Single Image Reflection Separation with Perceptual Losses
Supplementary Material

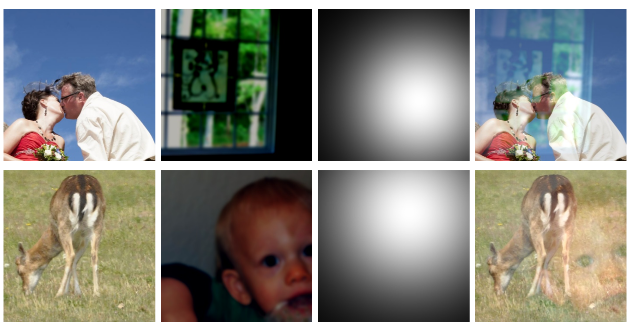
Xuaner Zhang
UC Berkeley

Ren Ng
UC Berkeley

Qifeng Chen
Intel Labs

A. Synthetic data generation

We illustrate how we generate our synthetic dataset. We followed a similar forward model described in CEILNet [1] with modification stated in Section 5.1 of the main paper. The vignette mask is applied to the reflection layer to simulate reflection captured from an oblique angle of view. In Figure 1, a synthetic image I is generated by $I = T + R \odot V$, where V denotes the vignette mask and $\odot$ denotes element-wise multiplication.



T R V I

Figure 1: Two examples on synthetic image generation. From left to right: transmission layer, reflection layer, vignette mask, and the blended synthetic image.

B. User study

This section provides more details on how we conducted the user study on Amazon Mechanical Turk. For each comparison, we present three images in a row: an input image with reflection, our transmission result, and the transmission result from a baseline we compared with. There are 160 real image comparisons in total; the test images are from real images provided by CEILNet dataset without ground truth, and our collected real-world test data. Among all 160 comparisons, 80 are compared against CEILNet [1] and 80 are compared against Li and Brown [2]. We randomly divided the 160 comparisons into two Human Intelligence Tasks (HITs), each contains 80 comparisons. In each HIT, we added 5 sentinel tests, where we replaced one of the results with ground truth. If a user fails more than 1 out of the 5 sanity check comparisons, we discard all the responses of that user.

C. Reflection layer for real dataset.

This section explains why we only collect reflection images on synthetic data but not on real-world data. There are two ways to get the ground-truth reflection layer R: to capture it with the portable glass covered with black cloth and use it as a mirror, or to use the captured I and T to compute R through $R = I - T$. However, we end up not using R to constrain $f_R(I;\theta)$ on real images for the following two reasons:

- Captured R can be hardly aligned at pixel precision: we tried to capture the ground-truth reflection layer by covering the portable glass with a black cloth and using it as a mirror. However, the reflection layer, usually not in focus and thus blurred, can be hardly aligned precisely with the input image (see Figure 2 top row).
- $R = I - T$ is unreliable: we also tried to obtain the ground-truth R from $R = I - T$, computed in the linear color space before gamma correction. However, we observed residual of T appear in the computed R (see Figure 2 bottom row), and artifacts at saturated pixels.

D. Qualitative results on real data with ground truth.

More results on real data with ground truth are shown in Figure 3 (3 pages). We compare our method with CEILNet [1] on our collected real test data. This extends Figure 5 in the main paper.

E. Qualitative results on real data without ground truth.

More results on real data without ground truth are shown in Figure 4 (3 pages) and Figure 5. We compare our method with CEILNet [1], Li and Brown [2] on publicly available real images. This extends Figure 6 in the main paper.

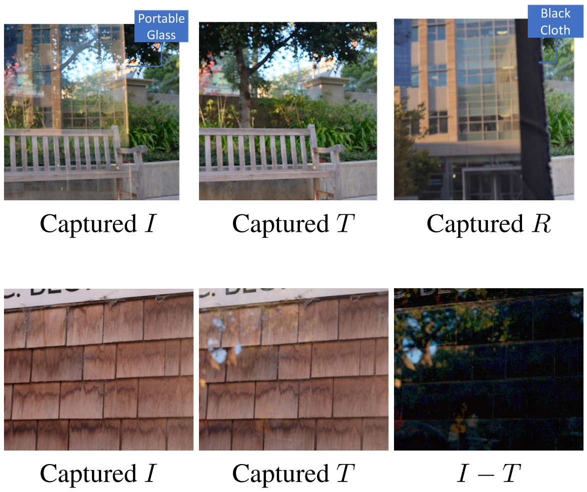


Figure 2: Reflection layer in real data. Top row left to right: captured I with a portable glass, captured T with the glass removed, and captured R with a black cloth covering the back of the glass. Bottom row left to right: capture I, captured T and computed R from $R = I - T$ in linear space before gamma correction.

F. Qualitative results on image dehazing.

Figure 6 shows more results on the extension application of image dehazing using our network. Note that we directly apply our trained network without training or fine-tuning on any dehazing dataset. This extends Figure 7 in the main paper.

References

[1] Q. Fan, J. Yang, G. Hua, B. Chen, and D. Wipf. A generic deep architecture for single image reflection removal and image smoothing. In *ICCV*, 2017. 1, 6, 7, 8

[2] Y. Li and M. S. Brown. Single image layer separation using relative smoothness. In *CVPR*, 2014. 1

Figure 3: Qualitative comparison on our collected real dataset.

Figure 3 (Cont.): Qualitative comparison on our collected real dataset.

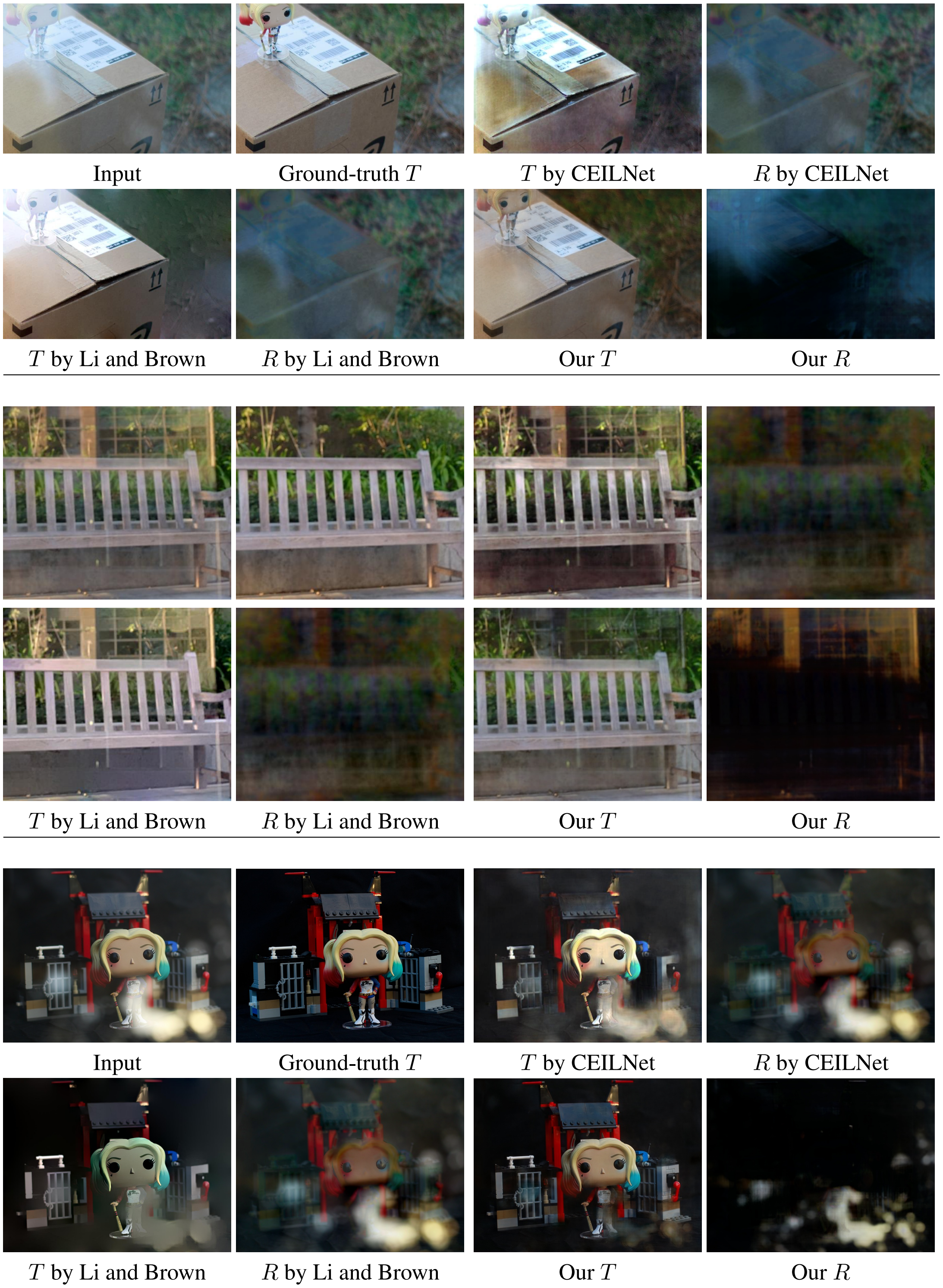


Figure 3 (Cont.): Qualitative comparison on our collected real dataset.

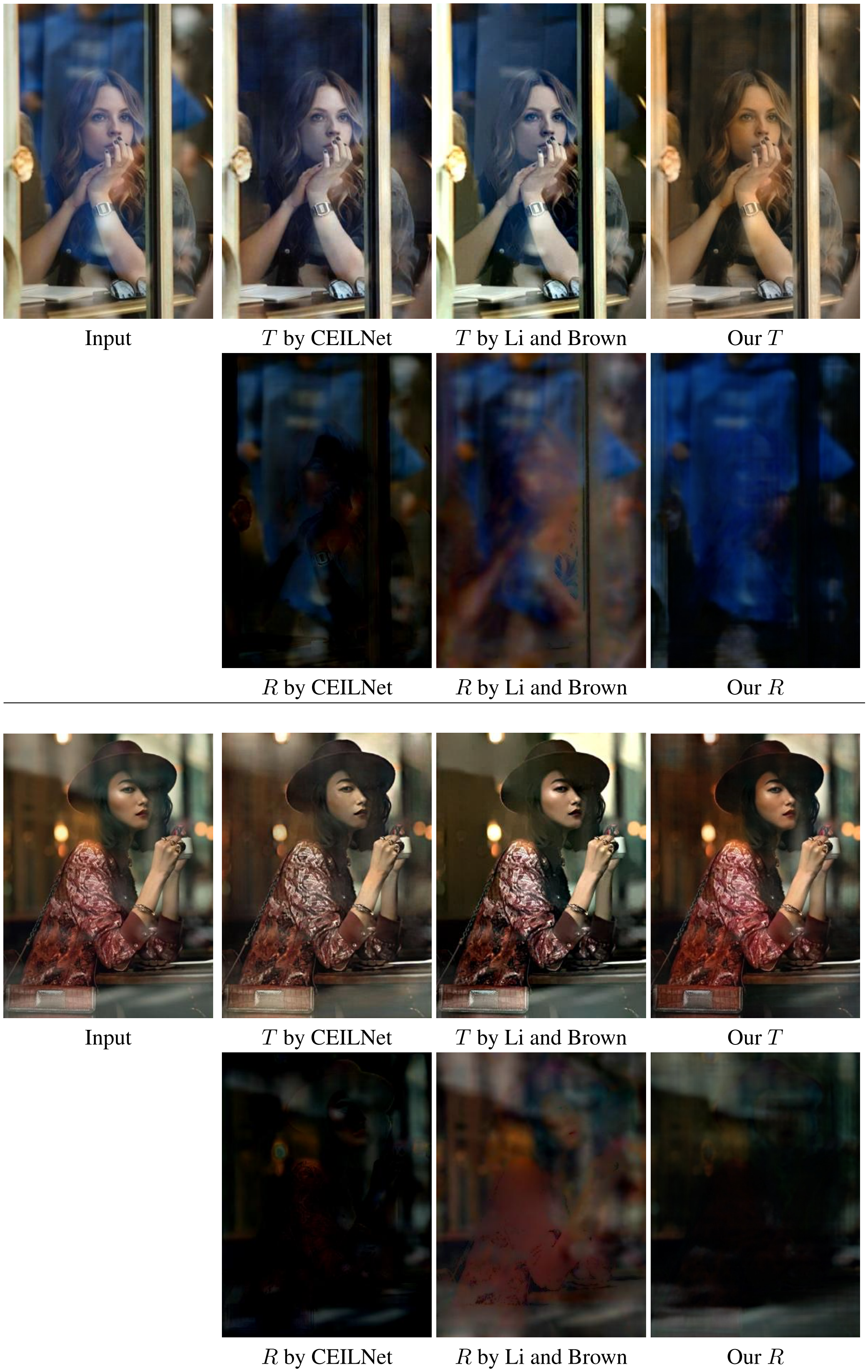


Figure 4: Qualitative comparisons on real data from CEILNet [1].

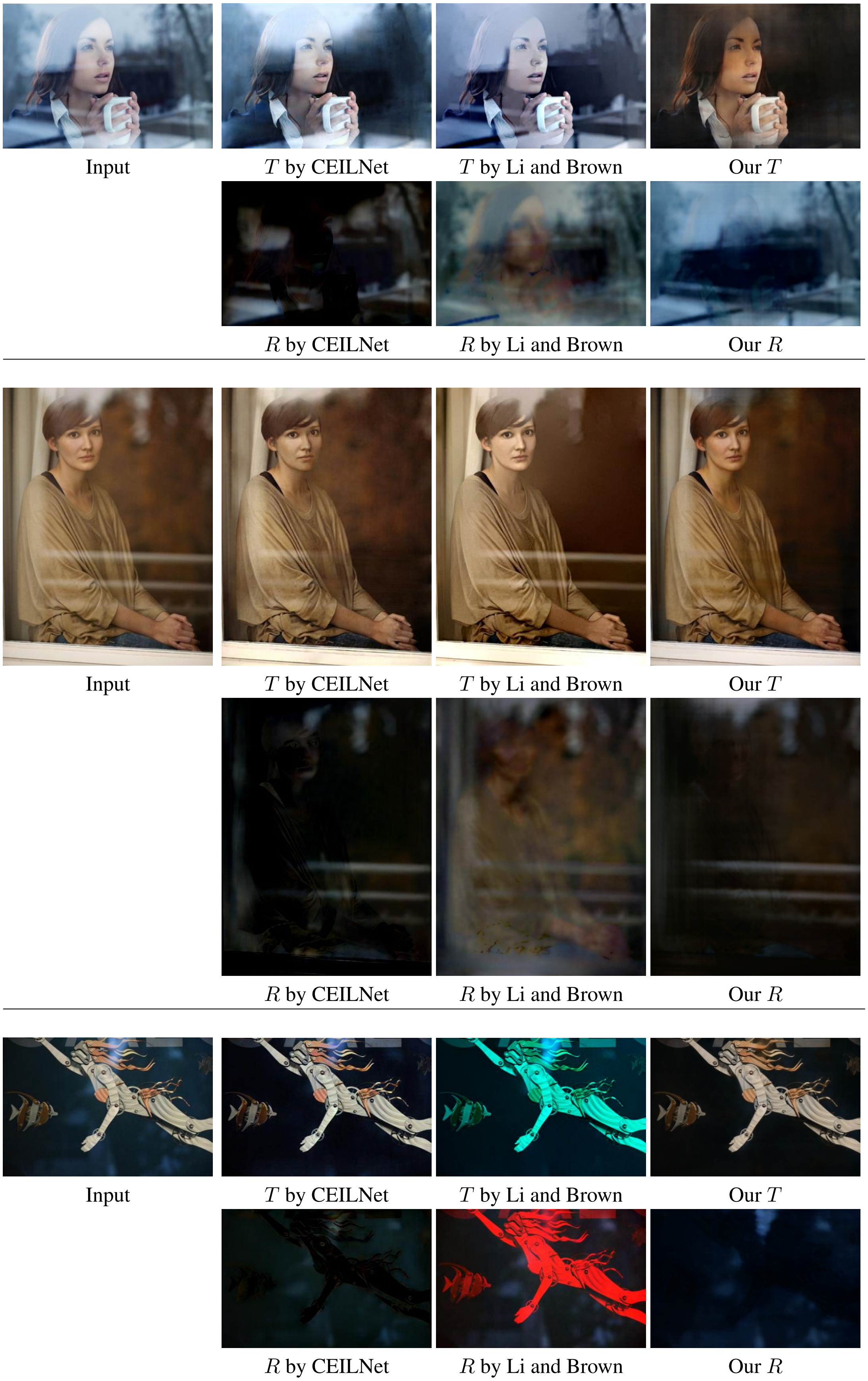
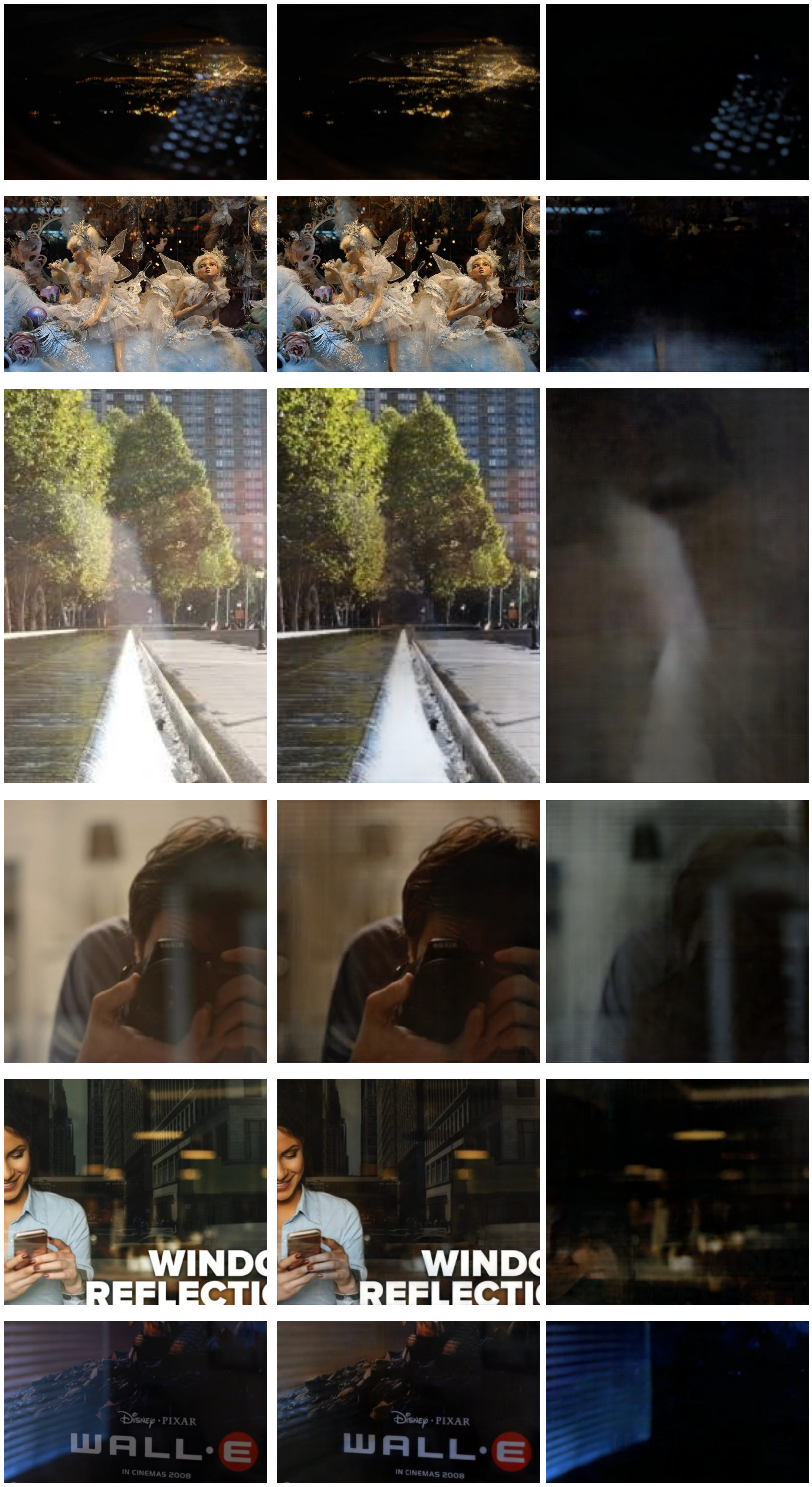


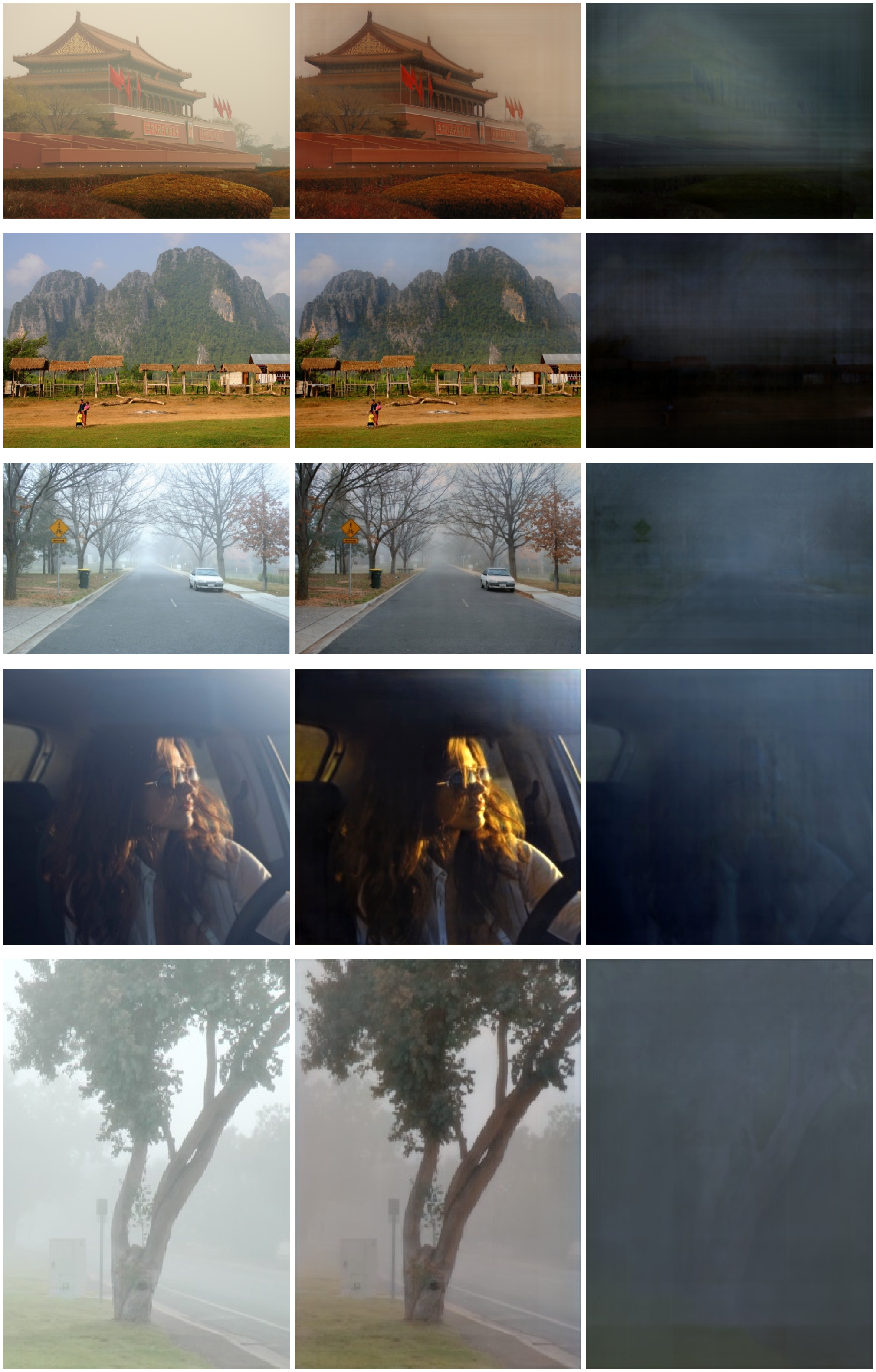
Figure 4 (Cont.): Qualitative comparisons on real data from CEILNet [1].

Figure 4 (Cont.): Qualitative comparisons on real data from CEILNet [1].



Input | Our T | Our R

Figure 5: Visual results on more real images.



Input | Our T | Our R

Figure 6: Visual results on image dehazing.