1 More comparisons on the MIT-Adobe 5K dataset

In the paper, we proposed three models for enhancing images, SL (supervised learning using the proposed generator trained on the MIT-Adobe 5K dataset), UL (the proposed 2-way GAN trained on the MIT-Adobe 5K dataset) and HDR (the proposed 2-way GAN trained on the collected HDR dataset). We compare the proposed models with five state-of-the-art methods, including CycleGAN [10], DPED [5], CLHE [8], NPEA [9] and FLLF [1]. Note that our model can be taken as an enhanced CycleGAN with three proposed improvements, a better generator, a better WGAN model with the adaptive weighting scheme and a better 2-way GAN model with individual batch normalization layers. The CycleGAN was trained on the collected HDR dataset. The DPED models are tailored with different mobile phones. Here, we show results of the DPED models for iPhone6, iPhone7 and Nexus 5x.

This section compares these methods on some images from the MIT-Adobe-5K [2] testing dataset. In general, we have the following observations. Our models trained with the photographer’s labels approximate the labels reasonably well. DPED models vary a lot with different phone models. Though trained on the same HDR dataset, CycleGAN cannot capture the characteristics of the dataset as well as our HDR model. It shows that the proposed improvements are effective and important, at least for this application. CLHE, NPEA and FLLF are not robust and could generate unnatural enhanced images at times. Our HDR model captures the characteristics of the collected HDR dataset well and generates the most natural enhanced images.
Fig. 1: Comparisons of different methods on a MIT-adobe-5K testing image, a3552.
Supplementary Material: Deep Photo Enhancer

Fig. 2: Comparisons of different methods on a MIT-adobe-5K testing image, a0212
Fig. 3: Comparisons of different methods on a MIT-adobe-5K testing image, a0481
Fig. 4: Comparisons of different methods on a MIT-adobe-5K testing image, a3203
Fig. 5: Comparisons of different methods on a MIT-adobe-5K testing image, a0535
Fig. 6: Comparisons of different methods on a MIT-adobe-5K testing image, a1305
Fig. 7: Comparisons of different methods on a MIT-adobe-5K testing image, a4963
2 More comparisons on images from the Internet

We show results of different methods on enhancing images collected from the Internet. We compare our HDR model with DPED and CLHE. For DPED, we show the results using all three phone models. Although with good enhancement in general, the results of CLHE sometimes look unnatural, particularly on colors. The results of DPED vary among phone models, showing its dependence to phone models. In general, our results give natural results with enhanced color, contrast and details. Note that the model was trained on HDR images. Thus, the results are "HDR-like". Sometime, it could look too prominent. It is however possible to train a modest model with a set of images with that style.

We also provide an accompanying video showing the results of our HDR model on a video. Each frame is processed independently. However, there is no obvious temporal flick. It shows that our model is quite stable. The video also demonstrates that our model can be used for a wide variety of images.
Fig. 8: Comparisons of different methods on an Internet image.
Fig. 9: Comparisons of different methods on an Internet image.
Fig. 10: Comparisons of different methods on an Internet image.
Fig. 11: Comparisons of different methods on an Internet image.
Fig. 12: Comparisons of different methods on an Internet image.
Fig. 13: Comparisons of different methods on an Internet image.
Fig. 14: Comparisons of different methods on an Internet image.
Fig. 15: Comparisons of different methods on an Internet image.
Fig. 16: Comparisons of different methods on an Internet image.
Fig. 17: Comparisons of different methods on an Internet image.
Fig. 18: Comparisons of different methods on an Internet image.
Fig. 19: Comparisons of different methods on an Internet image.
3 The collected HDR dataset

We show sample images of the collected HDR dataset below.

Fig. 20: Sample images from the collected HDR dataset
4 Training

This section shows some figures about the training process. Figure 21 shows the training progress of different GAN models, the proposed A-WGAN, WGAN-GP, DRAGAN, LSGAN (local D), LSGAN and GAN. It can be seen that the proposed A-WGAN generates the best results while some GAN models could totally collapse. Figure 22 shows the discriminator loss along the training process for different one-way GAN models trained on the MIT-Adobe 5K dataset. For the proposed model, the discriminator loss can be used as a good indicator for convergence. Figure 23 shows the discriminator loss for the proposed two-way GAN model with and without the individual batch normalization layers on the MIT-Adobe 5K and the collected HDR datasets. Although training on the MIT-Adobe 5K dataset is effective without individual batch normalization, individual batch normalization layers play an important role on the training with the HDR dataset.

Fig. 21: PSNR values of testing on the MIT-Adobe-5K dataset with different one-way GAN architectures which use different GAN formulas, GAN [3], LSGAN [7], DRAGAN [6] and WGAN-GP [4]. (Local D: using the local discriminator proposed by CycleGAN [10].)
Fig. 22: Discriminator loss of training on the MIT-Adobe-5K dataset for several one-way GAN architectures, GAN [3], LSGAN [7], DRAGAN [6] and WGAN-GP [4]. The value can be used as a good indicator of convergence for our model.

Fig. 23: Discriminator loss of training on the MIT-Adobe-5K dataset and our HDR dataset for the proposed two-way GAN architecture with and without individual BN. It shows that individual BN is crucial for training on the HDR dataset.
References