Deterministic Neural Illumination Mapping for Efficient Auto-White Balance Correction

Furkan Kınlı[1] Doğa Yılmaz[2] Barış Özcan[3] Furkan Kıraç[4]
Vision and Graphics Lab, Özyeğin University, Türkiye
{furkan.kinli[1], furkan.kirac[4]}@ozyegin.edu.tr,
{doga.yilmaz.11481[2], baris.ozcan.10097[3]}@ozu.edu.tr

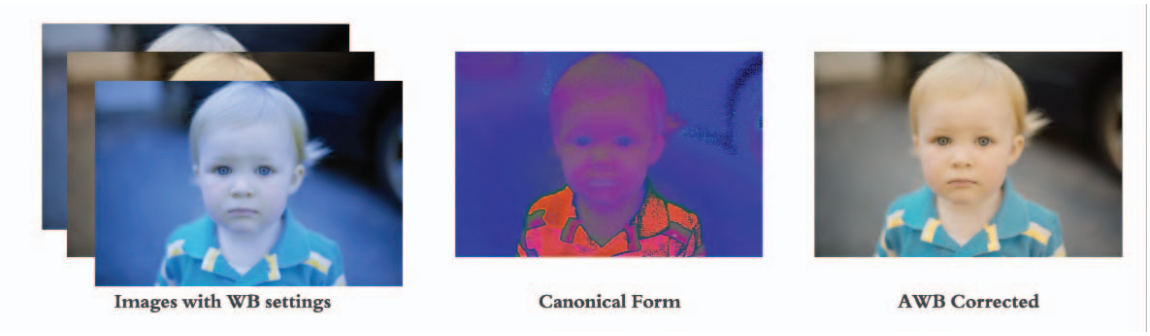

Figure 1: Learning deterministic illumination color mappings from images with different WB settings for both canonical illumination form and AWB corrected version. Pixel intensities in the canonical form are amplified for better visualization.

Abstract

Auto-white balance (AWB) correction is a critical operation in image signal processors for accurate and consistent color correction across various illumination scenarios. This paper presents a novel and efficient AWB correction method that achieves at least 35 times faster processing with equivalent or superior performance on high-resolution images for the current state-of-the-art methods. Inspired by deterministic color style transfer, our approach introduces deterministic illumination color mapping, leveraging learnable projection matrices for both canonical illumination form and AWB-corrected output. It involves feeding high-resolution images and corresponding latent representations into a mapping module to derive a canonical form, followed by another mapping module that maps the pixel values to those for the corrected version. This strategy is designed as resolution-agnostic and also enables seamless integration of any pre-trained AWB network as the backbone. Experimental results confirm the effectiveness of our approach, revealing significant performance improvements and reduced time complexity compared to state-of-the-art methods. Our method provides an efficient deep learning-based AWB correction solution, promising real-time, high-quality color correction for digital imaging applications. Source code is available at https://github.com/birdortyedi/DeNIM/

1. Introduction

In the realm of digital imaging, auto-white balance (AWB) correction is one of the most critical operations in image signal processors (ISPs). The colors presented in the final sRGB image should be somehow aligned with the colors perceived by the human eye. This operation mainly aims to ensure accurate and consistent color correction across a variety of illumination scenarios. Due to the effect of differing light sources in real-world scenarios, which possess continuous range of color temperatures, AWB correction task still remains challenging. Recent studies on AWB correction generally introduce a method to model leading illumination settings and undesired color casts in the scene, and then subsequently adjust the color balance.

A number of AWB correction methods have been introduced, which employ various strategies (*e.g.*, low-level statistical methods, gamut-based methods, and learning-based methods). Earlier studies [12, 11, 18, 42, 24, 14, 29, 39, 38] benefit from low-level statistics of images or patches to infer the illumination, and employ a simple diagonal-based

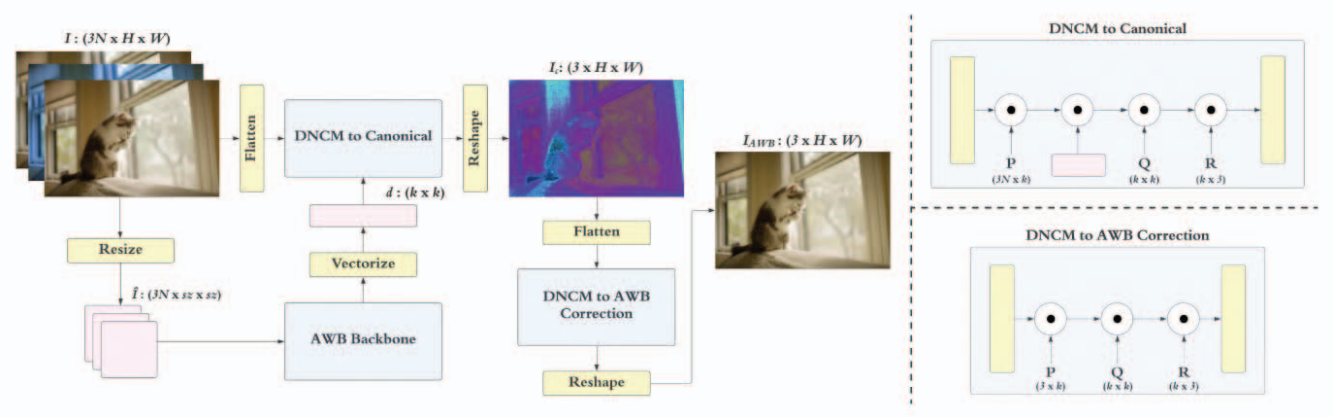

Figure 2: Overall design of our proposed illumination mapping strategy. We first reduce the resolution of the input images to a compatible size with the AWB correction backbone (*i.e.*, Mixed WB [4], Style WB [32]). Then, high-resolution input images and the latent representations of low-resolution versions, extracted by AWB correction backbone, are fed into a deterministic color mapping module (DNCM) [30] to obtain a canonical form. Another DNCM module (without fusion capability) takes the canonical form as input and learns to map the pixel values to the ones for AWB corrected version. This strategy ensures that the AWB correction model is resolution-agnostic.

correction matrix [23] of predicted illumination to rectify the color casts in the scene. In addition to low-level statistical methods, gamut-based methods [19, 16, 17, 22] mainly introduce models that aims to learn mappings from the images captured under unknown lighting conditions to the reference colors captured under known lighting conditions. Learning-based methods [9, 20, 10, 21, 26] have become more popular when compared to their ancestors, due to their better capability of representing the illumination in real-world scenarios.

With the advancements in computational photography, deep learning-based methods [36, 40, 8, 27, 43, 34, 1, 4, 37, 32, 33] have demonstrated an outstanding performance edge over all previous AWB correction strategies. However, the high computational requirements and significant power demands of these approaches restrict their direct integration within a camera pipeline. Especially, the recent approaches suffer from the computational complexity mostly leading to better performance without considering the time efficiency and their practical usage. Addressing this issue, we propose a novel, deep learning-based AWB correction method, which makes the current state-of-the-art methods at least 35 times faster, while delivering equivalent or better performance on high-resolution images.

The main contributions of this study can be summarized as follows:

- We propose a novel and efficient strategy for AWB correction, which learns deterministic color mappings for both canonical illumination and AWB-corrected forms with the help of learnable projection matrices.
- Our design allows the input to be resolution-agnostic and any pre-trained AWB network can be integrated into this design as the backbone network.
- We demonstrate that employing deterministic illumination color mapping for AWB correction yields a substantial improvement in the performance of existing state-of-the-art methods, while significantly reducing the time complexity, achieving a speedup of at least 35 times faster.

2. Methodology

Given a set of high-resolution images with different white balance (WB) settings I, our proposed strategy learns to achieve a deterministic illumination color mapping for efficient AWB correction. Prior works [4, 32] focus on learning the weighting maps for all different WB settings in low-resolution space. Then, they render the AWB-corrected version in high-resolution space by linearly combining images with different WB settings and their corresponding weighting maps. Although this approach can produce quite well outputs, it essentially requires multi-scale inference and smoothing after resizing the weighting maps to the original resolution to significantly improve the results. However, these post-processing steps make this approach challenging to use in practical scenarios.

Inspired by deterministic color style transfer [30], we developed an idea of deterministic illumination color mapping for AWB correction. The overall design of our proposed illumination mapping strategy is shown in Figure 2. First, we reduce the resolution of the input images I to make them compatible with the architectures of prior works (*i.e.*, 256px). By using only the encoder part of one of these ar-

chitectures, we feed low-resolution images $\hat{I}$ into the encoder to extract rich information from different WB settings. To obtain the latent representations, we use 1×1 convolutional layer followed by GeLU activation [25] to vectorize the feature maps. This provides an image-adaptive color mapping matrix d for DNCM module [30] to generate a canonical form.

$$d^{(k \times k)} = V(E(\hat{\mathbf{I}})) \tag{1}$$

where E refers to the AWB encoder (*i.e.*, [4] or [32]), V stands for the vectorization operation by 1×1 convolutional layer and activation. Note that we use pre-trained weights for E and freeze its weights during our training.

For *DNCM to canonical* module, the first step involves unfolding high-resolution image I into a 2D matrix of dimensions ($HW \times 3N$), where N refers to the number of WB settings, H and W represent height and width, respectively. Each pixel in I is then transformed into a k-dimensional vector using a projection matrix P ($3N \times k$). k can be any number depending on the computational power, but we set it to 32 in our design. The extracted image-adaptive color mapping matrix d is multiplied with k-dimensional vector to inject the rich information into the projected space. Q ($k \times k$) and R ($k \times 3$) are the following learnable projection matrices to form the canonical form in this module. We can formulate this module, namely *DNCMc*, as follows

$$DNCMc(\mathbf{I}, d) = \mathbf{I}^{(HW \times 3)} \cdot \mathbf{P}^{(3 \times k)} \cdot d^{(k \times k)} \cdot \mathbf{Q}^{(k \times k)} \cdot \mathbf{R}^{(k \times 3)} \tag{2}$$

where $\cdot$ denotes the matrix multiplication.

Next, we feed the canonical form into *DNCM to AWB correction* module (*DNCMa*). It does not have any fusion capability but learns to directly map the pixel values in the canonical form to the correct ones for the AWB version. Each pixel in the canonical form I_c is transformed into a k-dimensional vector by a projection matrix P ($3 \times k$). By using a similar design to *DNCMc*, Q ($k \times k$) and R ($k \times 3$) are responsible for converting the embedded k-dimensional vector back to the RGB color space, which finally forms the output I_{AWB}. The formal definition of *DNCMa* can be seen in Equation 3.

$$DNCMa(\mathbf{I_c}) = \mathbf{I_c}^{(HW \times 3)} \cdot \mathbf{P}^{(3 \times k)} \cdot \mathbf{Q}^{(k \times k)} \cdot \mathbf{R}^{(k \times 3)} \tag{3}$$

Apart from the self-supervised learning mechanism for DNCM, followed in [30], the learning objective is to minimize the reconstruction error between the ground truth and the AWB corrected output, as shown in Equation 4.

$$\mathcal{L} = ||\mathbf{I_{GT}} - \mathbf{I_{AWB}}||_F^2 \tag{4}$$

where I_{GT} and I_{AWB} denote the ground truth image and the output. To keep the training process simple and tractable, we did not include the smoothing loss [4] or perceptual loss [28] in our final objective function.

Our design removes the decoder part that generates the weighting maps in the prior works, and instead, it directly computes the illumination color mapping with two distinct DNCM modules for the canonical form and AWB-corrected version. This design mitigates the need for further post-processing of the weighting maps, which leads to reducing the time complexity without compromising the performance. Moreover, due to the one-by-one pixel value mapping characteristic delivered by matrix multiplications, it gives AWB correction model the ability to be resolution-agnostic. Lastly, any AWB correction method can be easily plugged into this design for extracting rich information in low-resolution space from different WB settings, which makes our design also model-agnostic.

3. Experiments

3.1. Experimental Details

For our training, we have employed the RenderedWB dataset [4], which contains 65,000 sRGB images with pre-defined WB settings and corresponding white-balanced versions, captured by different cameras. Following the experimental setup in the prior works, we have two sets of pre-defined WB settings, which are {`t,f,d,c,s`} and {`t,d,s`}. The color temperatures used for pre-defined WB settings are as follows: Tungsten (`t`, 2850K), Fluorescent (`f`, 3800K), Daylight (`d`, 5500K), Cloudy (`c`, 6500K), and Shade (`s`, 7500K). We did not apply any data augmentation technique to the images during our training.

For all experiments, we freeze the weights of the AWB backbone and only trained DNCM modules in our proposed strategy from scratch. We set the size of the low-resolution space to 256. We used AdamW optimizer [35] ($\beta_1 = 0.9$, $\beta_2 = 0.999$) with batch size of 16. The learning rate is set to $1e-4$ and we did not employ any scheduling strategy. We did not apply any post-processing operations after obtaining the output.

3.2. Evaluation

Following the prior works [2, 4, 32], we evaluate the AWB correction quality in terms of the mean-squared error (MSE), mean angular error (MAE) and color difference (ΔE 2000). We report the mean, first (Q1), second (Q2), and third (Q3) quantile averages for all metrics.

For qualitative and quantitative evaluation scenarios, we have used three different evaluation sets: Cube+ [7] and MIT-Adobe FiveK [13], along with the night photography rendering set [41]. The Cube+ dataset consists of 1,707 single illumination color-calibrated images, captured with

(a) Mixed WB (b) Style WB (c) DeNIM + Mixed WB (d) DeNIM + Style WB

Figure 3: Comparison of the qualitative results of our efficient AWB correction method, namely *DeNIM*, with the prior works on the selected samples from MIT-Adobe FiveK dataset [13]. We compare our results with Mixed WB [4] and Style WB [32]. Image indices from top to bottom: 323, 606, 2431, 2808, 2838.

a Canon EOS 550D camera during various seasons. The MIT-Adobe FiveK dataset comprises 5,000 images captured by different DSLR cameras, with each image manually retouched by multiple experts to correct the white balance.

4. Results and Discussion

This section presents a detailed review of notable findings in our experiments. We primarily focus on three aspects while analyzing the results obtained in our experiments: visual quality, numeric evaluation, and efficiency. Qualitative analysis is conducted by comparing the results obtained by Mixed WB [4], Style WB [32] and our strategy built on top of both methods on MIT-Adobe FiveK dataset and night photography rendering set. Following the literature, the evaluation of performance using quantitative metrics, analysis of model complexity, and comparison of efficiencies are all conducted using the Cube+ dataset.

Qualitative analysis: To use the images in MIT-Adobe FiveK dataset for our experiments, we first render the linear raw DNG images with different WB settings (*e.g.,* Daylight, Tungsten, Shade) by using the method presented in [6]. Figure 3 demonstrates the qualitative comparison of our AWB correction results and the prior works' on selected samples from the dataset. The indices of selected samples in the dataset are as follows: 323, 606, 2431, 2808, and

Table 1: Benchmark on single-illuminant Cube+ dataset [7]. Following the prior works [4, 32], we reported the mean, first (**Q1**), second (**Q2**) and third (**Q3**) quantile of mean-squared error (**MSE**), mean angular error (**MAE**) and color difference (Δ**E 2000**) metrics. Different WB settings are denoted as {t, f, d, c, s}, which refers to tungsten, fluorescent, daylight, cloudy, and shade, respectively. p refers to the patch size. The top results are indicated with colored cells as, the best: **green**, the second: yellow, the third: red.

Method	MSE				MAE				ΔE 2000				Size
	Mean	Q1	Q2	Q3	Mean	Q1	Q2	Q3	Mean	Q1	Q2	Q3	
FC4 [27]	371.90	79.15	213.41	467.33	6.49°	3.34°	5.59°	8.59°	10.38	6.60	9.76	13.26	5.89 MB
Quasi-U CC [8]	292.18	15.57	55.41	261.58	6.12°	1.95°	3.88°	8.83°	7.25	2.89	5.21	10.37	622 MB
KNN WB [5]	194.98	27.43	57.08	118.21	4.12°	1.96°	3.17°	5.04°	5.68	3.22	4.61	6.70	21.8 MB
Interactive WB [3]	159.88	21.94	54.76	125.02	4.64°	2.12°	3.64°	5.98°	6.20	3.28	5.17	7.45	**38 KB**
Deep WB [2]	80.46	15.43	33.88	74.42	3.45°	1.87°	2.82°	4.26°	4.59	2.68	3.81	5.53	16.7 MB
MIMT [33]	-	-	-	-	2.52°	0.98°	1.38°	2.96°	2.88	1.94	2.42	2.87	-
Mixed WB [4]													
$p = 64$, WB={t, d, s}	168.38	8.97	19.87	105.22	4.20°	1.39°	2.18°	5.54°	5.03	2.07	3.12	7.19	5.09 MB
$p = 64$, WB={t, f, d, c, s}	161.80	9.01	19.33	90.81	4.05°	1.40°	2.12°	4.88°	4.89	2.16	3.10	6.78	5.10 MB
$p = 128$, WB={t, f, d, c, s}	176.38	16.96	35.91	115.50	4.71°	2.10°	3.09°	5.92°	5.77	3.01	4.27	7.71	5.10 MB
Style WB [32]													
$p = 64$, WB={t, d, s}	92.65	**6.52**	**14.23**	35.01	2.47°	0.82°	1.44°	2.49°	2.99	**1.36**	2.04	3.32	61.0 MB
$p = 64$, WB={t, f, d, c, s}	151.38	29.49	56.35	125.33	4.18°	2.13°	3.03°	4.81°	5.42	3.11	4.42	6.76	61.1 MB
$p = 128$, WB={t, d, s}	88.03	7.92	17.73	45.01	2.61°	0.93°	1.58°	2.85°	3.24	1.50	2.30	3.95	61.2 MB
$p = 128$, WB={t, f, d, c, s}	100.24	10.77	37.74	70.18	3.09°	1.15°	2.61°	3.87°	3.96	1.59	3.55	5.51	61.3 MB
DeNIM + Mixed WB [4] (ours)													
$p = 64$, WB={t, d, s}	120.14	36.39	77.40	152.96	2.57°	1.53°	2.17°	3.19°	5.26	3.38	4.71	6.64	28.7 MB
$p = 64$, WB={t, f, d, c, s}	129.01	14.39	27.69	57.90	2.67°	0.99°	1.45°	2.29°	3.96	2.10	2.85	4.24	28.7 MB
$p = 128$, WB={t, d, s}	158.58	60.14	115.66	198.59	4.20°	2.38°	3.77°	5.63°	5.69	3.91	5.41	7.10	28.8 MB
$p = 128$, WB={t, f, d, c, s}	99.70	13.89	24.71	43.88	2.49°	1.07°	1.62°	2.41°	3.44	1.95	2.74	3.78	28.8 MB
DeNIM + Style WB [32] (ours)													
$p = 64$, WB={t, d, s}	**65.80**	10.06	16.98	**28.82**	2.03°	0.88°	1.23°	1.93°	2.95	1.79	2.33	3.18	196.3 MB
$p = 64$, WB={t, f, d, c, s}	83.41	13.23	21.46	37.44	**1.93°**	**0.77°**	**1.09°**	**1.70°**	**2.73**	1.62	**2.03**	**2.71**	196.3 MB
$p = 128$, WB={t, d, s}	80.53	17.59	27.80	44.35	2.16°	0.88°	1.34°	2.16°	3.08	1.86	2.37	3.30	196.4 MB
$p = 128$, WB={t, f, d, c, s}	89.10	11.27	19.34	43.01	2.49°	1.24°	1.64°	2.92°	3.16	1.87	2.53	3.35	196.4 MB

Table 2: Comparison of the complexity of our method and the prior methods with their post-processing tricks. ms: multi-scale weighting maps, eas: edge-aware smoothing.

Model Architecture	Time (s)	Param (M)	FLOPS (G)
Mixed WB [4] + ms + eas	10.390	**1.32**	82.68
Mixed WB [4] + ms	0.228		
Mixed WB [4] + eas	10.279		
Mixed WB [4]	0.212		
Style WB [32] + ms + eas	10.342	15.31	76.80
Style WB [32] + ms	0.232		
Style WB [32] + eas	10.307		
Style WB [32]	0.217		
Ours w/ Mixed WB [4]	**0.006**	1.67	**2.14**
Ours w/ Style WB [32]	**0.010**	16.19	26.89

2838. These results indicate that our proposed strategy performs comparably well to the prior works on a per-pixel basis for AWB correction in the sRGB space. Utilizing per-pixel color mapping seems to result in color casts that are closer to human perception by more accurately representing the lighting conditions within the scene.

Night photography rendering [15, 41] is an emerging topic in digital imaging. Night image capturing poses significant challenges due to its inherent nature, characterized by low light conditions, diverse illuminant sources, and hardware limitations. In night image capturing, AWB correction plays a pivotal role in preserving the realistic perspective of the output, ensuring that it aligns with human perception and avoids distortions. As practiced in [32], we integrate our AWB correction strategy into the camera ISP for processing night images given in the evaluation part of Night Photography Challenge 23' [41]. In our pipeline, we incorporate the same operations, including gamma correction, tone mapping, auto-contrast, and denoising [44], in the same order for all methods, but the only modification made is to the white-balancing strategies. Figure 4 illustrates the rendering results of various camera pipeline variants that encompass the prior works and our proposed strategy as the AWB correction method. The rendering results demonstrate that our strategy effectively produces more natural night images by mitigating undesired color casts commonly encountered in real-world scenarios.

Quantitative evaluation: The benchmark on single-illuminant Cube+ dataset [13] is presented in Table 1. Following the same experimental setup in the prior works [4, 32], we have used two different patch sizes for the back-

(a) Mixed WB (b) Style WB (c) DeNIM + Mixed WB (d) DeNIM + Style WB

Figure 4: Comparison of the night photography rendering results of our AWB correction strategy with Mixed WB [4] and Style WB [32] on the selected samples from Night Photography Rendering Challenge 23' evaluation set [41]. Image indices from top to bottom: 8678, 8210, 8817, 8894, 8941.

bone network (*i.e.*, 64 and 128), and we designed the input image with two sets of WB settings where the default choices include Tungsten, Daylight, and Shade, while we further incorporate Fluorescent and Cloudy color temperatures to enhance the versatility of the method. The quantitative results indicate that our strategy achieves not only increasing efficiency but also improving performance across all different patch sizes and WB settings, as evidenced by all evaluation metrics. The main observations extracted from these results are as follows: (1) In contrast to the results obtained with Style WB, the best-performing variant appears to be when using a patch size of 64 and incorporating all possible WB settings. This configuration leads to superior performance when compared to other settings. (2) The notable increase in performance, specifically observed on the third quantiles of all evaluation metrics, deserves highlighting. This observation suggests that our strategy can produce more robust results, particularly when dealing with challenging samples. (3) Confirming the findings in [4, 32], we observe that smaller patch sizes tend to lead to better modeling of the illuminant, and in our case, also learning color mappings. (4) We encountered difficulties in identifying a consistent pattern for the mean-squared error (MSE) metric when compared to the other two metrics, and this may suggest that MSE may not adequately capture the quality of color correction achieved by the different methods. We

(a) Mixed WB (b) Style WB (c) DeNIM + Mixed WB (d) DeNIM + Style WB

Figure 5: Failing to address unrealistic color casts, and not effectively handled by any AWB correction methods on MIT-Adobe FiveK dataset [13].

believe that this particular metric might not be suitable for accurately measuring the performance of AWB correction.

Efficiency: The results presented in Table 2 demonstrate the efficiency of our proposed strategy when compared to its prior works across different criteria. Specifically, we evaluated the efficiency based on the following criteria: the processing time (Time (s)), the model complexity in terms of parameter count (Param (M)), and computational load measured in Floating Point Operations Per Second (FLOPS (G)). In terms of processing time, our strategy significantly reduces the time required to process the images for AWB correction. The reduction in processing time is accomplished by designing a model that allows discarding the post-processing operations (*i.e.*, multi-scale inference, and edge-aware smoothing) and adopting simple learnable projection matrices in place of the decoder. DeNIM shows a remarkable speed advantage, being at least 35 times faster than previous models (up to 1700 times faster when post-processing is included).

Next, the model complexity is an essential factor to consider. DeNIM leads to a slight increase in the number of parameters compared to the prior works, even though it discards the decoder of the baseline models. The reason behind the increasing number of parameters lies in the decision to use fully-connected layers as projection matrices, as opposed to convolutional layers in the decoder. Fully-connected layers require more parameters, due to their dense connections between all input and output neurons. This design choice may have led to a slightly higher model complexity, however, it is important to note that this decision does not significantly impact the processing time. Lastly, we measure the computational load of all methods in terms of FLOPS. Lower FLOPS values imply less computational resources required, hence better efficiency. When DeNIM is trained with the Mixed WB backbone, it achieves a remarkable reduction in FLOPS by approximately 97%. Similarly, when trained with the Style WB backbone, the FLOPS are reduced by approximately 65%. This substantial decrease in computational load highlights the remarkable efficiency of our strategy compared to the prior works.

Limitations: Although deep-learning-driven AWB methods generally demonstrate significant resilience across various different scenarios, there are occasional examples where they yield unsatisfactory results. As shown in Figure 5, AWB correction operations may fail to address unrealistic color casts and produce poor results which do not align with human visual perception. At this point, our strategy also may not be able to handle the challenges effectively, primarily since it relies on the feature extraction part of the prior models. We can state that it may struggle to address certain complex and uncommon scenarios, which leads to sub-optimal results. Moreover, to further investigate the performance in handling more challenging cases, our strategy can be tested on multi-illuminant datasets [4, 31]. By subjecting this strategy to such datasets, we can gain valuable insights into its capabilities and limitations in handling diverse and complex lighting scenarios, and we left it as future work.

5. Conclusion

In this paper, we have introduced a novel and efficient deep learning-based AWB correction strategy built on top of the current state-of-the-art methods. This strategy incorporates the idea of deterministic color mapping by leveraging the encoder of existing AWB models and learnable projection matrices. Through extensive experiments, we showed the effectiveness of our strategy by achieving at least 35 times faster processing while surpassing the performance of state-of-the-art methods on high-resolution images. Our research provides a promising solution for real-time, high-quality color correction in practical scenarios, even in digital camera chipsets, addressing the challenges posed by increasing model complexities for better performance.

References

[1] Mahmoud Afifi, Jonathan T Barron, Chloe LeGendre, Yun-Ta Tsai, and Francois Bleibel. Cross-camera convolutional color constancy. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1981–1990, 2021. 2

[2] Mahmoud Afifi and Michael S Brown. Deep white-balance editing. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition*, pages 1397–1406, 2020. 3, 5

[3] Mahmoud Afifi and Michael S Brown. Interactive white balancing for camera-rendered images. In *Color and Imaging Conference*, volume 2020, pages 136–141. Society for Imaging Science and Technology, 2020. 5

[4] Mahmoud Afifi, Marcus A. Brubaker, and Michael S. Brown. Auto white-balance correction for mixed-illuminant scenes. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1210–1219, January 2022. 2, 3, 4, 5, 6, 7

[5] Mahmoud Afifi, Brian Price, Scott Cohen, and Michael S Brown. When color constancy goes wrong: Correcting improperly white-balanced images. In *CVPR*, 2019. 5

[6] Mahmoud Afifi, Abhijith Punnappurath, Abdelrahman Abdelhamed, Hakki Can Karaimer, Abdullah Abuolaim, and Michael S Brown. Color temperature tuning: Allowing accurate post-capture white-balance editing. In *Color and Imaging Conference*, volume 2019, pages 1–6. Society for Imaging Science and Technology, 2019. 4

[7] Nikola Banić, Karlo Koščević, and Sven Lončarić. Unsupervised learning for color constancy, 2019. 3, 5

[8] Simone Bianco and Claudio Cusano. Quasi-unsupervised color constancy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019. 2, 5

[9] David H Brainard and William T Freeman. Bayesian method for recovering surface and illuminant properties from photosensor responses. In *Human Vision, Visual Processing, and Digital Display V*, volume 2179, pages 364–376. SPIE, 1994. 2

[10] David H Brainard and William T Freeman. Bayesian color constancy. *JOSA A*, 14(7):1393–1411, 1997. 2

[11] David H Brainard and Brian A Wandell. Analysis of the retinex theory of color vision. *JOSA A*, 3(10):1651–1661, 1986. 1

[12] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980. 1

[13] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Fredo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. pages 97 – 104, 07 2011. 3, 4, 5, 7

[14] 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. 1

[15] Egor Ershov, Alex Savchik, Denis Shepelev, Nikola Banić, Michael S. Brown, Radu Timofte, Karlo Koščević, Michael Freeman, Vasily Tesalin, Dmitry Bocharov, Illya Semenkov, Marko Subašic, Sven Lončarić, Arseniy Terekhin, Shuai Liu, Chaoyu Feng, Hao Wang, Ran Zhu, Yongqiang Li, Lei Lei, Zhihao Li, Si Yi, Ling-Hao Han, Ruiqi Wu, Xin Jin, Chunle Guo, Furkan Kinli, Sami Menteş, Bariş Özcan, Furkan Kıraç, Simone Zini, Claudio Rota, Marco Buzzelli, Simone Bianco, Raimondo Schettini, Wei Li, Yipeng Ma, Tao Wang, Ruikang Xu, Fenglong Song, Wei-Ting Chen, Hao-Hsiang Yang, Zhi-Kai Huang, Hua-En Chang, Sy-Yen Kuo, Zhexin Liang, Shangchen Zhou, Ruicheng Feng, Chongyi Li, Xiangyu Chen, Binbin Song, Shile Zhang, Lin Liu, Zhendong Wang, Dohoon Ryu, Hyokyoung Bae, Taesung Kwon, Chaitra Desai, Nikhil Akalwadi, Amogh Joshi, Chinmayee Mandi, Sampada Malagi, Akash Uppin, Sai Sudheer Reddy, Ramesh Ashok Tabib, Ujwala Patil, and Uma Mudenagudi. Ntire 2022 challenge on night photography rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1287–1300, June 2022. 5

[16] Graham Finlayson and Steven Hordley. Improving gamut mapping color constancy. *IEEE Transactions on Image Processing*, 9(10):1774–1783, 2000. 2

[17] Graham D Finlayson, Steven D Hordley, and Ingeborg Tastl. Gamut constrained illuminant estimation. *International journal of computer vision*, 67:93–109, 2006. 2

[18] Graham D Finlayson and Elisabetta Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41. Society for Imaging Science and Technology, 2004. 1

[19] David A Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, 1990. 2

[20] Brian Funt, Vlad Cardei, and Kobus Barnard. Learning color constancy. In *Fourth Color Imaging Conference: Color Science, Systems and Applications*, pages 58–60. Simon Fraser University, 1996. 2

[21] Peter Vincent Gehler, Carsten Rother, Andrew Blake, Tom Minka, and Toby Sharp. Bayesian color constancy revisited. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8. IEEE, 2008. 2

[22] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Generalized gamut mapping using image derivative structures for color constancy. *International Journal of Computer Vision*, 86(2-3):127, 2010. 2

[23] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Computational color constancy: Survey and experiments. *IEEE transactions on image processing*, 20(9):2475–2489, 2011. 2

[24] Arjan Gijsenij, Theo Gevers, and Joost Van De Weijer. Improving color constancy by photometric edge weighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(5):918–929, 2011. 1

[25] Dan Hendrycks and Kevin Gimpel. Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*, 2016. 3

[26] Daniel Hernandez-Juarez, Sarah Parisot, Benjamin Busam, Ales Leonardis, Gregory Slabaugh, and Steven McDonagh. A multi-hypothesis approach to color constancy. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2270–2280, 2020. 2

[27] Yuanming Hu, Baoyuan Wang, and Stephen Lin. Fc4: Fully convolutional color constancy with confidence-weighted pooling. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017. 2, 5

[28] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pages 694–711. Springer, 2016. 3

[29] Hamid Reza Vaezi Joze, Mark S Drew, Graham D Finlayson, and Perla Aurora Troncoso Rey. The role of bright pixels in illumination estimation. In *Color and Imaging Conference*, volume 2012, pages 41–46. Citeseer, 2012. 1

[30] Zhanghan Ke, Yuhao Liu, Lei Zhu, Nanxuan Zhao, and Rynson W.H. Lau. Neural preset for color style transfer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14173–14182, June 2023. 2, 3

[31] Dongyoung Kim, Jinwoo Kim, Seonghyeon Nam, Dongwoo Lee, Yeonkyung Lee, Nahyup Kang, Hyong-Euk Lee, ByungIn Yoo, Jae-Joon Han, and Seon Joo Kim. Large scale multi-illuminant (lsmi) dataset for developing white balance algorithm under mixed illumination. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2410–2419, 2021. 7

[32] Furkan Kınlı, Doğa Yılmaz, Barış Özcan, and Furkan Kıraç. Modeling the lighting in scenes as style for auto white-balance correction. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 4903–4913, January 2023. 2, 3, 4, 5, 6

[33] Shuwei Li, Jikai Wang, Michael S. Brown, and Robby T. Tan. Mimt: Multi-illuminant color constancy via multi-task learning, 2023. 2, 5

[34] Yi-Chen Lo, Chia-Che Chang, Hsuan-Chao Chiu, Yu-Hao Huang, Chia-Ping Chen, Yu-Lin Chang, and Kevin Jou. Clcc: Contrastive learning for color constancy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8053–8063, 2021. 2

[35] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017. 3

[36] Zhongyu Lou, Theo Gevers, Ninghang Hu, Marcel P Lucassen, et al. Color constancy by deep learning. In *BMVC*, pages 76–1, 2015. 2

[37] Taishi Ono, Yuhi Kondo, Legong Sun, Teppei Kurita, and Yusuke Moriuchi. Degree-of-linear-polarization-based color constancy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19740–19749, 2022. 2

[38] Yanlin Qian, Joni-Kristian Kamarainen, Jarno Nikkanen, and Jiri Matas. On finding gray pixels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8062–8070, 2019. 1

[39] Yanlin Qian, Said Pertuz, Jarno Nikkanen, Joni-Kristian Kämäräinen, and Jiri Matas. Revisiting gray pixel for statistical illumination estimation. *arXiv preprint arXiv:1803.08326*, 2018. 1

[40] Wu Shi, Chen Change Loy, and Xiaoou Tang. Deep specialized network for illuminant estimation. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pages 371–387. Springer, 2016. 2

[41] Alina Shutova, Egor Ershov, Georgy Perevozchikov, Ivan Ermakov, Nikola Banić, Radu Timofte, Richard Collins, Maria Efimova, Arseniy Terekhin, Simone Zini, Claudio Rota, Marco Buzzelli, Simone Bianco, Raimondo Schettini, Chunxia Lei, Tingniao Wang, Song Wang, Shuai Liu, Chaoyu Feng, Guangqi Shao, Hao Wang, Xiaotao Wang, Lei Lei, Lu Xu, Chao Zhang, Yasi Wang, Jin Guo, Yangfan Sun, Tianli Liu, Dejun Hao, Furkan Kınlı, Barış Özcan, Furkan Kıraç, Hyerin Chung, Nakyung Lee, Sung Keun Kwak, Marcos Conde, Tim Seizinger, Florin Vasluianu, Omar Elezabi, Chia-Hsuan Hsieh, Wei-Ting Chen, Hao-Hsiang Yang, Zhi-Kai Huang, Hua-En Chang, I-Hsiang Chen, Yi-Chung Chen, and Yuan-Chun Chiang. Ntire 2023 challenge on night photography rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1981–1992, June 2023. 3, 5, 6

[42] Joost Van De Weijer, Theo Gevers, and Arjan Gijsenij. Edge-based color constancy. *IEEE Transactions on image processing*, 16(9):2207–2214, 2007. 1

[43] Bolei Xu, Jingxin Liu, Xianxu Hou, Bozhi Liu, and Guoping Qiu. End-to-end illuminant estimation based on deep metric learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3616–3625, 2020. 2

[44] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5728–5739, 2022. 5