MLI-NeRF: Multi-Light Intrinsic-Aware Neural Radiance Fields

Supplementary Material

In this supplementary material, we present the following:
1. More results on the Synthetic Dataset.
2. More results on the Real Object Dataset.
3. More results on the ReNe Dataset.

We have also submitted a supplementary video to showcase the results of our method on all datasets.

1. More results on the Synthetic Dataset

From Fig. 1 to Fig. 4, we present additional qualitative results with all settings including Single Light, Multiple Lights, and Random Lights on the Synthetic Dataset.

In the Synthetic Dataset, we conduct experiments on four scenes: Hotdog, FurBall, Drums, and Lego. Across all scenes, we observe that under the original Random Lights setting, our method achieves best results, closely matching the GT. The predicted reflectance and shading are significantly better than those of PIE-Net [2] and Ordinal [1], and the relighting results are comparable to NRHints [5].

Under the Single Light and Multiple Light settings, we compare our method with additional approaches. In the Single Light setting, all methods struggle to consistently produce good intrinsic images, particularly in separating cast shadows. However, our method successfully removes most cast shadows in the predicted reflectance for scenes like FurBall, Drums, and Lego, outperforming other methods. This underscores the importance of multi-light information for intrinsic decomposition.

In the Multiple Lights setting (with four different light positions), our method achieves satisfactory results, very close to the GT and the Random Lights setting, demonstrating that four different light positions are sufficient for our method to perform well. Under the same setting, TensoIR [4] fails to produce satisfactory results, with residual cast shadows mixing into the reflectance.

Overall, since the camera view used in comparisons is the same, the reflectance should have the same GT across all settings. Our predicted reflectance consistently outperforms others in comparisons within the same rows, while also showing improvements as the number of light positions increases.

2. More results on the Real Object Dataset

We present different results for the 4 scenes of the Real Object Dataset, including Fish, Pikachu, Pixiu, and FurScene [3, 5]. In Fig. 5, we show additional results on further scenes. In Fig. 6, we present the results of reflectance and shading from multiple camera views and light positions to demonstrate the coherence of the approach.

In Fig. 7 and Fig. 8, we zoom in on certain areas to show more details of the results. For reflectance, we want to highlight rows (c) and (d), where our reflectance shows very promising results. In (c), our approach correctly estimates the reflectance, likely due to the robust 3D information considered in the pseudo-shading. In (d), the high quality of the result is probably due to the pseudo-reflectance labels that effectively represent the reflectance in low-light areas. However, a limitation of our approach is evident in the shading of row (a). Although the global surface shading is satisfactory, the edges are not sharp, and there is some confusion in the areas below the object, likely due to none of the viewpoints or light sources providing adequate information for proper reconstruction.

For completeness of the comparison and potential needs, we also compared the relighting results with NRHints [5]. A quantitative evaluation is provided in Tab. 1. As shown, metrics show our method presents slightly lower PSNR and SSIM but better LPIPS.

	PSNR↑	SSIM↑	LPIPS↓	MSE ↓
NRHints	**31.62**	**0.9623**	0.0997	—
Ours	30.55	0.9341	**0.0797**	0.0012

Table 1. Quantitative results of the novel view synthesis and relighting on the **Real Object Dataset**.

3. More results on the ReNe Dataset

Fig. 9 to Fig. 12 present more detailed results on the ReNe dataset, where the four scenes are labeled as apple, cube, garden, and savannah in the original dataset. Similar to our observations on the Synthetic Dataset, our method outperforms previous methods across all settings, including the original ALL Lights, as well as the additional Single Light and Multiple Lights settings. Furthermore, the performance of our method improves as the number of light sources increases. In all scenes under the ALL Lights and Multiple Lights settings, our method consistently produces satisfactory intrinsic images.

The ReNe dataset poses significant challenges for 3D reconstruction and intrinsic decomposition due to the camera views and light views being concentrated within a limited area. As shown in Fig. 9, TensoIR fails to produce valid outputs in the Apple scene under both the Single Light and Multiple Lights settings. Despite these challenges, our method consistently delivers satisfactory results.

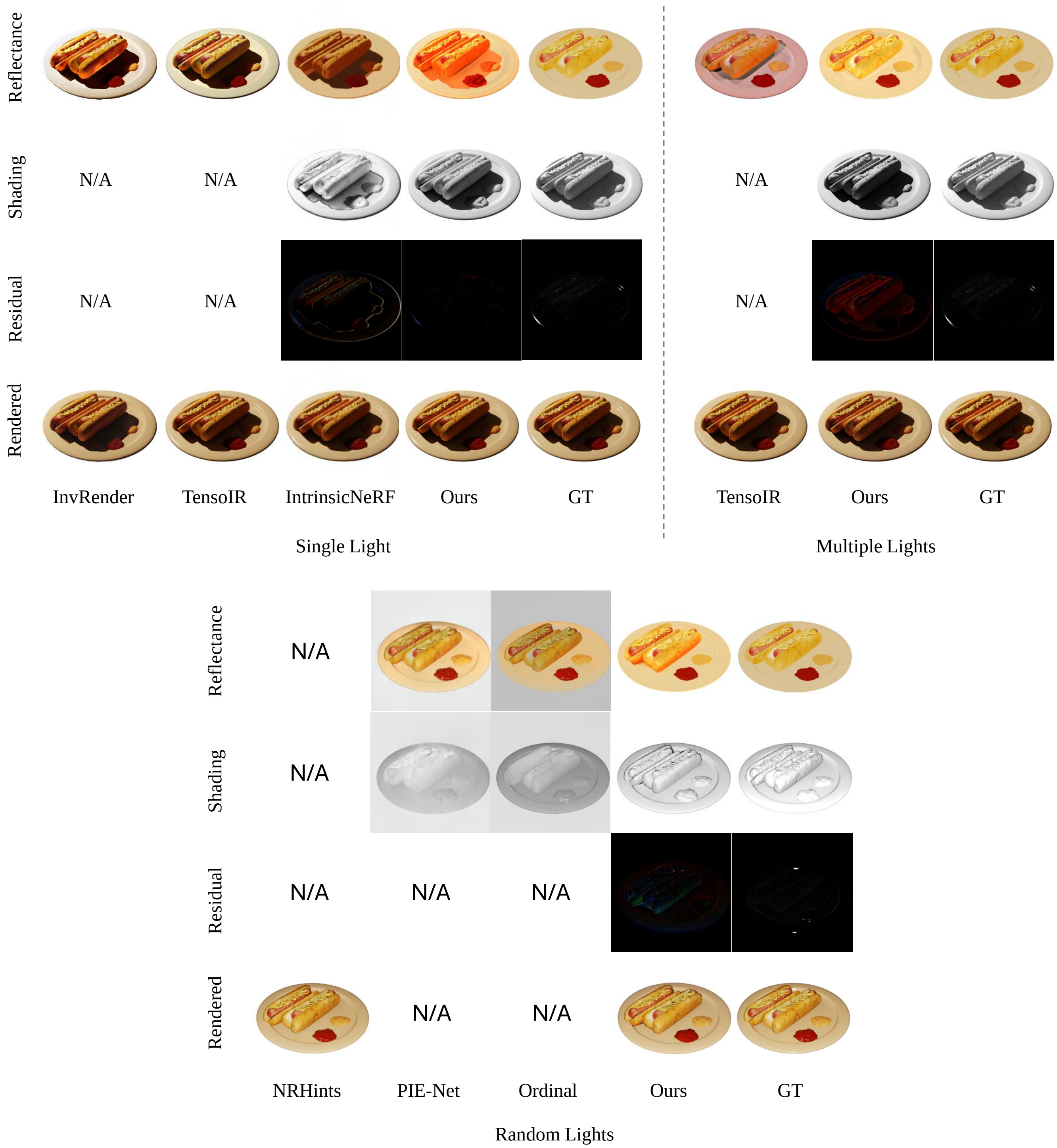


Figure 1. Additional Qualitative Results on the Synthetic Dataset. (Hotdog)

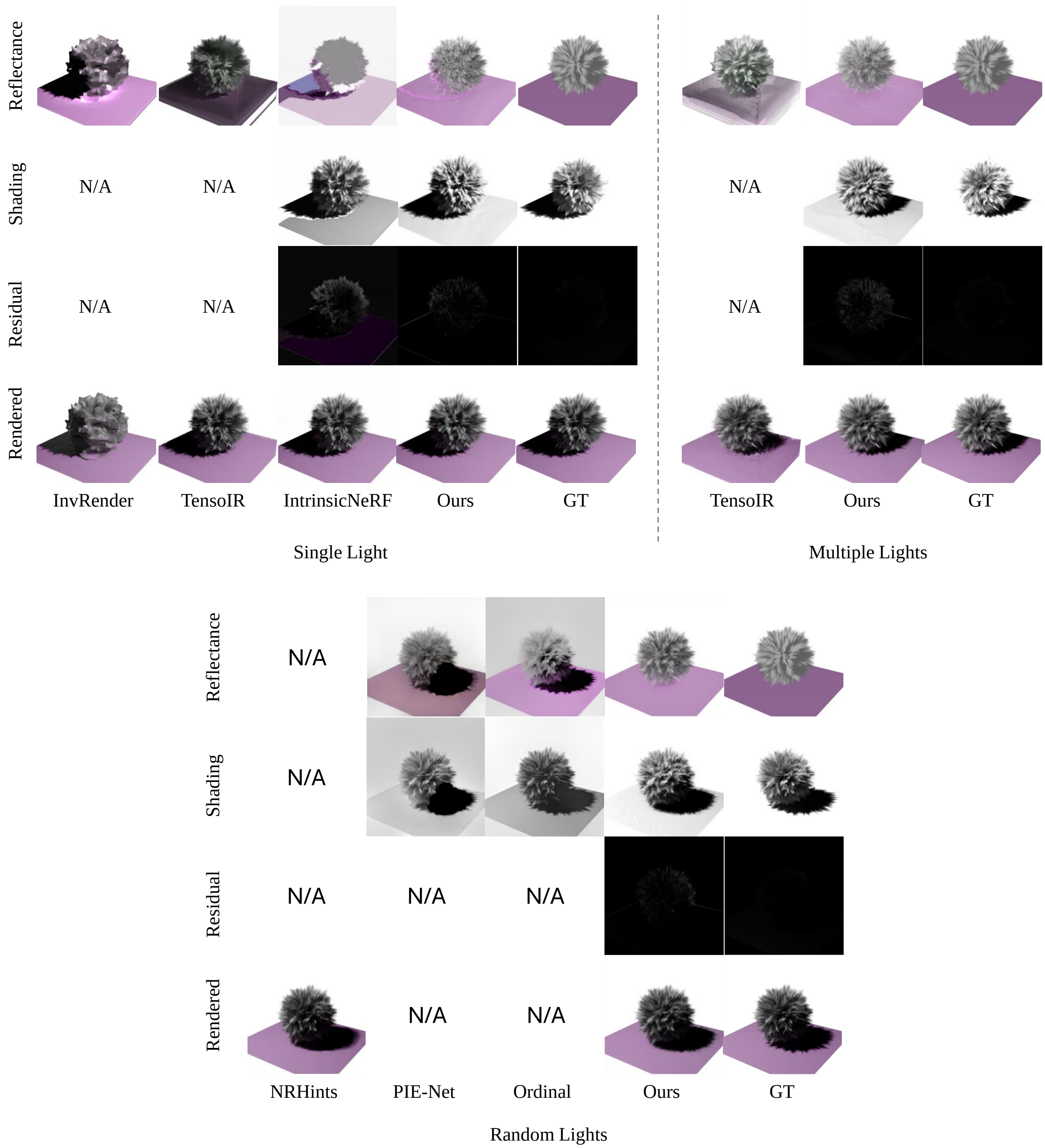


Figure 2. Additional Qualitative Results on the Synthetic Dataset (FurBall). However, it is worth noting that in the GT, the shading of the ground in this model was not correctly rendered in Blender. Our method under the Random Lights setting achieved the most reasonable results.

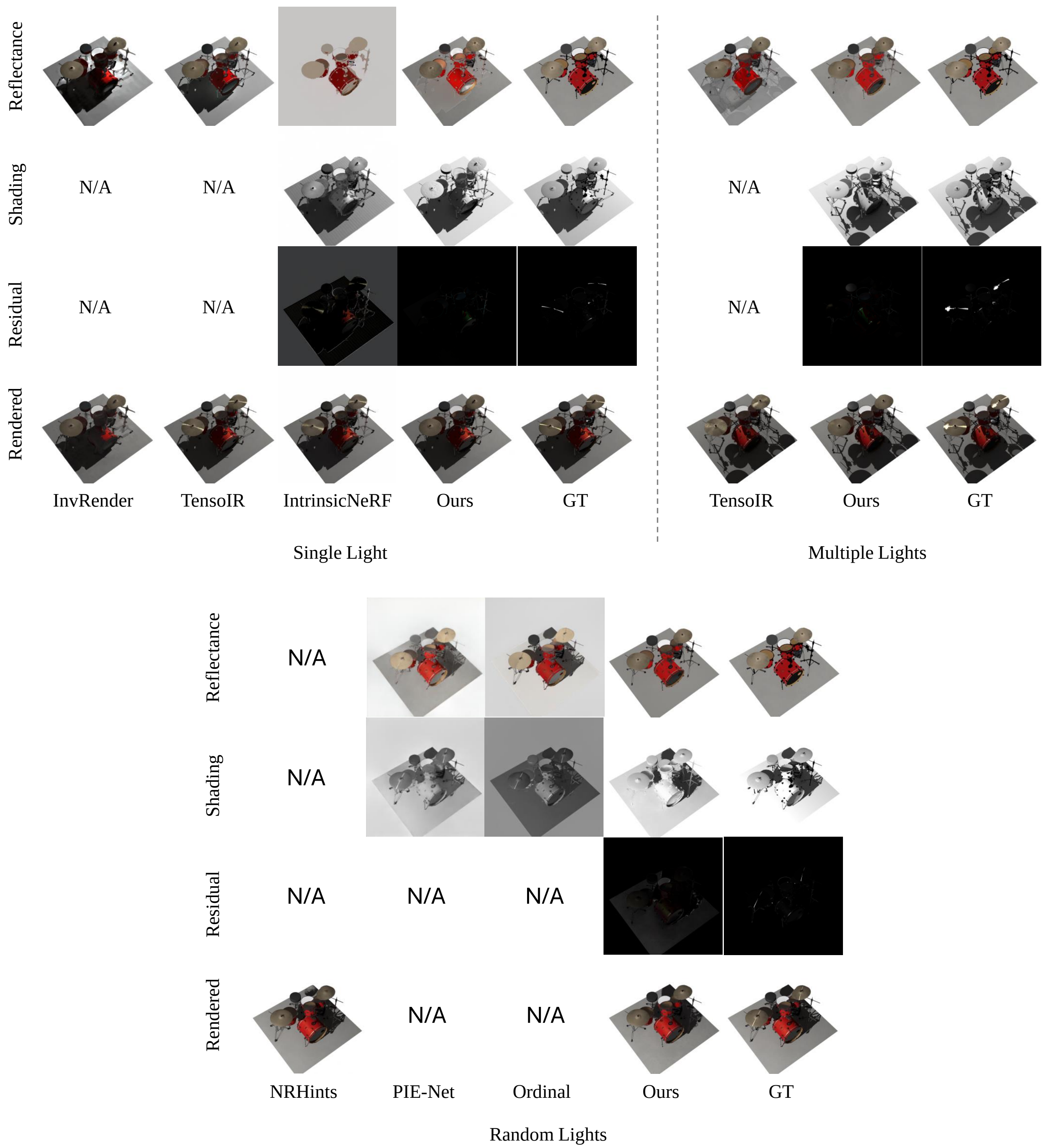


Figure 3. Additional Qualitative Results on the Synthetic Dataset. (Drums)

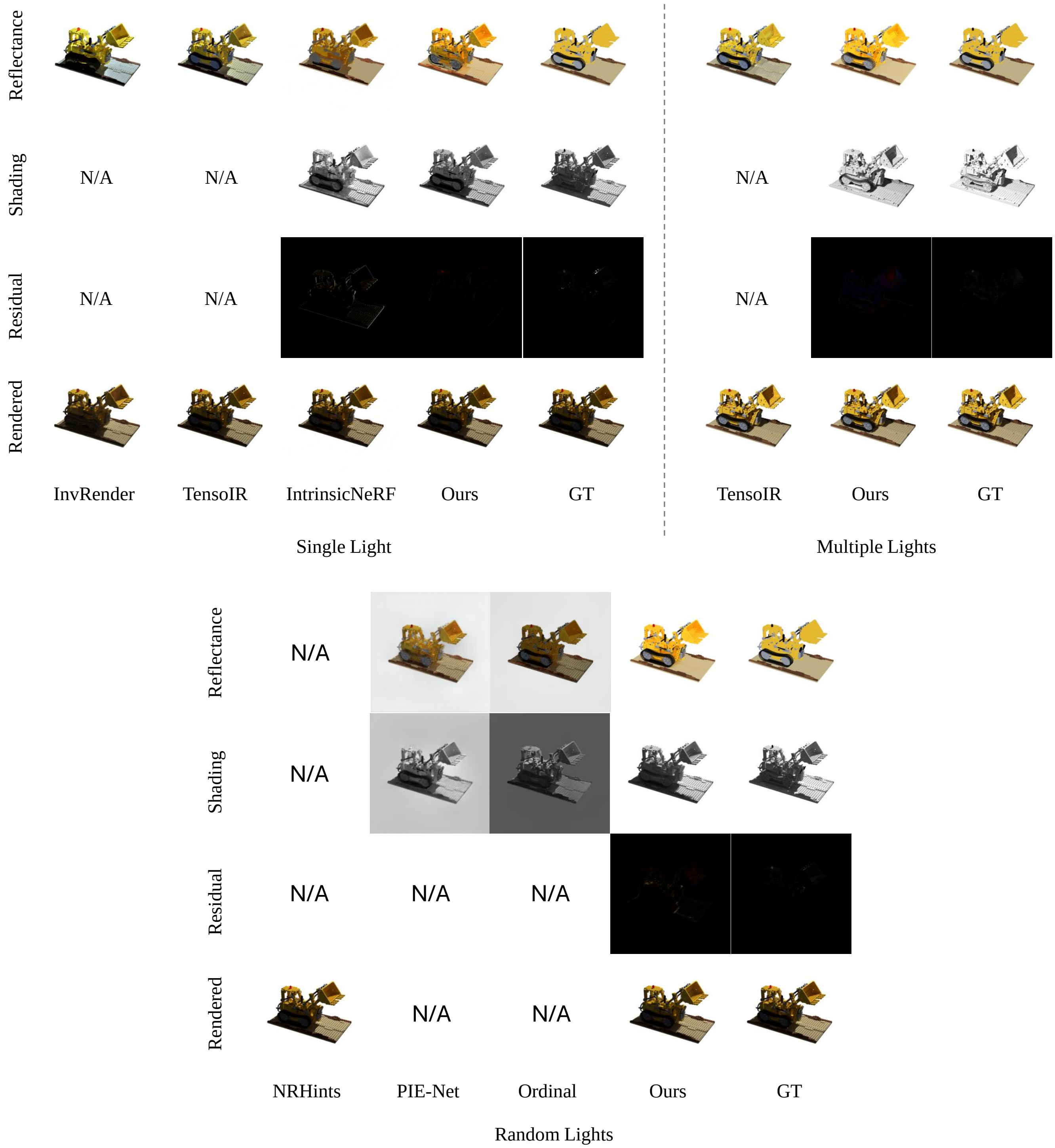


Figure 4. Additional Qualitative Results on the Synthetic Dataset. (Lego)

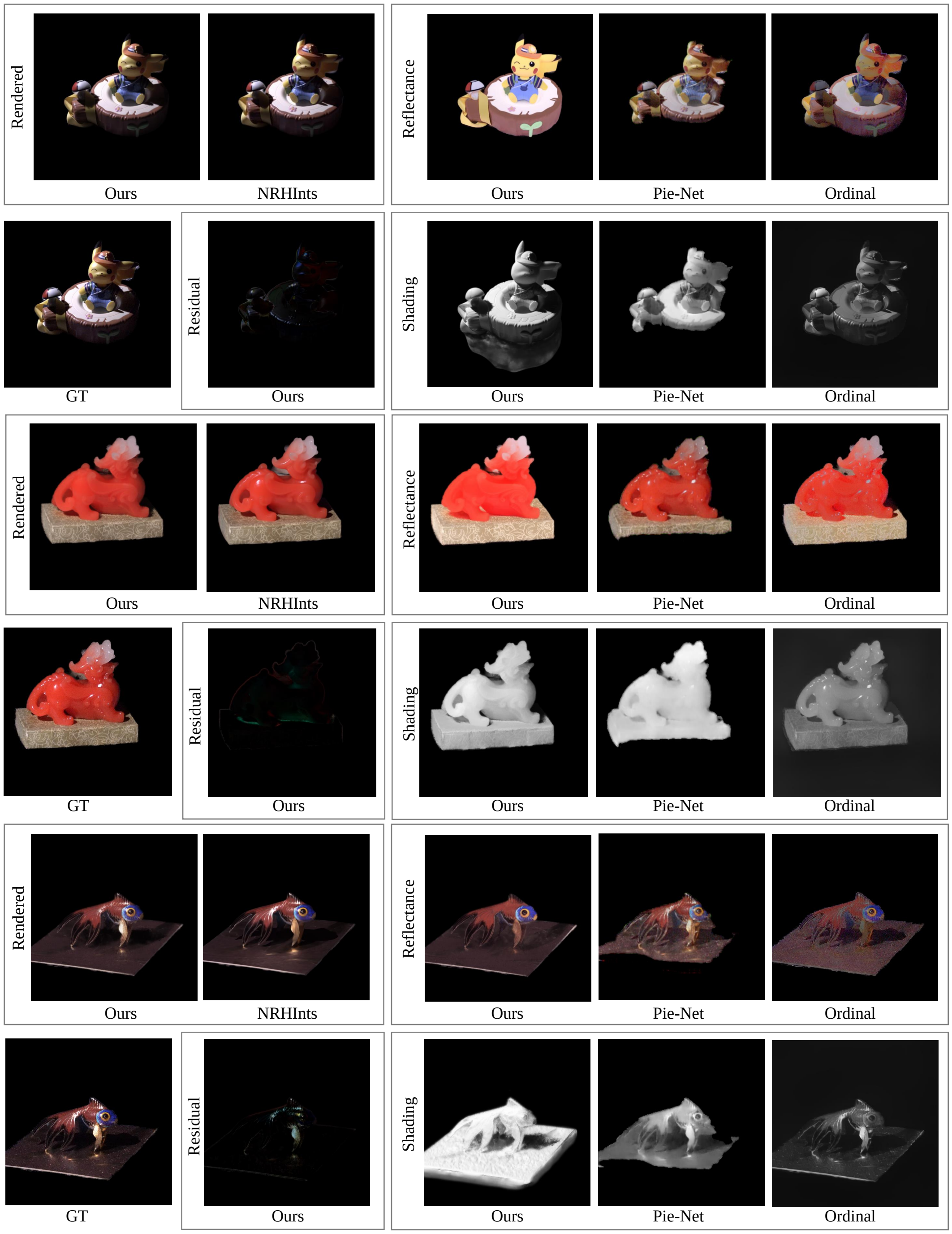


Figure 5. Additional Qualitative Results on the Real Object Dataset.

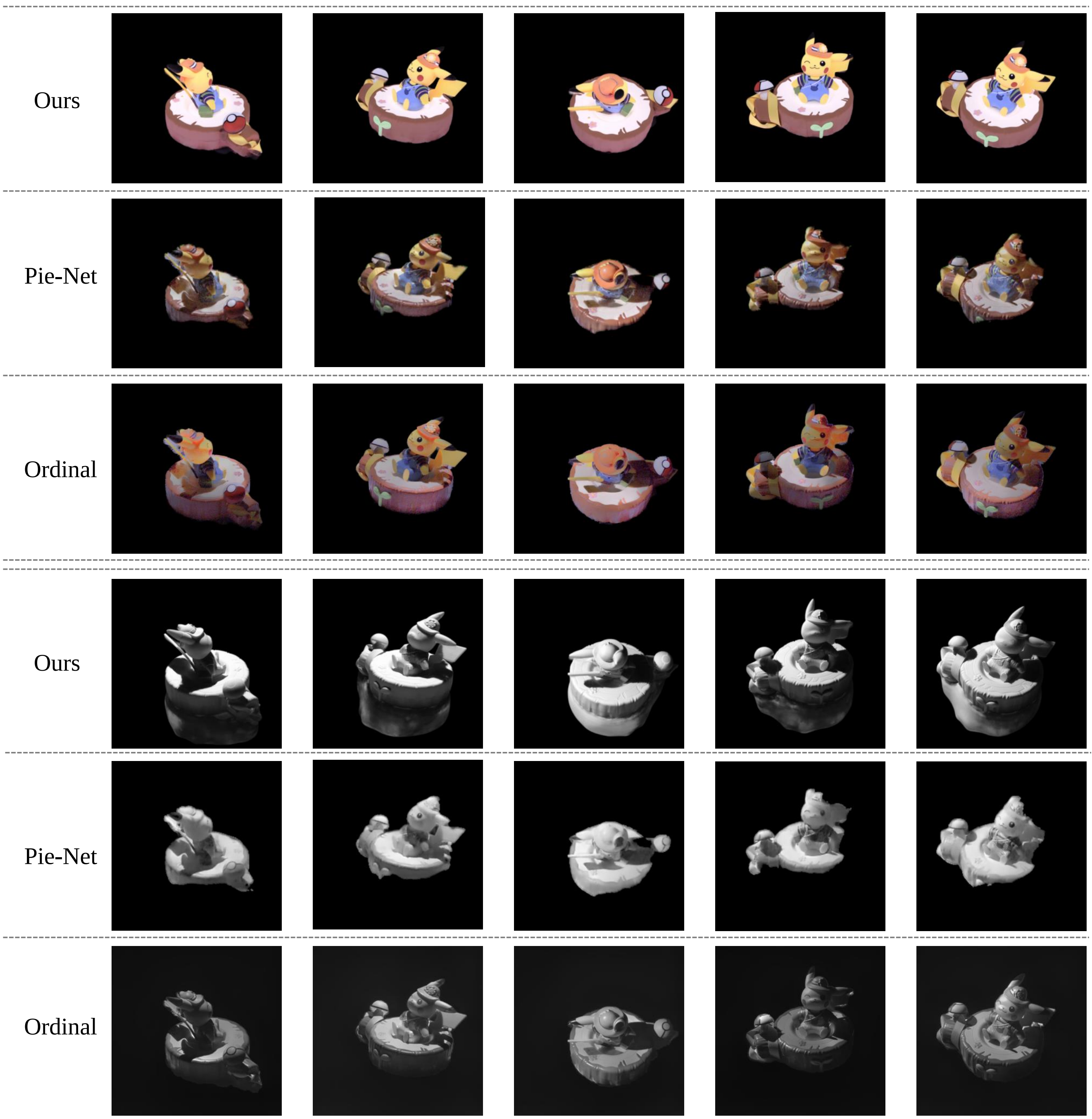


Figure 6. Reflectance and Shading estimation by our method for different points of view of the same scene on the Real Object Dataset.

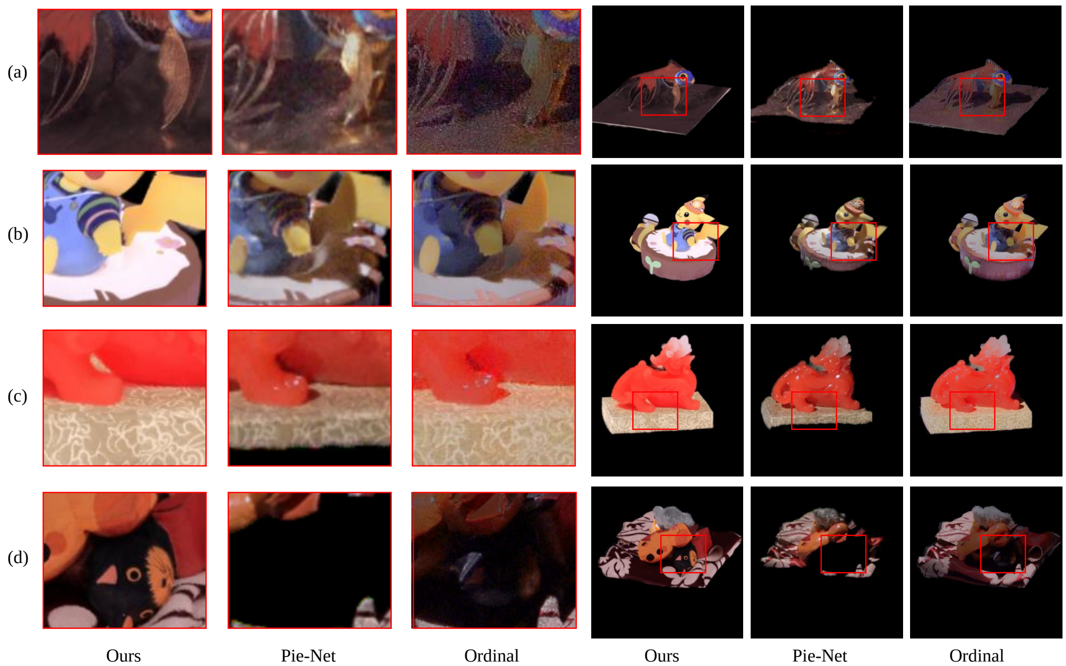


Figure 7. Reflectance Estimation details for different Scenes of the Real Object Dataset.

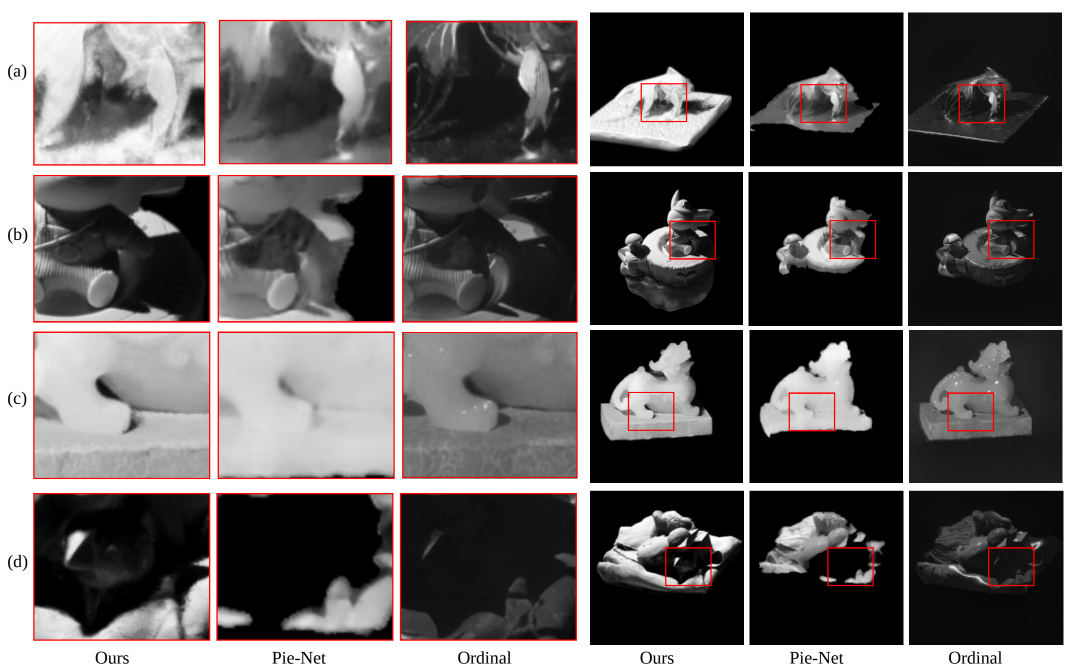


Figure 8. Shading Estimation details for different Scenes of the Real Object Dataset.

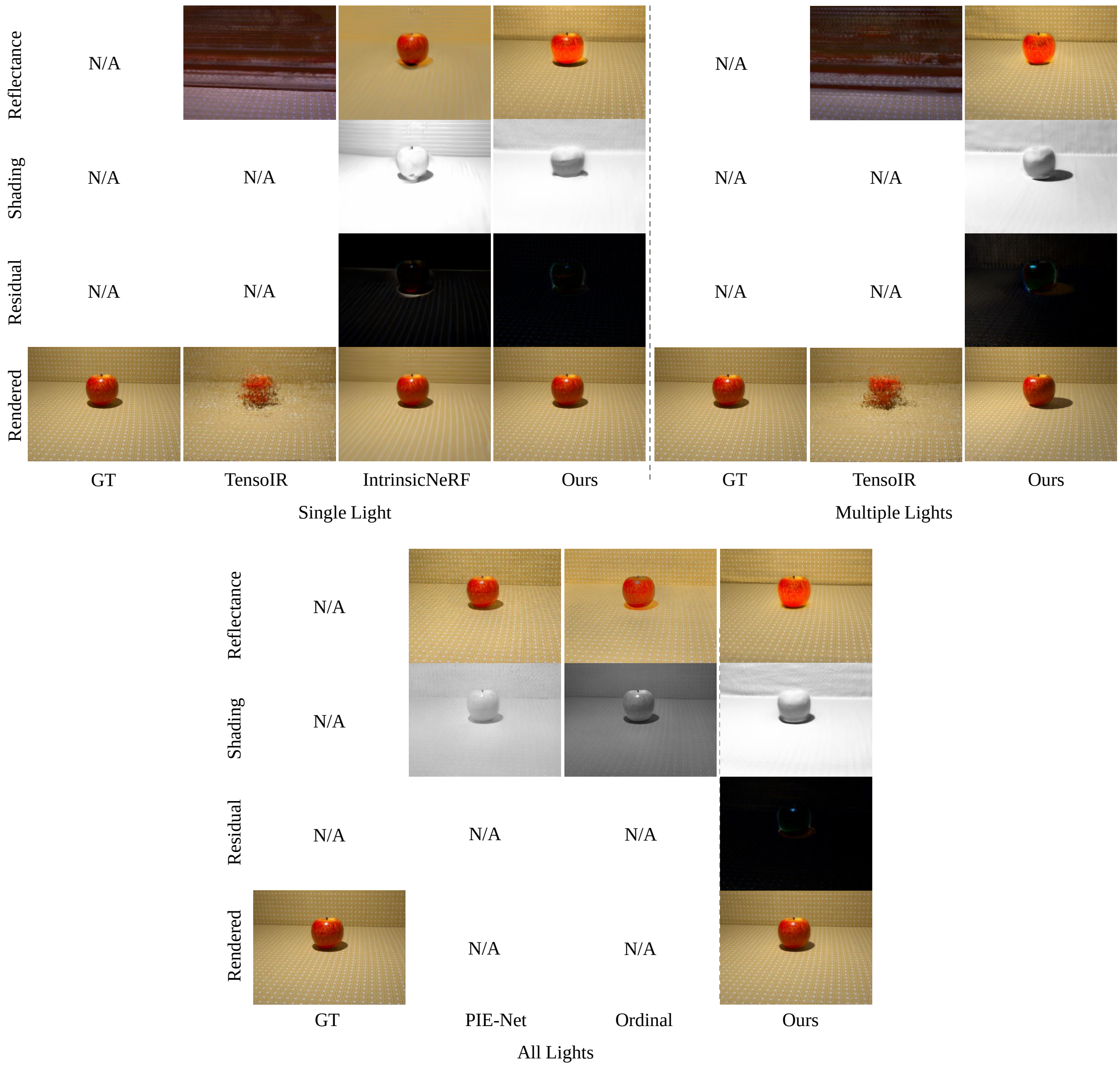
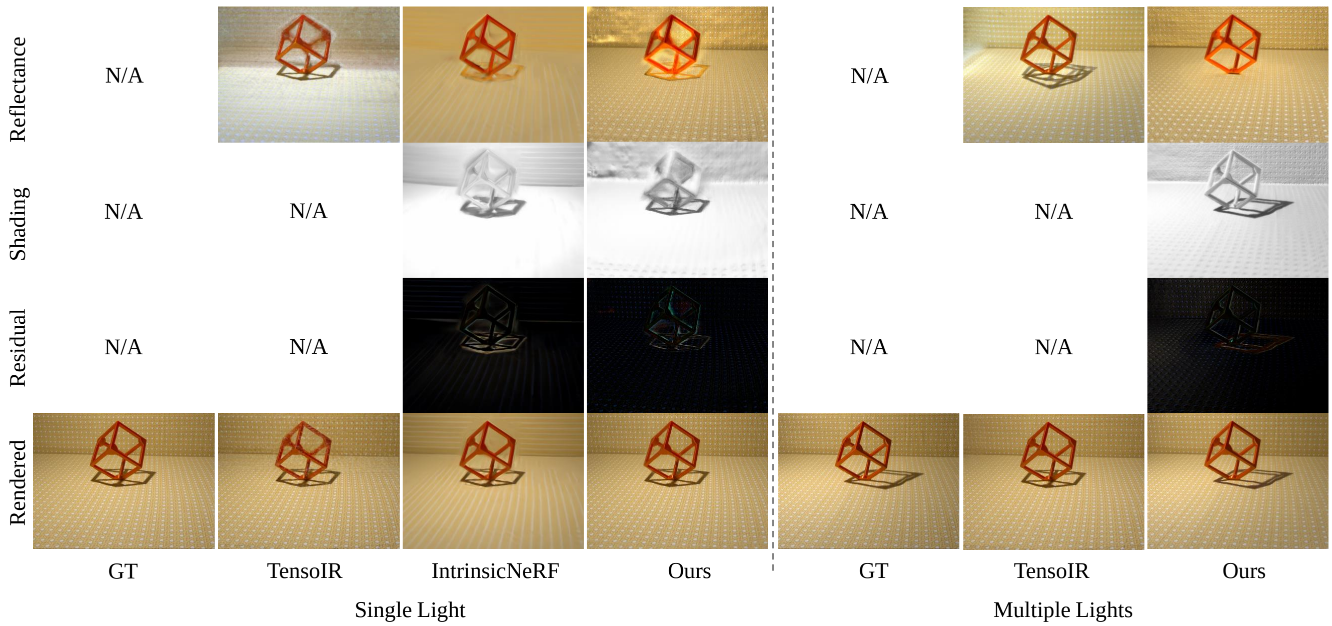


Figure 9. Additional Qualitative Results on the ReNe dataset. (Apple).



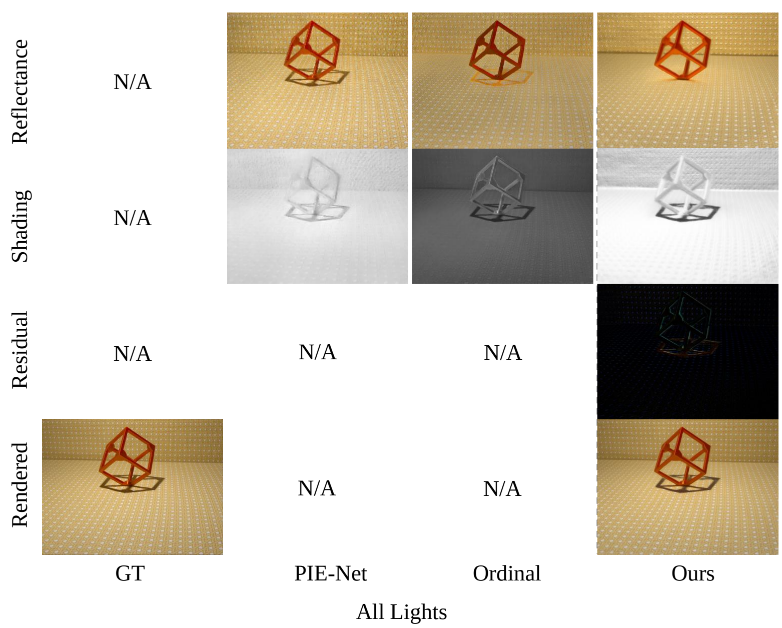


Figure 10. Additional Qualitative Results on the ReNe dataset. (Cube).

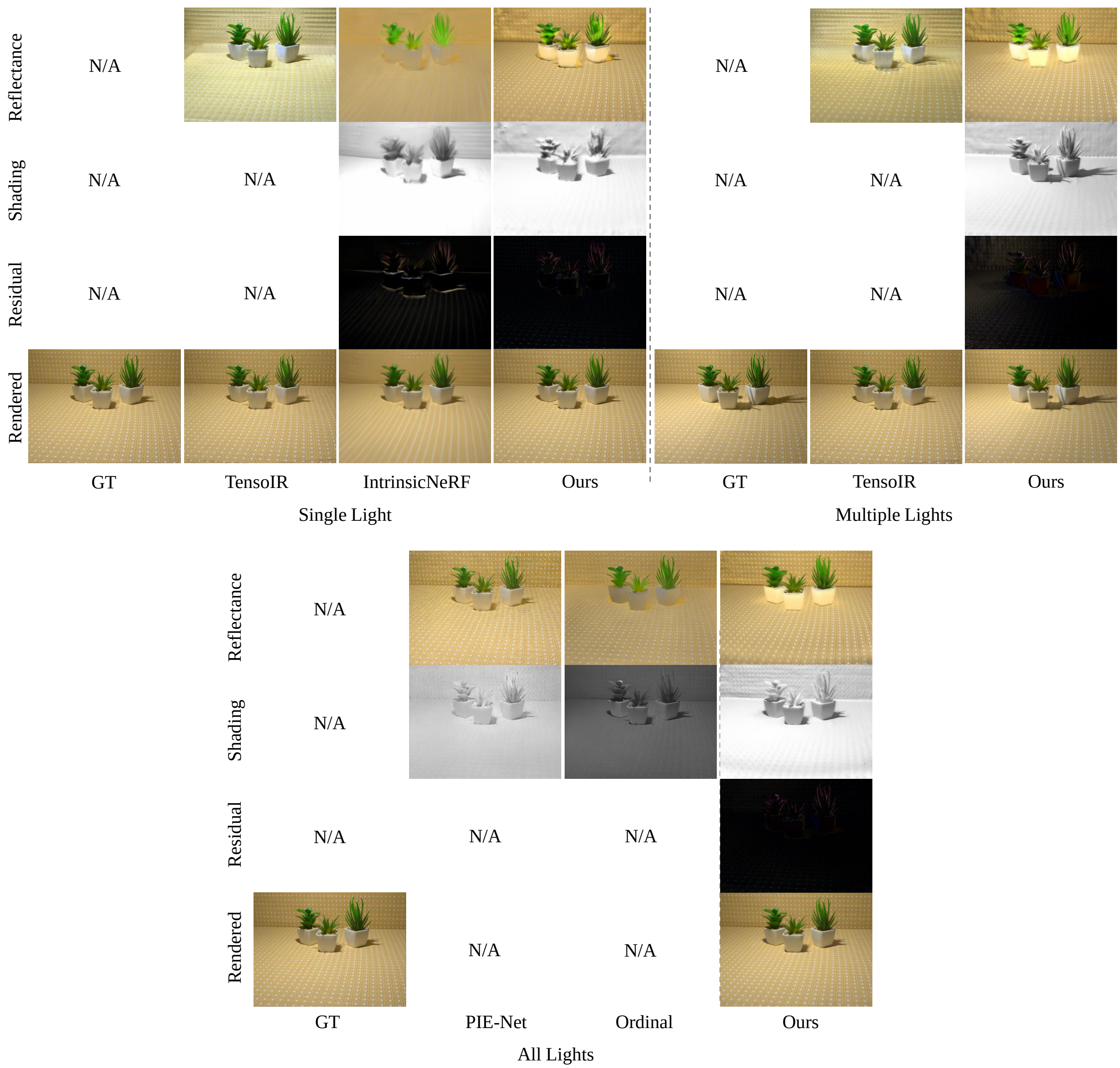
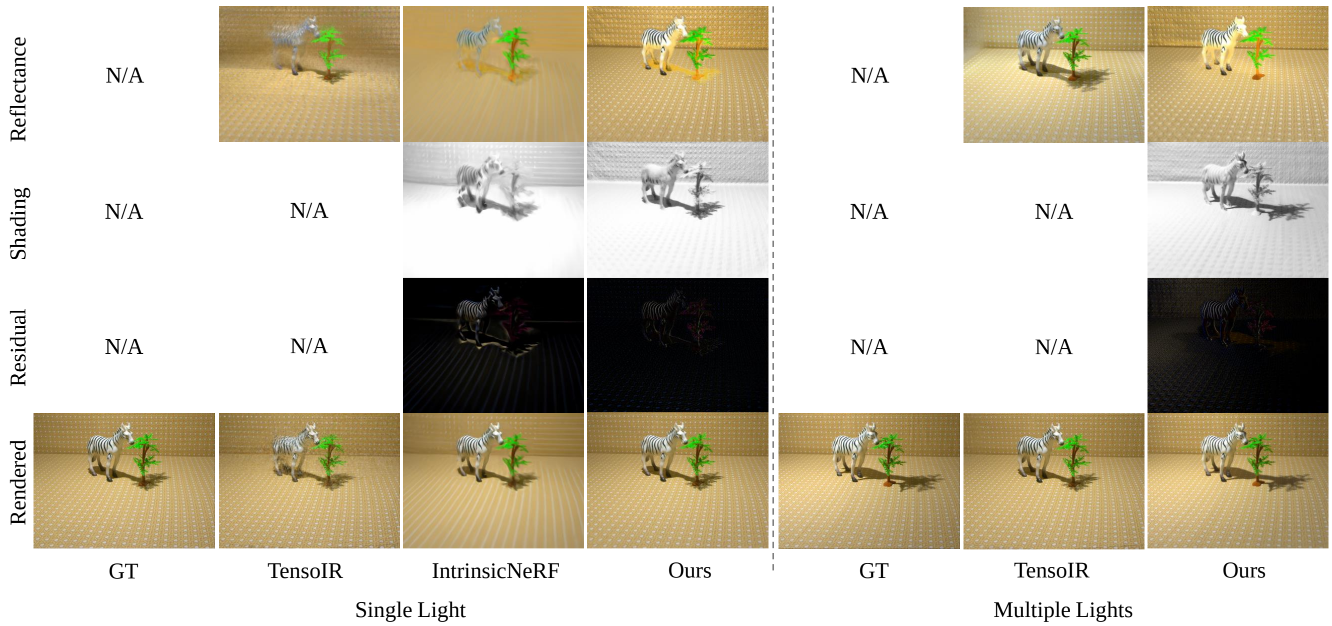


Figure 11. Additional Qualitative Results on the ReNe dataset. (Garden).



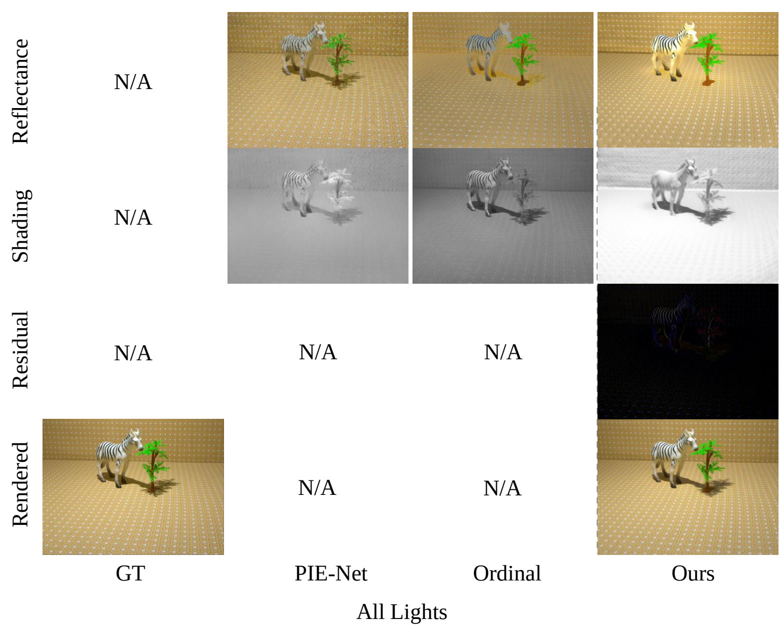


Figure 12. Additional Qualitative Results on the ReNe dataset. (savannah).

References

[1] Chris Careaga and Yağız Aksoy. Intrinsic image decomposition via ordinal shading. *ACM Trans. Graph.*, 2023. 1

[2] Partha Das, Sezer Karaoglu, and Theo Gevers. Pie-net: Photometric invariant edge guided network for intrinsic image decomposition. In *IEEE Conference on Computer Vision and Pattern Recognition, (CVPR)*, 2022. 1

[3] Duan Gao, Guojun Chen, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. Deferred neural lighting: free-viewpoint relighting from unstructured photographs. *ACM Transactions on Graphics (TOG)*, 39(6):258, 2020. 1

[4] Haian Jin, Isabella Liu, Peijia Xu, Xiaoshuai Zhang, Songfang Han, Sai Bi, Xiaowei Zhou, Zexiang Xu, and Hao Su. Tensoir: Tensorial inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 1

[5] Chong Zeng, Guojun Chen, Yue Dong, Pieter Peers, Hongzhi Wu, and Xin Tong. Relighting neural radiance fields with shadow and highlight hints. In *ACM SIGGRAPH 2023 Conference Proceedings*, 2023. 1