Unsupervised Learning for Intrinsic Image Decomposition from a Single Image
(Supplementary Material)

Yunfei Liu[1] Yu Li[2] Shaodi You[3] Feng Lu[1,4]
[1] State Key Laboratory of VR Technology and Systems, School of CSE, Beihang University
[2] Applied Research Center (ARC), Tencent PCG [3] University of Amsterdam
[4] Peng Cheng Laboratory, Shenzhen, China
{lyunfei,lufeng}@buaa.edu.cn ianyli@tencent.com s.you@uva.nl

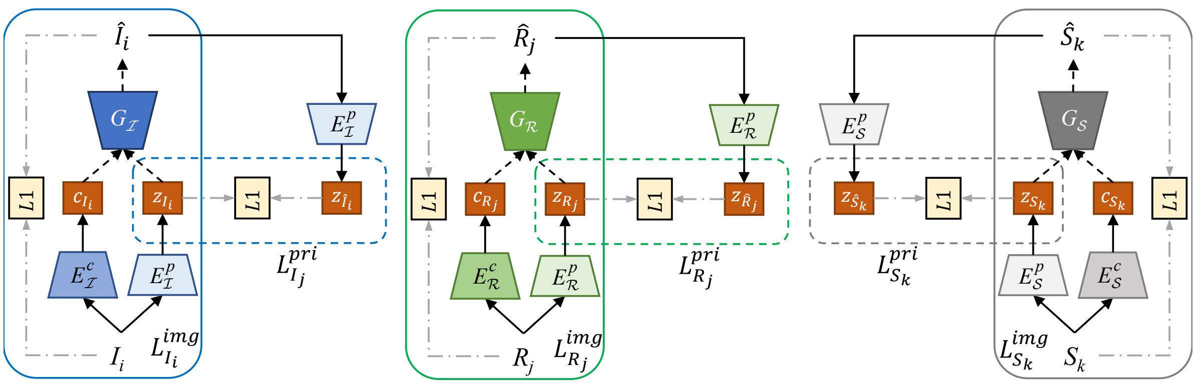

Figure 1. The architecture of auto-encoders for natural image domain $\mathcal{I}$ (left), reflectance domain $\mathcal{R}$ (middle) and shading domain $\mathcal{S}$ (right).

We provide more implementation detail and experimental results in this supplementary material. In detail, we prove more network architecture detail in Section A; we provide more training detail in Section B. Section C shows more visual results on MPI Sintel intrinsic benchmark [1]. Section D shows more visual results on MIT intrinsic dataset [4]. Section E shows additional qualitative results on IIW benchmark [7]. Section F shows more visual results on ShapeNet intrinsic dataset [2].

A. More Network Architecture Detail

The three auto encoders for three different domains are silimar. We only shown one of them in the main manuscript and in this supplementary material, we provide the implementation detail for all three auto-encoders in Figure 1.

Also, the detail of content encoder and shading encoder is illustrated in Figure 2. The content encoder consists of several strided convolutional layers followed by residual blocks. The style encoder contains several strided convolutional layers followed by a global average pooling layer and a fully connected layer.

B. Training details

The proposed framework is implemented with PyTorch. We adopt ADAM solver for optimizing. The learning rate is initialized at 1e-4, and decreased to 1e-5 to generate further improvement. The parameters of generators (including encoders, decoders and mapping) are initialized with the *Gaussian* method, the parameters of discriminators are initialized with the *Kaiming* method. The training samples are randomly shuffled at the beginning of each epoch. The model are trained and tested on an NVIDIA Titan X GPU.

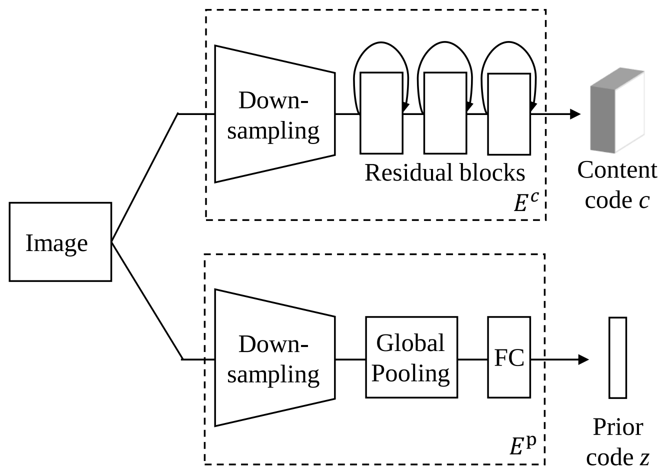

Figure 2. The detailed architecture of content encoder E^c and distribution prior encoder E^p.

C. Extra results on MPI Sintel

In Figure 3, we show more qualitative results on the MPI Sintel intrinsic benchmark. Additional visual results of test patches are shown in Fig. 4.

D. Extra results on MIT intrinsic dataset

Figure 5 shows more visual results on the MIT intrinsic dataset. The reference results are also presented for better comparison of image intrinsic decomposition methods. From the final output, it can be observed that our proposed method generates better results than the existing unsupervised methods [5] and even comparable to recent supervised method FY18 [3].

E. Extra results on IIW

Figure 6 shows more visual results on the IIW dataset. The reference results are also presented for better comparison of image intrinsic decomposition methods. From the final output, it can be observed that our proposed method generates better results than the existing unsupervised methods [6] and even comparable to recent supervised method FY18 [3].

F. Extra results on ShapeNet intrinsic dataset

Figure 7 shows more visual results on the ShapeNet intrinsic dataset. It can be observed that the natural images can be decomposed properly.

References

[1] Daniel J Butler, Jonas Wulff, Garrett B Stanley, and Michael J Black. A naturalistic open source movie for optical flow evaluation. In *ECCV*, 2012. 1

[2] Angel X Chang, Thomas Funkhouser, Leonidas Guibas, Pat Hanrahan, Qixing Huang, Zimo Li, Silvio Savarese, Manolis Savva, Shuran Song, and Hao Su. Shapenet: An information-rich 3d model repository. In *arXiv preprint arXiv:1512.03012*, 2015. 1

[3] Qingnan Fan, Jiaolong Yang, Gang Hua, Baoquan Chen, and David Wipf. Revisiting deep intrinsic image decompositions. In *CVPR*, 2018. 2, 3, 4, 5, 6

[4] Roger Grosse, Micah K Johnson, Edward H Adelson, and William T Freeman. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In *ICCV*, 2011. 1

[5] Yu Li and Michael S. Brown. Single image layer separation using relative smoothness. In *CVPR*, 2014. 2

[6] Zhengqi Li and Noah Snavely. Learning intrinsic image decomposition from watching the world. In *CVPR*, 2018. 2, 3, 4, 5, 6

[7] Bell Sean, Bala Kavita, and Snavely Noah. Intrinsic images in the wild. In *SIGGRAPH*, 2014. 1

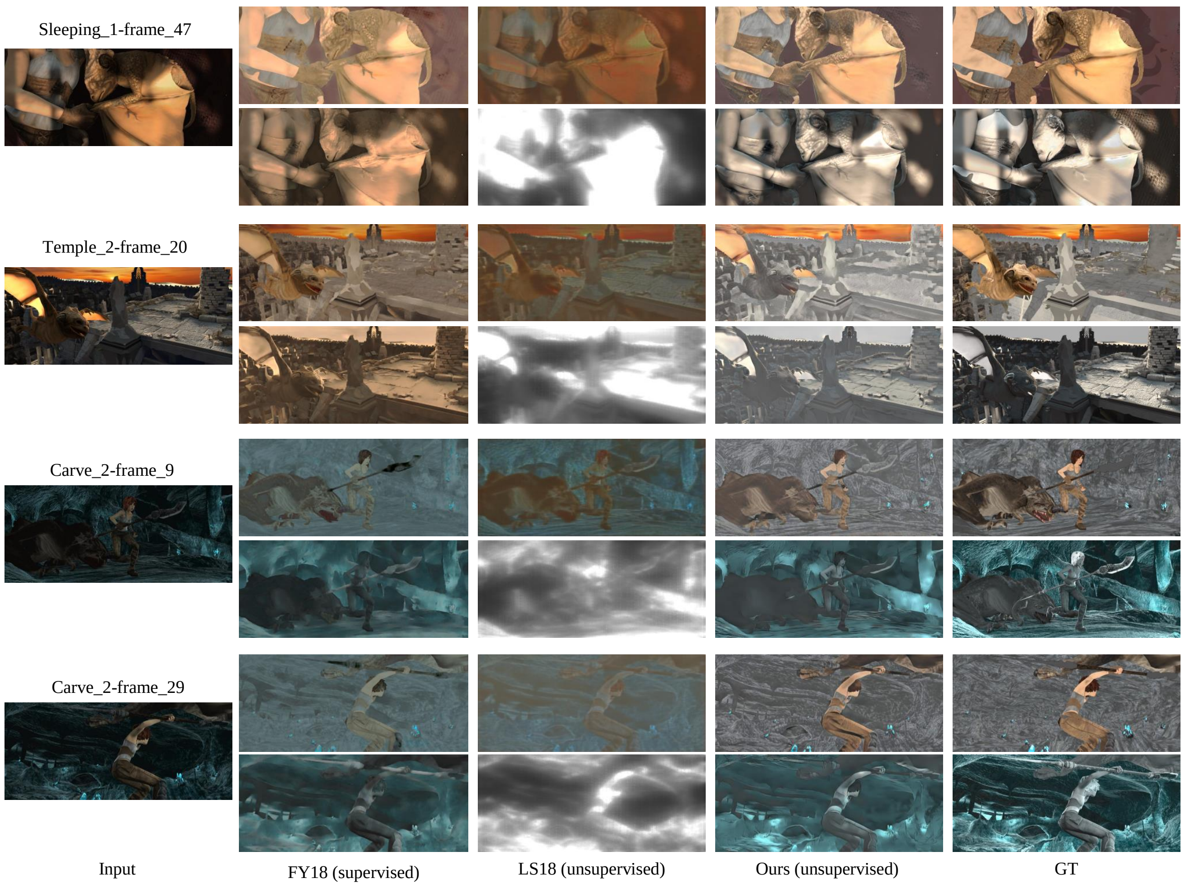

Figure 3. Extra visual results on MPI Sintel benchmark. Compared with state-of-the-art supervised method FY18 [3] and unsupervised method LS18 [6].

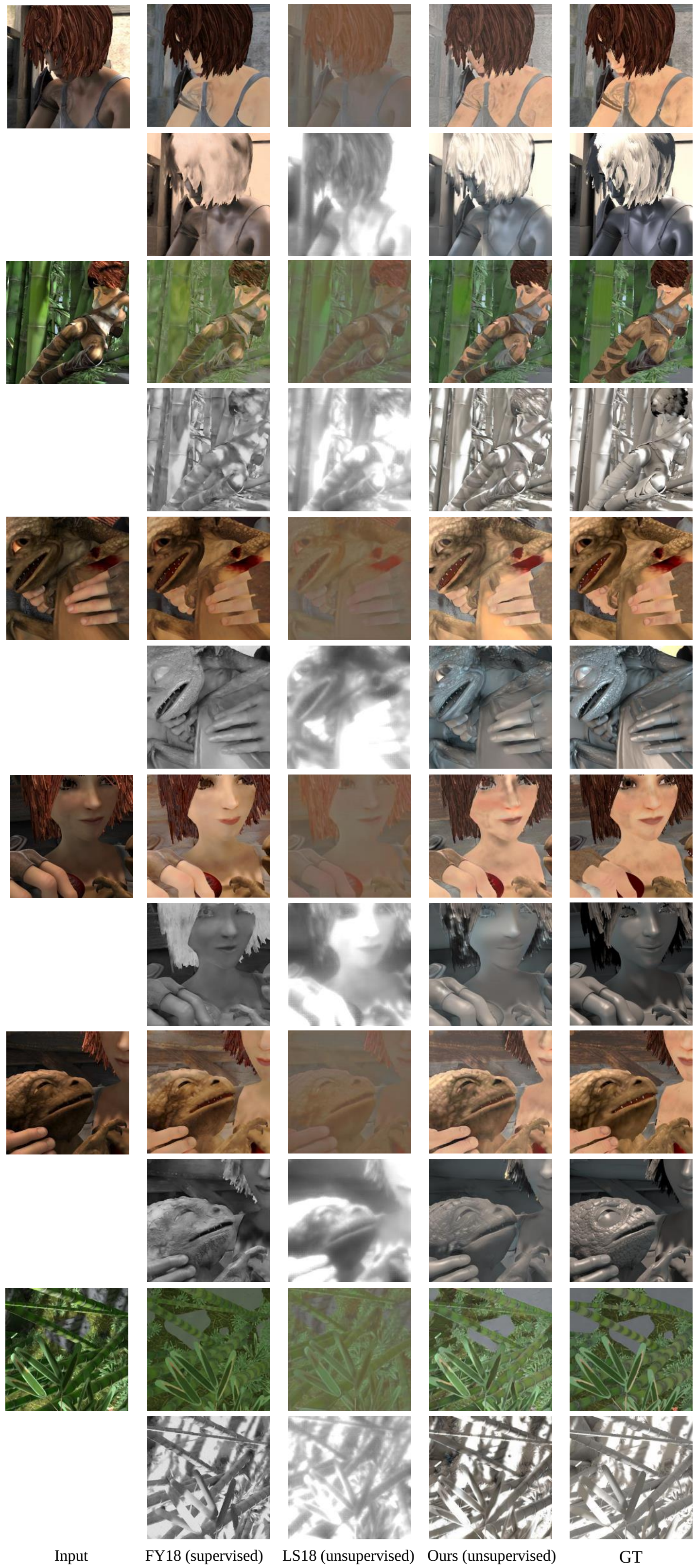

Figure 4. Extra visual results on patches of MPI Sintel benchmark. Compared with state-of-the-art supervised method FY18 [3] and unsupervised method LS18 [6].

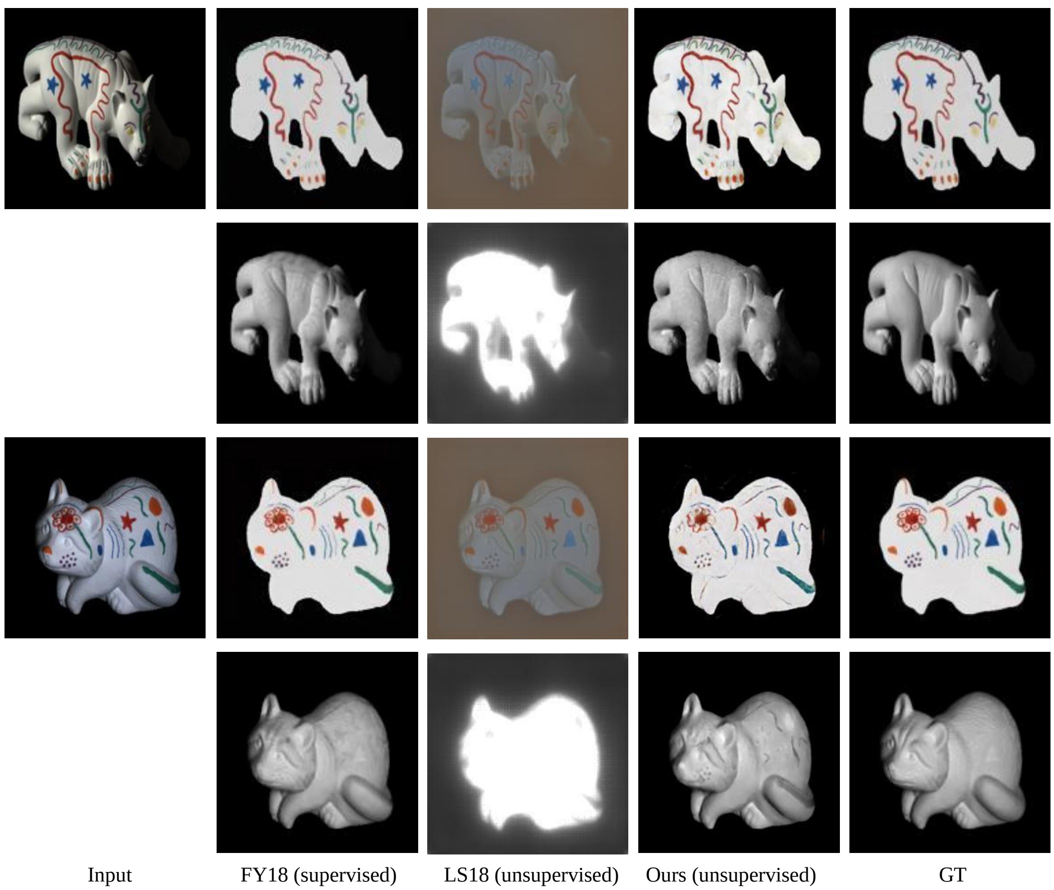

Figure 5. Extra qualitative comparison on the MIT test set. FY18 [3] is supervised method. LS18 [6] is unsupervised.

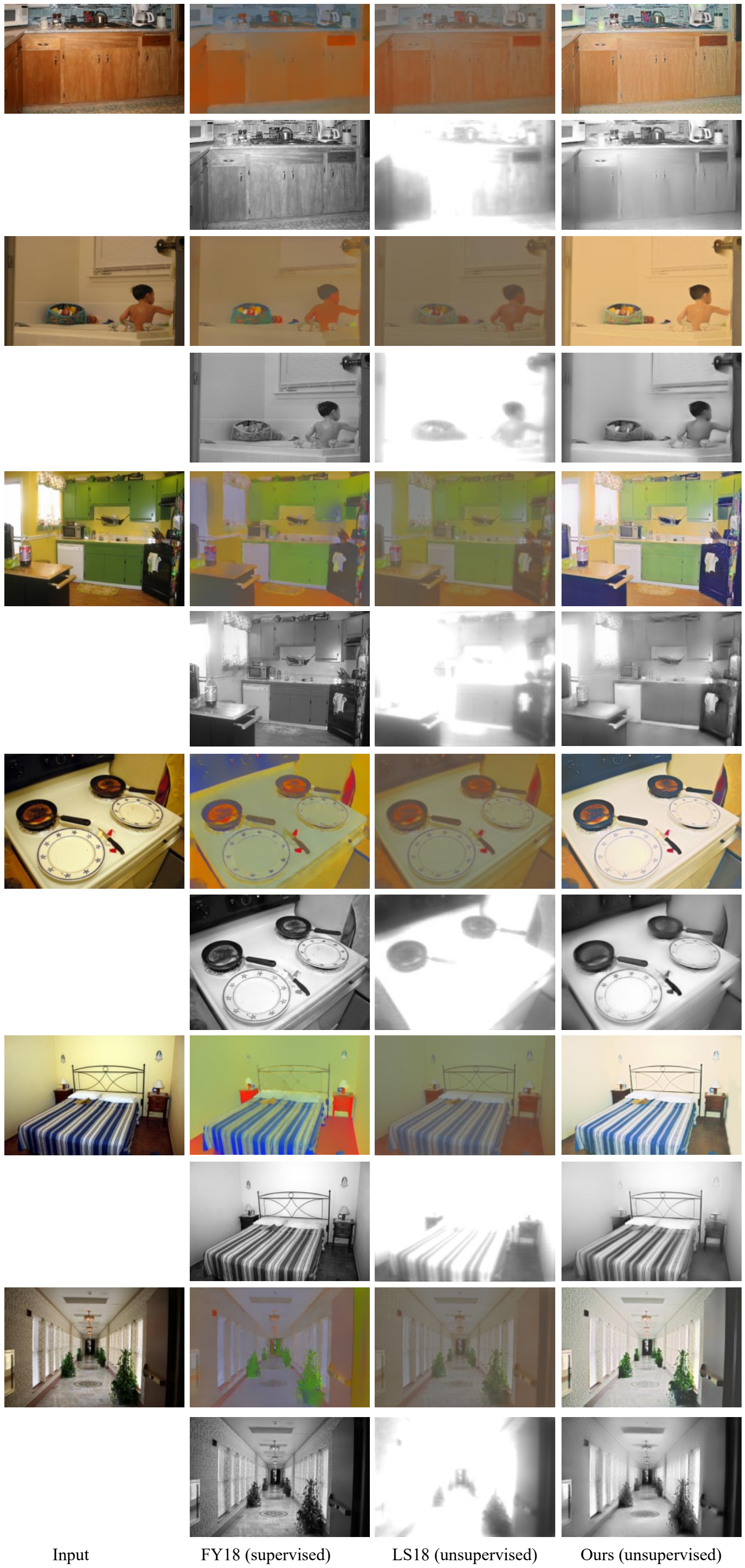

Figure 6. Extra qualitative comparison on the IIW test set. FY18 [3] is supervised method. LS18 [6] is unsupervised.

.

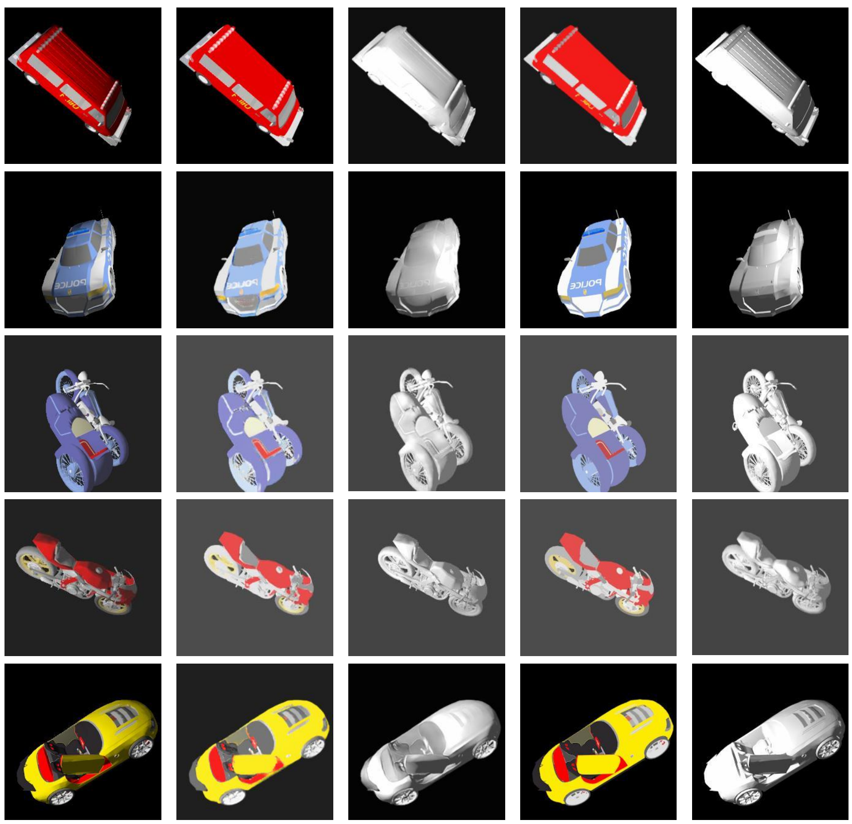

Figure 7. More visual results on the ShapeNet intrinsic dataset.

.