Does Image Anonymization Impact Computer Vision Training?

Håkon Hukkelås Frank Lindseth
Deparment of Computer Science, Norwegian University of Science and Technology
hakon.hukkkelas@ntnu.no

A. Additional Results

A.1. Cityscapes All Metrics

Table 1, Table 2, and Table 3 includes all quantitative metrics for Cityscapes, COCO, and BDD100k.

A.2. COCO Instance Segmentation

Table 4 includes experimental results training a Mask R-CNN for general instance segmentation on the COCO datasets. Specifically, we train a Mask R-50 FPN R-CNN. Note that we follow the experimental details from the Keypoint R-CNN experiment in the main paper.

A.3. Cityscapes - Ignoring Person Annotations

Table 5 show experimental results for Mask R-CNN [2] R-50 FPN on the Cityscapes dataset [1] without the person class. It is important to note that we measure the performance drop to the original dataset with person annotations removed.

B. Qualitative Anonymization Examples

The following figures include qualitative examples from Cityscapes [1] and BDD100k [4].

- Cityscapes body: Figure 1, 2, 3.
- Cityscapes Body Histogram matching: Figure 7, 8, 9.
- Cityscapes face: Figure 4, 5, 6.
- BDD100k face: Figure 10, 11, 12, 13.
- BDD100k body: Figure 14, 15, 16, 17.

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 *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3213–3223. IEEE, jun 2016. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

[2] Kaiming He, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. Mask R-CNN. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2980–2988. IEEE, oct 2017. 1, 2

[3] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: Common Objects in Context. In *European conference on computer vision*, volume 8693 LNCS, pages 740–755. Springer, Cham, 2014. 2

[4] Fisher Yu, Haofeng Chen, Xin Wang, Wenqi Xian, Yingying Chen, Fangchen Liu, Vashisht Madhavan, and Trevor Darrell. BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2633–2642. IEEE, jun 2020. 1, 2, 12, 13, 14, 15, 16, 17, 18, 19

Table 1. Instance segmentation AP on the Cityscapes [1] validation set with a Mask R-CNN [2] R-50 FPN. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization.

	Anonymization Method Original	AP ↑ 36.7 ± 0.1 (Δ)	AP@50 ↑ 62.8 ± 0.2 (Δ)	AP_{person} 35.0 ± 0.2 (Δ)	AP_{bus} 57.6 ± 0.6 (Δ)	$AP_{bicycle}$ 23.6 ± 0.2 (Δ)	AP_{car} 53.6 ± 0.0 (Δ)	$AP_{motorcycle}$ 21.9 ± 0.1 (Δ)	AP_{rider} 28.8 ± 0.2 (Δ)	AP_{train} 37.2 ± 0.4 (Δ)	AP_{truck} 36.1 ± 0.4 (Δ)
Face	Blur	36.4 ± 0.2 (-0.3)	62.5 ± 0.2 (-0.3)	34.9 ± 0.1 (-0.1)	**58.0** ± 0.5 (0.4)	**23.3** ± 0.1 (-0.3)	53.1 ± 0.1 (-0.5)	20.8 ± 0.5 (-1.1)	28.6 ± 0.1 (-0.2)	37.0 ± 1.3 (-0.2)	35.5 ± 0.5 (-0.6)
	Mask-out	**36.7** ± 0.2 (0.0)	**63.1** ± 0.2 (0.3)	34.9 ± 0.1 (-0.1)	57.5 ± 0.5 (-0.1)	23.2 ± 0.2 (-0.4)	53.2 ± 0.0 (-0.4)	21.4 ± 0.1 (-0.5)	28.7 ± 0.1 (-0.1)	**39.5** ± 0.2 (2.3)	35.5 ± 0.3 (-0.6)
	Realistic	36.6 ± 0.1 (-0.1)	62.8 ± 0.3 (0.0)	**35.0** ± 0.1 (0.0)	57.2 ± 0.3 (-0.4)	23.2 ± 0.2 (-0.4)	**53.4** ± 0.1 (-0.2)	**21.7** ± 0.3 (-0.2)	**28.8** ± 0.0 (0.0)	36.6 ± 1.2 (-0.6)	**36.7** ± 0.5 (0.6)
Body	Blur	31.4 ± 0.2 (-5.3)	54.5 ± 0.4 (-8.3)	2.1 ± 0.1 (-32.9)	56.7 ± 0.6 (-0.9)	22.7 ± 0.1 (-0.9)	52.9 ± 0.1 (-0.7)	20.9 ± 0.2 (-1.0)	25.9 ± 0.2 (-2.9)	34.4 ± 0.5 (-2.8)	36.0 ± 0.7 (-0.1)
	Mask-out	31.2 ± 0.1 (-5.5)	53.2 ± 0.1 (-9.6)	0.7 ± 0.1 (-34.3)	55.6 ± 0.4 (-2.0)	22.9 ± 0.2 (-0.7)	52.9 ± 0.1 (-0.7)	21.7 ± 0.7 (-0.2)	25.3 ± 0.2 (-3.5)	35.5 ± 0.5 (-1.7)	35.1 ± 0.3 (-1.0)
	Realistic	34.6 ± 0.1 (-2.1)	59.0 ± 0.3 (-3.8)	20.3 ± 0.2 (-14.7)	**58.5** ± 0.2 (0.9)	**23.2** ± 0.1 (-0.4)	**53.4** ± 0.1 (-0.2)	21.2 ± 0.2 (-0.7)	27.9 ± 0.5 (-0.9)	**36.1** ± 1.5 (-1.1)	35.8 ± 0.6 (-0.3)
	Realistic + HM	34.3 ± 0.2 (-2.4)	58.9 ± 0.2 (-3.9)	21.3 ± 0.3 (-13.7)	57.9 ± 0.2 (0.3)	22.8 ± 0.1 (-0.8)	**53.4** ± 0.1 (-0.2)	22.0 ± 0.2 (0.1)	27.6 ± 0.1 (-1.2)	34.0 ± 1.1 (-3.2)	35.7 ± 0.2 (-0.4)
	Realistic + HM-LO	**34.8** ± 0.2 (-1.9)	**60.0** ± 0.3 (-2.8)	**21.5** ± 0.1 (-13.5)	57.7 ± 0.8 (0.1)	**23.2** ± 0.2 (-0.4)	53.2 ± 0.0 (-0.4)	**22.2** ± 0.3 (0.3)	**28.1** ± 0.1 (-0.7)	34.9 ± 1.4 (-2.3)	**37.3** ± 0.7 (1.2)

Table 2. Keypoint (Kp.) AP on the COCO [3] validation set with a Keypoint R-50 FPN R-CNN [2].

	Anonymization Method Original	Box AP ↑ 55.7 ± 0.0 (Δ)	Box AP@50 ↑ 83.3 ± 0.0 (Δ)	Box AP@75 ↑ 60.7 ± 0.1 (Δ)	Box AP_l ↑ 73.0 ± 0.1 (Δ)	Box AP_m ↑ 62.9 ± 0.1 (Δ)	Box AP_s ↑ 37.8 ± 0.1 (Δ)	Kp. AP ↑ 65.2 ± 0.0 (Δ)	Kp. AP@50 ↑ 86.3 ± 0.2 (Δ)	Kp. AP@75 ↑ 71.1 ± 0.2 (Δ)	Kp. AP_l ↑ 73.0 ± 0.0 (Δ)	Kp. AP_m ↑ 61.2 ± 0.1 (Δ)
Face	Blur	50.3 ± 0.2 (-5.4)	77.0 ± 0.2 (-6.3)	54.6 ± 0.1 (-6.1)	67.8 ± 0.4 (-5.2)	52.3 ± 0.2 (-10.6)	37.1 ± 0.0 (-0.7)	53.5 ± 0.2 (-11.7)	74.5 ± 0.4 (-11.8)	57.6 ± 0.2 (-13.5)	63.5 ± 0.3 (-9.5)	48.6 ± 0.1 (-12.6)
	Mask-out	49.9 ± 0.2 (-5.8)	76.6 ± 0.2 (-6.7)	54.1 ± 0.3 (-6.6)	66.9 ± 0.6 (-6.1)	52.3 ± 0.1 (-10.6)	36.8 ± 0.1 (-1.0)	52.0 ± 0.3 (-13.2)	73.5 ± 0.3 (-12.8)	56.2 ± 0.4 (-14.9)	61.4 ± 0.3 (-11.6)	47.7 ± 0.2 (-13.5)
	Realistic	54.3 ± 0.1 (-1.4)	81.7 ± 0.1 (-1.6)	59.0 ± 0.1 (-1.7)	**72.7** ± 0.1 (-0.3)	60.0 ± 0.1 (-2.9)	37.3 ± 0.2 (-0.5)	60.6 ± 0.1 (-4.6)	**82.9** ± 0.3 (-3.4)	65.9 ± 0.1 (-5.2)	69.9 ± 0.1 (-3.1)	**56.1** ± 0.2 (-5.1)
	Realistic refined	**54.4** ± 0.0 (-1.3)	**81.8** ± 0.1 (-1.5)	**59.1** ± 0.1 (-1.6)	**72.7** ± 0.1 (-0.3)	**60.1** ± 0.2 (-2.8)	**37.5** ± 0.1 (-0.3)	**60.8** ± 0.2 (-4.4)	**82.9** ± 0.2 (-3.4)	**66.2** ± 0.4 (-4.9)	**70.2** ± 0.1 (-2.8)	**56.1** ± 0.3 (-5.1)
Body	Blur	17.8 ± 0.0 (-37.9)	35.1 ± 0.1 (-48.2)	16.3 ± 0.1 (-44.4)	2.6 ± 0.1 (-70.4)	10.5 ± 0.1 (-52.4)	33.3 ± 0.1 (-4.5)	4.4 ± 0.1 (-60.8)	9.1 ± 0.2 (-77.2)	3.7 ± 0.1 (-67.4)	0.4 ± 0.1 (-72.6)	10.4 ± 0.1 (-50.8)
	Mask-out	17.4 ± 0.1 (-38.3)	34.5 ± 0.1 (-48.8)	15.5 ± 0.2 (-45.2)	2.1 ± 0.1 (-70.9)	10.6 ± 0.2 (-52.3)	32.5 ± 0.1 (-5.3)	2.0 ± 0.1 (-63.2)	4.9 ± 0.2 (-81.4)	1.4 ± 0.1 (-69.7)	0.1 ± 0.1 (-72.9)	4.1 ± 0.1 (-57.1)
	Realistic	**24.0** ± 0.1 (-31.7)	**46.1** ± 0.2 (-37.2)	**22.4** ± 0.1 (-38.3)	**8.2** ± 0.3 (-64.8)	**26.4** ± 0.3 (-36.5)	**34.1** ± 0.1 (-3.7)	**15.6** ± 0.1 (-49.6)	**29.4** ± 0.2 (-56.9)	**14.2** ± 0.1 (-56.9)	**13.0** ± 0.1 (-60.0)	**22.5** ± 0.1 (-38.7)

Table 3. Instance segmentation AP on the BDD100K [4] validation set with a Mask R-CNN [2] R-50 FPN.

	Method Original	AP ↑ 20.2 ± 0.2	AP@50 ↑ 34.9 ± 0.4	$AP_{pedestrian}$ 32.0 ± 0.0	AP_{bus} 30.2 ± 0.2	$AP_{bicycle}$ 6.0 ± 0.3	AP_{car} 45.4 ± 0.1	$AP_{motorcycle}$ 11.0 ± 0.9	AP_{rider} 9.7 ± 0.3	AP_{train} 0.0 ± 0.0	AP_{truck} 26.9 ± 0.4
Face	Blur	20.5 ± 0.1 (0.3)	**35.9** ± 0.1 (1.0)	**31.7** ± 0.1 (-0.3)	30.1 ± 0.3 (-0.1)	**6.9** ± 0.2 (0.9)	45.4 ± 0.1 (0.0)	13.8 ± 1.0 (2.8)	**9.4** ± 0.3 (-0.3)	0.0 ± 0.0 (0.0)	26.5 ± 0.2 (-0.4)
	Mask-out	20.3 ± 0.1 (0.1)	35.3 ± 0.3 (0.4)	31.4 ± 0.1 (-0.6)	29.9 ± 0.5 (-0.3)	5.8 ± 0.3 (-0.2)	45.5 ± 0.1 (0.1)	14.4 ± 0.6 (3.4)	8.8 ± 0.4 (-0.9)	0.0 ± 0.0 (0.0)	26.3 ± 0.2 (-0.6)
	Realistic	**20.6** ± 0.1 (0.4)	35.8 ± 0.3 (0.9)	31.6 ± 0.2 (-0.4)	**30.7** ± 0.6 (0.5)	6.7 ± 0.4 (0.7)	**45.6** ± 0.1 (0.2)	**14.7** ± 0.7 (3.7)	8.7 ± 0.2 (-1.0)	0.0 ± 0.0 (0.0)	**26.7** ± 0.0 (-0.2)
Body	Blur	15.4 ± 0.1 (-4.8)	26.3 ± 0.2 (-8.6)	0.5 ± 0.0 (-31.5)	29.5 ± 0.6 (-0.7)	5.4 ± 0.2 (-0.6)	**45.6** ± 0.0 (0.2)	12.1 ± 0.9 (1.1)	4.2 ± 0.3 (-5.5)	0.0 ± 0.0 (0.0)	25.9 ± 0.6 (-1.0)
	Mask-out	15.3 ± 0.0 (-4.9)	25.5 ± 0.1 (-9.4)	0.0 ± 0.0 (-32.0)	**30.8** ± 0.2 (0.6)	5.5 ± 0.2 (-0.5)	45.5 ± 0.1 (0.1)	10.9 ± 0.2 (-0.1)	3.9 ± 0.5 (-5.8)	0.0 ± 0.0 (0.0)	**26.0** ± 0.4 (-0.9)
	Realistic	**17.0** ± 0.1 (-3.2)	**28.9** ± 0.4 (-6.0)	**12.8** ± 0.1 (-19.2)	29.7 ± 0.4 (-0.5)	**6.7** ± 0.3 (0.7)	45.2 ± 0.2 (-0.2)	10.3 ± 0.6 (-0.7)	**5.8** ± 0.4 (-3.9)	0.0 ± 0.0 (0.0)	25.9 ± 0.4 (-1.0)

Table 4. Instance segmentation AP on the COCO [3] validation set with a Mask R-50 FPN R-CNN [2].

	Anonymization Method Original	Box AP ↑ 40.9 ± 0.0 (Δ)	Segm. AP ↑ 37.0 ± 0.0 (Δ)	Bbox. AP_{person} ↑ 55.3 ± 0.1 (Δ)
Face	Blur	40.7 ± 0.0 (-0.2)	36.9 ± 0.1 (-0.1)	51.9 ± 0.1 (-3.4)
	Mask-out	40.6 ± 0.1 (-0.3)	36.9 ± 0.0 (-0.1)	51.6 ± 0.1 (-3.7)
	Realistic	**40.8** ± 0.1 (-0.1)	**37.0** ± 0.0 (0.0)	**54.6** ± 0.1 (-0.7)

Table 5. Instance segmentation AP on the Cityscapes [1] validation set with a Mask R-CNN [2] R-50 FPN. Note that the "person" class is removed from all experiments in this table, including the original dataset.

	Anonymization Method Original	AP ↑ 30.9 ± 0.3 (Δ)	AP@50 ↑ 52.7 ± 0.7 (Δ)	AP_{bus} 56.1 ± 0.3 (Δ)	$AP_{bicycle}$ 22.2 ± 0.1 (Δ)	AP_{car} 51.3 ± 0.1 (Δ)	$AP_{motorcycle}$ 20.5 ± 0.3 (Δ)	AP_{rider} 25.9 ± 0.2 (Δ)	AP_{train} 36.3 ± 1.4 (Δ)	AP_{truck} 34.7 ± 0.8 (Δ)
Body	Blur	30.4 ± 0.1 (-0.5)	51.6 ± 0.1 (-1.1)	56.3 ± 0.1 (0.2)	21.9 ± 0.1 (-0.3)	50.7 ± 0.1 (-0.6)	21.1 ± 0.2 (0.6)	24.4 ± 0.2 (-1.5)	35.8 ± 0.6 (-0.5)	33.3 ± 0.5 (-1.4)
	Mask-out	30.6 ± 0.2 (-0.3)	51.8 ± 0.3 (-0.9)	55.3 ± 0.8 (-0.8)	21.5 ± 0.0 (-0.7)	51.1 ± 0.1 (-0.2)	20.7 ± 0.3 (0.2)	23.7 ± 0.2 (-2.2)	38.3 ± 0.6 (2.0)	34.4 ± 0.6 (-0.3)

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 1. Random anonymization examples from Cityscapes [1]. **HM=**Histogram matching. **HM-LO=**Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 2. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 3. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 4. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 5. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 6. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Body - Realistic (b) Body - Realistic (HM) (c) Body - Realistic (HM-LO)

Figure 7. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Body - Realistic (b) Body - Realistic (HM) (c) Body - Realistic (HM-LO)

Figure 8. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Body - Realistic (b) Body - Realistic (HM) (c) Body - Realistic (HM-LO)

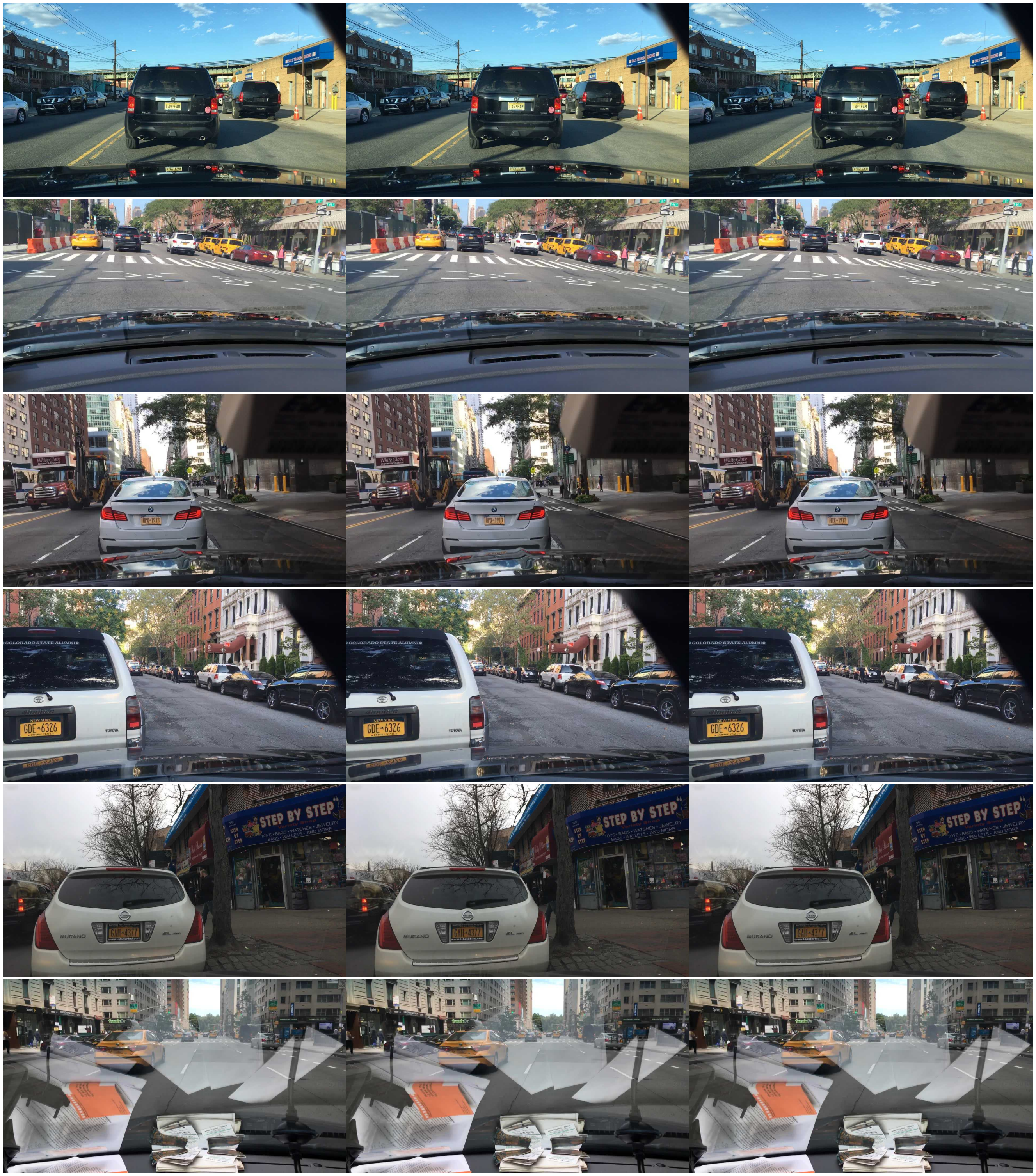
Figure 9. Random anonymization examples from Cityscapes [1]. **HM**=Histogram matching. **HM-LO**=Histogram matching via Latent Optimization. Note that the images are compressed.

(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 10. Random anonymization examples from BDD100K [4]. Note that the images are compressed.

(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 11. Random anonymization examples from BDD100K [4]. Note that the images are compressed.



(a) Face - Gaussian (b) Face - Mask Out (c) Face - Realistic

Figure 12. Random anonymization examples from BDD100K [4]. Note that the images are compressed.

(a) Face - Gaussian　　(b) Face - Mask Out　　(c) Face - Realistic

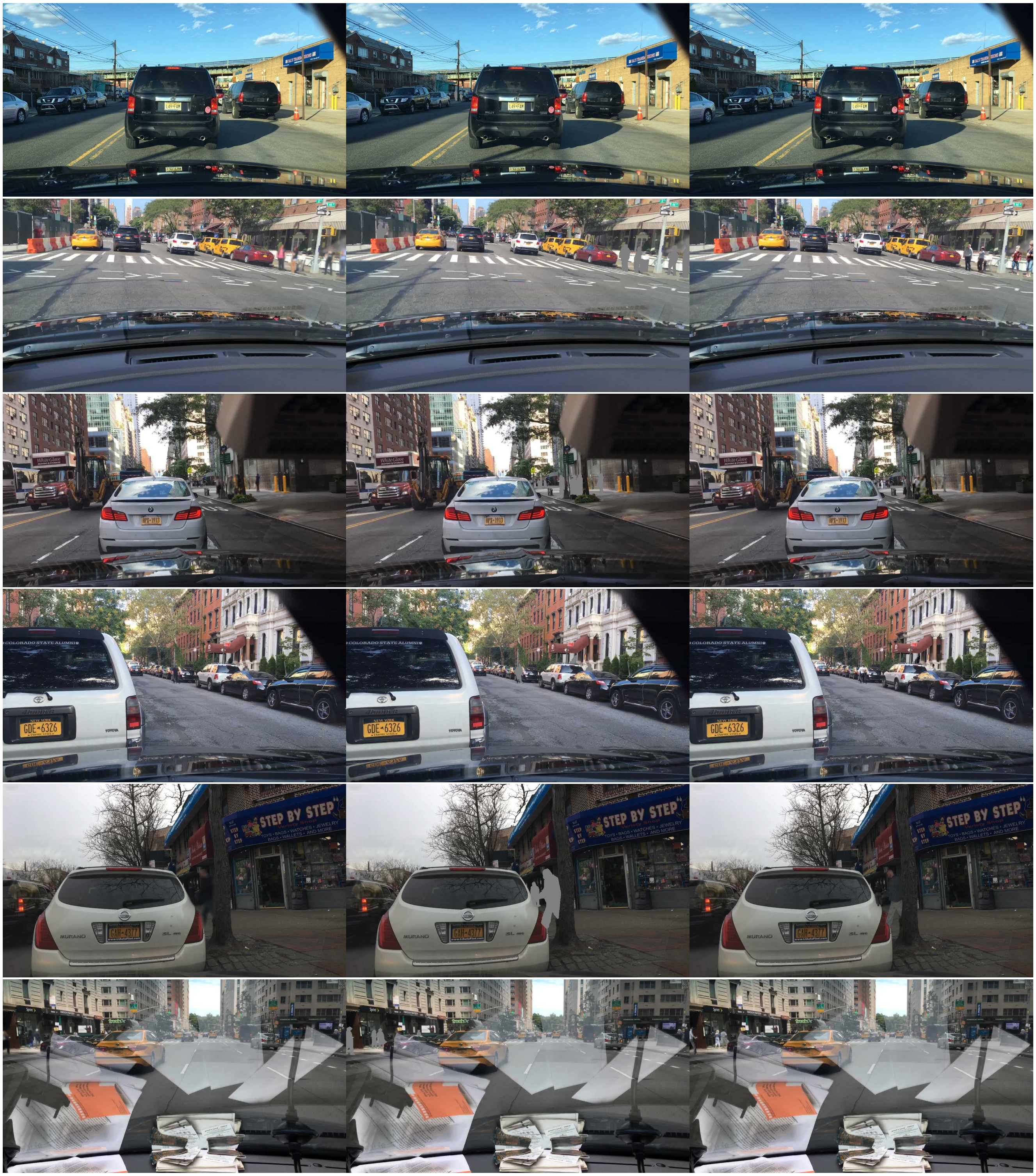
Figure 13. Random anonymization examples from BDD100K [4]. Note that the images are compressed.

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 14. Random anonymization examples from BDD100K [4]. Note that the images are compressed.

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 15. Random anonymization examples from BDD100K [4]. Note that the images are compressed.



(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 16. Random anonymization examples from BDD100K [4]. Note that the images are compressed.

(a) Body - Gaussian (b) Body - Mask Out (c) Body - Realistic

Figure 17. Random anonymization examples from BDD100K [4]. Note that the images are compressed.