Supplementary Material: Image Harmonization with Attention-based Deep Feature Modulation

Guoqing Hao, Satoshi Iizuka, Kazuhiro Fukui

University of Tsukuba

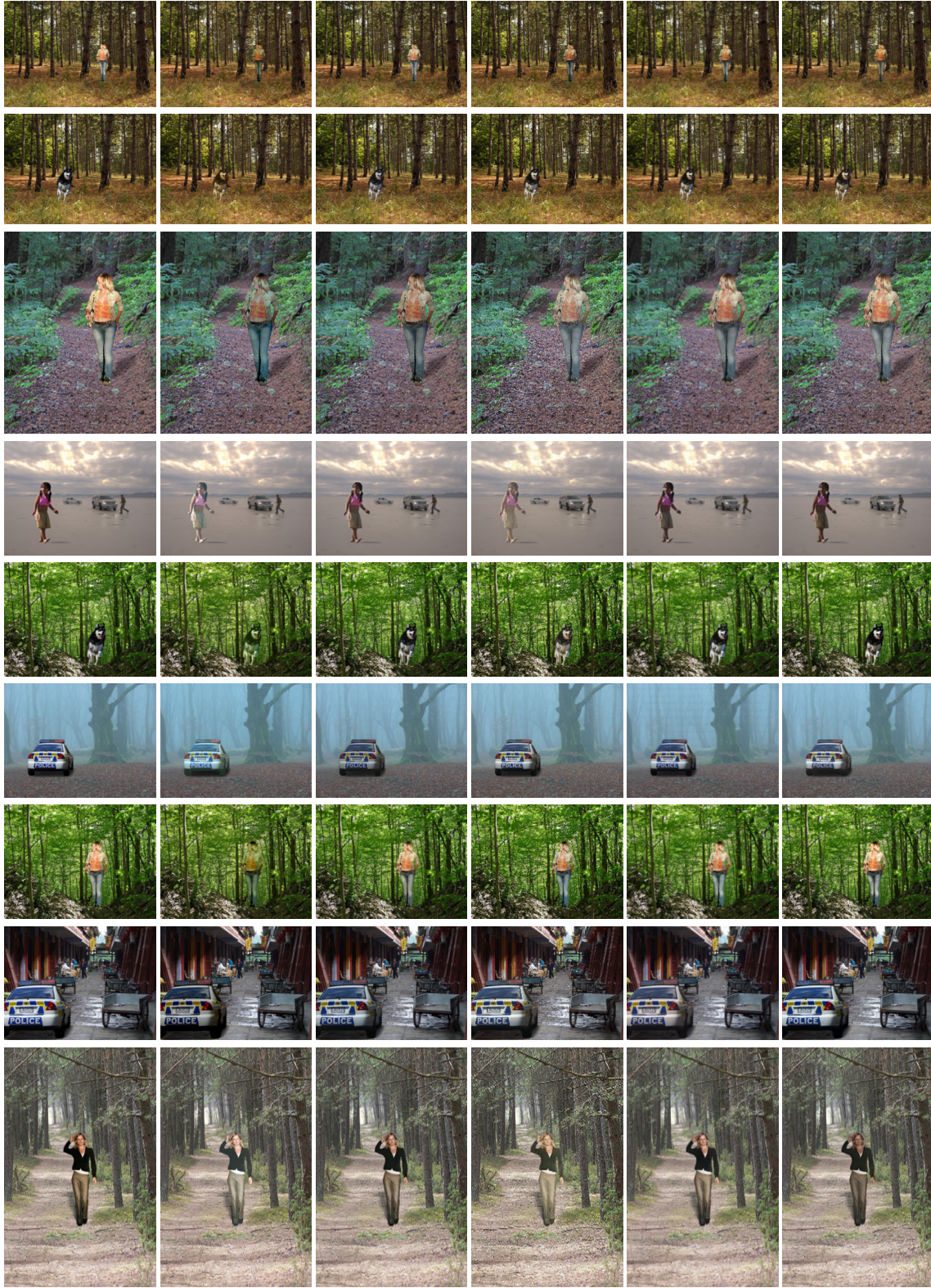## Comparisons of different methods using mean-squared errors (MSE)

Table 1. Comparison between our method against existing methods, on the iHarmony4 dataset[1], based on criteria of MSE.

Method	HCOCO	HAdobe5k	HFlickr	Hday2night	All
Lalonde and Efros [3]	110.10	158.90	329.87	199.93	150.53
Xue *et al*. [5]	77.04	274.15	249.54	190.51	155.87
Zhu *et al*. [6]	79.82	414.31	315.42	136.71	204.77
DIH [4]	51.85	92.65	163.38	82.34	76.77
S^2AM [2]	41.07	63.40	143.45	76.61	59.67
DoveNet [1]	36.72	52.32	133.14	**54.05**	52.36
Ours	**21.05**	**28.26**	**68.75**	85.14	**30.37**

References

[1] W. Cong, J. Zhang, L. Niu, L. Liu, Z. Ling, W. Li, and L. Zhang. Dovenet: Deep image harmonization via domain verification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23

[2] X. Cun and C. Pun. Improving the harmony of the composite image by spatial-separated attention module. *IEEE Transactions on Image Processing*, 29:4759–4771, 2020. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23

[3] J.-F. Lalonde and A. A. Efros. Using color compatibility for assessing image realism. In *IEEE International Conference on Computer Vision (ICCV)*, 2007. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[4] Y.-H. Tsai, X. Shen, Z. Lin, K. Sunkavalli, X. Lu, and M.-H. Yang. Deep image harmonization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23

[5] S. Xue, A. Agarwala, J. Dorsey, and H. Rushmeier. Understanding and improving the realism of image composites. *ACM Transactions on graphics (TOG)*, 31(4):1–10, 2012. 1

[6] J.-Y. Zhu, P. Krähenbühl, E. Shechtman, and A. A. Efros. Learning a discriminative model for the perception of realism in composite images. In *IEEE International Conference on Computer Vision (ICCV)*, 2015. 1

Results on Real Composite Images

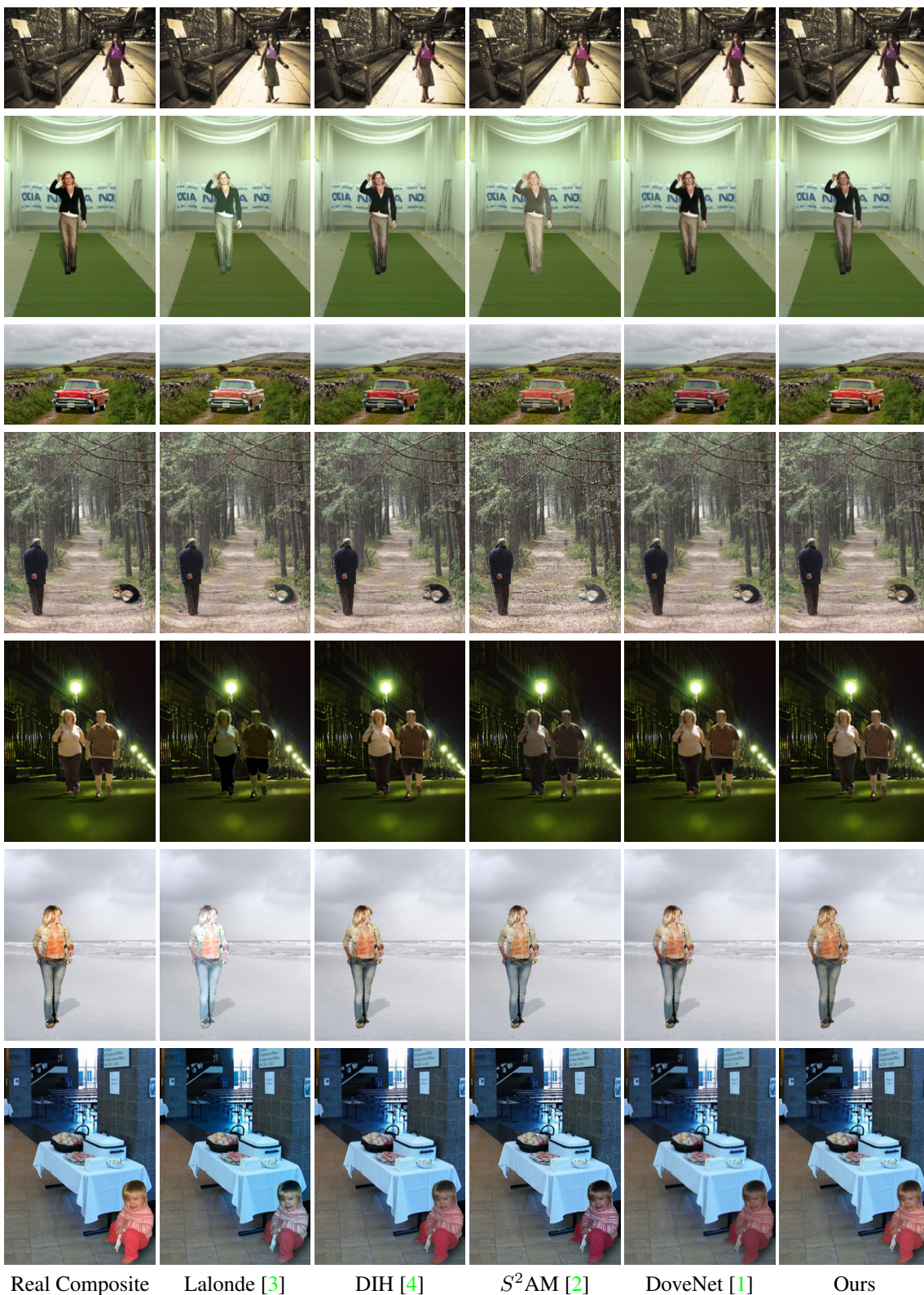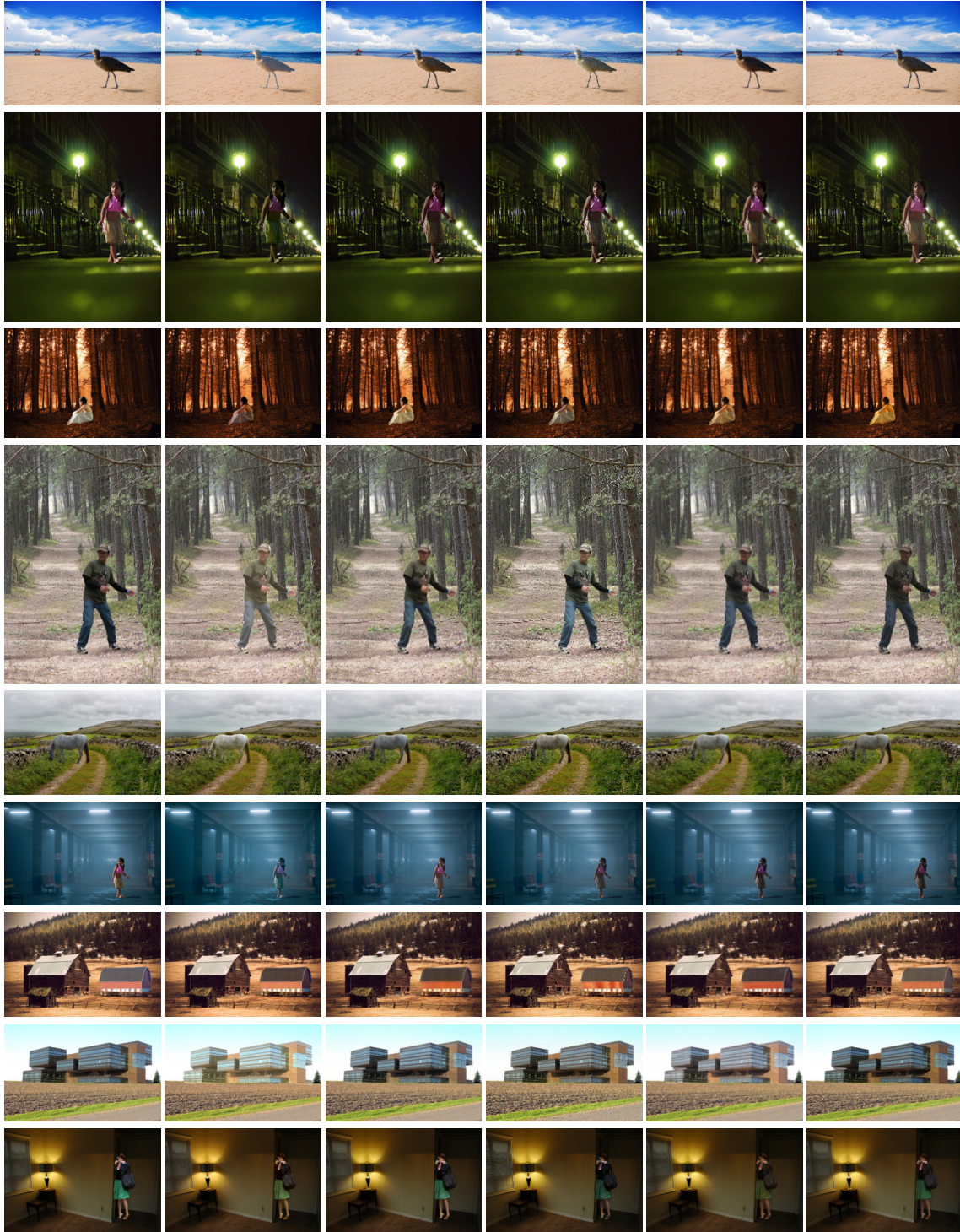| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 1. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 2. Results on real composite images, including the input composite, four existing methods, and our proposed method.

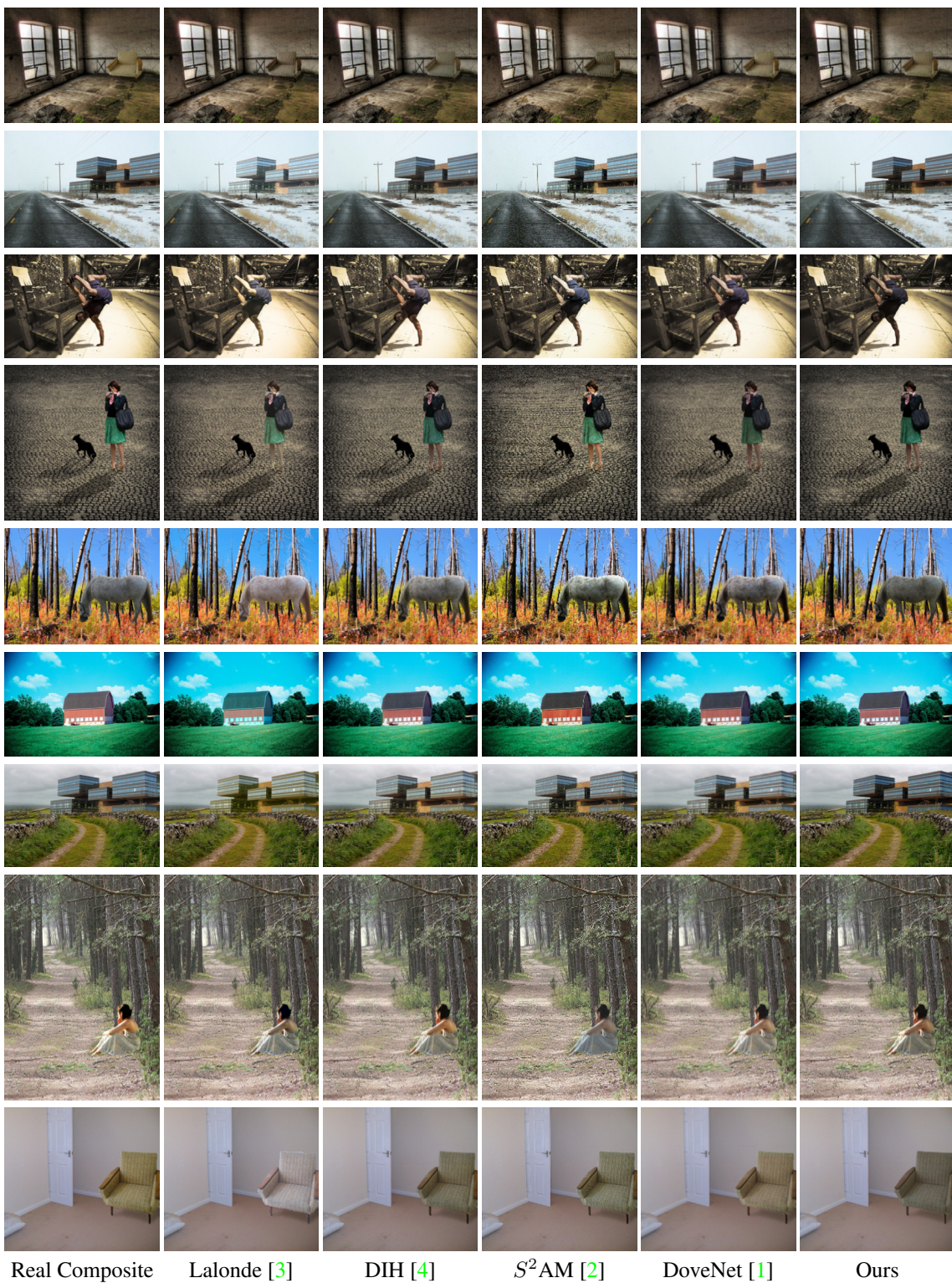| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 3. Results on real composite images, including the input composite, four existing methods, and our proposed method.

Figure 4. Results on real composite images, including the input composite, four existing methods, and our proposed method.
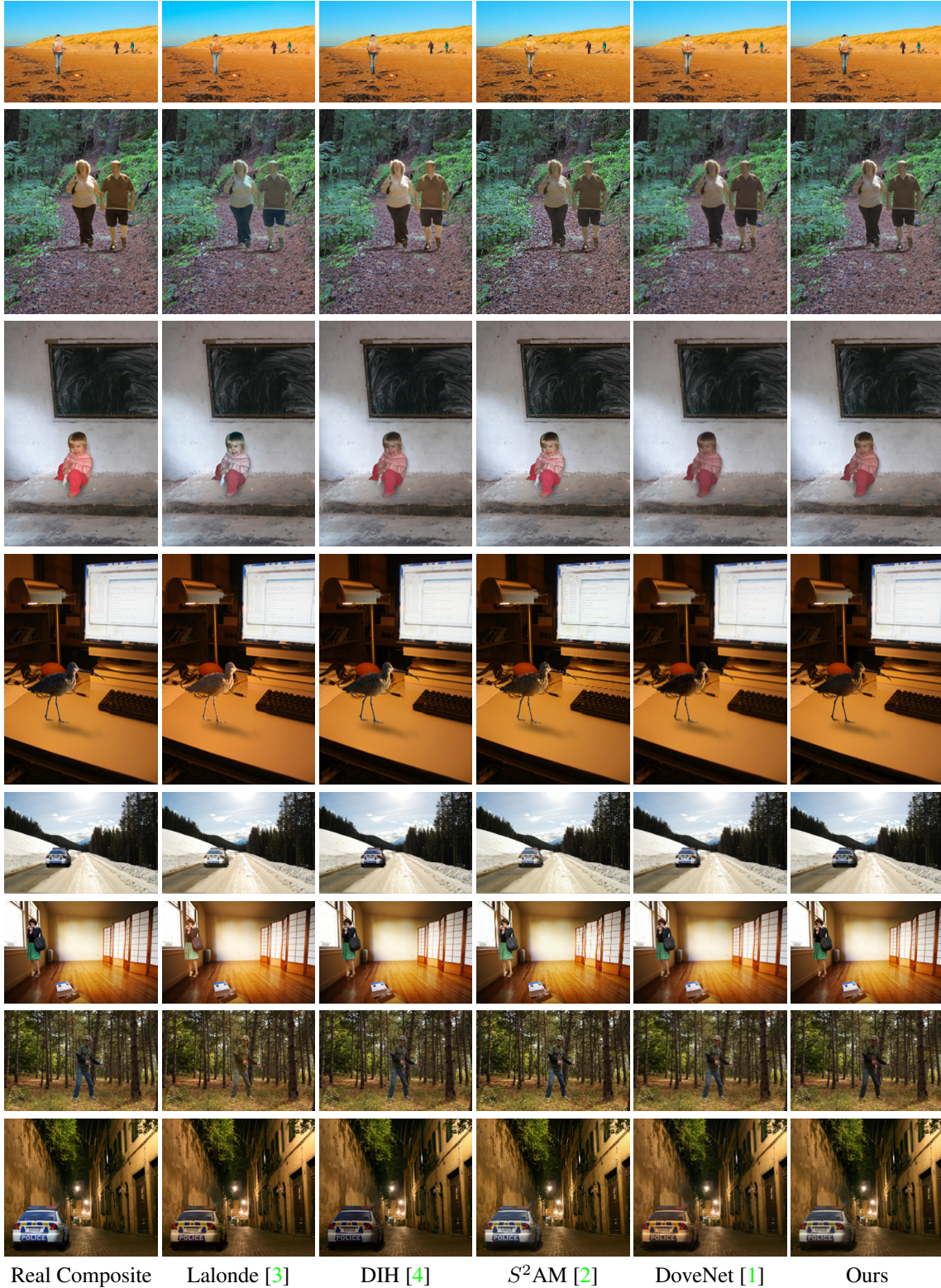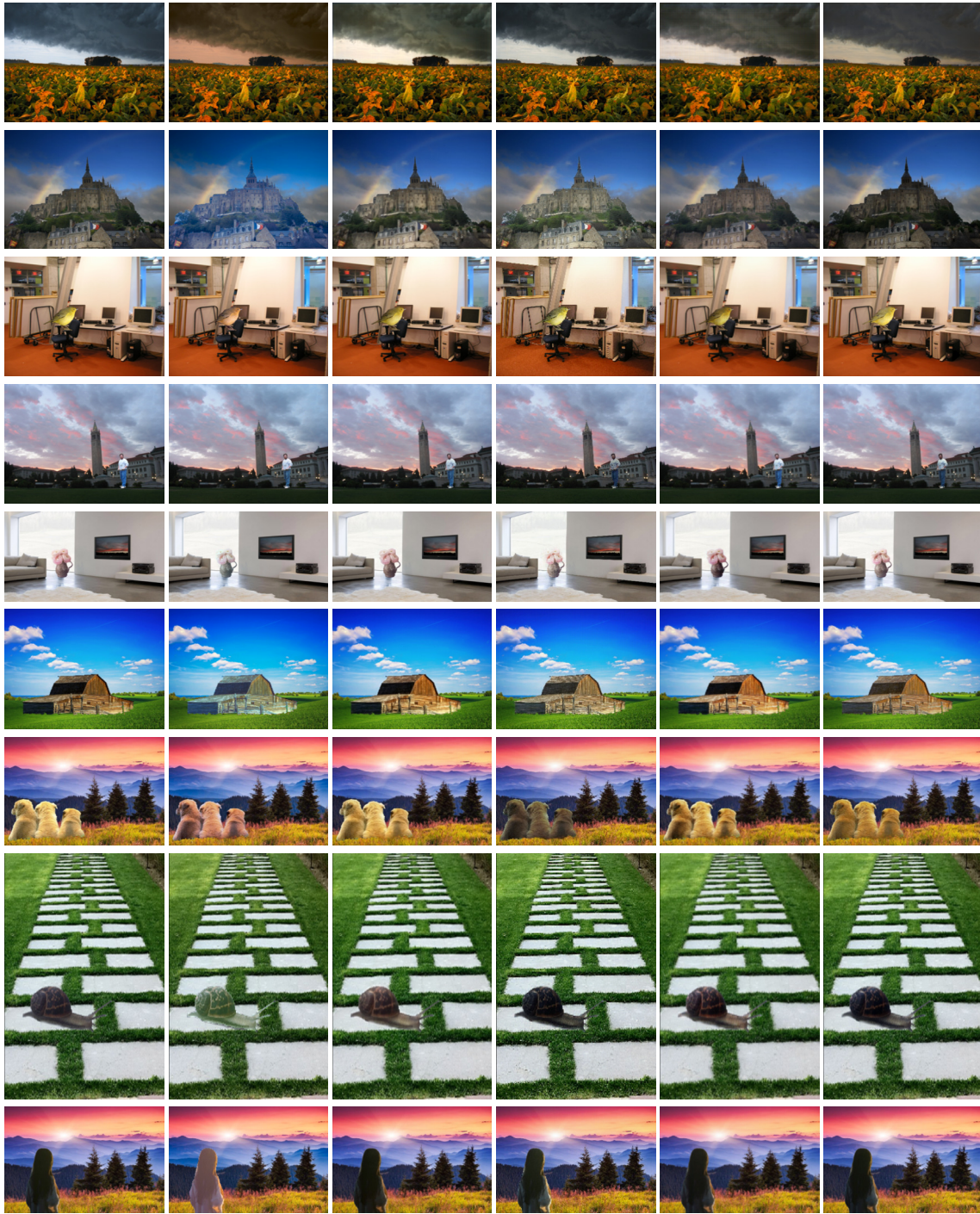
| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 5. Results on real composite images, including the input composite, four existing methods, and our proposed method.

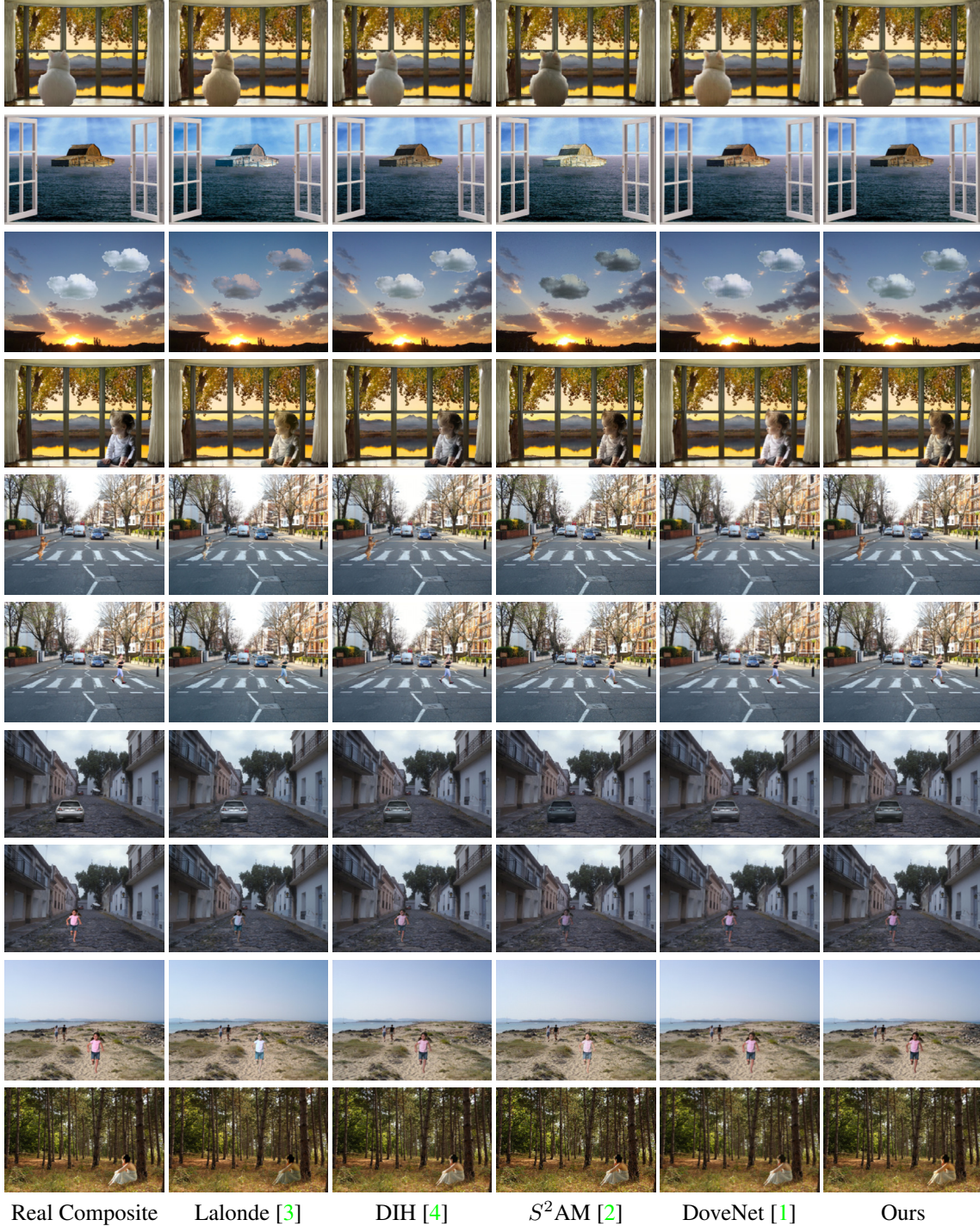| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 6. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 7. Results on real composite images, including the input composite, four existing methods, and our proposed method.

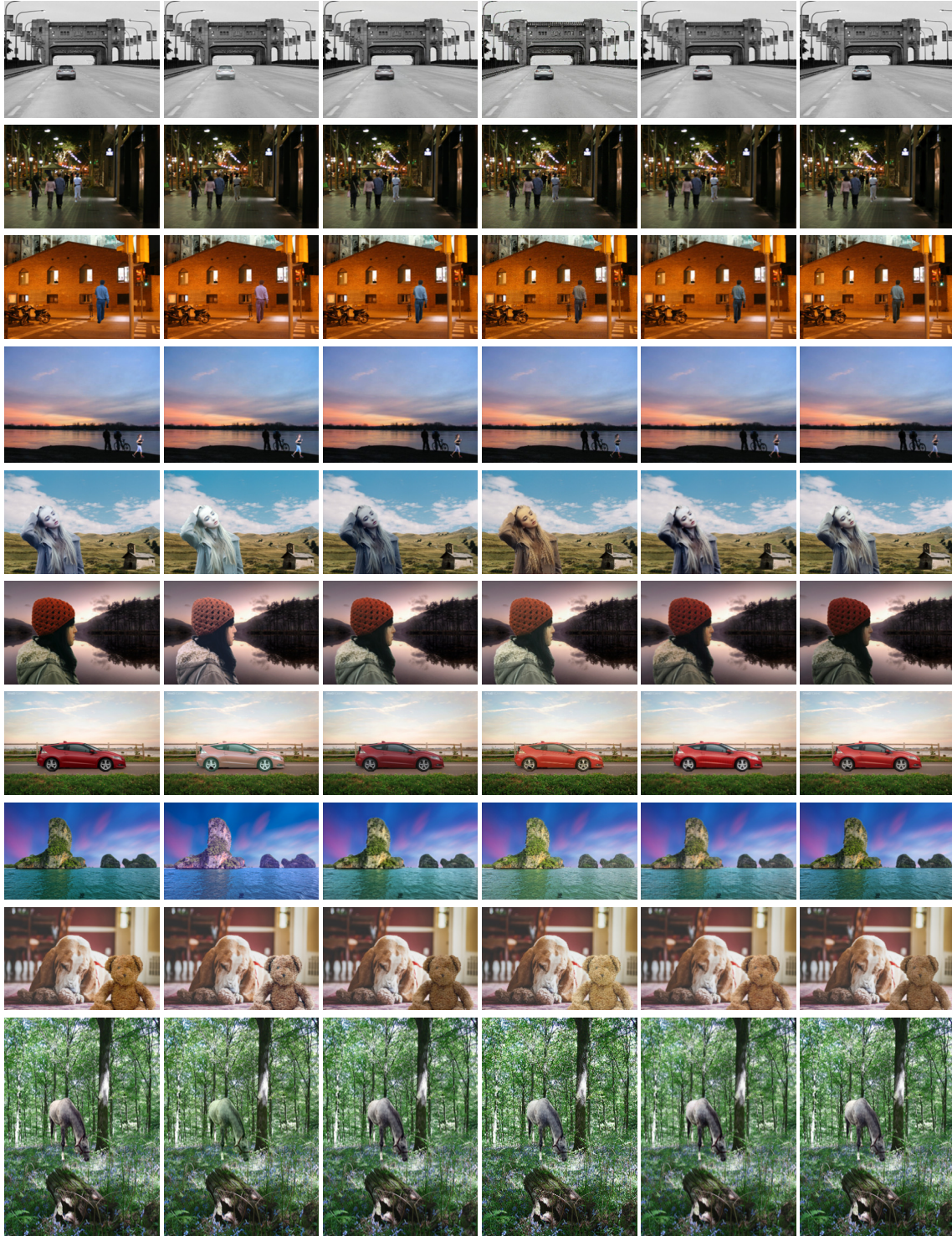| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 8. Results on real composite images, including the input composite, four existing methods, and our proposed method.
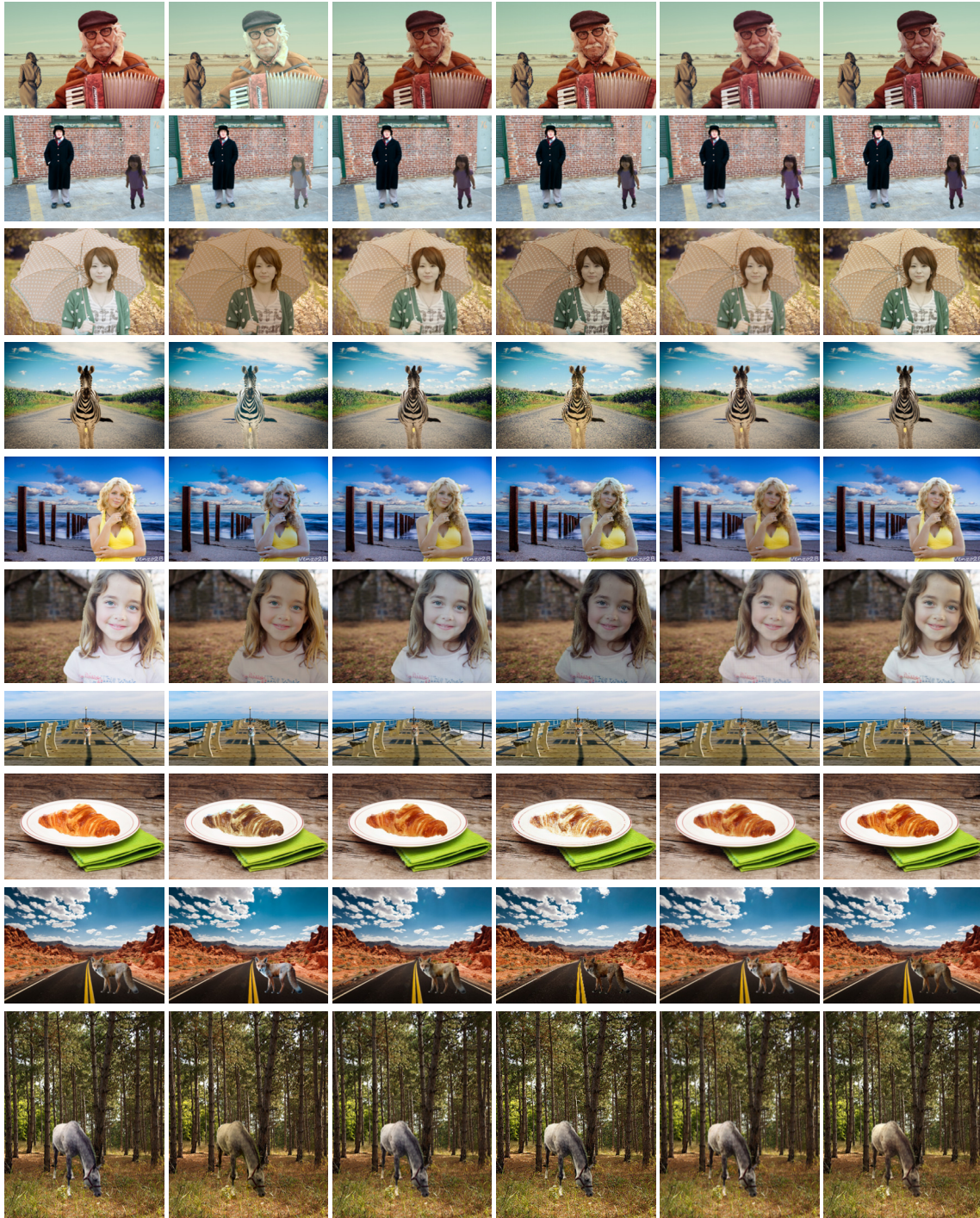
| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 9. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 10. Results on real composite images, including the input composite, four existing methods, and our proposed method.

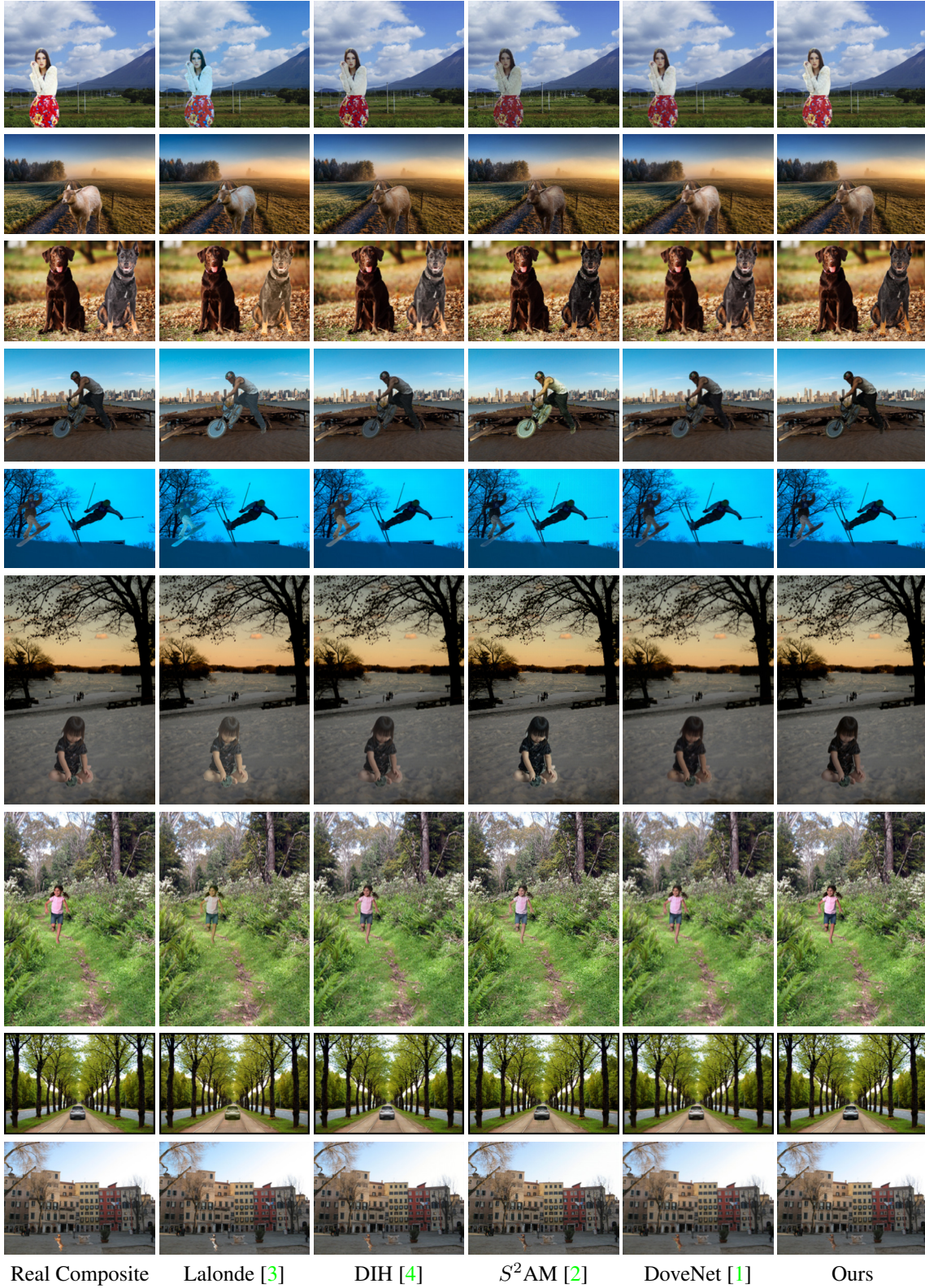| Real Composite | Lalonde [3] | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 11. Results on real composite images, including the input composite, four existing methods, and our proposed method.

Results on iHarmony4 [1] test set

| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 12. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 13. Results on real composite images, including the input composite, four existing methods, and our proposed method.

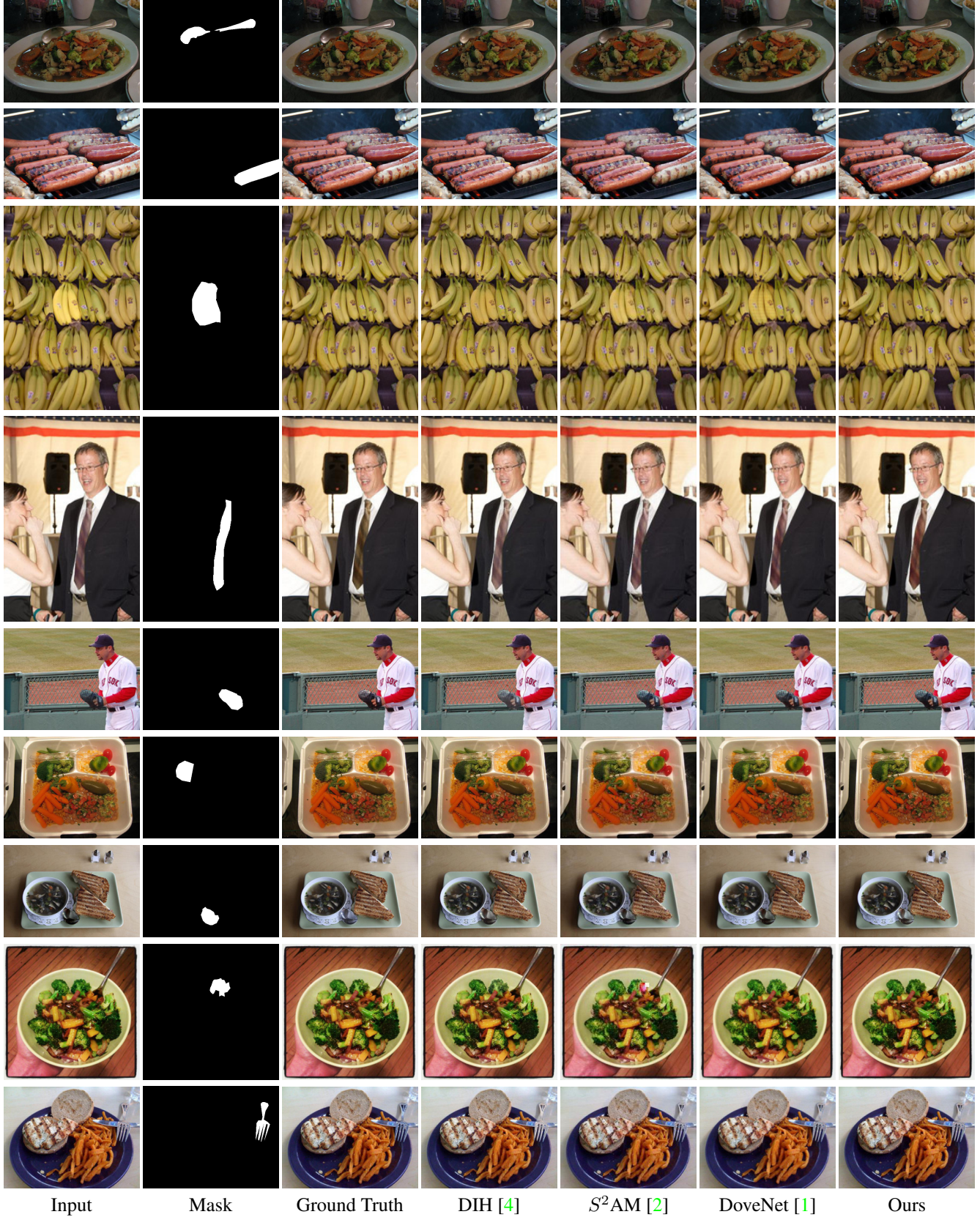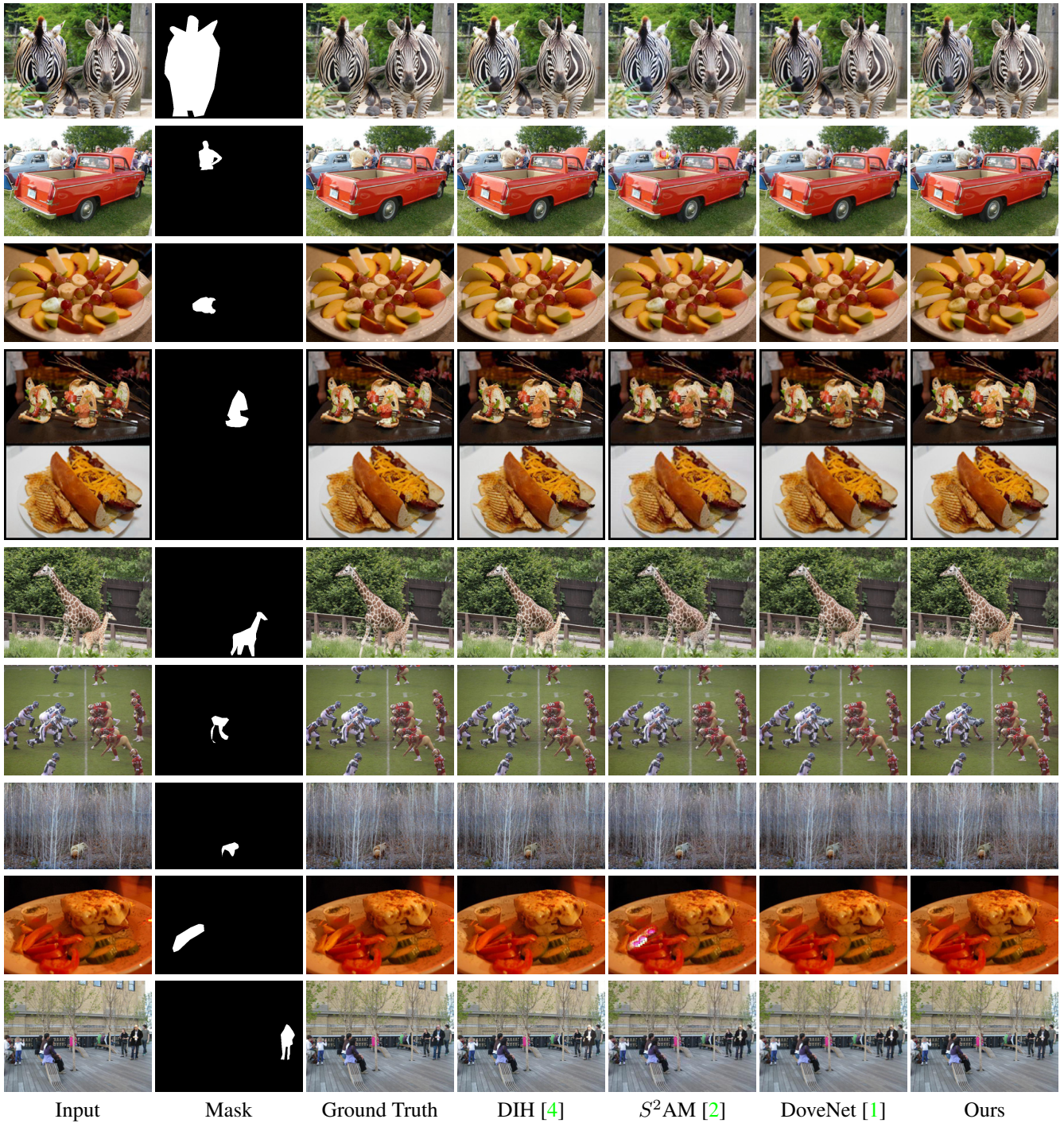| Input	Mask	Ground Truth	DIH [4]	S^2AM [2]	DoveNet [1]	Ours

Figure 14. Results on real composite images, including the input composite, four existing methods, and our proposed method.

Input	Mask	Ground Truth	DIH [4]	S^2AM [2]	DoveNet [1]	Ours

Figure 15. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 16. Results on real composite images, including the input composite, four existing methods, and our proposed method.
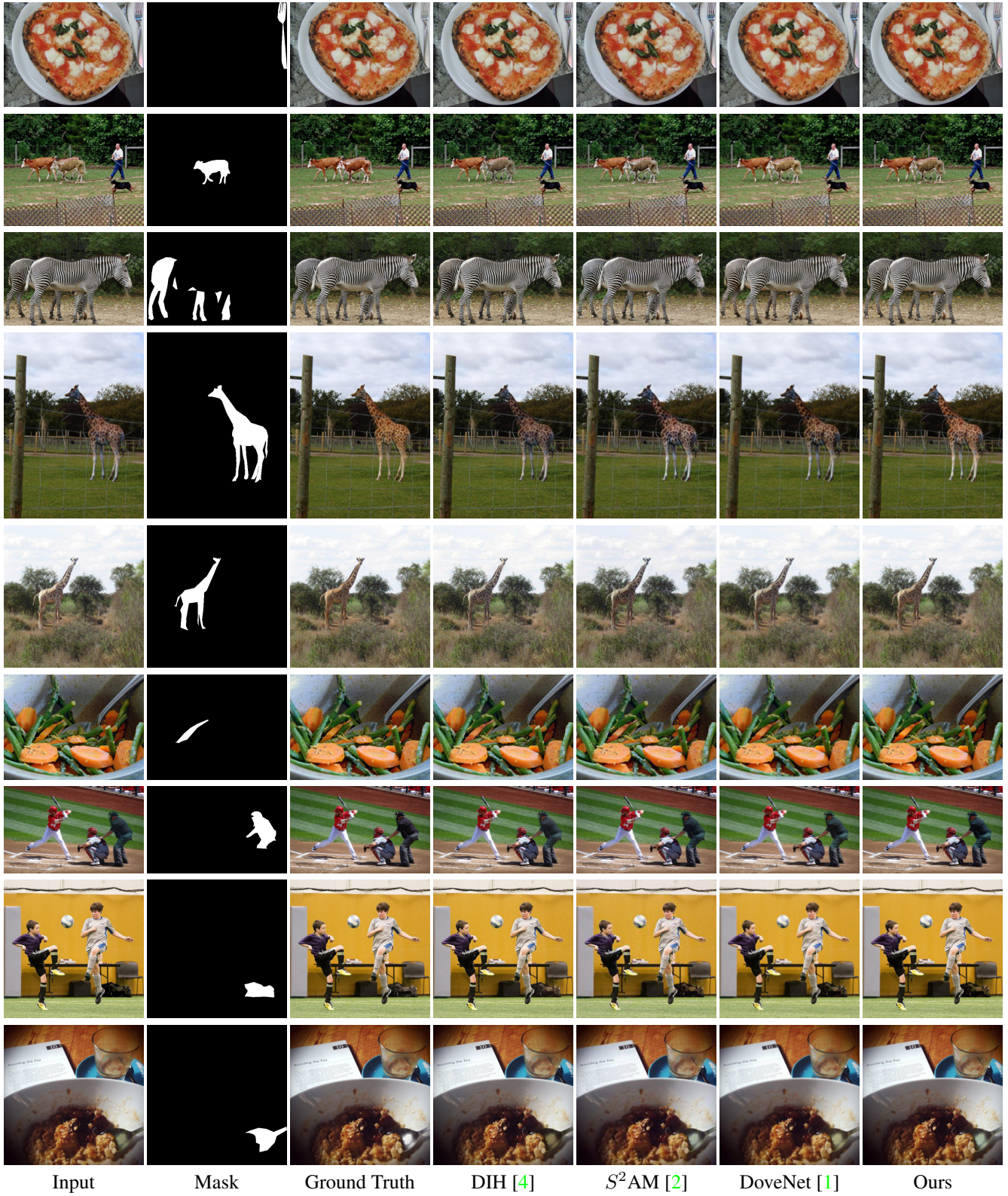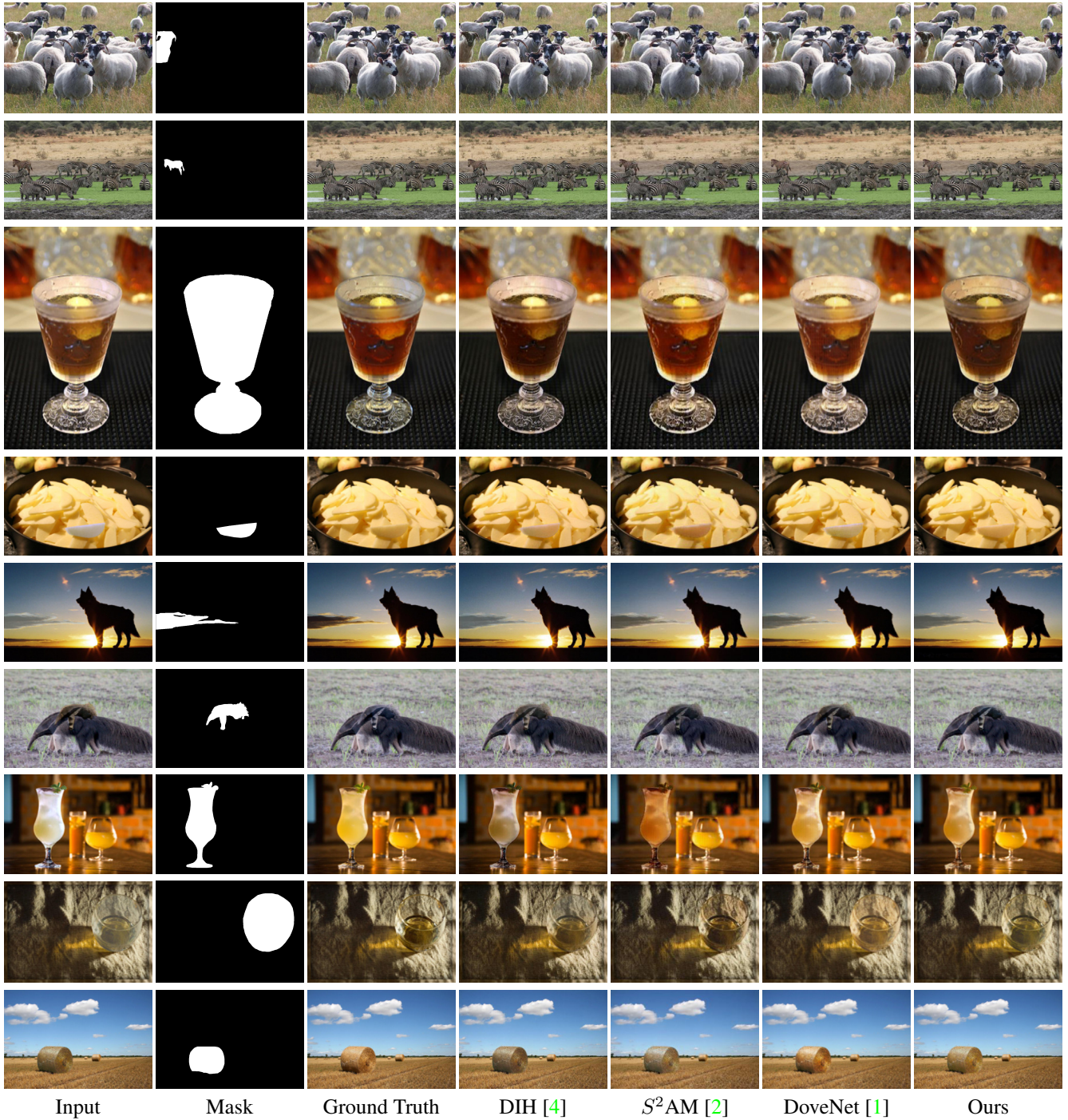
| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 17. Results on real composite images, including the input composite, four existing methods, and our proposed method.
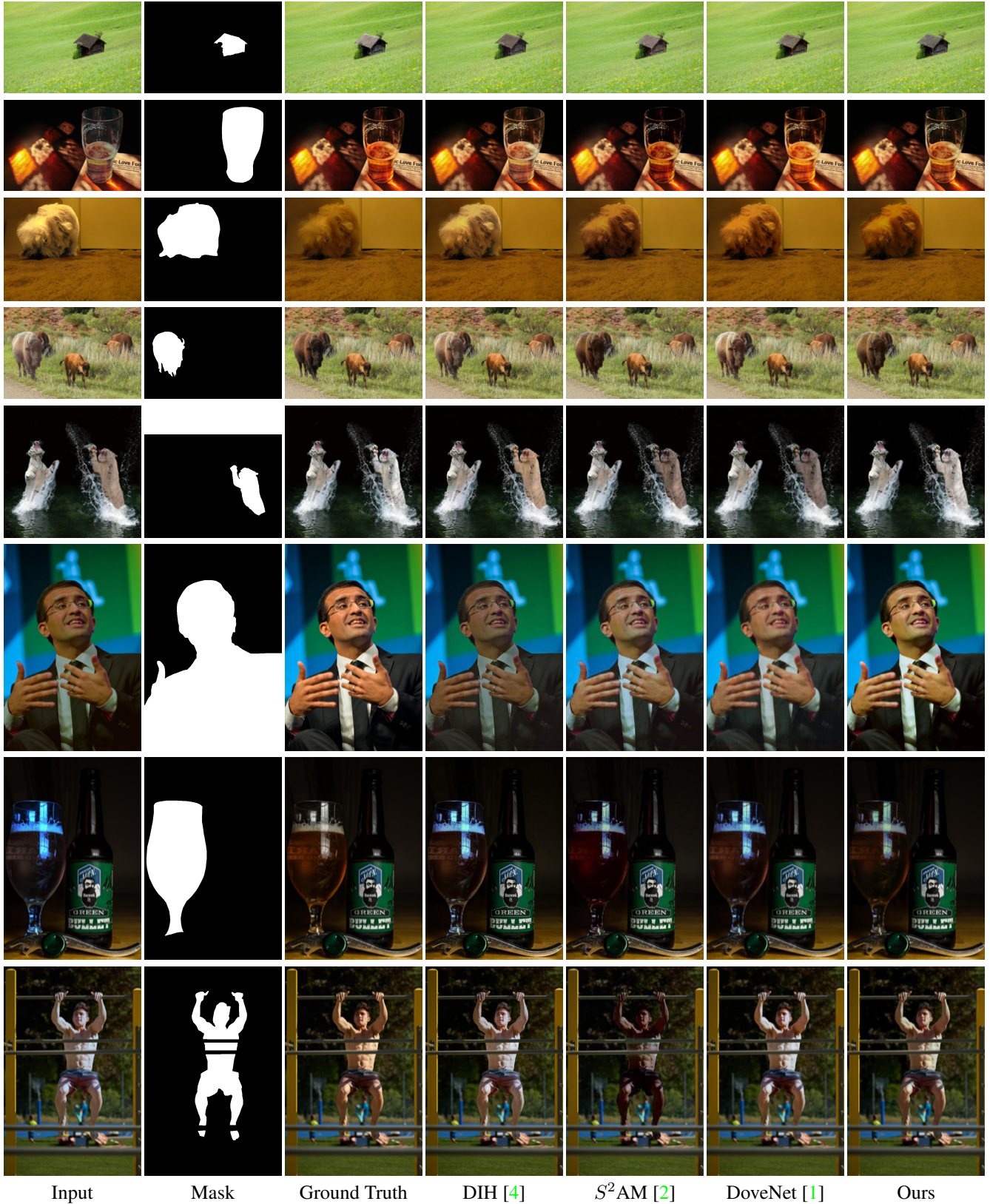
| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 18. Results on real composite images, including the input composite, four existing methods, and our proposed method.
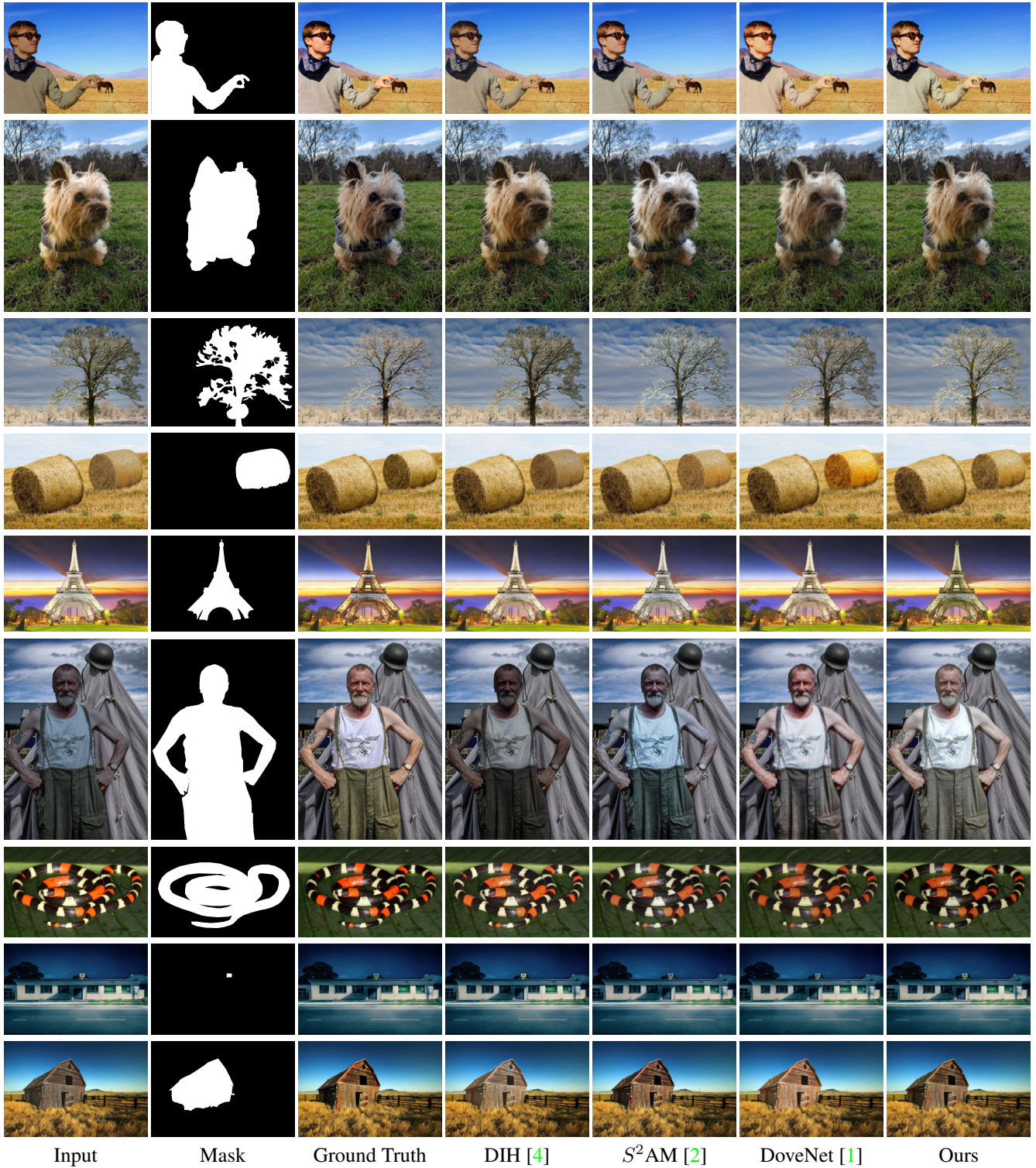
| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 19. Results on real composite images, including the input composite, four existing methods, and our proposed method.

| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 20. Results on real composite images, including the input composite, four existing methods, and our proposed method.
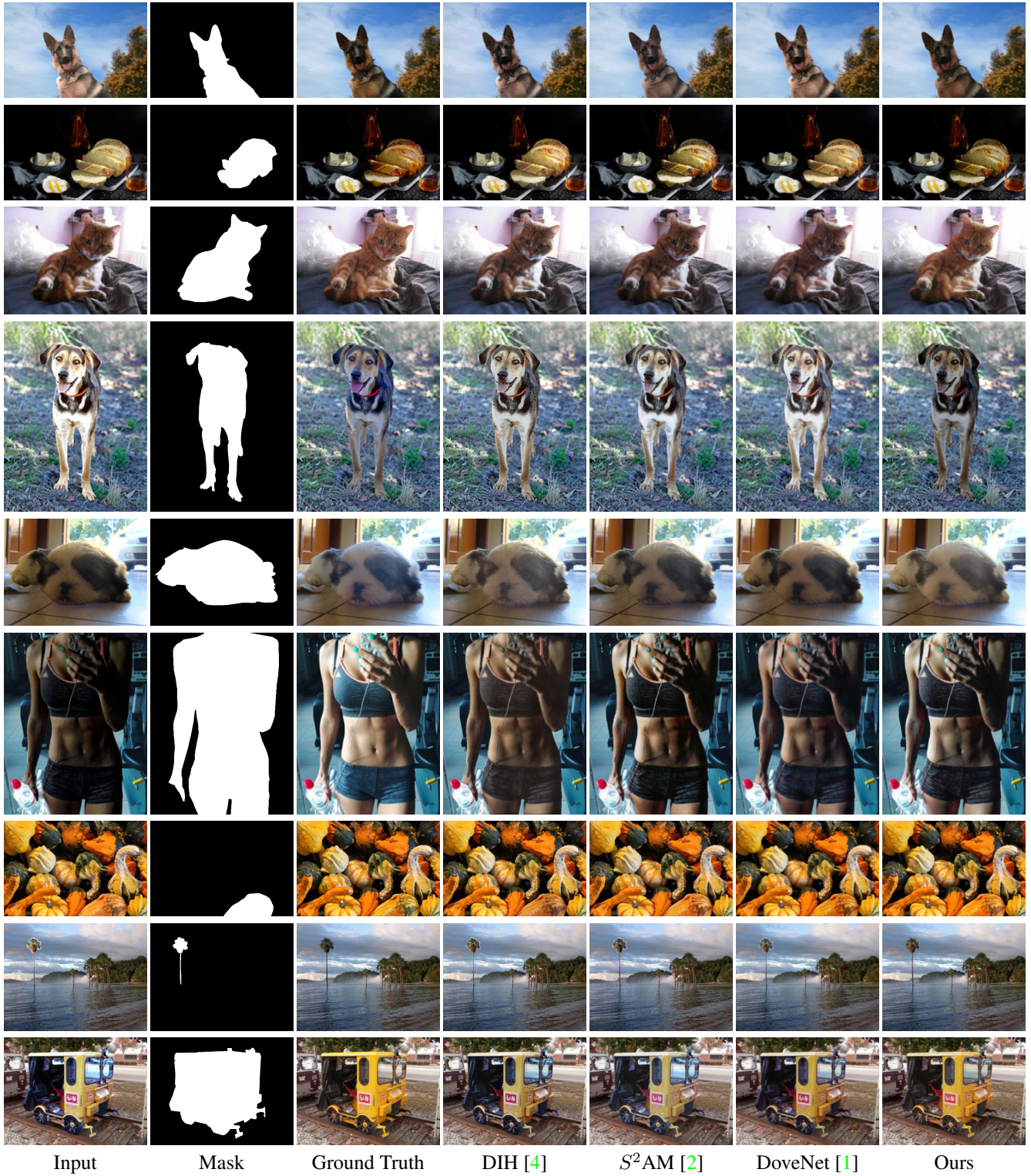
| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 21. Results on real composite images, including the input composite, four existing methods, and our proposed method.
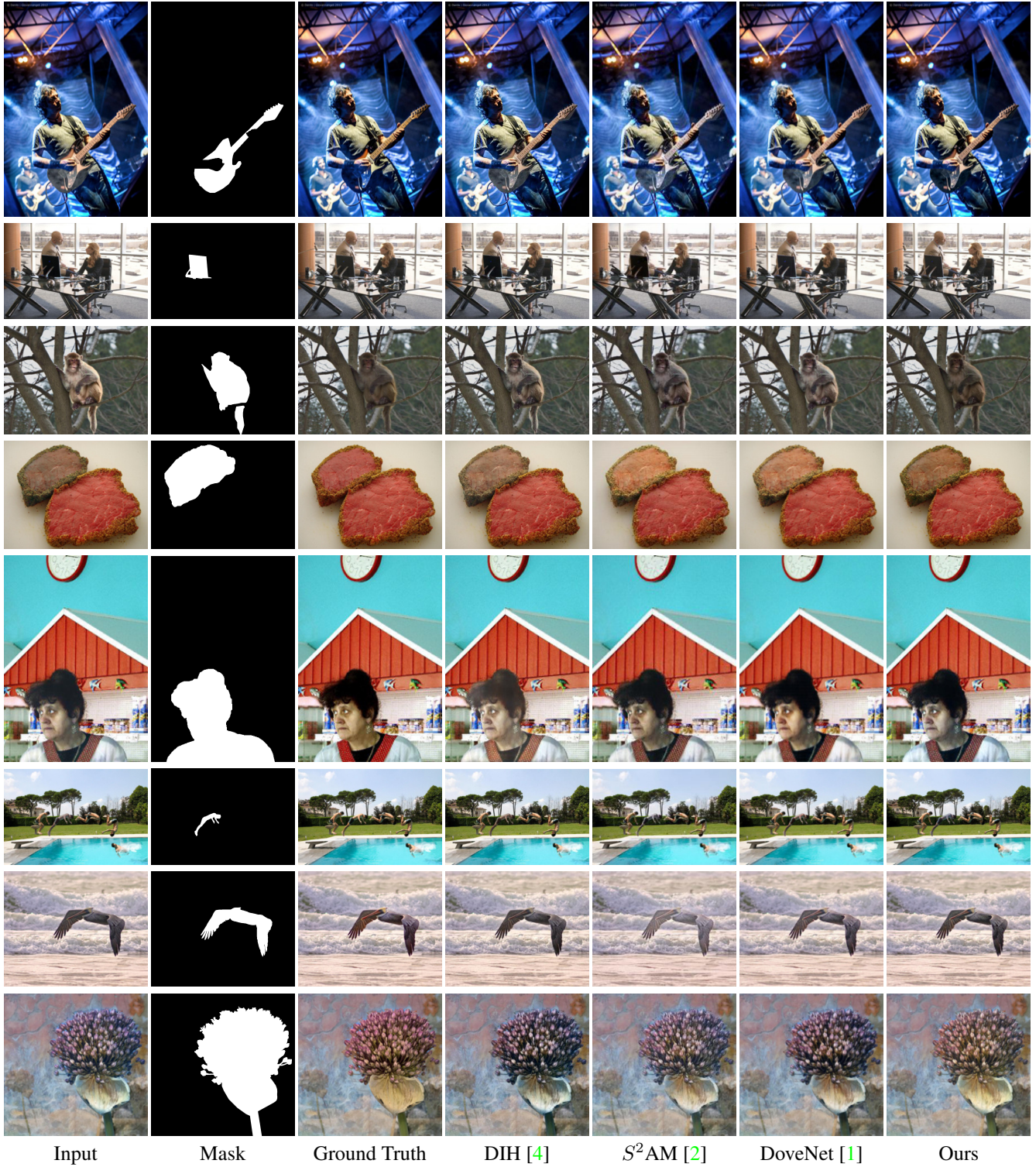
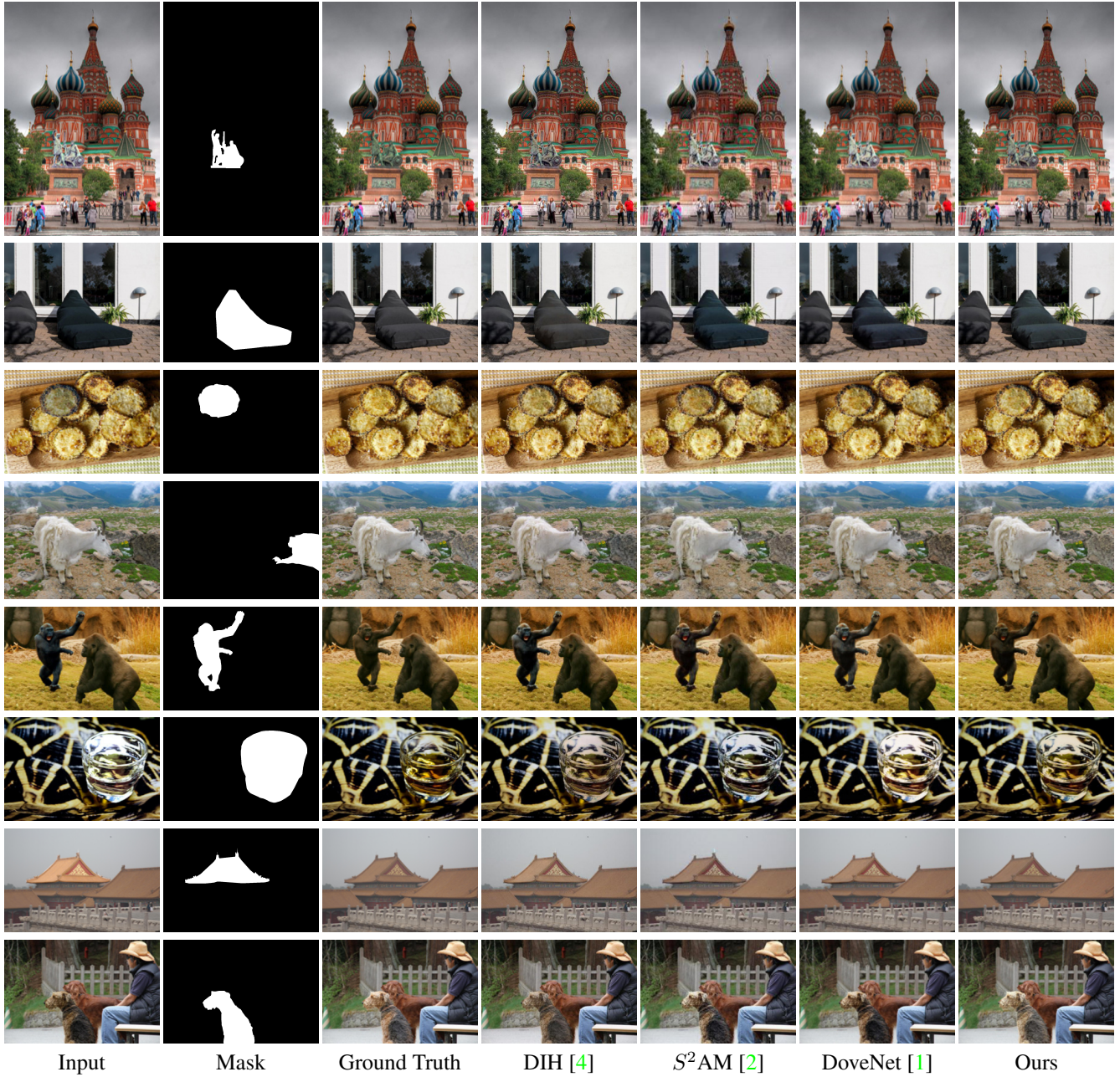| Input | Mask | Ground Truth | DIH [4] | S^2AM [2] | DoveNet [1] | Ours |

Figure 22. Results on real composite images, including the input composite, four existing methods, and our proposed method.