# GamutMLP: A Lightweight MLP for Color Loss Recovery —Supplemental Material—

Hoang M. Le[1] Brian Price[2] Scott Cohen[2] Michael S. Brown[1]
[1]York University [2]Adobe Research
{hminle,mbrown}@yorku.ca, {bprice,scohen}@adobe.com

In this document, we provide additional details and results that were not included in the main paper due to limited page length.

## 1. Mapping between color spaces

For our Eq. (1) in the main paper, we convert ProPhoto images to sRGB images using a single $3 \times 3$ matrix $\mathbf{M}$. This conversion using $\mathbf{M}$ is comprised of three parts: the transformation from ProPhoto colors to the profile connection space (i.e., CIE XYZ) $\mathbf{M}_{ProPhoto}^{-1}$, the chromatic adaptation $\mathbf{C}_{CAT02}$, and the conversion from CIE XYZ to sRGB with $\mathbf{M}_{sRGB}$. The full matrix can be written as:

$$\mathbf{M} = \mathbf{M}_{sRGB}\mathbf{C}_{CAT02}\mathbf{M}_{ProPhoto}^{-1} \tag{1}$$

Different color spaces use different white points. For example, sRGB and ProPhoto RGB use D65 and D50 respectively. If we want to transform ProPhoto colors to the sRGB space, first we convert ProPhoto colors to the profile connection space, then we chromatically adapt those tristimulus values to the white point of sRGB (D65), and finally convert them to sRGB. We choose the standard chromatic adaptation CIECAM02 [4]. The complete transformation is:

$$\mathbf{M} = \begin{bmatrix} 2.0365 & -0.7376 & -0.2993 \\ -0.2257 & 1.2232 & 0.0027 \\ -0.0105 & -0.1349 & 1.1452 \end{bmatrix} \tag{2}$$

## 2. Training details for baseline methods and inference time

The four baseline DNNs used different training sizes in their github repositories. We felt to be fair we should use a common dataset when training all methods. The $512 \times 512$ crop was the most common input size. The per-image optimization approaches were all significantly better than pretrained DNNs, so small improvements based on different training sizes might not matter.

Inference time over the test dataset of our lightweight MLPs is around **0.27 seconds**, while it takes DNN-based methods approximate **0.57 seconds** in average. The ProPhoto-Sampled [2] that uses a spatially varying mapping functions take up to 45 seconds per image.

We note that the MLP-based methods purposely overfit the MLP to the input image. Longer optimization gives better PSNR. However, to be practical we wanted the optimization to be within 2 seconds.

## 3. Ablations

**Encoding function** The feature encoding (Eq. (4) in our main paper) used has been shown to be effective for implicit neural methods. It makes a big difference in our task. See Table 1 below, where we show that our MLP with the encoding function is significantly better than the MLP without using the function.

Table 1. Our MLPs with and without using the encoding function.

| Method | RMSE↓ | RMSE OG↓ | PSNR↑ | PSNR OG↑ |
|---|---|---|---|---|
| MLP (23KB) w/o enc. | 0.1467 | 0.1547 | 16.67 | 16.21 |
| MLP (23KB) w/ enc. | 0.0021 | 0.0031 | 53.65 | 50.04 |

## 4. Qualitative results

We provide additional qualitative results as shown in the main paper. Figure 1 and Figure 2 compare our method with others on the test set. As shown in the figures, especially in per-pixel RMSE error maps, our approach achieve better qualitative results compared with others.

## References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2017. 2, 3

[2] Hoang Le, Mahmoud Afifi, and Michael S Brown. Improving color space conversion for camera-captured images via wide-gamut metadata. *Color and Imaging Conference*, 2020. 1, 2, 3

[3] Hoang Le, Taehong Jeong, Abdelrahman Abdelhamed, Hyun Joon Shin, and Michael S Brown. GamutNet: Restoring wide-gamut colors for camera-captured images. *Color and Imaging Conference*, 2021. 2, 3


Figure 1. Qualitative comparisons between the predicted ProPhoto full-size output of Clip, Soft Clip, Pix2pix [1], Pix2PixHD [7], ASAP-Net [5], GamutNet [3], PP-sampled [2], SIREN [6]-residual, and our optimized GamutMLP. Error maps of per-pixel RMSE and plots of out-of-gamut (OG) colors on CIE-xy chromaticity diagram with the gamuts of sRGB and ProPhoto are shown.

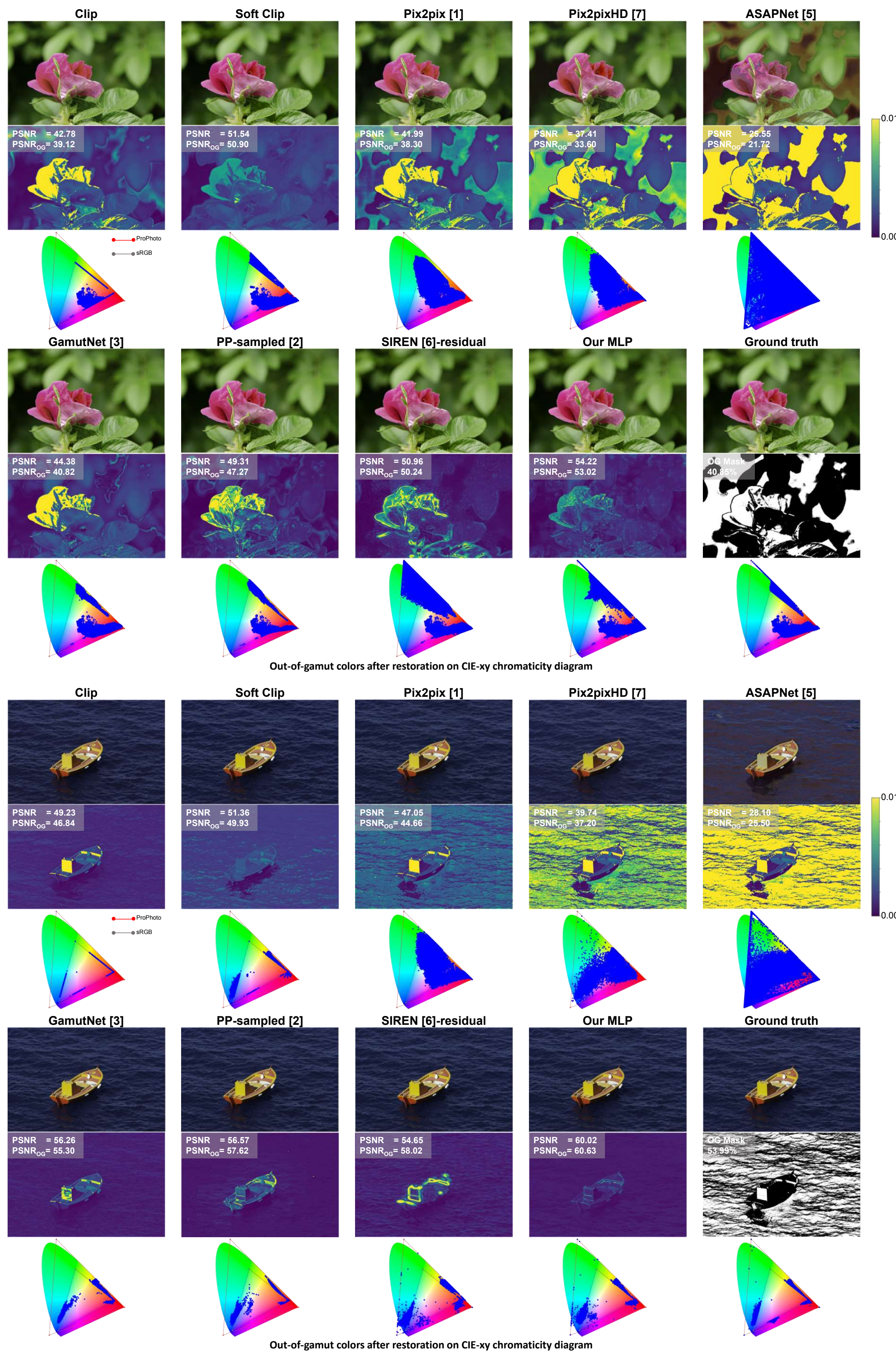


Figure 2. Qualitative comparisons between the predicted ProPhoto full-size output of Clip, Soft Clip, Pix2pix [1], Pix2PixHD [7], ASAP-Net [5], GamutNet [3], PP-sampled [2], SIREN [6]-residual, and our optimized GamutMLP. Error maps of per-pixel RMSE and plots of out-of-gamut (OG) colors on CIE-xy chromaticity diagram with the gamuts of sRGB and ProPhoto are shown.

[4] Nathan Moroney, Mark D Fairchild, Robert W.G. Hunt, Changjun Li, M Ronnier Luo, and Todd Newman. The CIECAM02 color appearance model. In *Color and Imaging Conference*, 2002. 1

[5] Tamar Rott Shaham, Michaël Gharbi, Richard Zhang, Eli Shechtman, and Tomer Michaeli. Spatially-adaptive pixelwise networks for fast image translation. In *CVPR*, 2021. 2, 3

[6] Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *NeurIPS*, 2020. 2, 3

[7] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional gans. In *CVPR*, 2018. 2, 3