powerful than DeeplabV2, the performance of our method on the Synthia→Cityscapes task is similar using both backbones. This is because Synthia mostly consists of large objects where high-resolution feature maps of DeeplabV3+ provide limited improvement.

We also compare the mean cross-category margins (CCM) of different methods on both Deeplab V2 and Deeplab V3+ models in the following table. We compare CCM with other SOTAs([7, 9, 6]). Our proposed method improves CCM on different segmentation models. Note that the magnitude of CCM highly depends on the choice of models, and we still use mIoU as our main evaluation metric.

Seg. Model	DTST [7]	FGAN [6]	CAG [9]	ours
DeeplabV2(ResNet101)	0.043	0.040	-	0.069
DeeplabV3+(ResNet101)	-	-	0.223	0.327

4. More Qualitative Samples

More segmentation results are given in this section in addition to the qualitative results in the paper. In the following part, we demonstrate (1) our photometric alignment results for the GTA5 [4]→Cityscapes [1] task with frequency alignment results from FDA [8] as a comparison (Figure 2-Figure 17); (2) our photometric alignment results for the Synthia [5]→Cityscapes [1] task (Figure 18-Figure 33); (3) our segmentation results for the GTA5 [4]→Cityscapes [1] task with segmentation results from CAG [9] as a comparison (Figure 34-Figure 49); (4) our segmentation results for the Synthia [5]→Cityscapes [1] task (Figure 50-Figure 65).

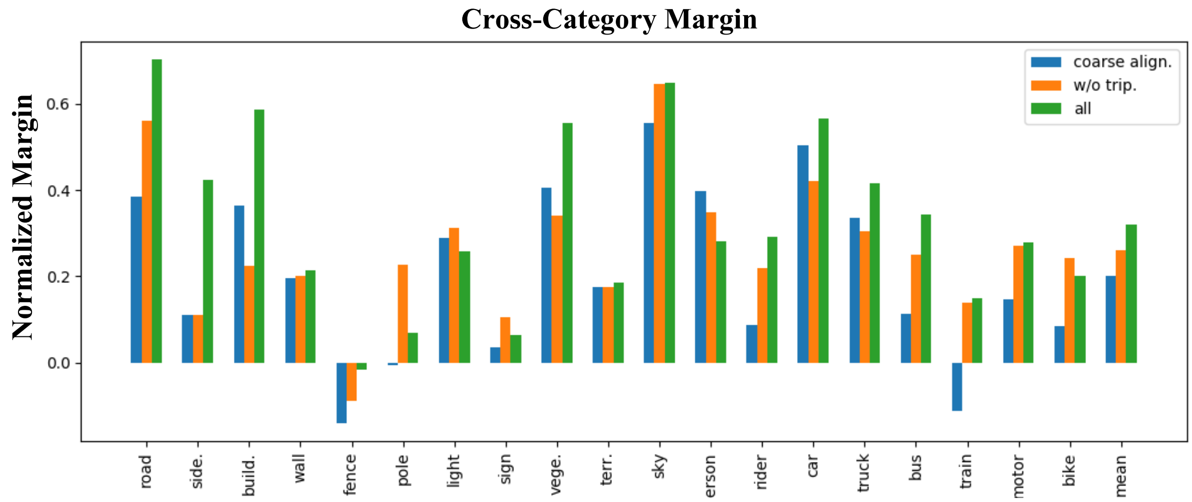


Figure 1: Our proposed category-oriented triplet loss exploits hard samples and further enlarge category margins.

References

[1] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. The cityscapes dataset for semantic urban scene understanding. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3213–3223, 2016. 2, 5

[2] Yunsheng Li, Lu Yuan, and Nuno Vasconcelos. Bidirectional learning for domain adaptation of semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6936–6945, 2019. 1, 2

[3] Fei Pan, Inkyu Shin, Francois Rameau, Seokju Lee, and In So Kweon. Unsupervised intra-domain adaptation for semantic segmentation through self-supervision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3764–3773, 2020. 2

[4] Stephan R Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun. Playing for data: Ground truth from computer games. In *European conference on computer vision*, pages 102–118. Springer, 2016. 2, 5

[5] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M Lopez. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3234–3243, 2016. 2, 22

[6] Haoran Wang, Tong Shen, Wei Zhang, Lingyu Duan, and Tao Mei. Classes matter: A fine-grained adversarial approach to cross-domain semantic segmentation. *arXiv preprint arXiv:2007.09222*, 2020. 2

[7] Zhonghao Wang, Mo Yu, Yunchao Wei, Rogerio Feris, Jinjun Xiong, Wen-mei Hwu, Thomas S Huang, and Honghui Shi. Differential treatment for stuff and things: A simple unsupervised domain adaptation method for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12635–12644, 2020. 2

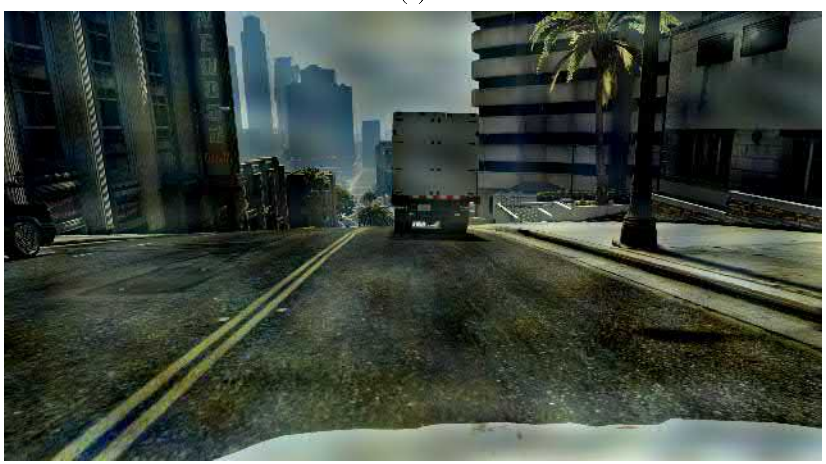
[8] Yanchao Yang and Stefano Soatto. Fda: Fourier domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4085–4095, 2020. 2, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37

[9] Qiming Zhang, Jing Zhang, Wei Liu, and Dacheng Tao. Category anchor-guided unsupervised domain adaptation for semantic segmentation. In *Advances in Neural Information Processing Systems*, pages 435–445, 2019. 2, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71

Section 4.1
Global Photometric Alignment on GTA2Cityscapes

(a)

(b)



(c)

(d)

Figure 2: Aligned sample from GTA5 [4]→Cityscapes [1] task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 3: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Figure 4: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a) (b)

(c) (d)

Figure 5: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 6: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a) (b) (c) (d)

Figure 7: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 8: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a) (b) (c) (d)

Figure 9: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a) (b)

(c) (d)

Figure 10: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 11: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 12: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Figure 13: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a) (b) (c) (d)

Figure 14: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Figure 15: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 16: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Figure 17: Aligned sample from GTA5→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Section 4.2
Global Photometric Alignment on Synthia2Cityscapes

(a)

(b)

(c)

(d)

Figure 18: Aligned sample from Synthia [5]→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

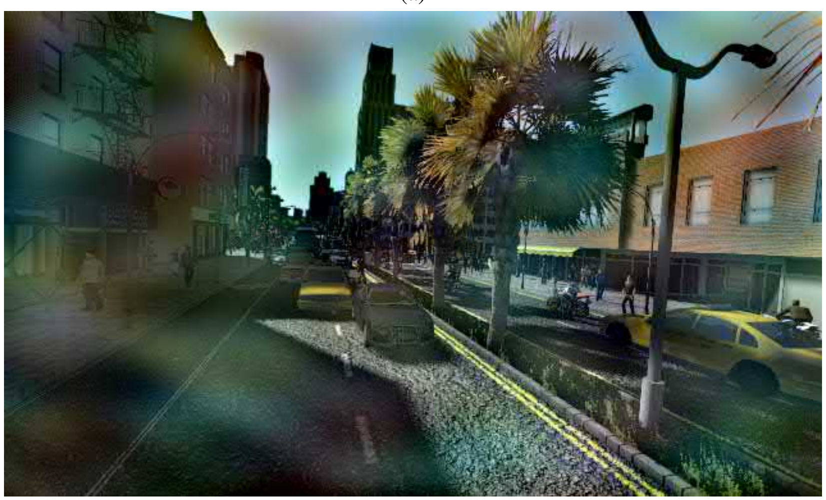
(c)

(d)

Figure 19: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)



(c)

(d)

Figure 20: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 21: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

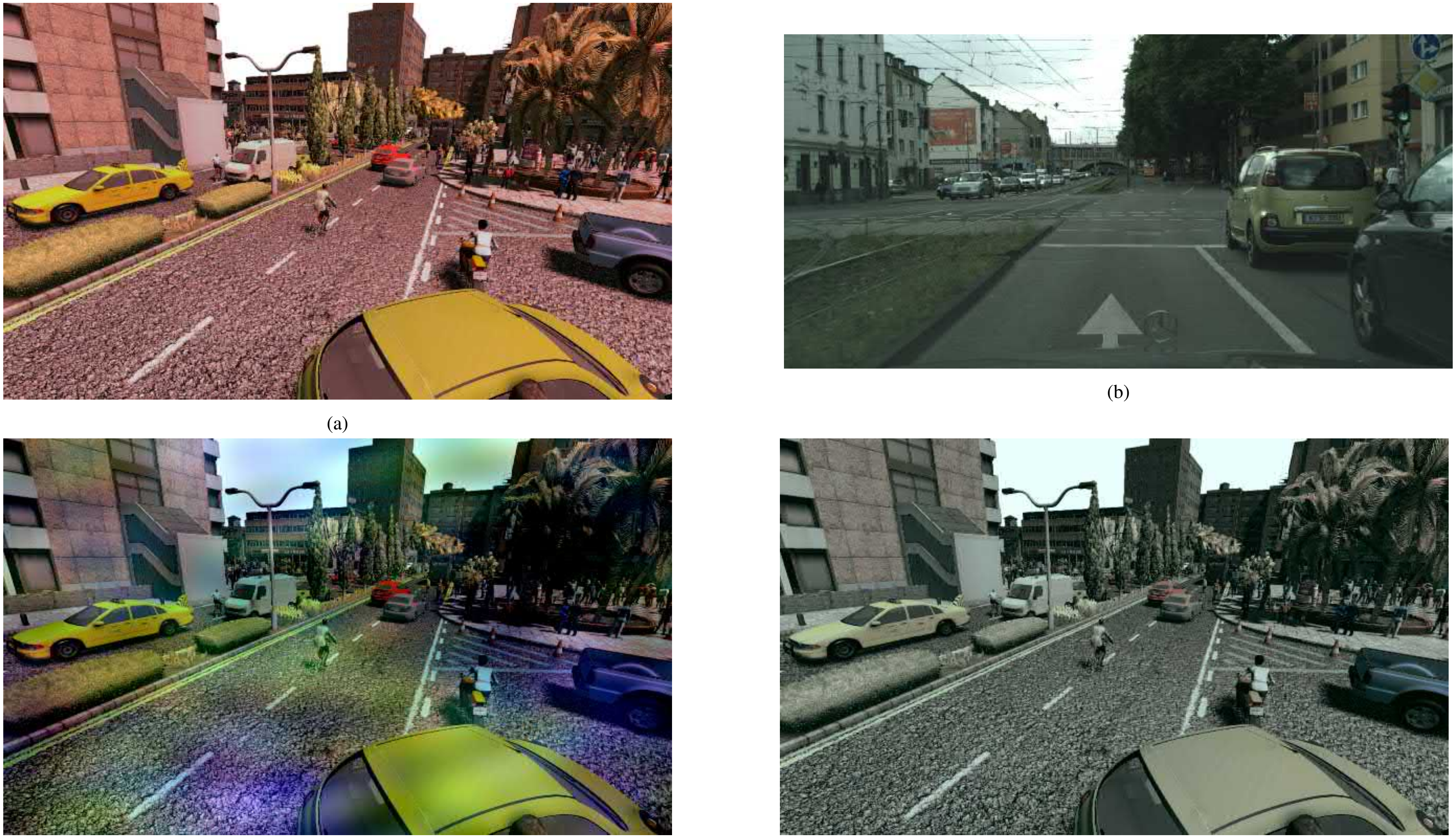
(a)

(b)

(c)

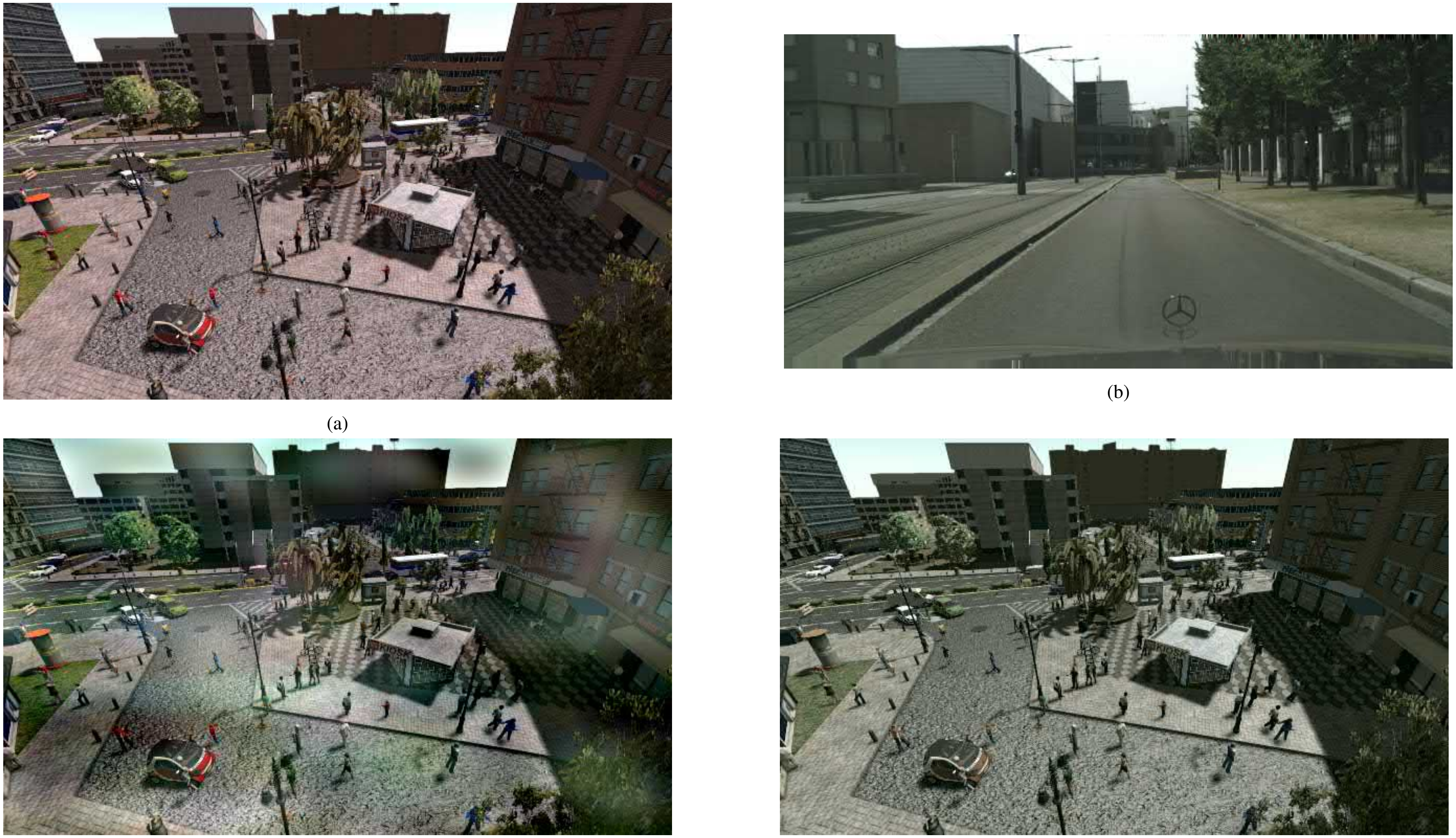
(d)

Figure 22: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



(a) (b) (c) (d)

Figure 23: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



(a) (b) (c) (d)

Figure 24: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 25: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

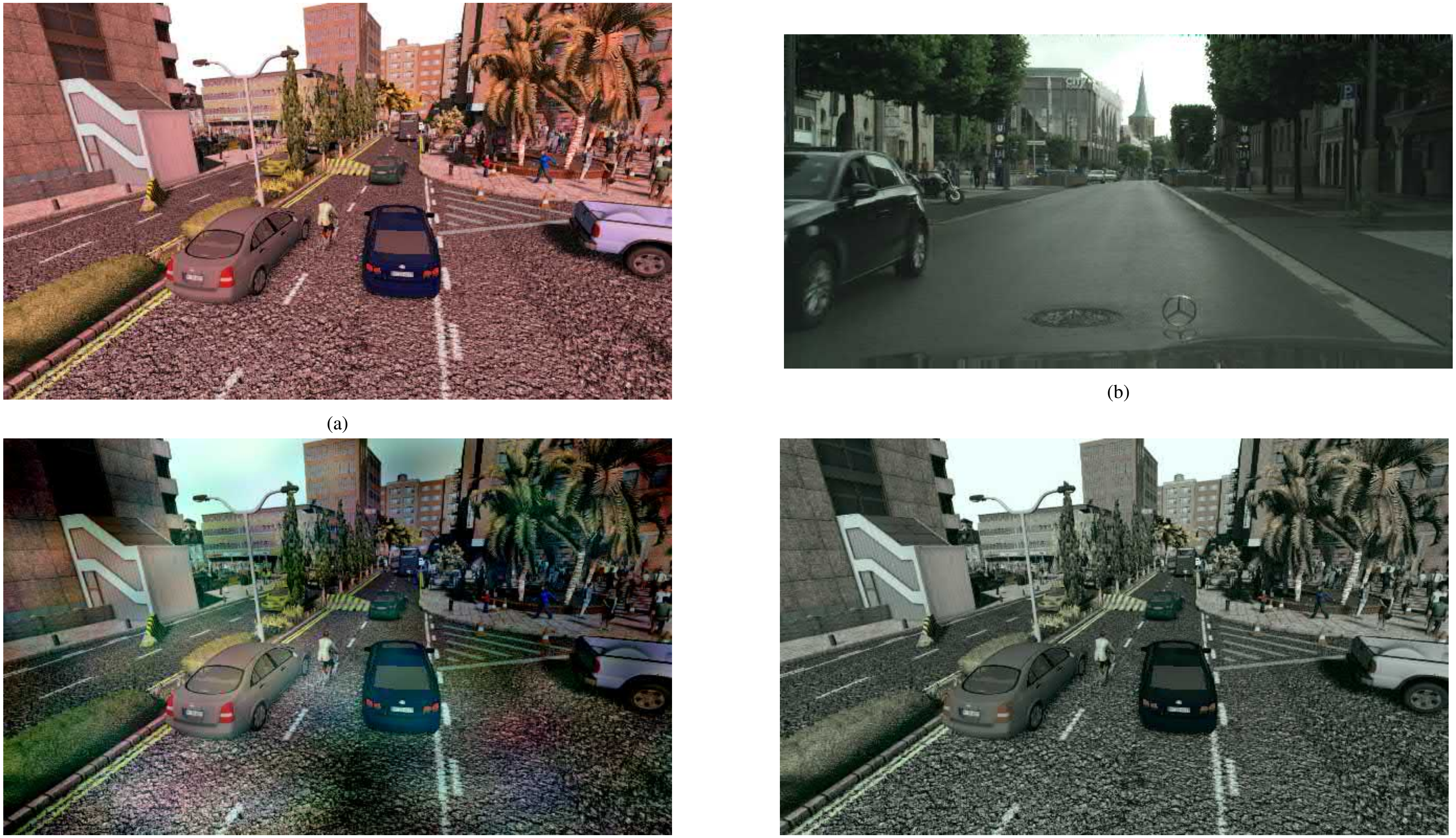
(a)

(b)

(c)

(d)

Figure 26: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



(a) (b) (c) (d)

Figure 27: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 28: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

(a)

(b)

(c)

(d)

Figure 29: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

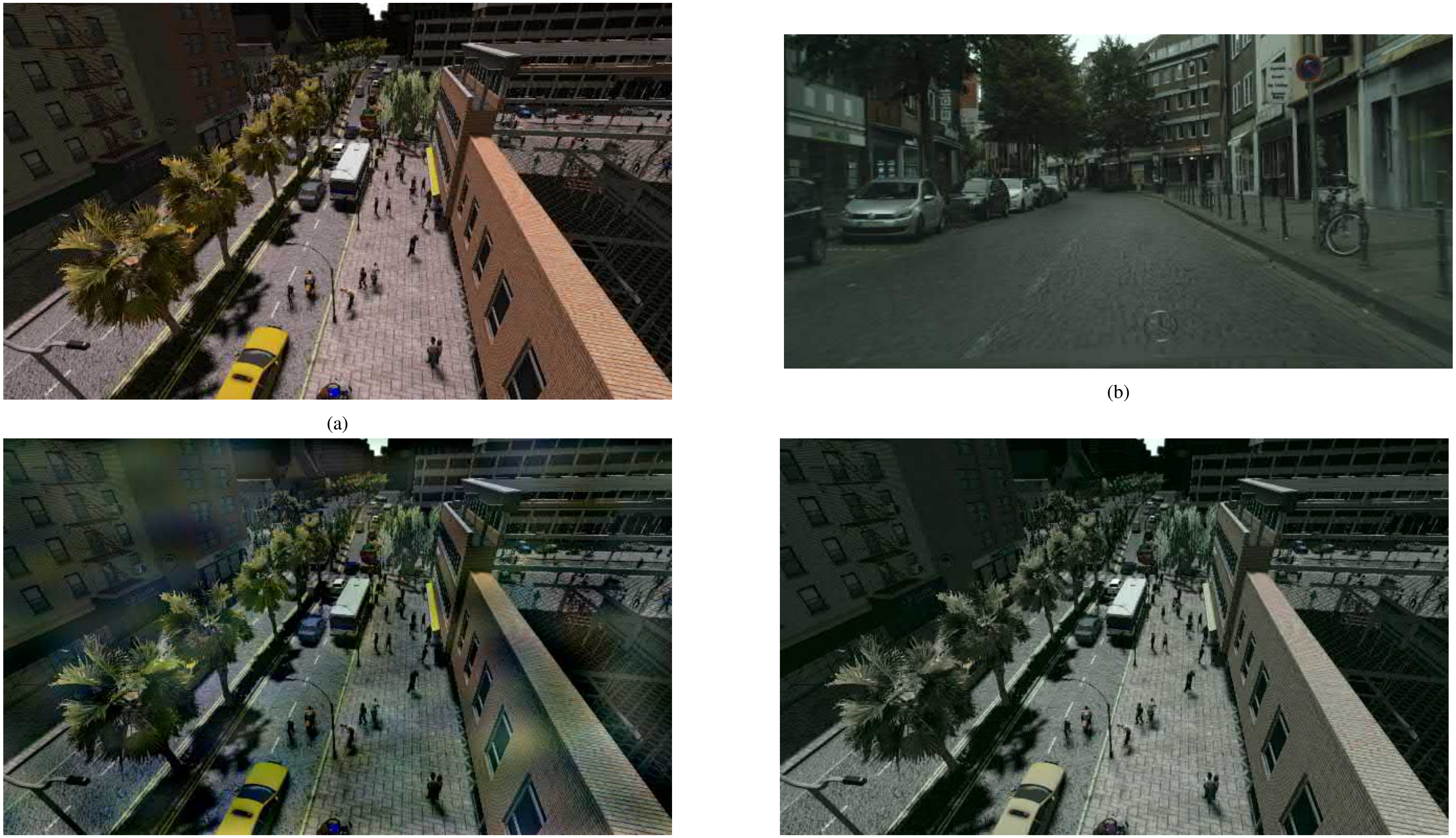
(a)

(b)

(c)

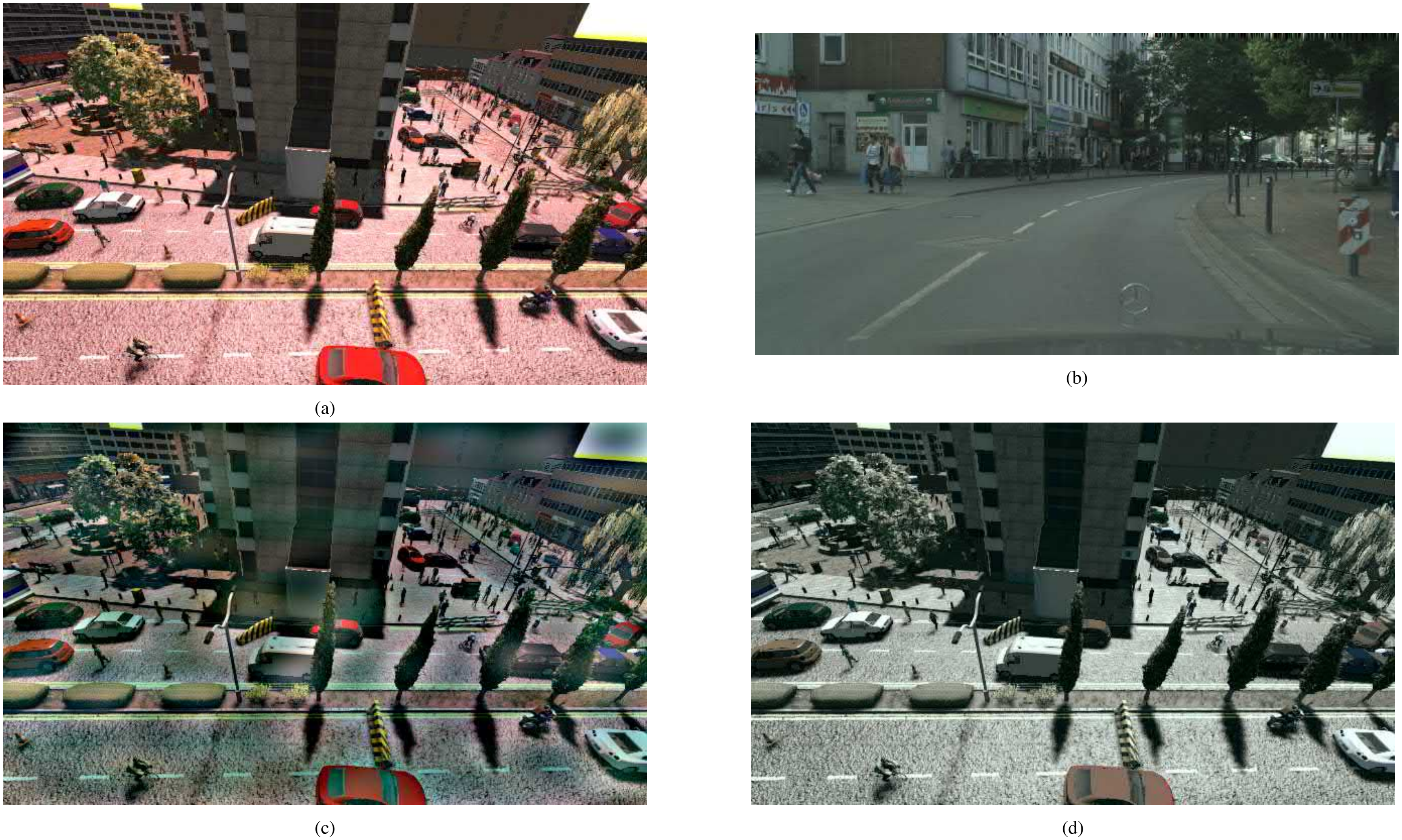
(d)

Figure 30: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



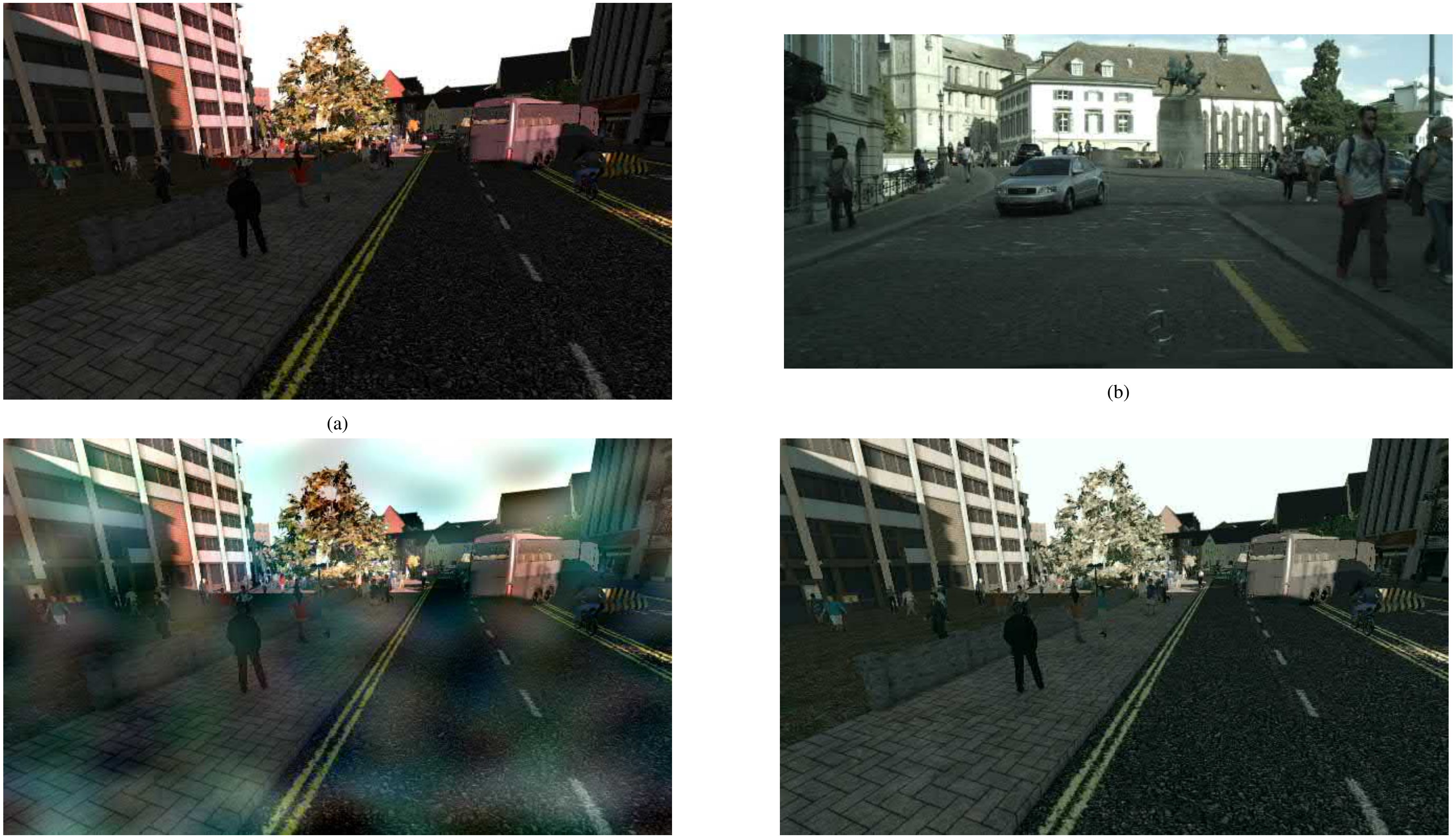
(a) (b) (c) (d)

Figure 31: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



(a) (b) (c) (d)

Figure 32: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment



(a) (b) (c) (d)

Figure 33: Aligned sample from Synthia→Cityscapes task: (a) source domain image, (b) target domain reference image, (c) FDA [8], (d) our global photometric aligment

Section 4.3
Sample Segmentation Results on GTA2Cityscapes

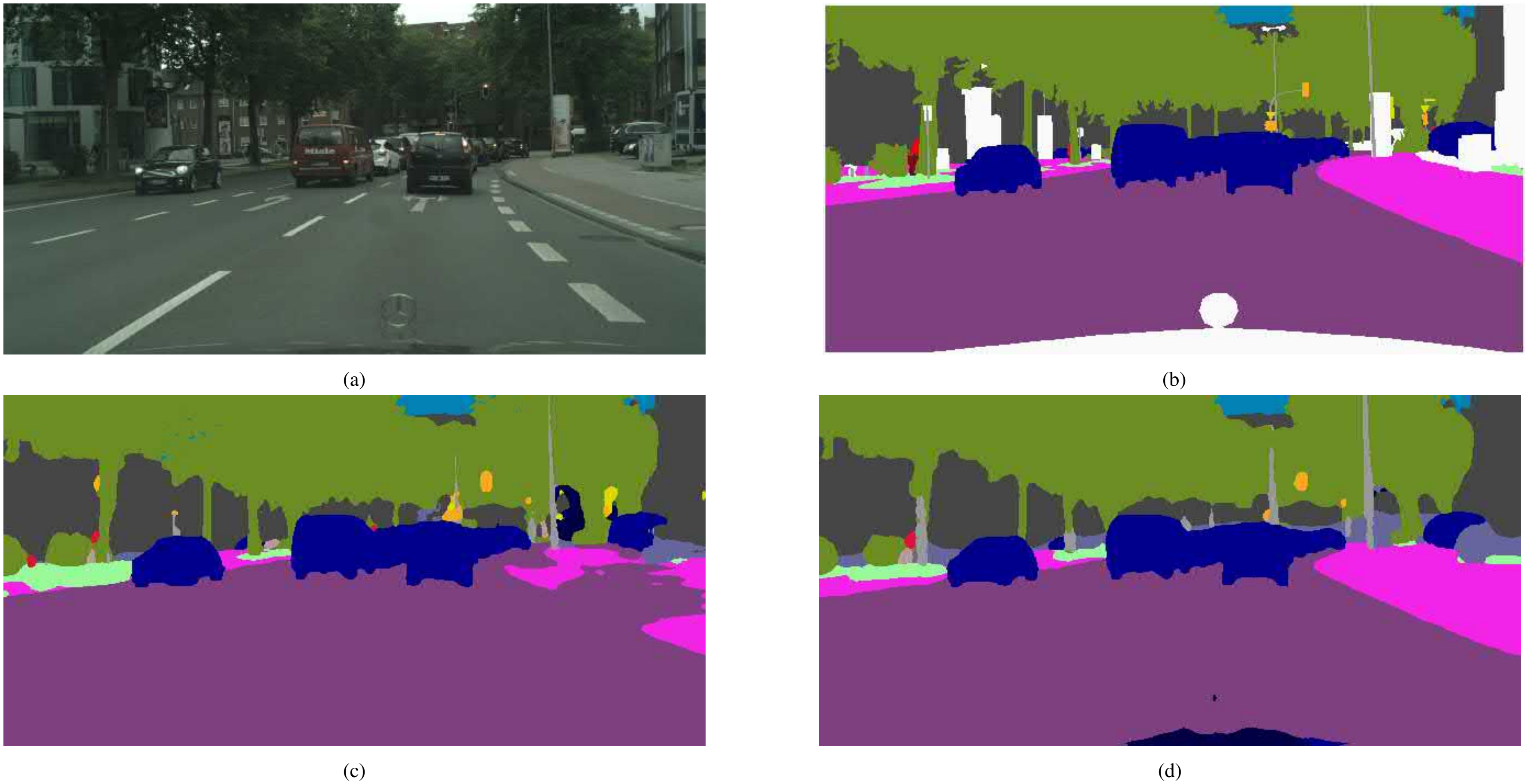
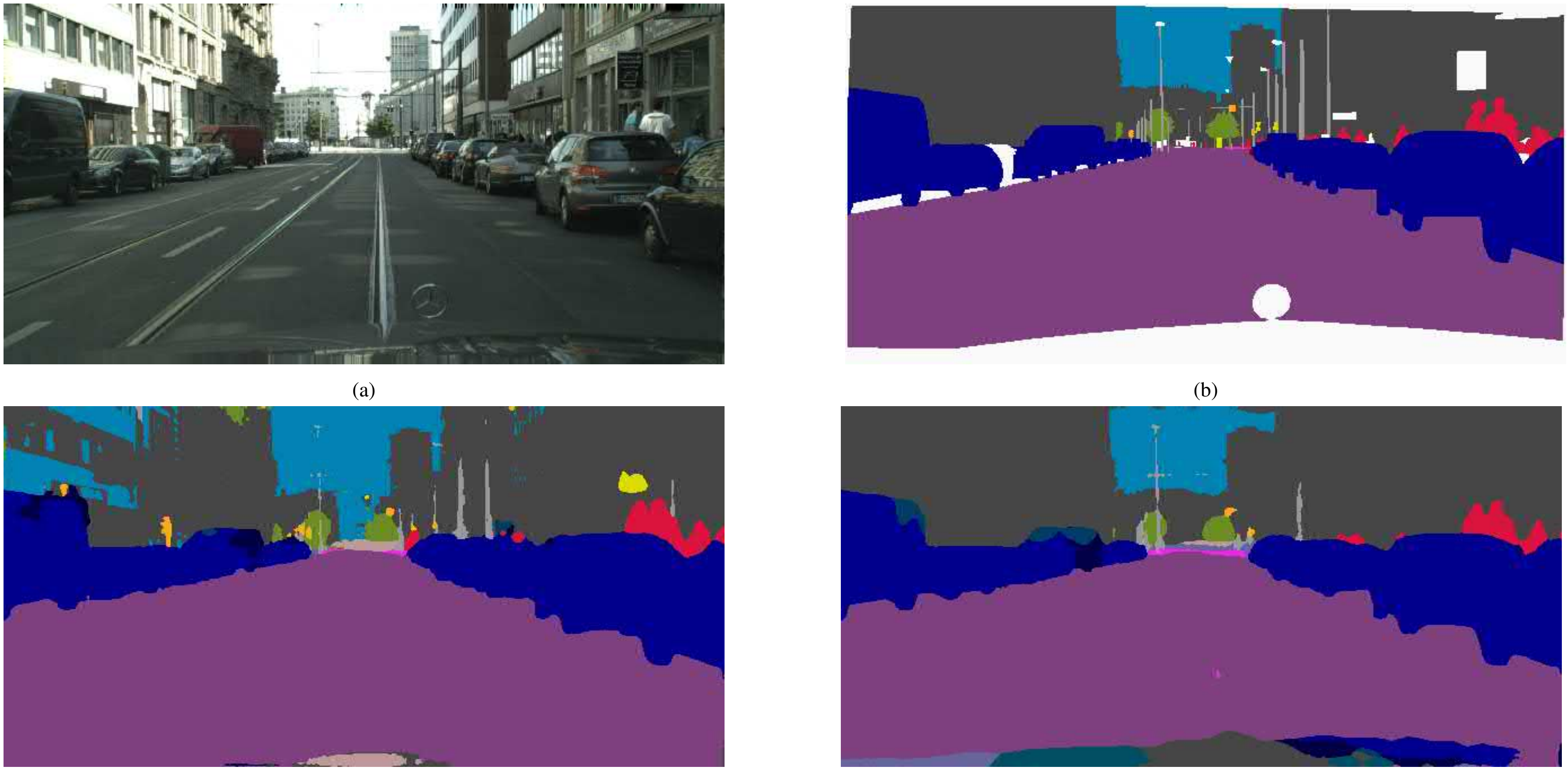
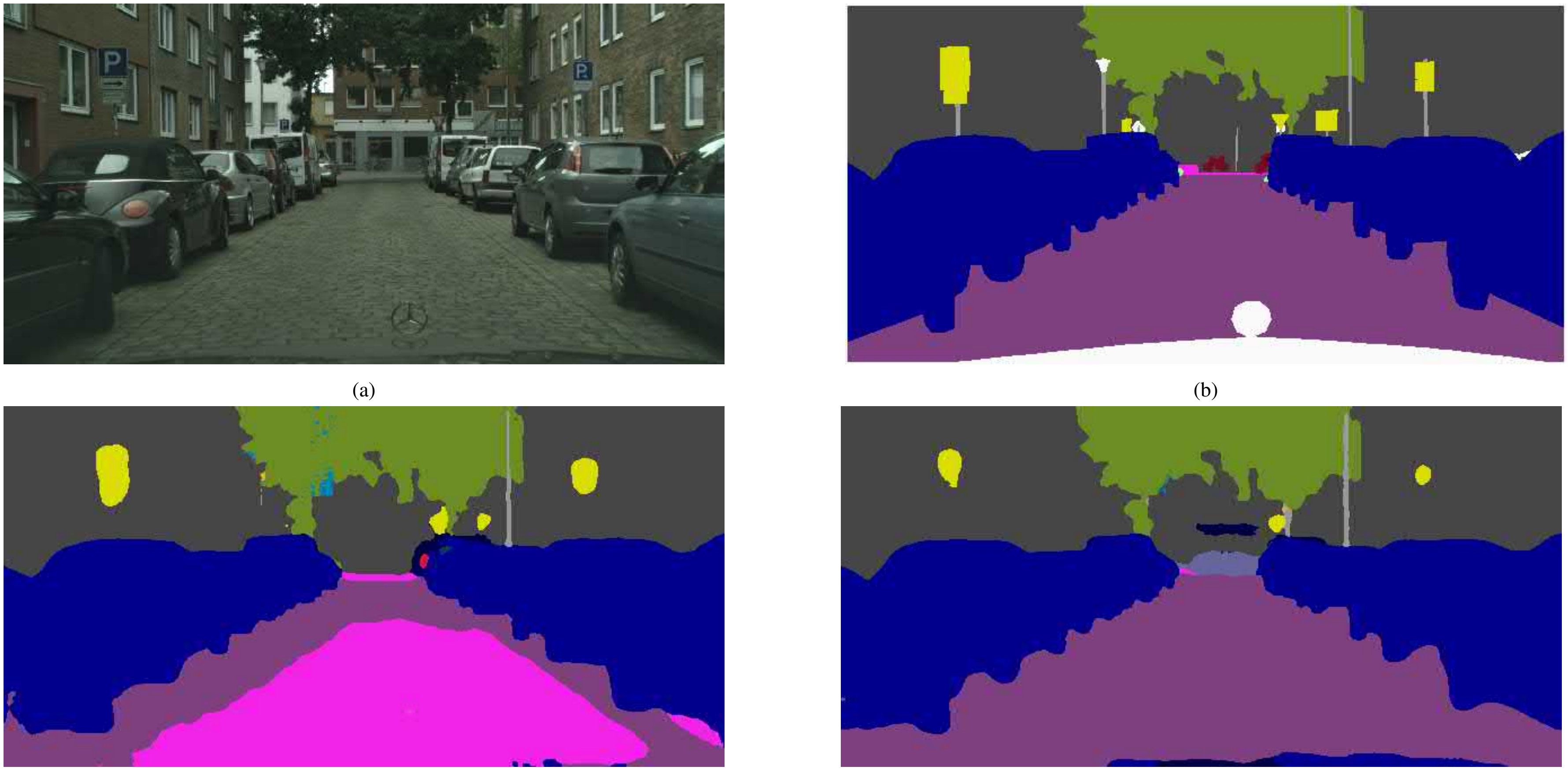


Figure 34: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b) (c) (d)

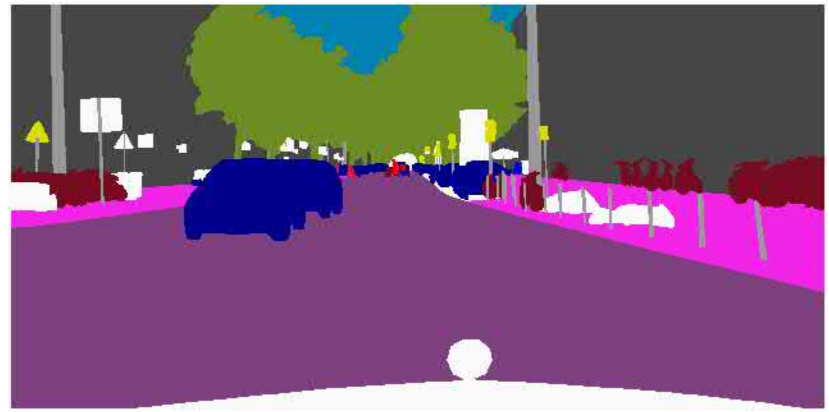
Figure 35: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



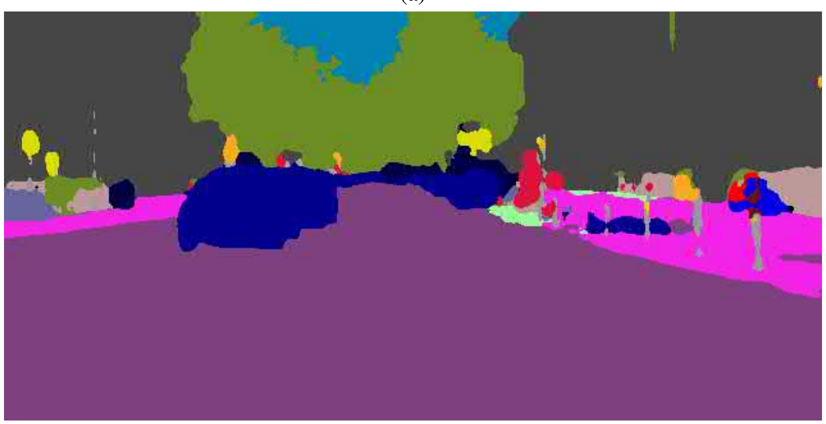
(a) (b) (c) (d)

Figure 36: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

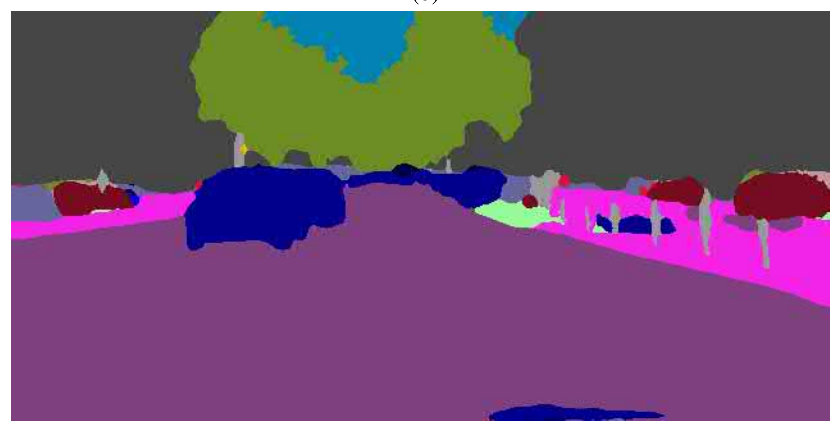
(a)



(b)

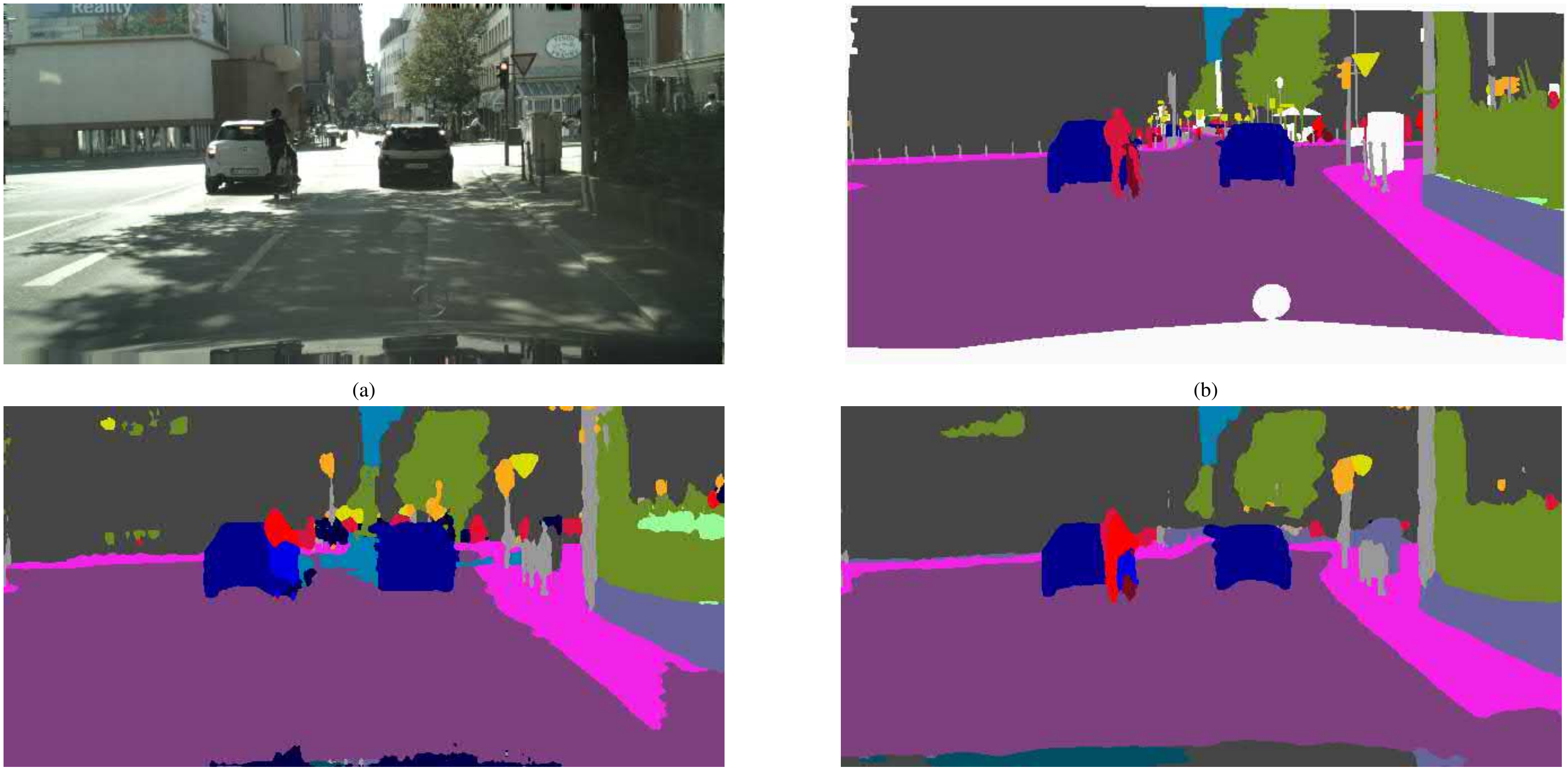


(c)



(d)

Figure 37: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

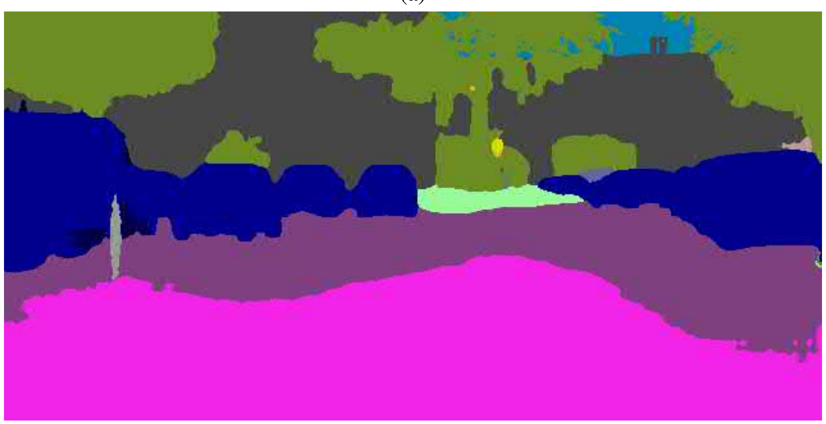


(a) (b) (c) (d)

Figure 38: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)

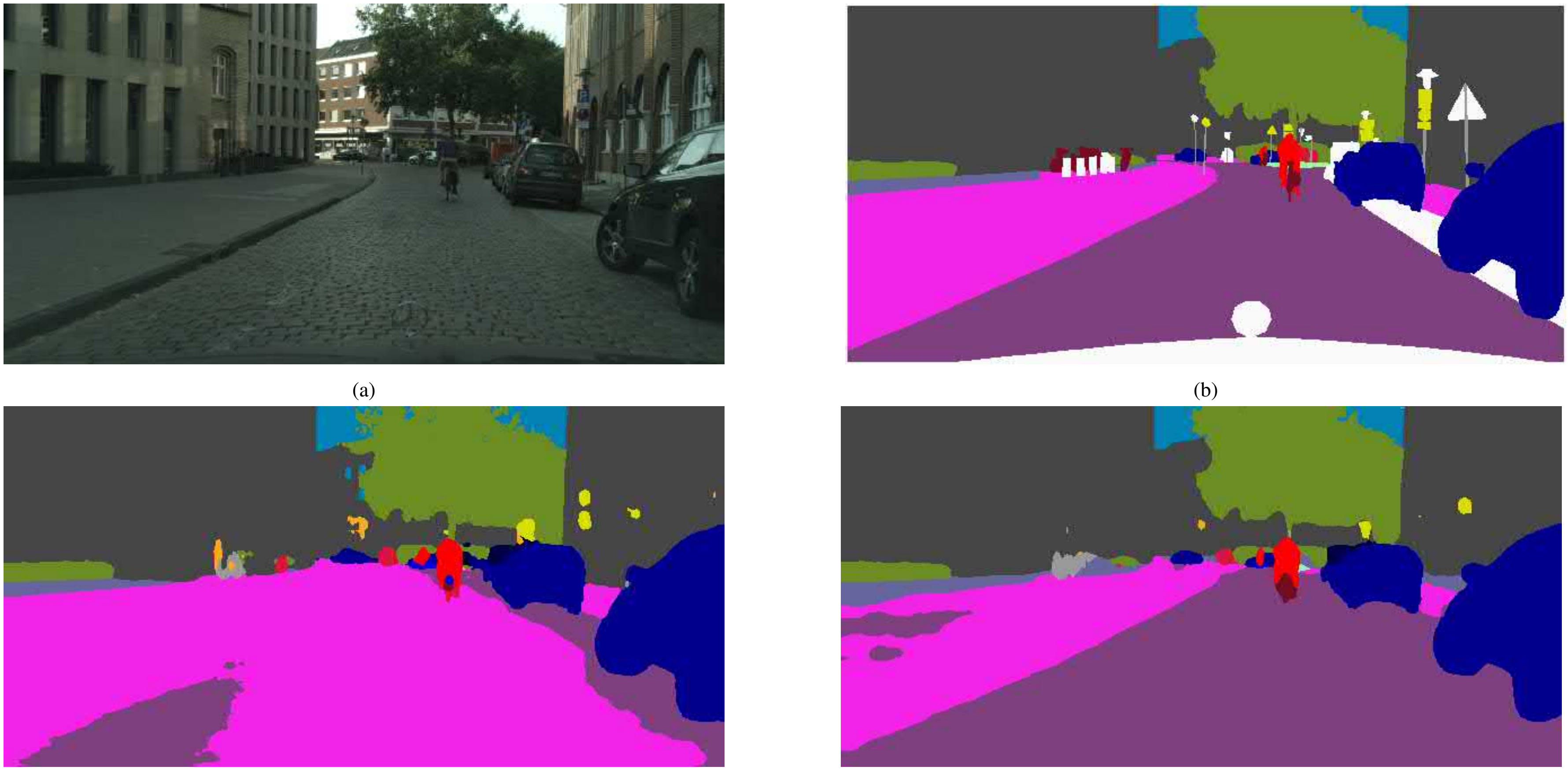
(b)



(c)

(d)

Figure 39: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b) (c) (d)

Figure 40: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

Figure 41: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

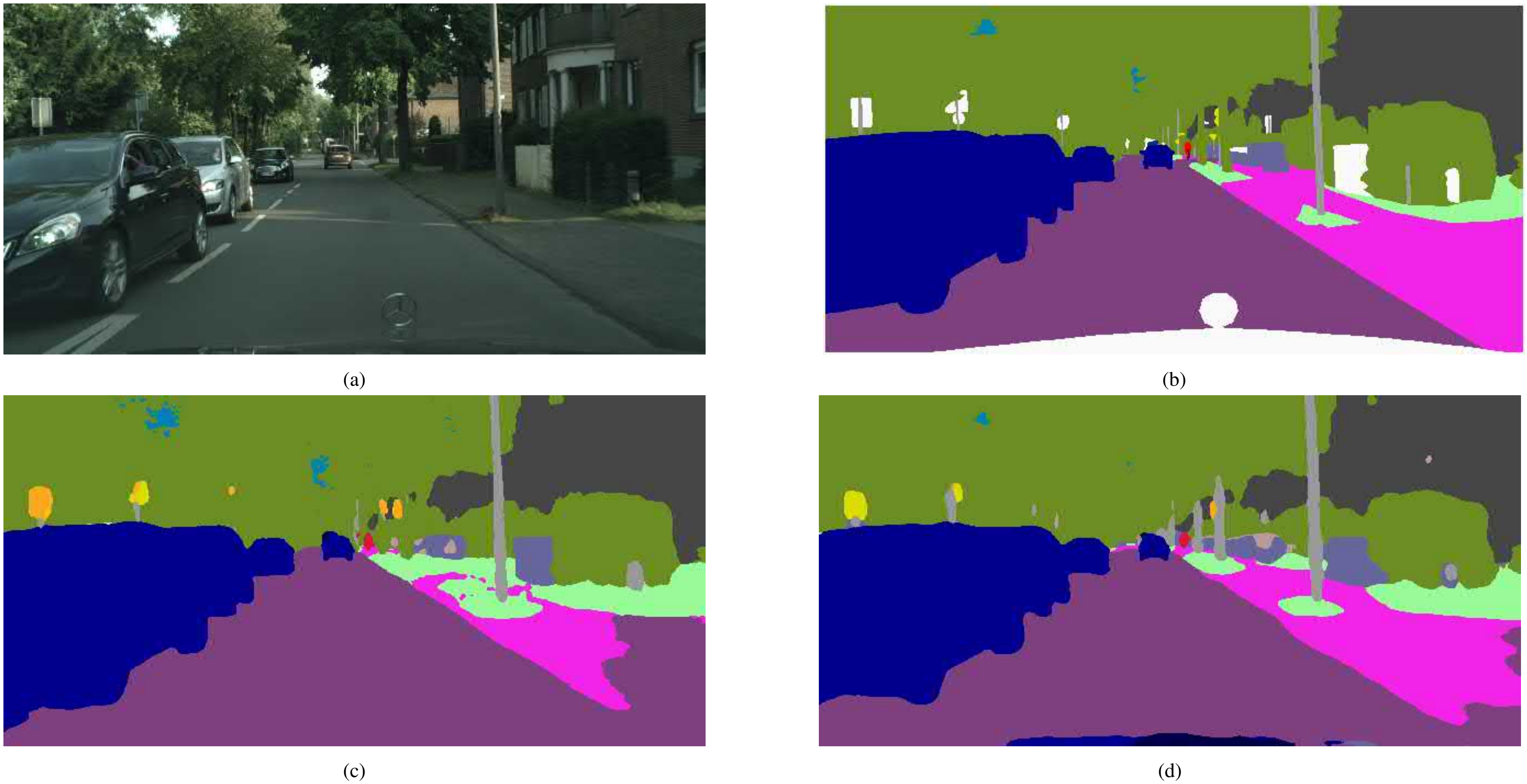


Figure 42: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a) (b) (c) (d)

Figure 43: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

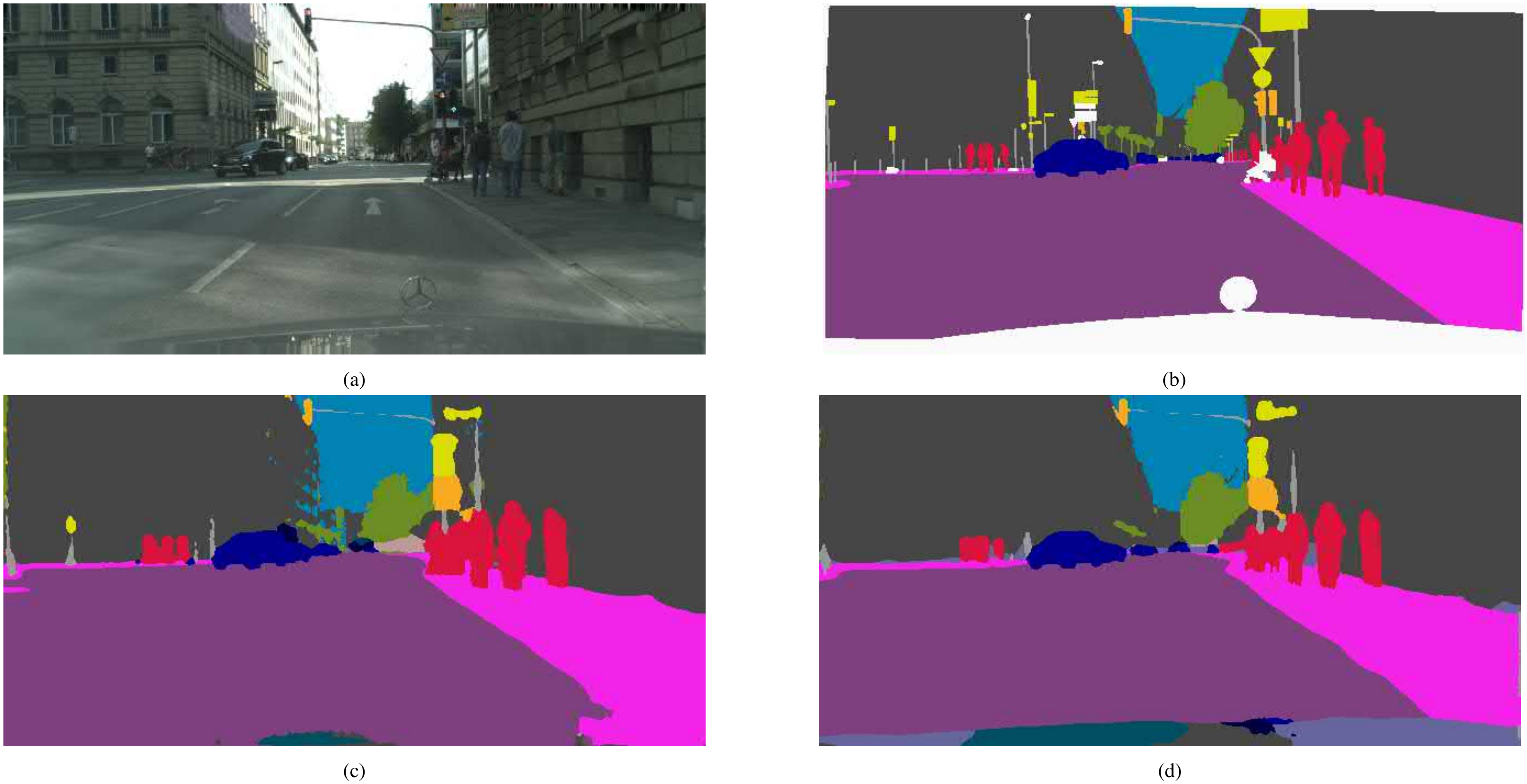


Figure 44: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

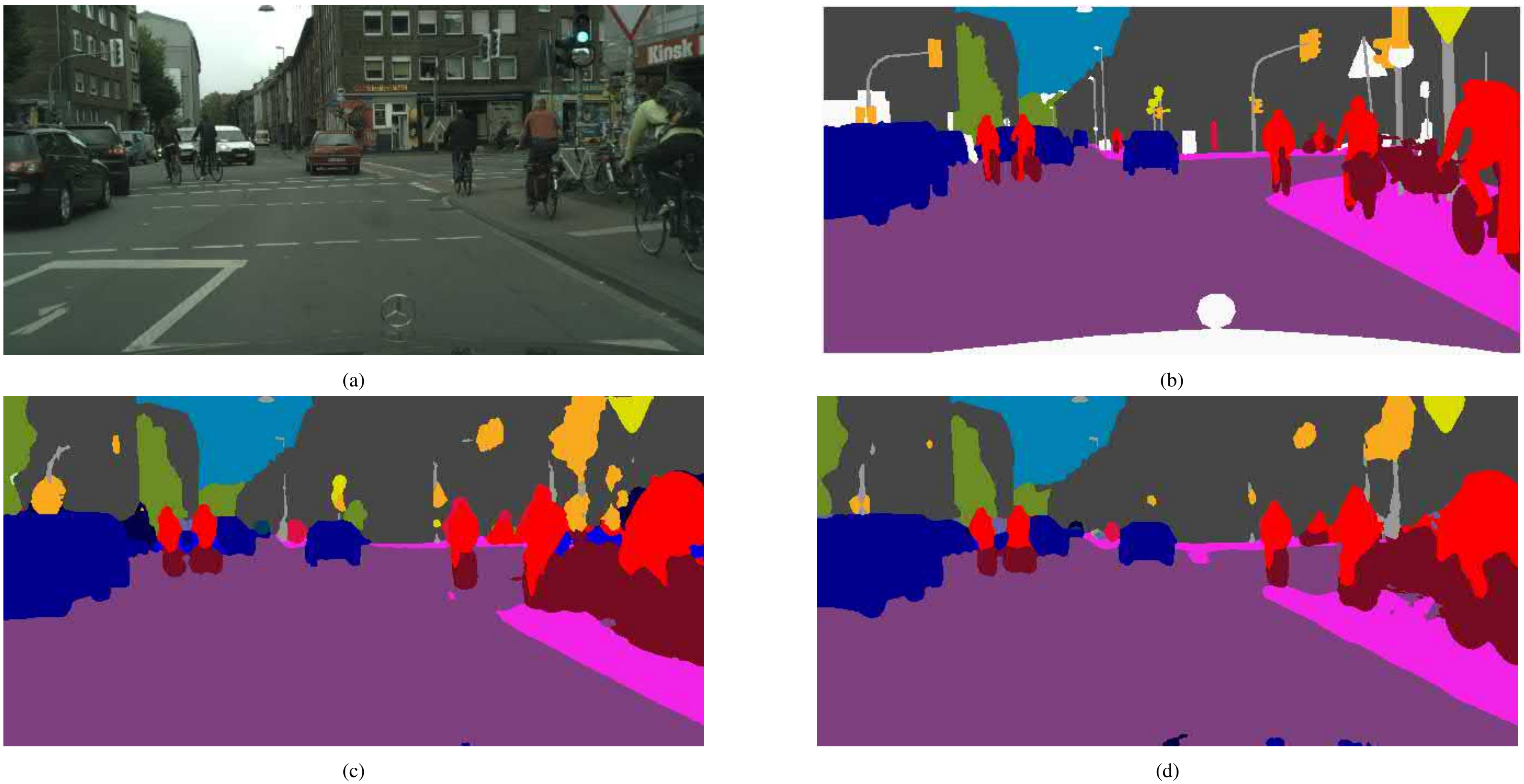
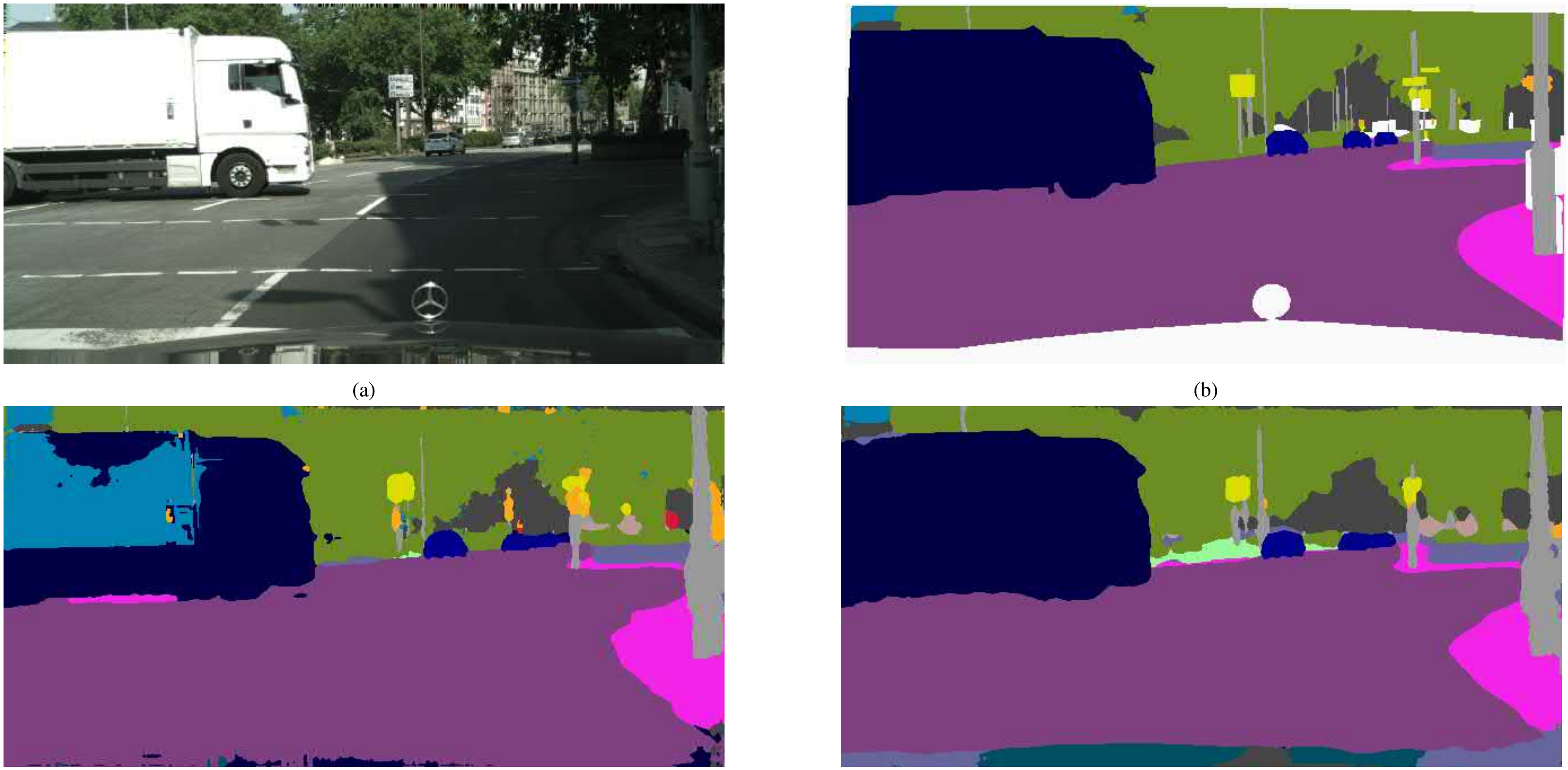


Figure 45: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



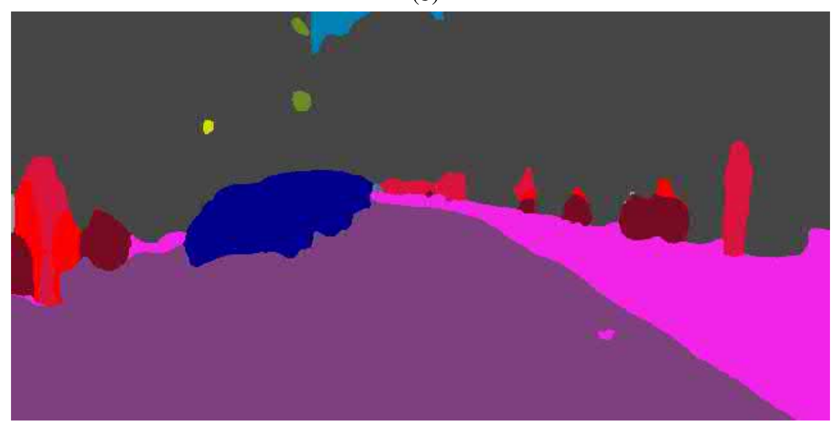
(a) (b) (c) (d)

Figure 46: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)

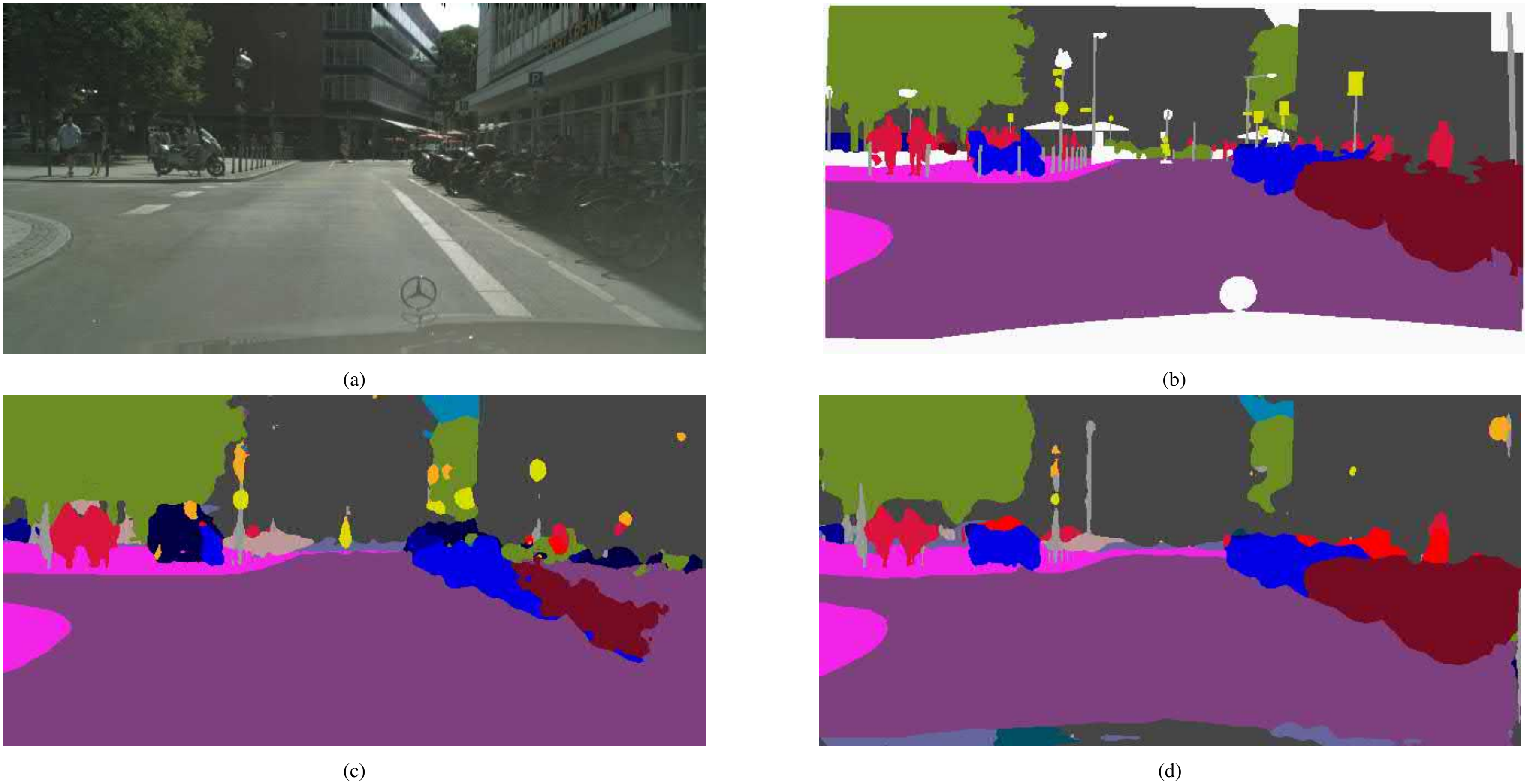
(b)

(c)



(d)

Figure 47: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

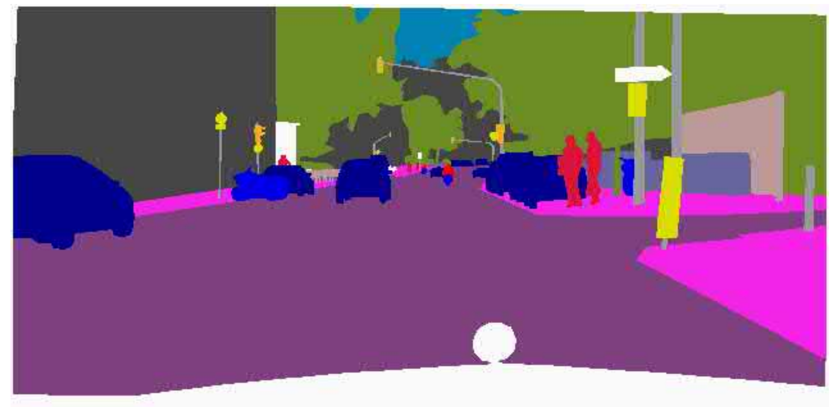


(a) (b)

(c) (d)

Figure 48: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)



(b)

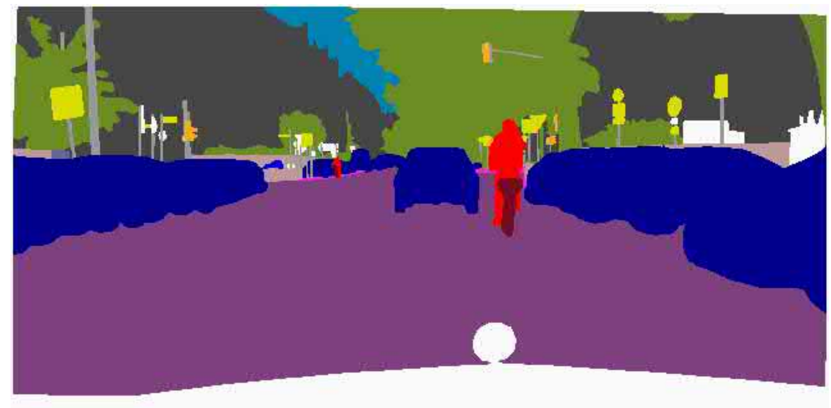
(c)

(d)

Figure 49: Sample segmentation result from GTA5→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

Section 4.4
Sample Segmentation Results on Synthia2Cityscapes

(a)

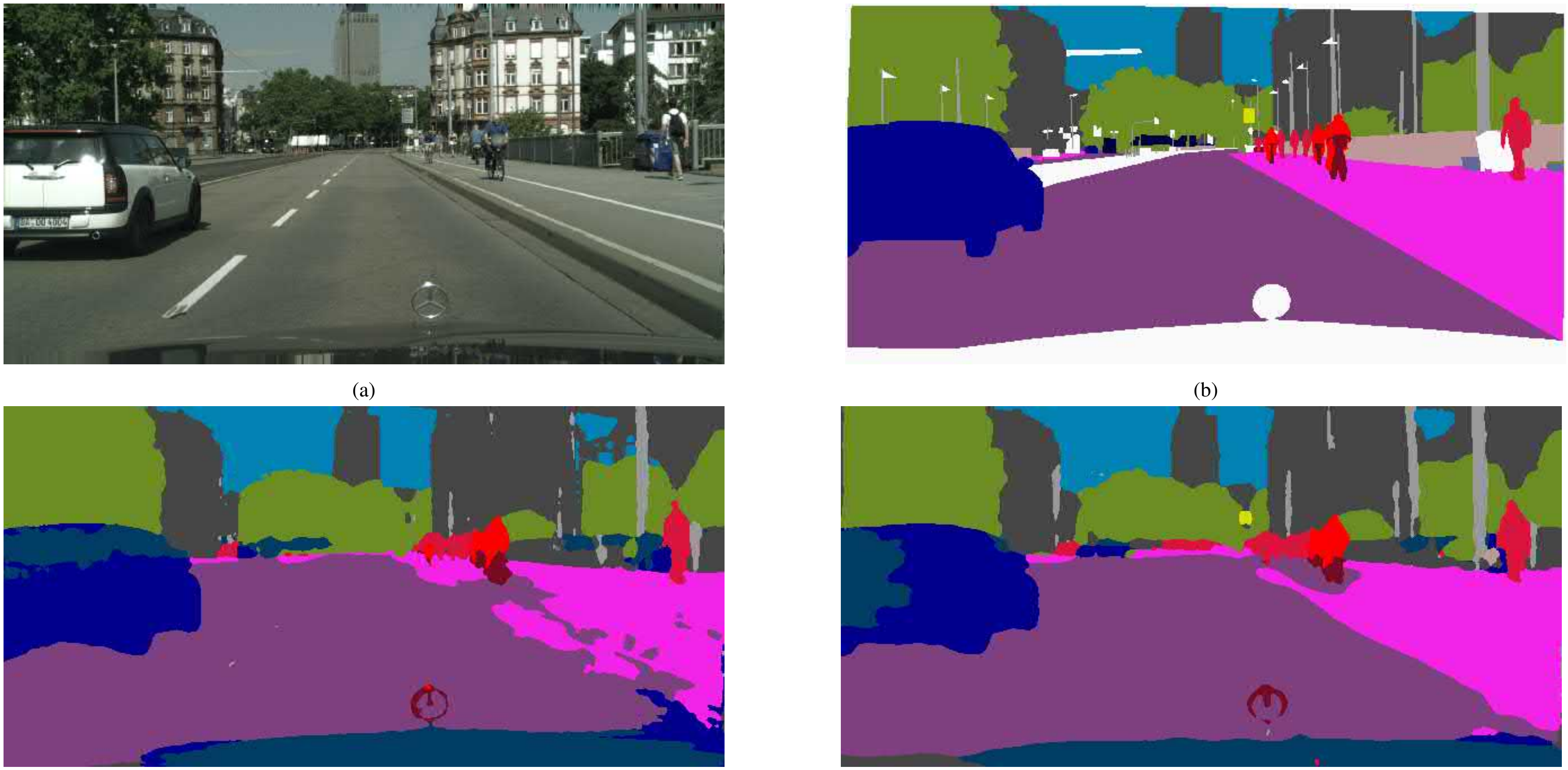


(b)

(c)

(d)

Figure 50: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b) (c) (d)

Figure 51: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

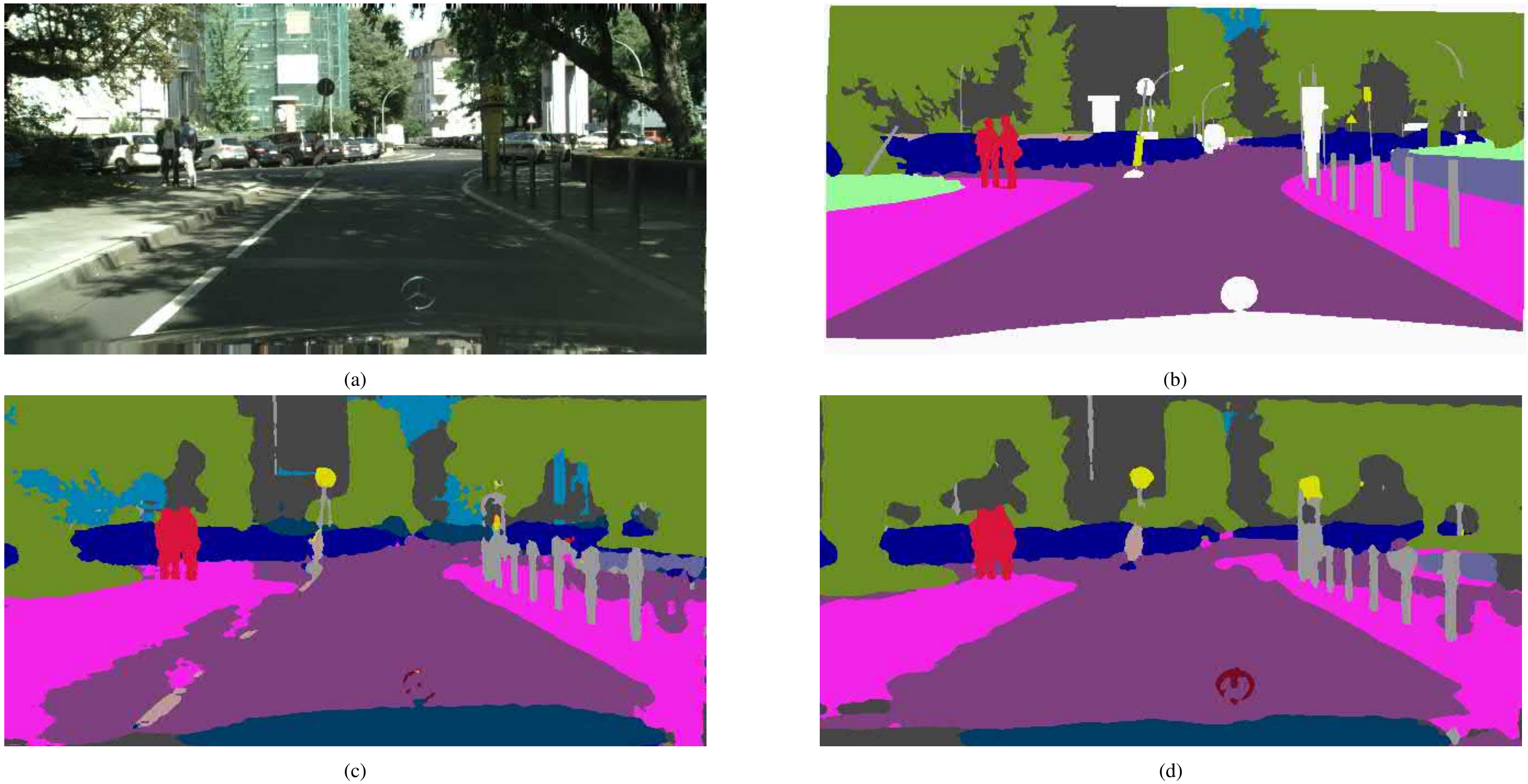


Figure 52: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

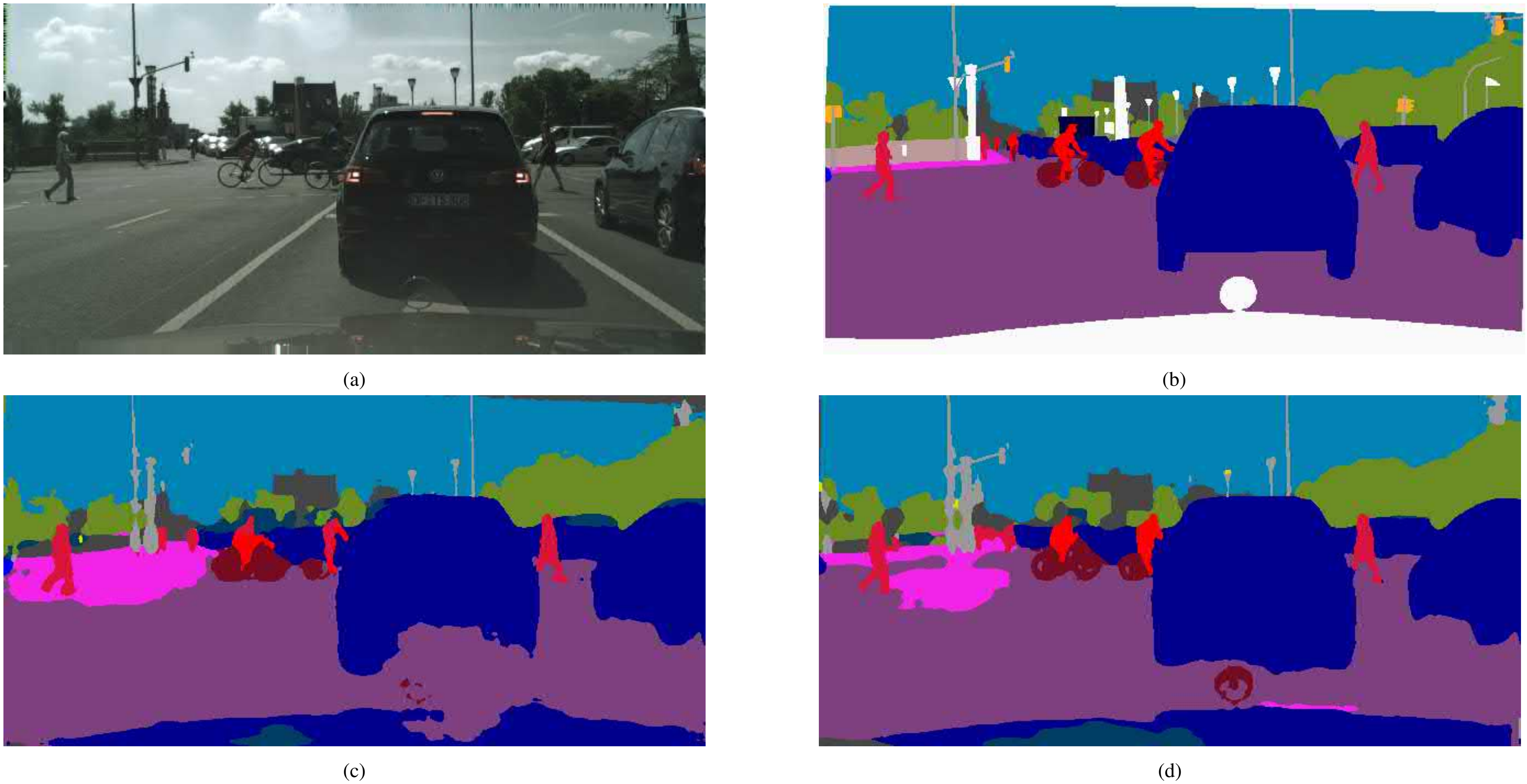


Figure 53: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)

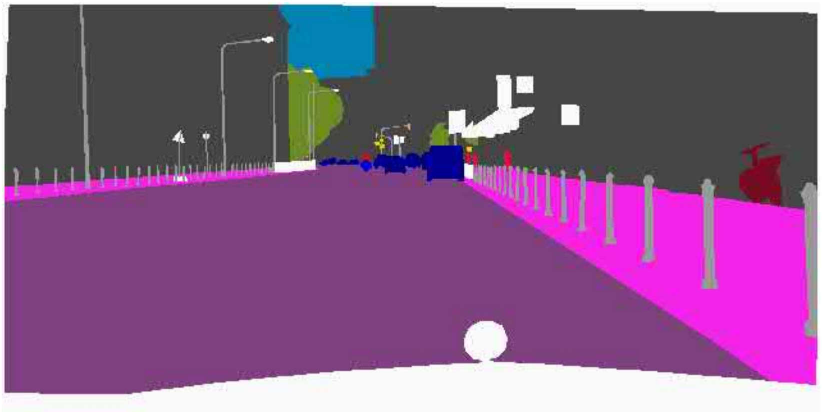
(b)

(c)

(d)

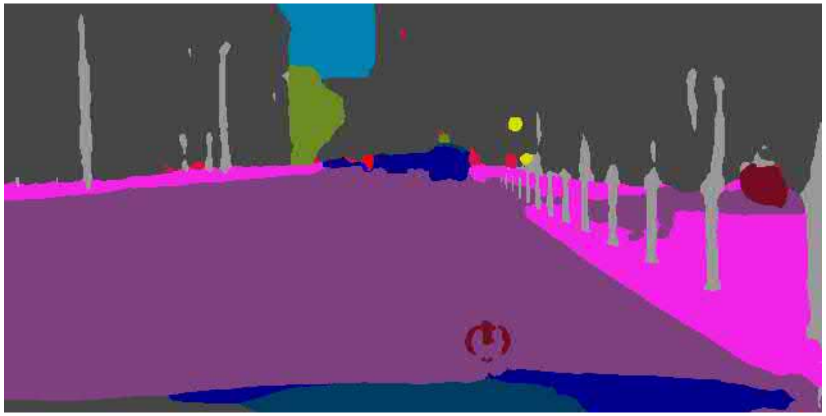
Figure 54: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)



(b)

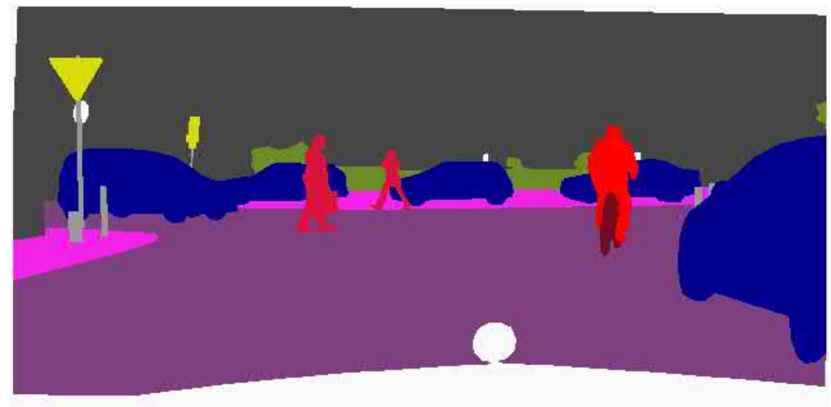
(c)



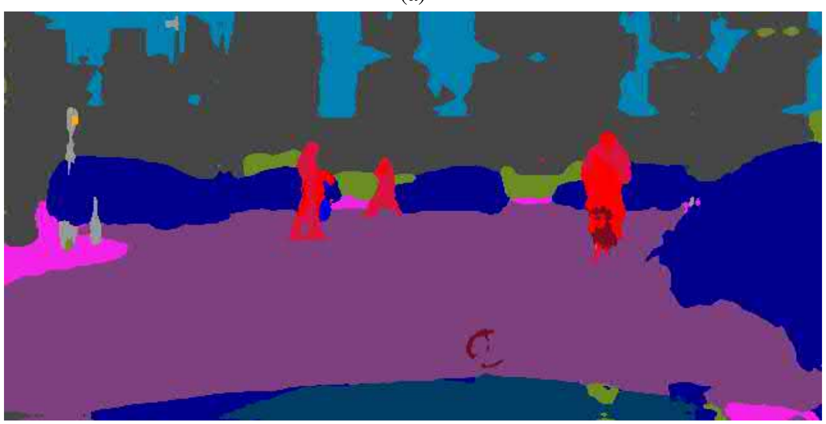
(d)

Figure 55: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

(a)



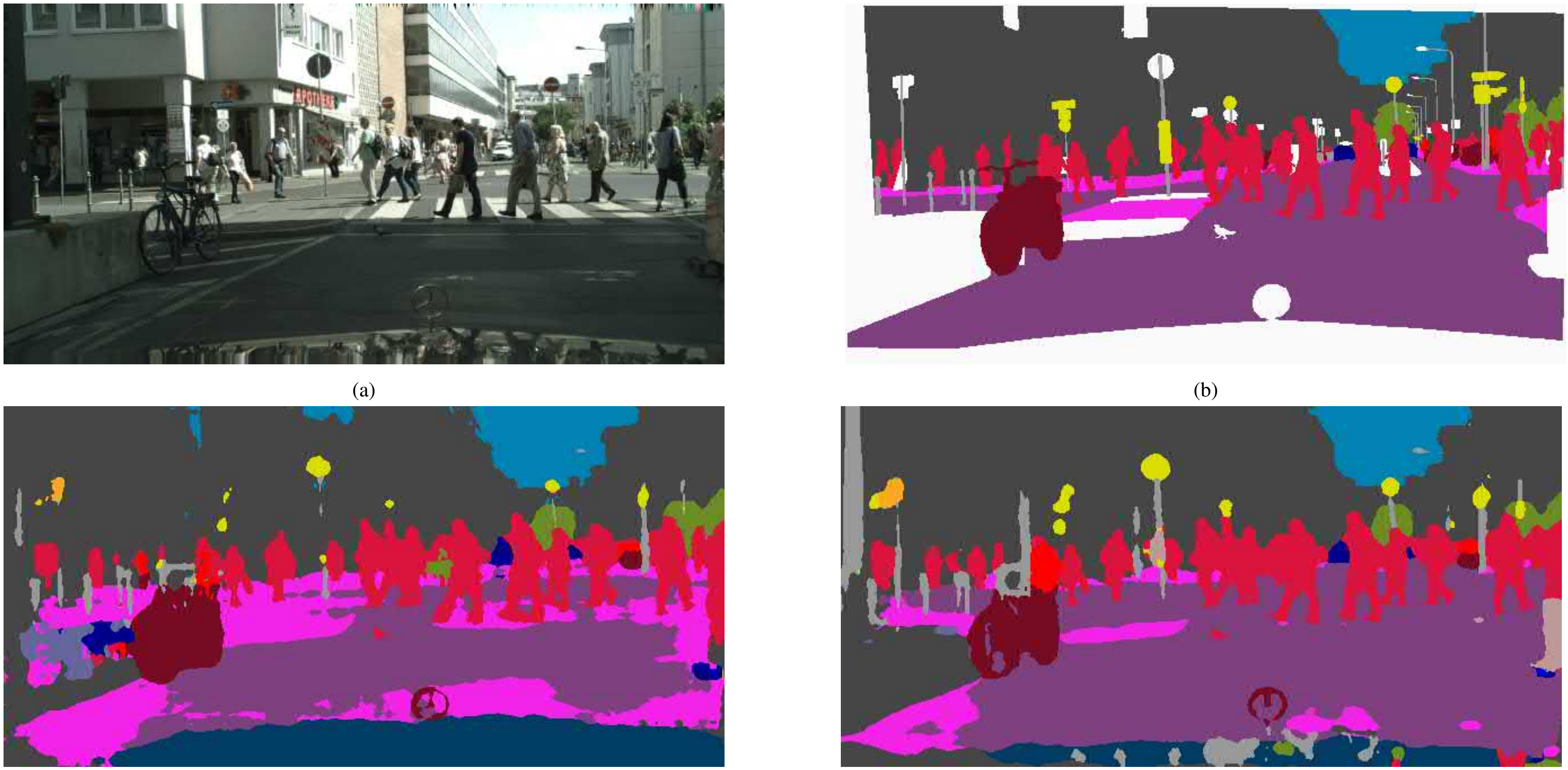
(b)



(c)

(d)

Figure 56: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b)

(c) (d)

Figure 57: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

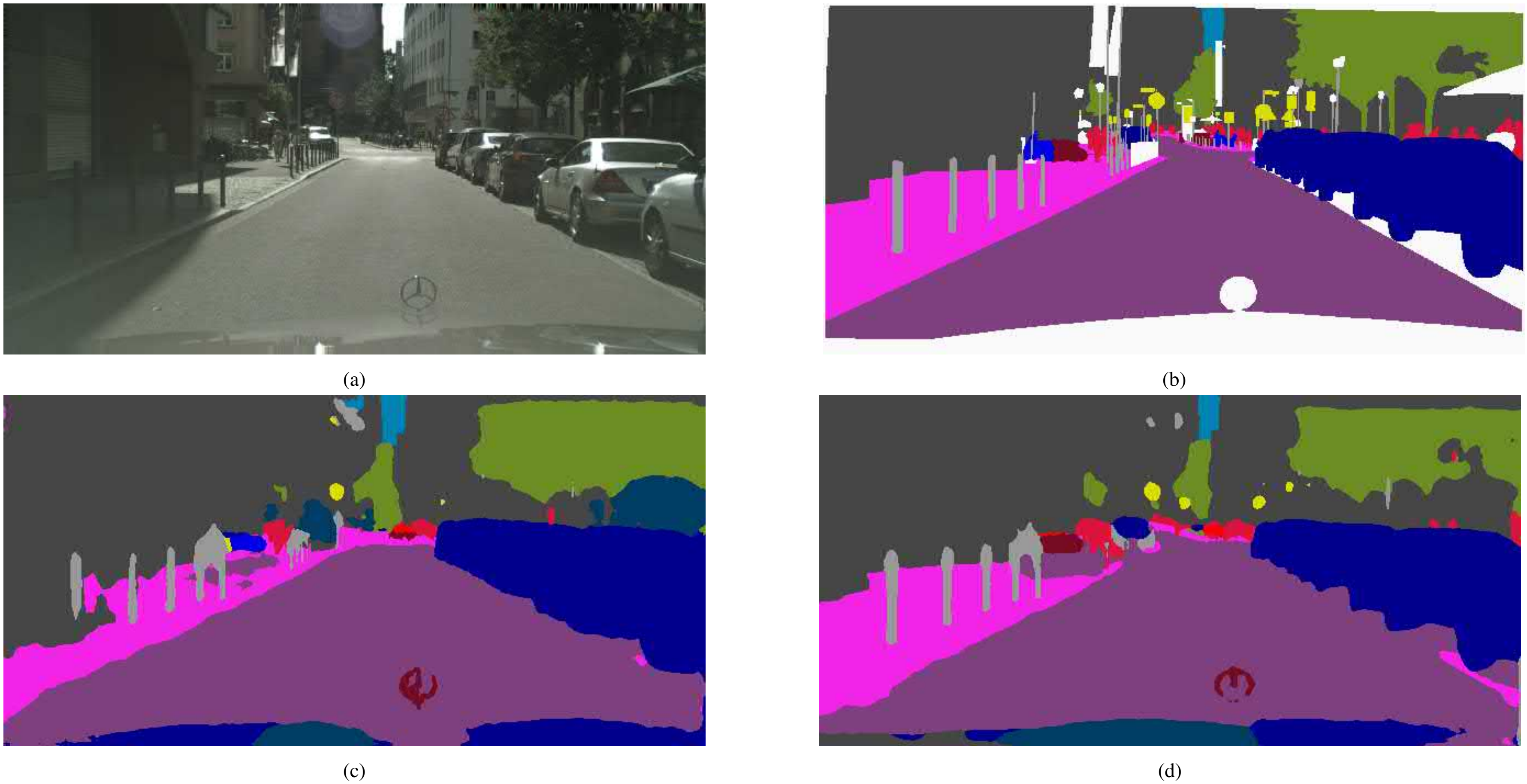
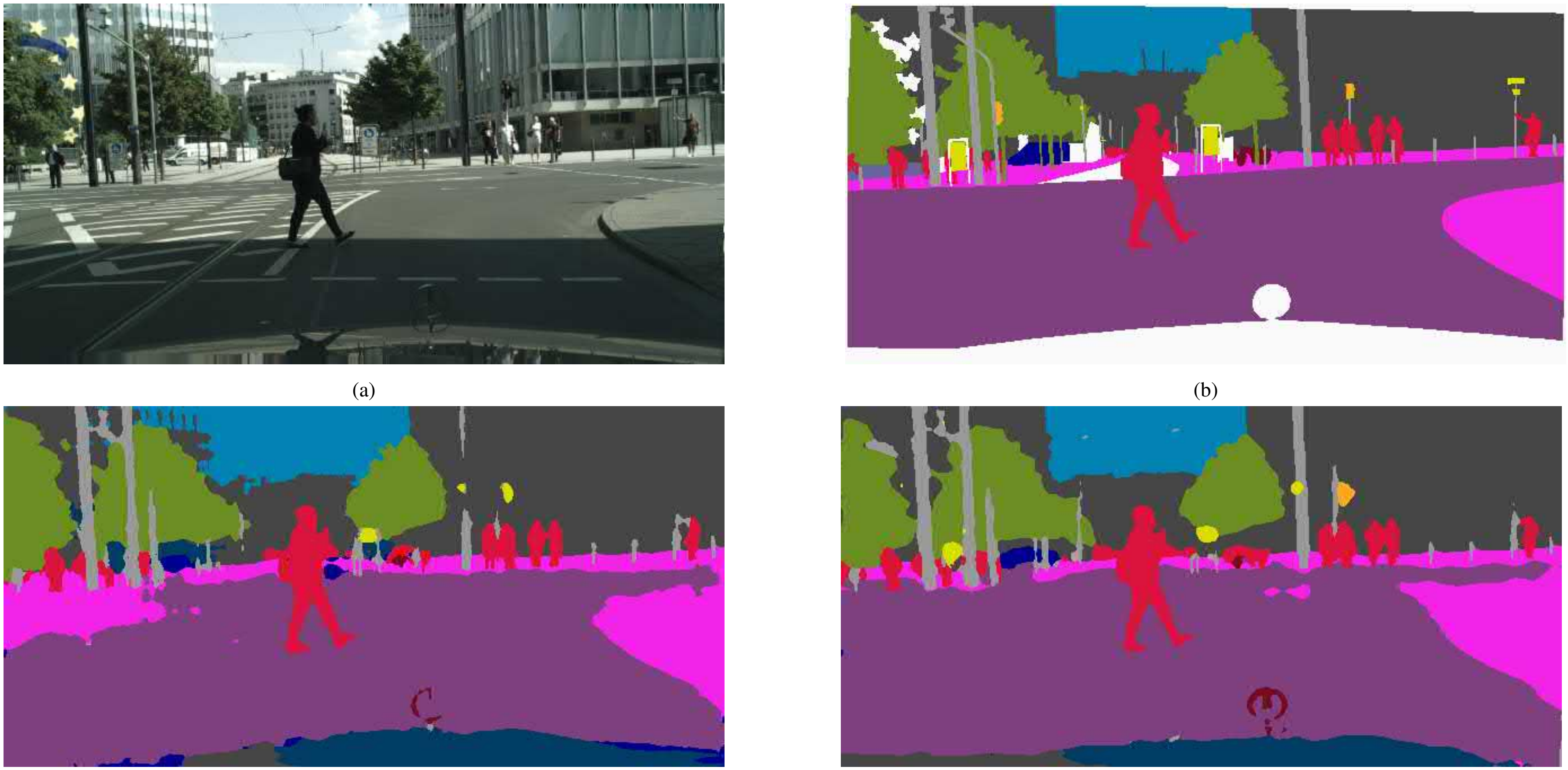


Figure 58: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b) (c) (d)

Figure 59: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

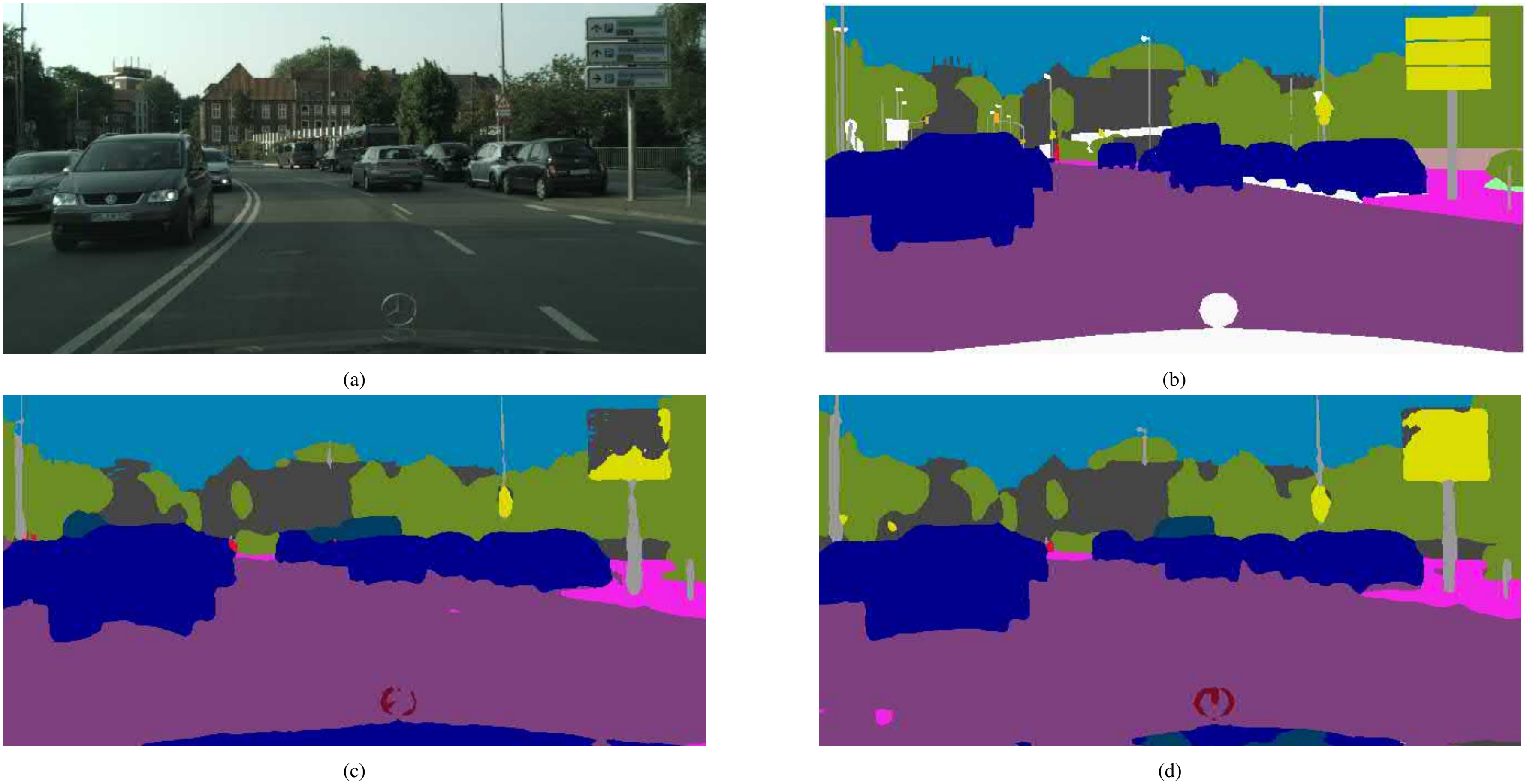
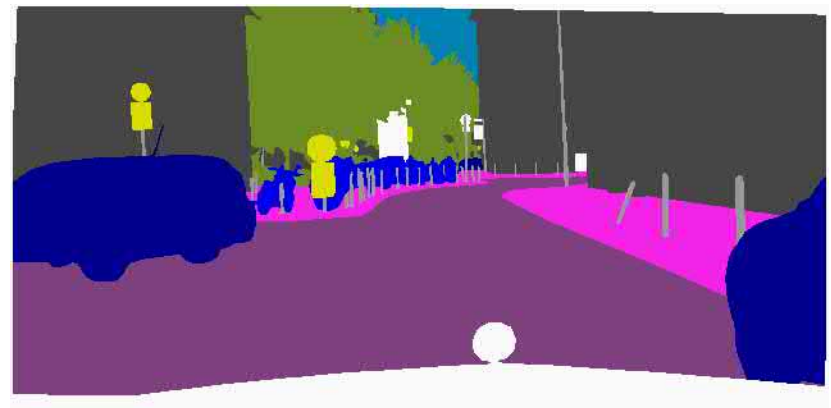
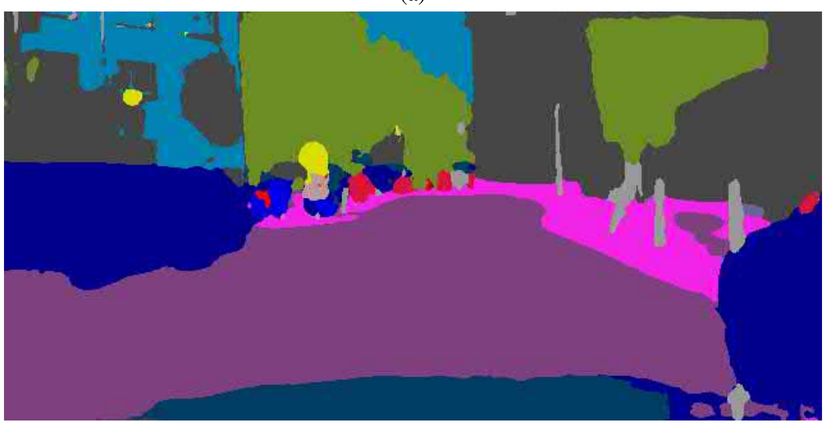


Figure 60: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

Figure 61: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b)



(c) (d)

Figure 62: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

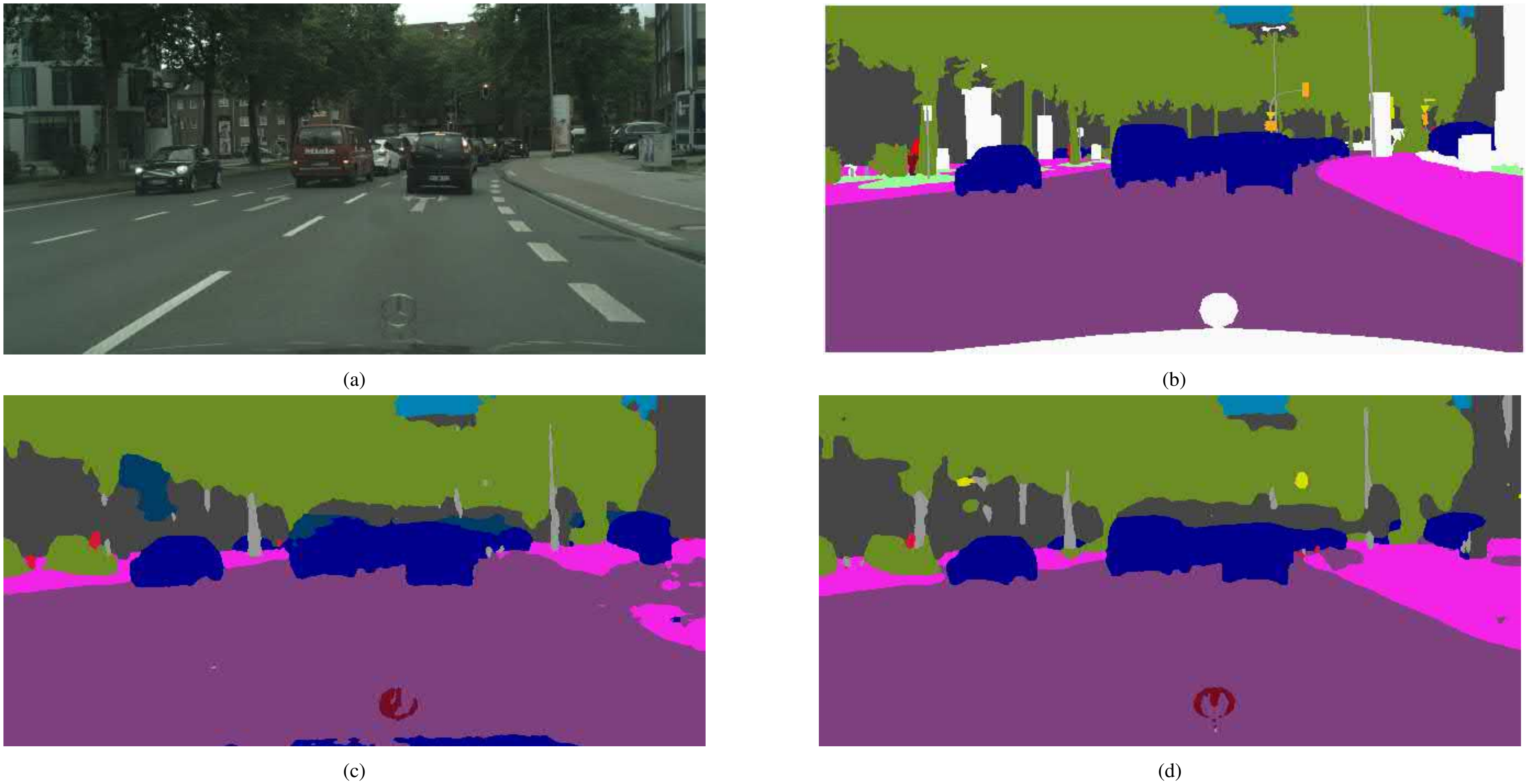


Figure 63: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours

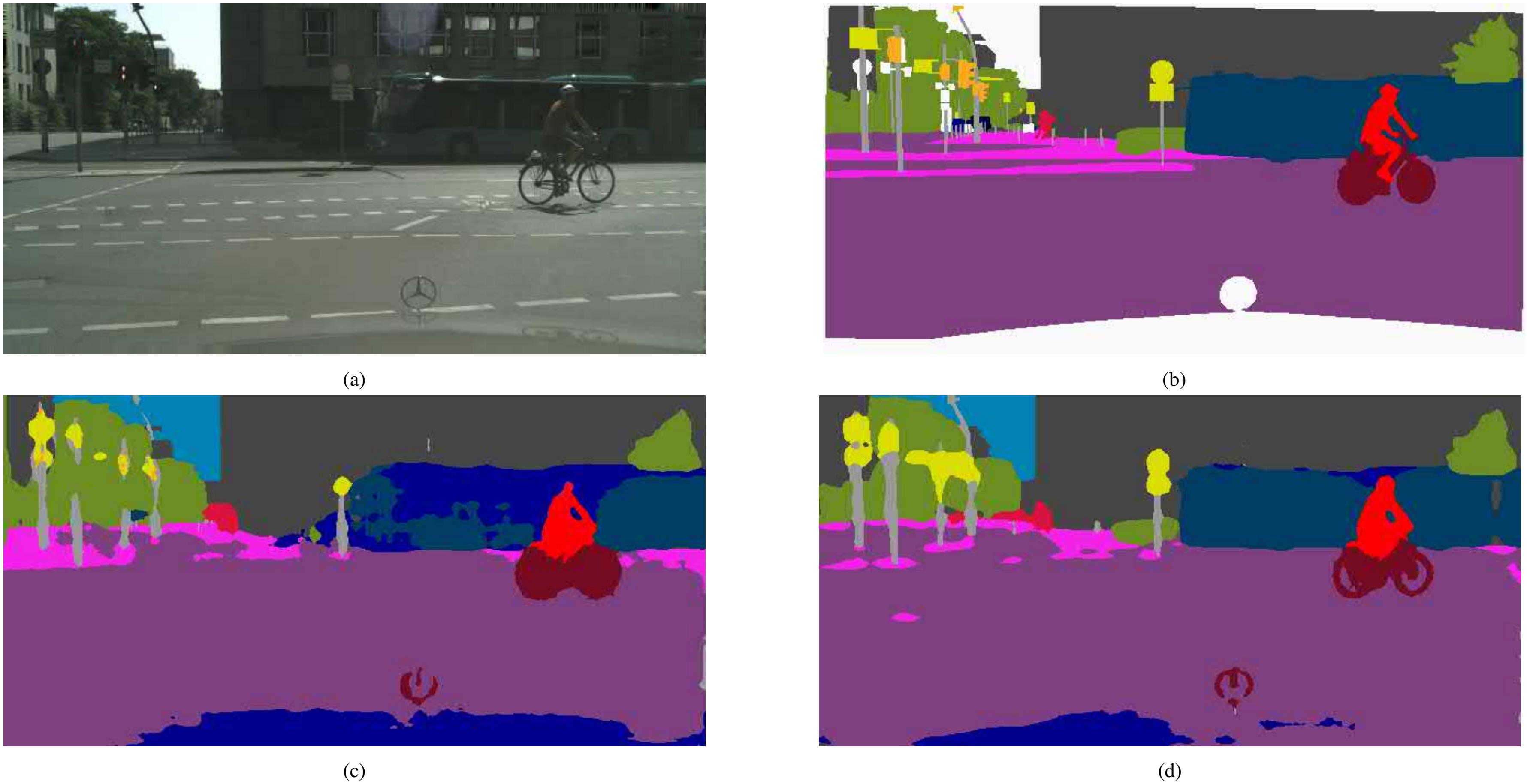
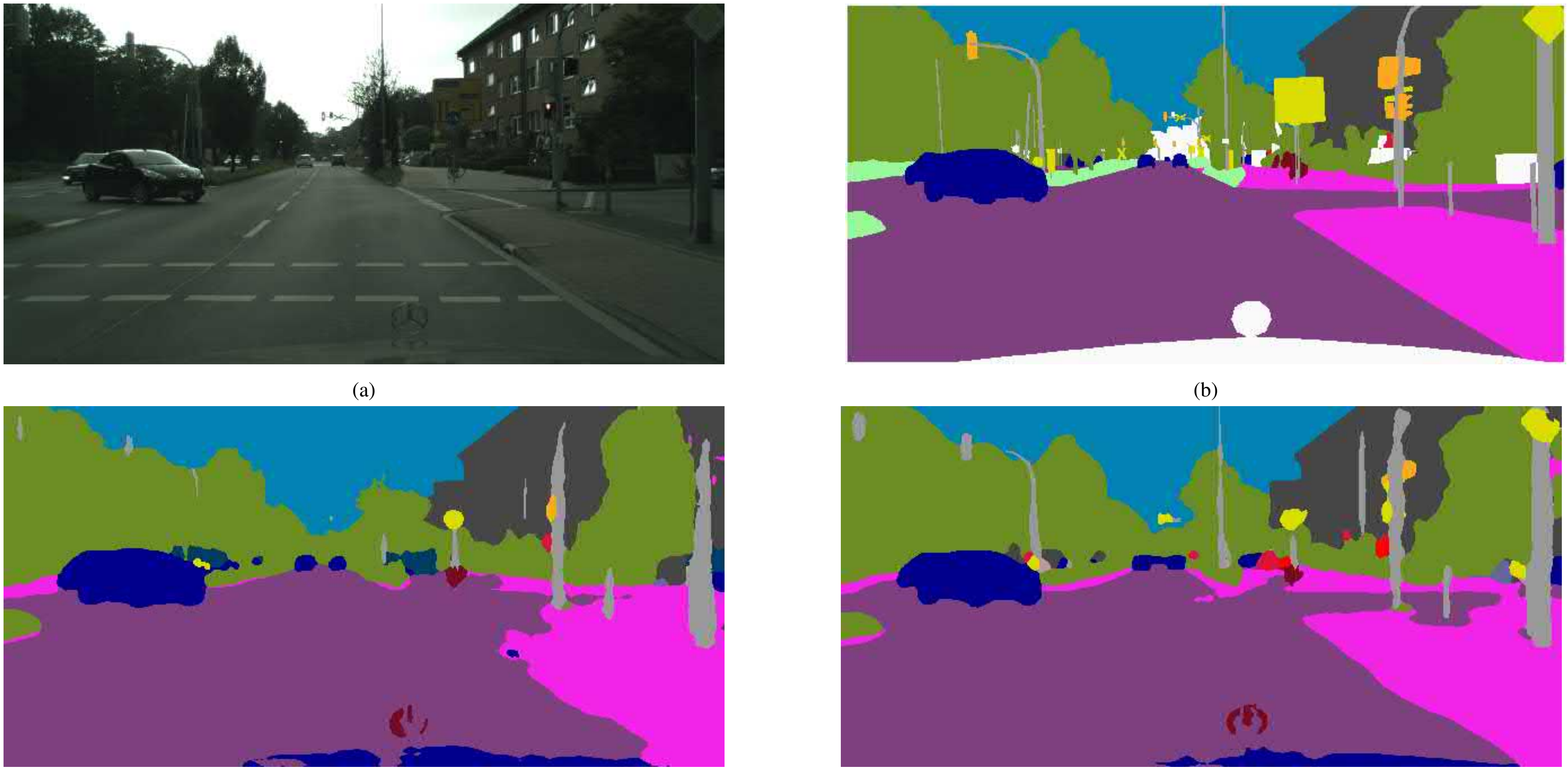


Figure 64: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours



(a) (b) (c) (d)

Figure 65: Sample segmentation result from Synthia→Cityscapes task: (a) input image, (b) ground-truth, (c) CAG[9], (d)ours