Color Matching Using Hypernetwork-Based Kolmogorov-Arnold Networks

Supplementary Material

In this supplementary material, we first discuss the mathematical background of applying Kolmogorov-Arnold Networks (KANs) to the color-matching task in Sec. 1. Next, we provide additional ablation studies in Sec. 2. Then, we elaborate on our proposed dataset in Sec. 3. In Sec. 4, we present and results of our method for raw-to-raw, raw-to-sRGB, and sRGB-to-sRGB tasks. In Sec. 5, we describe and analyze the user study conducted to validate our method from a human guidance perspective. Finally, we discuss the limitations of the proposed method and future works (Sec. 6).

1. KANs and Color-Matching Problem

This section details the mathematical reasoning for using a single KAN layer in our color-matching task.

An abstract camera Image Signal Processing (ISP) pipeline can be represented as linear and non-linear transformations [4, 11, 13, 19, 20]. Specifically, the image formation process can be represented as:

$$I_{\text{rgb}} = T(I_{\text{raw}}\mathbf{L}), \tag{1}$$

where I_{raw} is the input image in the camera raw space, I_{rgb} is the output in one of the display standard spaces (*e.g.*, sRGB), $\mathbf{L}$ represents the linear component of the ISP pipeline, and T implements the non-linear transformations (tone mapping, gamut mapping, image enhancement, etc.) [30].

For color matching between two images of the same scene processed through different ISPs (ISP_x and ISP_y), the color matching of two pixels $\mathbf{x}$ and $\mathbf{y}$ in the corresponding images, $X = ISP_x(I_{\text{raw}})$ and $Y = ISP_y(I_{\text{raw}})$, is:

$$\hat{\mathbf{y}} = T_y(T_x^{-1}(\mathbf{x})\mathbf{L}), \tag{2}$$

where $\mathbf{L}$ represents a linear transformation, and T_x and T_y are the non-linear color transformations of the respective ISPs. Note that T_x and T_y are applied element-wise; $\mathbf{x}$ and $\hat{\mathbf{y}}$ are row vectors.

Recent research in color matching [13] (Equation 7) shows that the non-linear transformations T_x and T_y can be represented by a single operation F, and the linear part $\mathbf{L}$ can be factored out. Thus, we can rewrite Equation 2 as:

$$\hat{\mathbf{y}} = F(\mathbf{x})\mathbf{L}, \tag{3}$$

where F is applied element-wise; $\mathbf{x}$ and $\hat{\mathbf{y}}$ are row vectors.

Accurate approximation of $F(\cdot)$ requires a plausible parametric space. Current state-of-the-art methods [11, 13] utilize polynomial approximations of Eq. 3. Although these methods may outperform deep CNNs/MLPs in color matching [13, 20], they can struggle to accurately represent color transformations in intricate cases [2, 13, 20]. To address this, we propose using a more complex model that incorporates separate functions $F_{ij}(\cdot)$ for each channel:

$$\hat{\mathbf{y}} = F(\mathbf{x})\mathbf{L} \rightarrow \hat{y}_j = \sum_{i=0}^{2} F_{ij}(x_i) \cdot l_{ij}, \tag{4}$$

where x_i, $\hat{y}_j$ are input and output color components, l_{ij} are elements of $\mathbf{L}$; $i, j = 0..2$.

To enable the use of KANs for color-matching, we propose parameterizing the non-linear components, $F_j(\cdot)$, with a B-spline approximation of Eq. 4:

$$\hat{y}_j = \sum_{i=0}^{2} \left(\sum_{m=0}^{G+k-1} c_{ijm} B_{ijm}(x_i) \right) \cdot l_{ij}, \tag{5}$$

where x_i, $\hat{y}_j$ are input and output color components, l_{ij} are elements of $\mathbf{L}$, c_{ijm} are spline coefficients and $B_{ijm}(\cdot)$ are B-spline basis functions of order k and grid size G; $i, j = 0..2$.

This B-spline approximation can be directly implemented with a *single KAN layer* [23], offering several advantages. KAN is comparable to MLP and easily integrates with neural networks. Unlike MLPs (based on the Universal Approximation Theorem), KANs use the Kolmogorov-Arnold Representation Theorem [18], with learnable activation functions on edges instead of weights, and the summation of the resultant learned function's output is performed at the nodes. Additionally, the KAN layer provides smooth spline approximation, crucial for accurate color matching [23, 29]. In our implementation, we use a single KAN layer [23] which uses a residual connection and expresses the activation function as a combination of $silu(x)$ and a B-spline, with 3 inputs and outputs, spline order $k = 3$, and grid size $G = 5$, represented as:

$$\hat{y}_j = \sum_{i=0}^{2} \left(u_{ij} silu(x_i) + \underbrace{v_{ij}}_{l_{ij}} \underbrace{\sum_{m=0}^{G+k-1} c_{ijm} B_{ijm}(x_i)}_{F_{ij}(x_i)} \right) \tag{6}$$

where x_i, $\hat{y}_j$ are input and output components, $silu(\cdot)$ is the residual activation, $B_{ijm}(.)$ are B-spline basis functions, u_{ij}, v_{ij}, c_{ijm} are KAN layer parameters, $i, j = 0..2$, and G and k are grid size and spline order. Note that unlike [23], our KAN layer has no trainable parameters (u_{ij}, v_{ij}, c_{ijm}); the hypernetwork (generator network) dynamically provides them for each pixel.

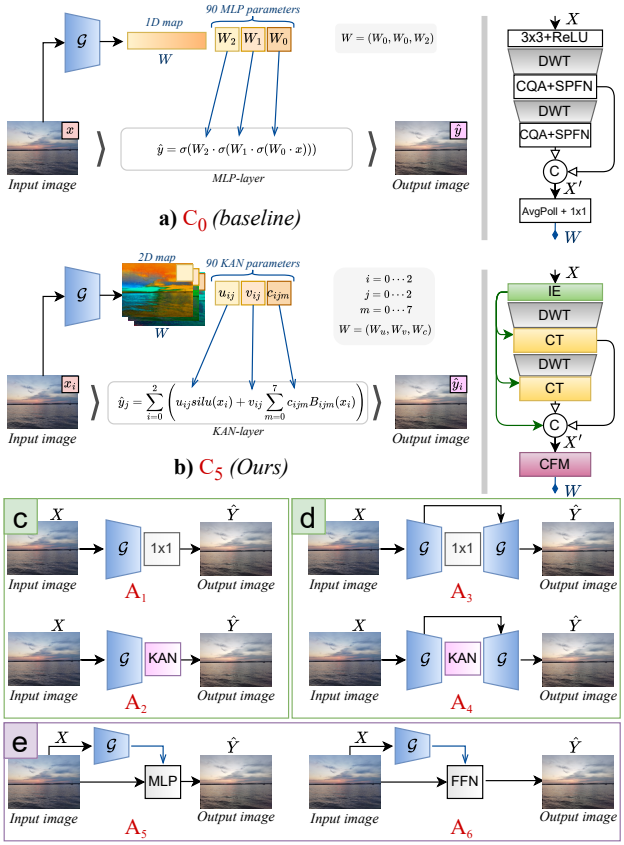

Figure 1. Visualization of various ablation studies architectures: (a) and (b) represent the baseline (C_0) and our approach (C_1); c) the generator with (A_2) and without (A_1) fixed learnable KAN layer; d) U-Net-like architecture with (A_3) and without (A_4) a KAN layer in the bottleneck; e) hyper-networks with MLP (A_5) and FFN [7] (A_6).

Table 1. *Ablation results* of the impact of the compression operation on vectors of attention on the proposed dataset.

Modification	PSNR	SSIM	ΔE	GFLOPs
w/o compression	25.92	0.89	4.57	7.99
w compression	25.94	0.89	4.51	2.94

2. Ablation Studies

In the main paper, we presented several ablation studies conducted to validate the decisions made in the proposed design of cmKAN (Fig. 1). Here, we present an additional ablation experiment performed to validate the operations of the Multi-Scale Color Attention (MCA) block and general architecture approach.

Table 1 demonstrates the impact of spatial-wise compression, on the attention vectors (Q, K, and A) of the MCA block. This ablation study demonstrates that incorporating compression into these vectors improves performance across all quantitative metrics with significantly lower computational complexity (2.94 vs 7.99) due to reduced matrix dimension.

Table 2. Ablation study on the impact of various architectures and the KAN layer. Results of sRGB-to-sRGB mapping on our dataset.

Method	PSNR	SSIM	ΔE	#Params
	a) Embedded KAN			
A_1 Generator	23.81	0.78	5.79	71.3K
A_2 Generator+KAN	24.67	0.84	5.11	76.8K
A_3 U-Net	24.33	0.81	5.30	117.7K
A_4 U-Net+KAN	24.89	0.85	5.09	107.7K
	b) Hyper Networks			
A_5 Generator+MLP	24.08	0.78	5.69	115.9K
A_6 Generator+FFN [7]	24.22	0.79	5.41	82.1K
Ours	**25.94**	**0.89**	**4.51**	**76.4K**

To validate the effectiveness of our KAN hyper-network approach, we compared our method against various architecture types (Table 2). *Direct RGB Prediction* (A_1): We modified the generator to directly predict a 3-channel RGB vector instead of emitting the KAN 90 parameters. *Generator with Fixed Learnable KAN Layer* (A_2): We applied a learnable KAN layer after the generator, without parameter substitution. *Base U-Net Architecture* (A_3): U-Net with Color Transformer blocks and DWT/IWT for down- and upsampling in the encoder and decoders. *U-Net with KAN Integration* (A_4): We replaced the U-Net bottleneck with KAN, using a 90-dimensional input to predict RGB values. Finally, we evaluate the impact of using an MLP instead of KAN by replacing KAN with two alternative configurations: a large MLP with five layers and dimensions (3, 12, 24, 12, 3) (A_5), and an FFN [7] (A_6). The results in Table 2 show that while KAN layers (A_2, A_4) improve results, our approach yields superior performance, because the generator provides parameters for the KAN layer, ensuring their smoothness and increasing the robustness of the KAN layer application to noisy data [23, 29].

Table 3. Ablation study results for color matching on our dataset.

Config	KAN	2D Map	IE	CT	CFM	PSNR	SSIM	ΔE
C_5 **(ours)**	✓	✓	✓	✓	✓	25.94	0.89	4.51
C_6	✗	✓	✓	✓	✓	24.02	0.78	5.57
C_7	✓	✗	✓	✓	✓	25.34	0.84	4.81
C_8	✓	✓	✗	✓	✓	25.63	0.87	4.58
C_9	✓	✓	✓	✗	✓	25.78	0.87	4.59
C_{10}	✓	✓	✓	✓	✗	25.59	0.87	4.58

Additionally, in Table 3, we present the ablation study, removing one module at a time from the complete architecture instead of adding them. This subtractive analysis consistently demonstrates the integral role of each compo-

nent, as the full model (C_5) achieves the highest performance across all metrics. The most substantial impact is seen when the KAN is replaced (C_6), causing a significant performance drop of 1.92 dB in PSNR, which underscores its critical importance. The second most crucial component is the use of 2D spatial maps; reverting to 1D global vectors (C_7) results in a noticeable 0.6 dB decrease in PSNR. The removal of the Illumination Estimator (C_8), Color Transformer (C_9), and Color Feature Modulator (C_{10}) individually leads to smaller but still consistent performance reductions. These findings reaffirm the conclusions from the additive ablation in the main paper, highlighting that while all components contribute positively, the KAN and the 2D spatial map generator are the most essential elements for effective color transformation.

3. Additional Details of Our Dataset

In the main paper, we presented our large-scale dataset captured using a Huawei P40 Pro phone. This device was specifically chosen because it features two distinct cameras with different sensor types: Quad-Bayer RGGB sensor (Sony IMX700) and RYYB sensor (Sony IMX608). These differences in sensor types result in varying image processing algorithms, with the RGGB and RYYB sensors requiring distinct demosaic methods, white balance (WB) corrections, and color calibration. The RYYB sensor, being more sensitive to light and affected by color lens shading (caused by small focal length and wide lens angle), also demands different tone-mapping techniques. Furthermore, the color gamut and saturation levels differ significantly between RGB and RYB, especially under low-light conditions. These variations in sensor behavior introduce a substantial domain gap between images captured by the two cameras, which is ideal for evaluating color-matching methods. The dataset thus simulates the real-world challenge of color matching between cameras.

Our dataset, spanning four years of data collection, includes images captured at over 20 locations across four countries. The images are well-aligned with minimal matching errors (see Fig. 2), and we provide detailed annotations for various scenes (see Table 4). Additionally, our dataset includes keypoint clouds generated using the SURF detector [3] for each image pair, along with binary matching masks to exclude misaligned regions. While our proposed method performs well even with unaligned data, this additional information is valuable for future research. To ensure ease of access and usability, we also offer a command-line interface (CLI) toolkit for processing, preparing, and cropping the dataset into smaller sections, making it more convenient for use in a variety of tasks. Our dataset and the CLI toolkit will be made available upon acceptance.

Table 4. Scene and image classes in our dataset.

Scene class	Count	Image class	Count
Indoor	114	City	324
Outdoor/spring	109	Countryside	164
Outdoor/summer	558	Forest	49
Outdoor/autumn	175	Mountains	166
Outdoor/winter	97	Seaside	161
Low-light	107	Sunset	92
Total	1260	Difficult illumination	21

Table 5. Results for *unsupervised* raw-to-raw mapping using NUS dataset [8]. The best results are highlighted in yellow.

Method	PSNR	SSIM	ΔE	PSNR	SSIM	ΔE
	Canon-to-Nikon			Nikon-to-Canon		
SSRM [1]	32.36	0.93	6.21	30.81	0.93	5.95
UVCGANv2 [27]	37.11	0.96	4.34	37.29	0.96	4.28
RawFormer [25]	41.89	0.98	2.04	41.37	0.98	2.53
cmKAN	41.93	0.98	1.61	41.62	0.98	1.54

4. Additional Results

4.1. Raw-to-Raw Mapping

In this section, we provide additional qualitative and quantitative results for *unsupervised* raw-to-raw mapping. Using the NUS dataset [8], we performed mapping tasks between Nikon D5200 and Canon EOS 600D DSLR cameras, following the evaluation protocol described in [1, 25]. These results are summarized in Table 5. As shown, our method achieves state-of-the-art performance across all metrics. Our method improves visual fidelity while preserving structural consistency, as confirmed by the superior SSIM and ΔE values. In addition, Fig. 3 shows visual comparisons.

4.2. Raw-to-sRGB Mapping

In addition to the reported results in the main paper, we evaluate our method on the *supervised* raw-to-sRGB mapping task using the ISPIW dataset [26]. This dataset comprises raw images captured with a Huawei Mate 30 Pro smartphone and their corresponding sRGB reference images taken with a Canon 5D Mark IV camera. For our experiments, we utilize all 192 full-resolution images available, splitting them into 90% for training and 10% for testing. This setup yields 5,152 patch pairs for training and 572 for testing. The results are presented in Table 6. Our method demonstrates superior performance, achieving state-of-the-art results across all evaluated metrics. In Fig. 4, we show qualitative results on the ZRR [15] and the ISPIW [26] datasets.

4.3. sRGB-to-sRGB Mapping

To further evaluate our method, we provide additional results for *supervised* sRGB-to-sRGB mapping using the

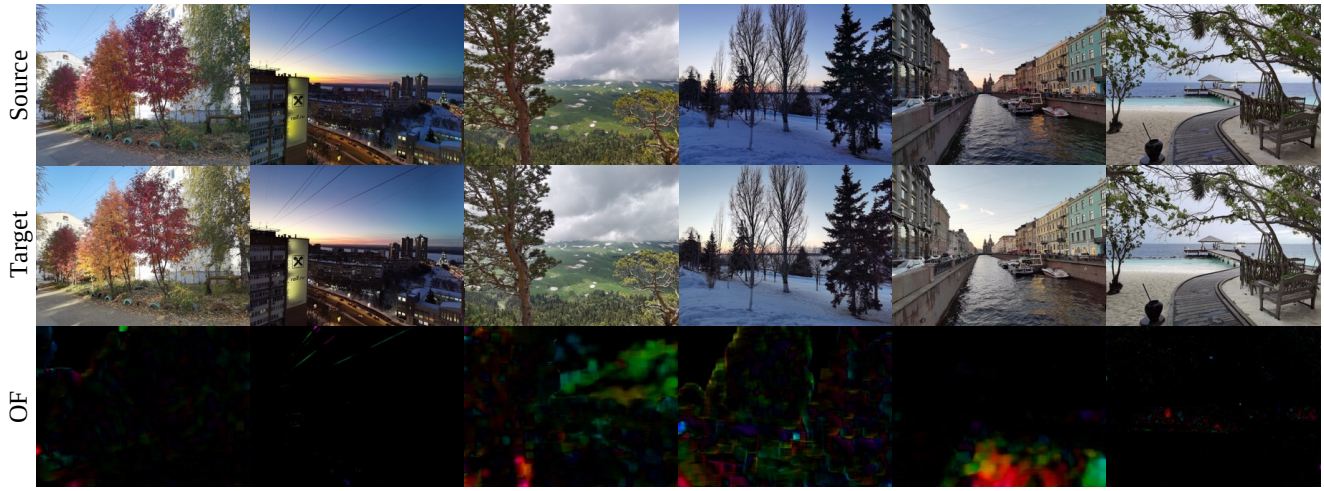

Figure 2. Example images from our dataset alongside their corresponding optical flow (OF) maps. The dataset spans diverse locations and lighting conditions, featuring well-aligned image pairs with minimal matching errors.

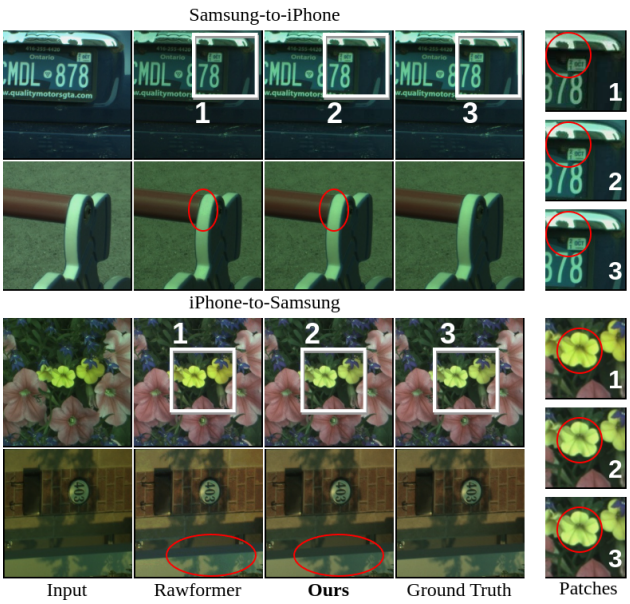

Figure 3. Qualitative results of raw translation on the Raw-to-Raw dataset [1]. Shown are images captured by Samsung S9 and iPhone X. We show the input raw patch from the corresponding camera and the corresponding ground-truth raw patch from the other camera, along with the results by RawFormer [25].

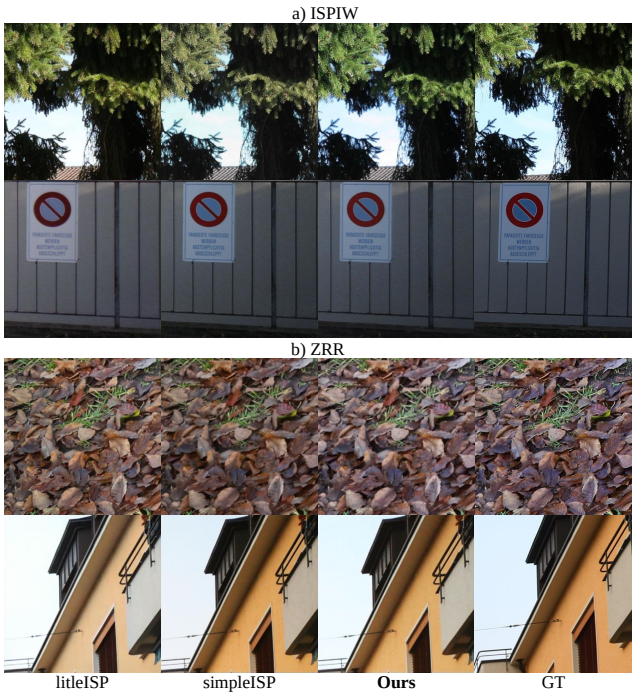

Figure 4. Qualitative comparison of raw-to-sRGB rendering on a) the ISPIW [26] and b) the Zurich raw-to-sRGB [15] datasets. Our cmKAN method achieves the most accurate color mapping than LiteISP [31] and SimpleISP [9]. Best viewed in the electronic version.

PPR10K dataset [22] and *paired-based optimization* on the R.G. Rodriguez *et al.* [12] dataset.

The PPR10K dataset consists of 10,000 raw images rendered by three expert photographers (Experts A, B, and C). In this experiment, we used 7,000 images for training and 3,000 for testing. Table 8 summarizes our results, where our method (cmKAN) achieves state-of-the-art performance across all metrics, significantly outperforming existing approaches.

The R.G. Rodriguez *et al.* [12] dataset contains 35 images captured using two camera models (Nikon D3100 and

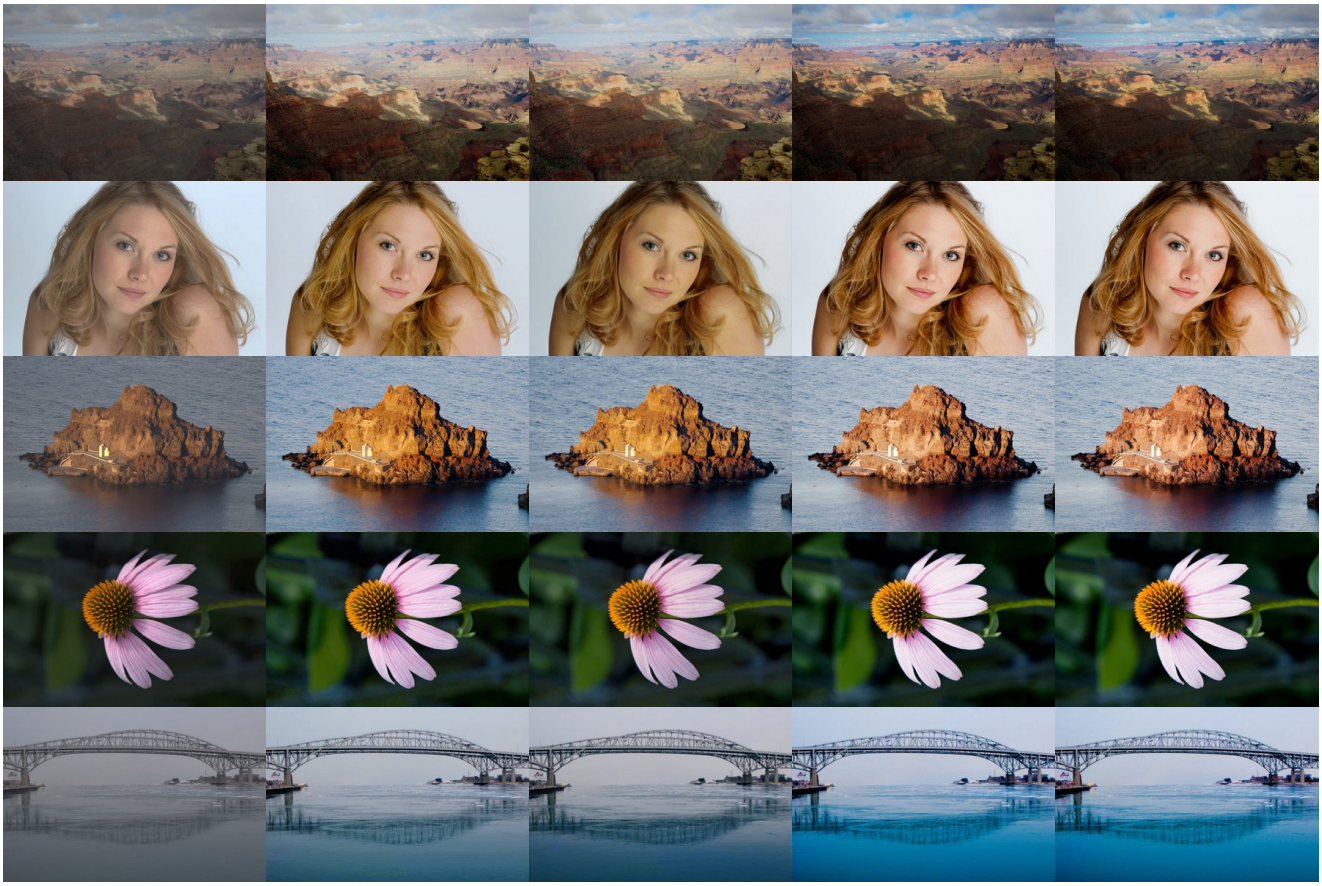

Figure 5. Qualitative results of sRGB-to-sRGB translation on the Adobe FiveK dataset [6]. Our method demonstrates SoTA results compared to other methods (SepLUT [28] and SIRLUT [21]).

Table 6. Results for *supervised* raw-to-sRGB mapping on the ISPIW raw-to-sRGB dataset [26].

Method	PSNR	SSIM	ΔE
MW-ISP [14]	21.90	0.81	7.03
LiteISP [31]	22.14	0.81	6.31
MicroISP [16]	20.70	0.77	6.92
SimpleISP [9]	23.67	0.82	5.91
cmKAN	24.22	0.83	5.29

Table 7. Additional results on R.G. Rodriguez *et al.* [12] dataset. We compare the results of *paired-based inference*.

Method	PSNR	SSIM	ΔE
R.G. Rodriguez *et al.* [12]	26.91	0.84	3.26
cmKAN-Light	28.21	0.86	3.07

Canon EOS80D). Due to the limited dataset size, we performed *paired-based optimization* for evaluation. Our results, detailed in Table 7, show that our lightweight model variant (cmKAN-Light) achieves the best scores in PSNR, SSIM, and ΔE, further validating the versatility and adaptability of our approach.

Additional qualitative results of sRGB-to-sRGB mapping are shown in Fig. 5, 6 and 7, showcasing our method's ability to handle diverse sRGB-to-sRGB mapping tasks effectively.

Table 8. Results of *supervised* sRGB-to-sRGB mapping on the PPR10K dataset [22]. We report training results using experts A, B, and C as ground-truth targets.

Method	PSNR / ΔE	PSNR / ΔE	PSNR / ΔE
	Expert A	Expert B	Expert C
SepLUT [28]	26.28 / 6.59	25.23 / 7.49	25.59 / 7.51
LYT-Net [5]	26.10 / 7.03	23.93 / 9.21	23.93 / 9.21
SIRLUT [21]	28.31 / 5.65	27.67 / 5.89	27.79 / 6.13
cmKAN	32.49 / 2.21	32.01 / 2.35	32.27 / 2.77

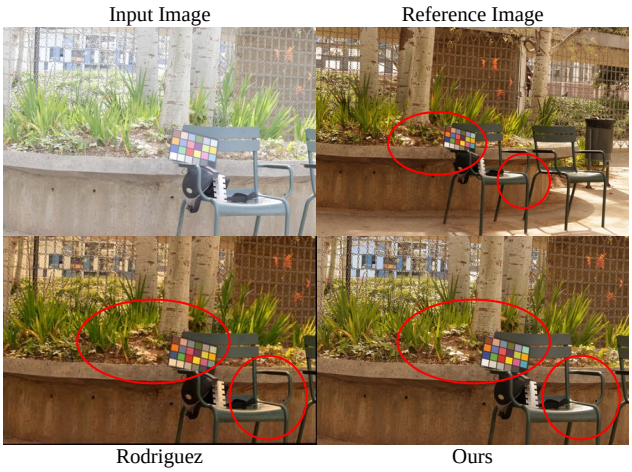

Figure 6. Qualitative comparison of sRGB-to-sRGB rendering on R.G. Rodriguez *et al.* [12] dataset for 19th image. Our cmKAN-Light method achieves the most accurate color mapping compared with the original R.G. Rodriguez *et al.* [12] method. Best viewed in the electronic version.

5. User Study Experiment

In addition to the quantitative and qualitative comparisons presented in the main paper and this supplementary material, we conducted a user study to validate our method from a human-guidance perspective.

We conducted the user study using the 'Toloka' crowdsourcing platform. Thirty scenes from a test sample of the Adobe FiveK dataset [6] were processed with four different methods: 1) SepLUT [28], 2) SIRLUT [21], 3) R.G. Rodriguez *et al.* [12], and 4) ours. Each scene generates 6 combinations of style pairs, resulting in $6 \times 30 = 180$ pairs for style comparison. It is important to note that the comparisons also included "filtering" pairs of identical images to exclude unscrupulous participants. On each Toloka webpage, five random image pairs were displayed, with the reference image placed in the center and the other images on the sides. For each, participants were asked to answer the question: 'Which image was more similar to the center image?' The images were displayed on a 50% gray background, as color comparisons should be conducted on a neutral background to eliminate bias [24]. Additionally, we ensured that the space between the two images contained no other information or controls, to prevent distractions for the participants. To facilitate the comparison process, a gap was maintained between the images. For each pair, participants were presented with three response options; an example of the interface is shown in Fig. 9):

- The right image is more like the center image
- The left image is more similar to the center image
- Both images are the same

A total of 431 participants took part in the experiments. Votes were not counted if participants marked images as similar. Additionally, if a participant failed the 'filter' pair, the entire set of votes from that participant was discarded, and they were banned from the study.

The evaluation methodology for obtaining mean opinion scores, shown in Fig. 8, uses a modified Bradley-Terry model [10] to rank image versions based on pairwise comparisons from participant ratings, simplifying the process by assuming negligible observer dependence and averaging favorable votes. *Our non-reference method outperforms existing methods by a factor of two*, which helps justify the significance of the results. Additionally, our method outperforms paired-based solutions [12] by$\times$1.5. The results of this study demonstrate that our proposed cmKAN significantly outperforms other methods in color matching.

6. Limitations and Future Work

While our *cmKAN* framework demonstrates significant versatility and high performance, its current implementation has some limitations that define clear directions for future research.

The first limitation lies in the performance of purely unsupervised training for tasks with a profound domain gap, most notably in *raw-to-sRGB* color matching. This transformation is not a simple color mapping but an approximation of a camera's entire Image Signal Processor (ISP) pipeline. In a supervised setting, a neural network can learn this end-to-end transformation from paired raw and sRGB images. However, in an unsupervised scenario, the model lacks the explicit guidance required to learn these intricate and ordered operations correctly. Standard unsupervised loss functions, such as cycle-consistency, often prove insufficient. The problem is severely ill-posed: there are infinite mappings that can transform a raw image into a visually plausible sRGB image while failing to match the target camera's specific color science. The second limitation is that the current *cmKAN* architecture is designed for transformations between a single source and a single target. This one-to-one approach is effective but not scalable for real-world scenarios involving many different cameras or styles.

Our future work will focus on extending *cmKAN* into a unified and scalable multi-device framework. We plan to investigate architectural solutions like using multiple specialized output heads or integrating parameter-efficient fine-tuning techniques, such as LoRA, to adapt a single base model to various targets. Furthermore, we will explore more robust loss functions to improve performance for challenging unsupervised domain adaptation tasks.

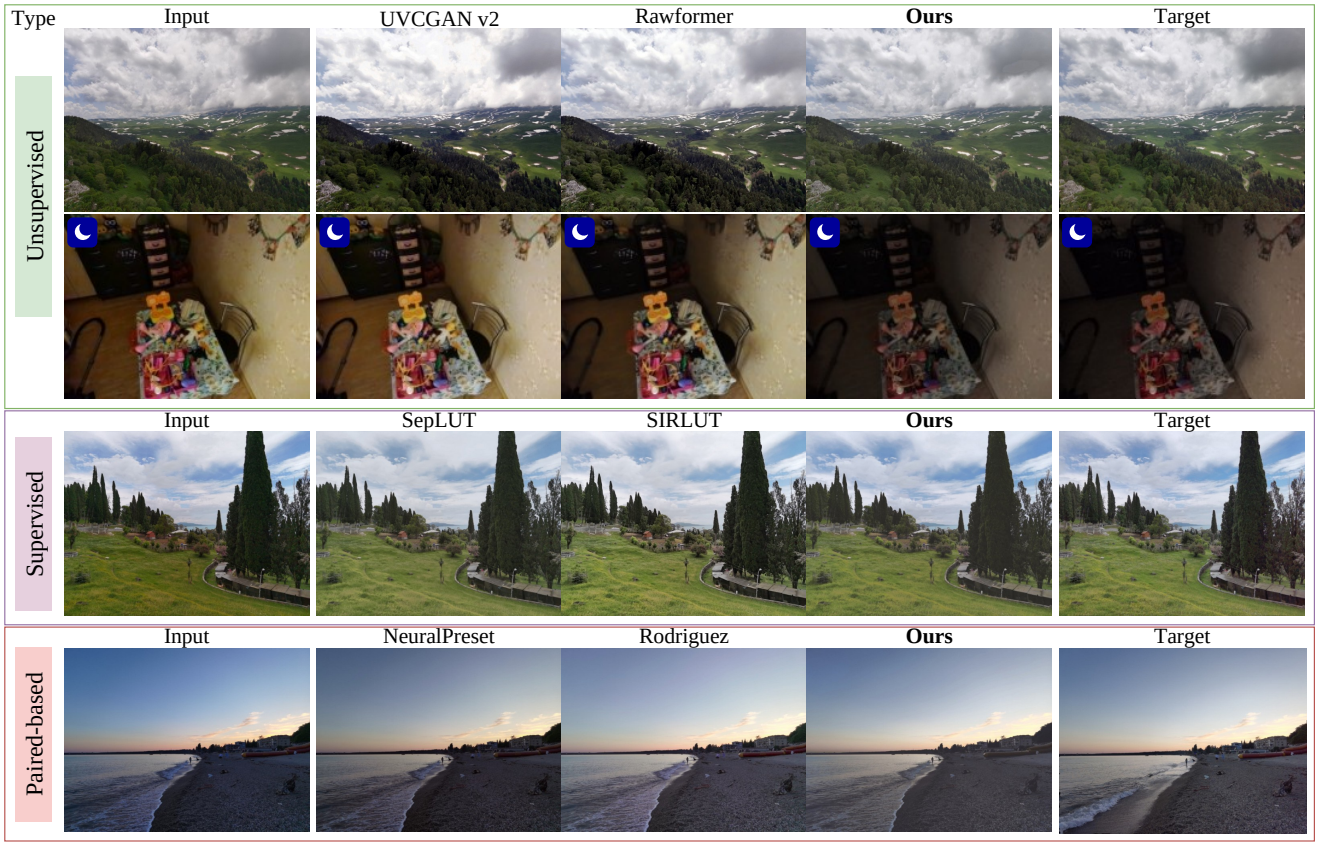

Figure 7. Additional results of sRGB-to-sRGB mapping on the proposed dataset. In (a), we show three scenarios using our method: unsupervised learning (green), supervised learning (purple), and paired-based optimization (red). Our method demonstrates superior color matching compared to other methods (UVCGAN v2 [27], Rawformer [25], SepLUT [28], SIRLUT [21], NeuralPreset [17], and Rodriguez *et al*.'s method [12]), even under low-light conditions (marked with a moon icon).

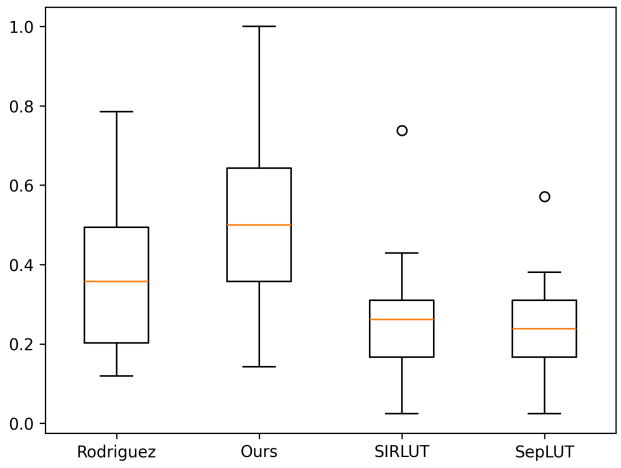

Figure 8. Normalized win count aggregated on all images.

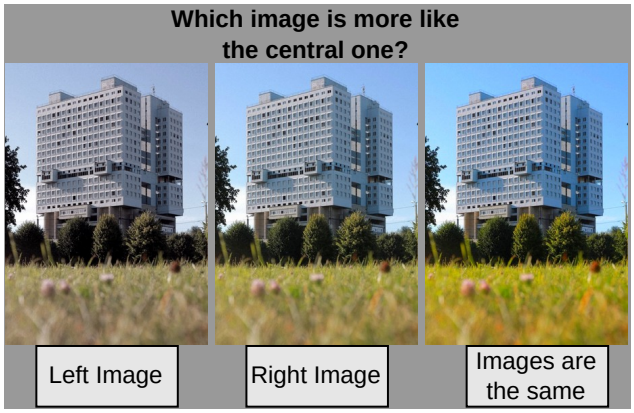

Figure 9. Illustration of a voting interface in the "Toloka" service.

References

[1] Mahmoud Afifi and Abdullah Abuolaim. Semi-supervised raw-to-raw mapping. In *BMVC*, 2021. 3, 4

[2] Mahmoud Afifi, Brian Price, Scott Cohen, and Michael S Brown. When color constancy goes wrong: Correcting improperly white-balanced images. In *CVPR*, 2019. 1

[3] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF:

Speeded up robust features. In *ECCV*, 2006. 3

[4] Simone Bianco, Arcangelo Bruna, Filippo Naccari, and Raimondo Schettini. Color space transformations for digital photography exploiting information about the illuminant estimation process. *J. Opt. Soc. Am. A*, 29(3):374–384, 2012. 1

[5] Alexandru Brateanu, Raul Balmez, Adrian Avram, and CC Orhei. LYT-NET: Lightweight YUV transformer-based network for low-light image enhancement. *arXiv preprint arXiv:2401.15204*, 2024. 5

[6] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *CVPR*, 2011. 5, 6

[7] Yuanhao Cai, Jing Lin, Zudi Lin, Haoqian Wang, Yulun Zhang, Hanspeter Pfister, Radu Timofte, and Luc Van Gool. MST++: Multi-stage spectral-wise transformer for efficient spectral reconstruction. In *CVPR*, 2022. 2

[8] Dongliang Cheng, Dilip K Prasad, and Michael S Brown. Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution. *JOSA A*, 31(5):1049–1058, 2014. 3

[9] Omar Elezabi, Marcos V Conde, and Radu Timofte. Simple image signal processing using global context guidance. In *ICIP*, 2024. 4, 5

[10] Egor Ershov, Artyom Panshin, Ivan Ermakov, Nikola Banic, Alex Savchik, and Simone Bianco. Reliability and stability of mean opinion score for image aesthetic quality assessment obtained through crowdsourcing. *Proceedings Copyright*, 365:372, 2024. 6

[11] Graham Finlayson, Han Gong, and Robert B Fisher. Color homography: theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(1):20–33, 2017. 1

[12] Raquel Gil Rodríguez, Javier Vazquez-Corral, and Marcelo Bertalmío. Color matching images with unknown non-linear encodings. *IEEE Transactions on Image Processing*, 29: 4435–4444, 2020. 4, 5, 6, 7

[13] Raquel Gil Rodríguez, Javier Vazquez-Corral, Marcelo Bertalmío, and Graham D. Finlayson. Color matching in the wild. *Pattern Recognition*, 154:110575, 2024. 1

[14] Andrey Ignatov, Radu Timofte, Zhilu Zhang, Ming Liu, Haolin Wang, Wangmeng Zuo, Jiawei Zhang, Ruimao Zhang, Zhanglin Peng, Sijie Ren, et al. AIM 2020 challenge on learned image signal processing pipeline. In *ECCVW*, 2020. 5

[15] Andrey Ignatov, Luc Van Gool, and Radu Timofte. Replacing mobile camera ISP with a single deep learning model. In *CVPRW*, 2020. 3, 4

[16] Andrey Ignatov, Anastasia Sycheva, Radu Timofte, Yu Tseng, Yu-Syuan Xu, Po-Hsiang Yu, Cheng-Ming Chiang, Hsien-Kai Kuo, Min-Hung Chen, Chia-Ming Cheng, et al. MicroISP: processing 32mp photos on mobile devices with deep learning. In *ECCV*, 2022. 5

[17] Zhanghan Ke, Yuhao Liu, Lei Zhu, Nanxuan Zhao, and Rynson WH Lau. Neural preset for color style transfer. In *CVPR*, 2023. 7

[18] A. K. Kolmogorov. On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition. *Doklady Akademii Nauk SSSR*, 114:369–373, 1957. 1

[19] Abdullah Kucuk, Graham Finlayson, Rafal Mantiuk, and Maliha Ashraf. Comparison of regression methods and neural networks for colour corrections. In *London Imaging Meeting*, pages 74–79, 2022. 1

[20] Abdullah Kucuk, Graham D. Finlayson, Rafal Mantiuk, and Maliha Ashraf. Performance comparison of classical methods and neural networks for colour correction. *Journal of Imaging*, 9(10), 2023. 1

[21] Kaijiang Li, Hao Li, Haining Li, Peisen Wang, Chunyi Guo, and Wenfeng Jiang. SIRLUT: Simulated infrared fusion guided image-adaptive 3D lookup tables for lightweight image enhancement. In *ACM MM*, 2024. 5, 6, 7

[22] Jie Liang, Hui Zeng, Miaomiao Cui, Xuansong Xie, and Lei Zhang. PPR10K: A large-scale portrait photo retouching dataset with human-region mask and group-level consistency. In *CVPR*, 2021. 4, 5

[23] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y Hou, and Max Tegmark. KAN: Kolmogorov-arnold networks. *arXiv preprint arXiv:2404.19756*, 2024. 1, 2

[24] Rafał K Mantiuk, Anna Tomaszewska, and Radosław Mantiuk. Comparison of Four Subjective Methods for Image Quality Assessment. In *Computer graphics forum*, pages 2478–2491. Wiley Online Library, 2012. 6

[25] Georgy Perevozchikov, Nancy Mehta, Mahmoud Afifi, and Radu Timofte. Rawformer: Unpaired raw-to-raw translation for learnable camera ISPs. In *ECCV*, 2024. 3, 4, 7

[26] Ardhendu Shekhar Tripathi, Martin Danelljan, Samarth Shukla, Radu Timofte, and Luc Van Gool. Transform your smartphone into a dslr camera: Learning the isp in the wild. In *European Conference on Computer Vision*, pages 625–641. Springer, 2022. 3, 4, 5

[27] Dmitrii Torbunov, Yi Huang, Haiwang Yu, Jin Huang, Shinjae Yoo, Meifeng Lin, Brett Viren, and Yihui Ren. UVCGAN v2: An improved cycle-consistent GAN for unpaired image-to-image translation. *arXiv preprint arXiv:2303.16280*, 2023. 3, 7

[28] Canqian Yang, Meiguang Jin, Yi Xu, Rui Zhang, Ying Chen, and Huaida Liu. SepLUT: Separable image-adaptive lookup tables for real-time image enhancement. In *ECCV*, 2022. 5, 6, 7

[29] Chen Zeng, Jiahui Wang, Haoran Shen, and Qiao Wang. Kan versus mlp on irregular or noisy functions. *arXiv preprint arXiv:2408.07906*, 2024. 1, 2

[30] Feng Zhang, Ming Tian, Zhiqiang Li, Bin Xu, Qingbo Lu, Changxin Gao, and Nong Sang. Lookup table meets local laplacian filter: Pyramid reconstruction network for tone mapping. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. 1

[31] Zhilu Zhang, Haolin Wang, Ming Liu, Ruohao Wang, Jiawei Zhang, and Wangmeng Zuo. Learning raw-to-sRGB mappings with inaccurately aligned supervision. In *ICCV*, 2021. 4, 5