



Revisiting Color Constancy Using CNNs: Including Recent Observations

Oguzhan Ulucan*, Diclehan Ulucan , and Marc Ebner 

University of Greifswald, Institute of Mathematics and Computer Science
17489 Greifswald, Germany
{oguzhan.ulucan, diclehan.ulucan, marc.ebner}@uni-greifswald.de

Abstract. In recent studies, alongside introducing new approaches for color constancy, we have also focused on improving existing techniques and introducing new perspectives. Our motivation is the idea that investigating different strategies, concepts, and their combinations that have not been analyzed in this field in detail yet, might help us to find simple, effective, and cost-efficient solutions. Thereupon, we utilized observations we obtained from our algorithms to analyze how we can enhance the performance of well-known learning-free methods. We demonstrated why using *salient* pixels, performing block-based operations, and carrying out scale-space computations benefit color constancy approaches significantly and provide a notable performance increase. In this study, we make use of our recent observations on learning-free algorithms to analyze if they are also beneficial for enhancing the performance of a learning-based color constancy model. According to our evaluation, all of our observations contribute to the performance of a convolutional neural network model and increase its effectiveness in estimating the illuminant. Thus, the contribution of these strategies in learning-based models should be further investigated to improve their performance with simple yet effective solutions.

Keywords: Color constancy · illuminant estimation · white-balancing.

1 Introduction

Our visual system has evolved in such a way that it can discount the illuminant of the environment, and perceive the true colors of the objects [61]. For instance, when we enter a room illuminated by purple light, we are able to recognize the color of a white object as white. This ability of unconsciously discounting the light source of a scene is called *color constancy*, and it is usually also referred to as a phenomenon since the biological mechanism behind this ability is not yet fully understood. While we can perform color constancy unintentionally, machine vision systems require guidance to perform this task since as a result of the interaction between the sensitivity of the camera sensors and the light source of the scene, a white object illuminated by purple light would be captured as purple [55]. The field analyzing this problem from the perspective

* Corresponding author

¹ This is the pre-print version of the manuscript accepted at the Computational Color Imaging Workshop (CCIW 2024)

of computer science is called *computational color constancy*. It benefits both digital photography and various computer vision tasks including but not limited to underwater image enhancement and image dehazing.

In computational color constancy, we utilize an image formation model where we mostly assume that (i) the surfaces are Lambertian, i.e., the surface is equally reflecting the light into all directions, and (ii) there is a point light source illuminating the scene. Furthermore, we presume that the camera capturing the scene consists of three sensors that detect the energy of incoming light by reacting to a particular segment of the visible spectrum, e.g., short-, middle-, and long-wavelengths. By taking these assumptions into account, we can mathematically express an image I based on the irradiance E falling onto the sensors of the capturing device, and the sensor sensitivity function S of the camera representing the responses of the sensors to light at a particular wavelength as follows:

$$I(x, y) = \int_w E(x, y; \lambda) S_i(\lambda) d\lambda, \quad (1)$$

where (x, y) is the spatial location, λ corresponds to the wavelength of the visible spectrum w , and $i \in \{\text{long, middle, short}\}$.

We can formulate the irradiance falling onto the sensors of the camera by making use of its relationship with the reflectance $R(x, y; \lambda)$, the scaling factor $G(x, y)$, and the point light source $L(x, y; \lambda)$ as follows:

$$E(x, y; \lambda) = G(x, y) R(x, y; \lambda) L(x, y; \lambda), \quad (2)$$

where $G(x, y)$ can be taken as $\cos \alpha$, where α is the angle between the surface normal vector and a vector pointing in the direction of the illumination source.

We can model an image by utilizing Eqn. 1, and Eqn. 2 as follows:

$$I(x, y) = G(x, y) \int_w R(x, y; \lambda) L(x, y; \lambda) S_i(\lambda) d\lambda. \quad (3)$$

In color constancy, we aim at computing L from a color cast image I by taking advantage of Eqn. 3. Yet, despite the aforementioned assumptions the ill-posed nature of the problem cannot be overcome due to the fact that the image is affected by both the lighting conditions and the often unknown sensor characteristics in real-world scenarios. Therefore, to relax the problem, we generally make additional assumptions such as a single light source is illuminating the scene, G does not affect the computation of the illuminant, and the responses of the capturing device are narrow-band, e.g., they can be approximated by Dirac's delta functions [21]. Consequently, a color-cast image can be represented as the Hadamard product of the (shaded) reflectance R and the single light source \mathbf{L} as follows:

$$I(x, y) = R(x, y) \circ \mathbf{L}. \quad (4)$$

Over the last decades, various single-illuminant color constancy algorithms have been introduced which depend on different strategies [21]. These algorithms can be categorized as traditional and data-driven methods. The former utilizes image statistics to compute the color vector of the light source. The most well-known traditional methods are white-patch Retinex [39] and gray world [13] whose assumptions lay the foundations of several other classical color constancy algorithms [10,16,24,29,35,45,55,57].

Traditional methods are mostly cost-efficient and easy to implement, however, their ability of computing the color vector of the light source may decrease significantly if only a limited number of different colors is available in the scene [14]. Close-up shootings and scenes with dominant sky or grass regions can be given as examples of scenes that are challenging for classical color constancy algorithms. On the other hand, data-driven methods generally achieve higher accuracy on images having a uniform color distribution. Data-driven models can be further grouped as gamut-based methods [9,22,23,25,30], Bayesian strategies [11,12,27], and neural network-based models [1,8,17,18,34,37]. Neural networks-based models estimate the illuminant by learning from the features present in their large-scale training sets [15]. Early neural network models make use of convolutional neural networks [4,7] which are still widely used in this field [1,6,18,34].

In our previous studies, we investigated the effects of utilizing block-based operations, *salient* pixels, and conducting scale-space computations in the field of color constancy. We showed that these operations can improve the performance of existing traditional color constancy algorithms, and we demonstrated that we can introduce new classical methods based on these strategies [54,55,56,57,58]. In this study, we utilize all our previous observations on learning-free algorithms and apply them to a convolutional neural networks (CNNs) model to ascertain whether they might also be useful for learning-based strategies. For our investigation, we adopt the CNNs model introduced by Bianco *et al.* [7] as our baseline architecture since (i) the simple structured model contains a low number of parameters, and (ii) it is specifically designed for estimating the color vector of light sources based on the image’s locally varying statistics, i.e., image patches, which overlaps with one of our previous observations. We modify this model with a mechanism to guide the model to only consider the salient pixels since not all regions are informative while estimating the illuminant [35,55]. Moreover, we extend this model to multiple scales since estimating the illuminant with a scale-space approach improves the performance [57,58]. These two observations also coincide with experiments on the human visual system which we often try to mimic in computer vision, i.e., the human visual system may benefit from regions with the highest luminance while discounting the effects of the light source [20,42,43,50], and the illuminant estimate is available in the stimulus at the proper scale [48,49]. In our experiments, we demonstrate that our simple yet effective operations are beneficial for a learning-based model as well.

This paper is organized as follows. In Sec. 2, we revisit the color constancy model introduced by Bianco *et al.* [7], while also introducing our simple modifications and detailing the training phase. In Sec. 3, we introduce our experimental setup and discuss the performance of the model. Finally, in Sec. 4, we summarize our work.

2 Modifications based on Recent Observations

As we discussed in Sec. 1, the modifications we proposed are based on our recent observations on learning-free color constancy algorithms. In order to investigate whether the observations are also beneficial for a learning-based approach, as our foundation model,

we utilize the approach introduced by Bianco *et al.* [7]. In this section, we revisit the architecture, introduce our modifications, and detail our training procedure.

2.1 Revisiting Color Constancy using CNNs

The network consists of only 5 layers. The input layer takes in $32 \times 32 \times 3$ RGB image patches. Afterwards, a convolutional layer utilizes a bank of $240 \ 1 \times 1 \times 3$ kernels with a stride of 1 to filter the input. This operation results in a feature map with a size of $32 \times 32 \times 240$ which is reduced to $4 \times 4 \times 240$ by using a max-pooling layer with a 8×8 kernel. The obtained feature maps are then reshaped into a 3840 dimensional vector and sent to a fully connected layer of 40 nodes. Subsequently, a ReLU activation function is utilized, and a 40×3 dense layer is added to obtain the illuminant estimate. Consequently, the model consists of a total 154,723 trainable parameters.

2.2 Adjusting the Baseline Model

We prefer to maintain the input requirements of the network since Bianco *et al.* [7] stated that a block size of 32×32 is sufficient while smaller ones are not, and since we recently investigated the impact of block sizes on color constancy and demonstrated that there is indeed a certain block size range to effectively perform color constancy [55,57,58]. Adjusting the block size appropriately is important since when we choose large image blocks we might neglect the varying surface orientations present at local regions which decreases the considered amount of information [55]. On the other hand, if we select small block sizes the possibility of obtaining uniformly colored inputs increases, thus, the information that can be extracted from a block decreases [55]. Therefore, to respect the local information varying throughout the scene, and to keep the number of parameters of the model adequately low, we utilize image blocks with the size of $32 \times 32 \times 3$.

Salient Pixel Detection Layer. We first add a mechanism to the model to assign larger weights to the brightest pixels in the image and to decrease the impact of the non-informative elements. We can explain the reason behind the choice of this mechanism based on biological findings and from the perspective of computational photography [59]. According to biological studies the human visual system might be discounting the illuminant of the scene by relying on the regions having the highest luminance rather than the darkest areas [40,42,43,50]. Furthermore, in digital photography it is known that estimating the illuminant can be easier when achromatic regions are used rather than colored ones [19,35,44,45,54,58]. In this study, to utilize the salient pixels, i.e., brightest pixels, we obtain the luma of the image block and determine the pixels having highest brightness by selecting the top brightest pixels in the block. It is important to stress that while selecting the brightest pixels, we exclude the over-saturated pixels to reduce possible noise since these elements negatively affect the performance of the color constancy algorithms [55]. According to our experiments, choosing 50% or more of the brightest pixels is sufficient to estimate the color vector of the light source.

We select this threshold since in our previous experiments we determined that approximately 3% of the brightest pixels for a whole image is adequate to effectively estimate the illuminant [55]. Since we consider much smaller image regions in this study, i.e., blocks with the size of $32 \times 32 \times 3$, we adjust the number of the brightest pixels accordingly. After highlighting the top brightest pixels, we mask the darkest regions and weigh the input block with the resulting saliency map. It is worth noting that achromatic pixels have already been used in learning-based models as salient pixels [6]. However, the model used in that study requires a significantly high number of trainable parameters due to the nature of the U-Net architecture [47]. On the other hand, our simple weighting process allows us to avoid raising the complexity of the model while providing a significant performance increase which we demonstrate in Sec. 3.

Scale-space Computations. Lastly, we consider scale-space computations in our approach since they allow us to capture local details and global context more effectively, and enable us to highlight discriminative patterns that are not apparent at an individual scale. Furthermore, we prefer to perform operations at multiple levels since for studies utilizing the color feature, the scale-space computations are beneficial thanks to their sensitivity to low-level features [28,41,51]. As color is a low-level feature, utilizing scale-space computations naturally benefits the task of illumination estimation [57,58]. This is also not surprising, considering that in the human visual system, the information we require for accurate local estimates of the illumination might be available in the stimulus at the appropriate scale [48,49]. To mimic the scale-space computations in the model, we downsample either the input block or the input block weighted by the salient pixels. We consider only the first 4 levels of the image pyramid since the spatial information degrades significantly at further coarser scales [53]. After forming the image pyramid, at each scale, we apply a convolutional layer with 240 filters followed by a max-pooling layer with distinct kernel sizes to pool the feature maps. For instance, we pool the features coming from the finest scale via a max-pooling operation with a kernel size of 8×8 , while for the consecutive scale, we select a kernel size of 4×4 . This ensures that we maintain a balance between the preservation of spatial information and the reduction of the dimensions of the feature maps. Thus, we optimize the feature extraction process for each scale and enhance the model’s ability to capture various patterns across different levels of the pyramid. Subsequently, we combine the features into a vector and forward them to the fully connected layer. The scale-space computations increase the number of trainable parameters to approximately four times that of the baseline model since we are considering the first four levels of the image pyramid.

2.3 Training Phase

Data Augmentation. For the training process, we utilize the ColorChecker RECom-mended dataset [27], which is detailed in Sec. 3.1, to stay consistent with the prior study [7]. However, we employ a different data augmentation strategy from the prior work based on the observations from our previous study in which we demonstrated that the performance of learning-based models tends to decrease significantly when only a limited number of illumination conditions are considered in their training sets [56].

Motivated by this observation, and the fact that 568 images and illumination conditions alone are inadequate for the effective training process, we augment this dataset as follows. First of all, by using the linear input images and their corresponding ground truth color vector of the light sources, we obtain ground truth white-balanced images via the inversion of Eqn. 4. Then, to augment the dataset by considering various single-illuminant light source conditions, we collect additional color vectors of the illuminant that are present in well-known benchmarks, i.e., the INTEL-TAU [38], and Cube+ [3] datasets. In case there is a repetitive illuminant vector after combining the ground truths of benchmarks, we discard the color vectors having the same RGB triplet [52]. Thus, we obtain an illumination set with unique color vectors of the light source. Afterwards, for each ground truth image, we randomly select 3 illuminants from this set, and by using Eqn. 4, we render a total of 1704 new scenes. It is important to note that, while rendering the images, we remove the randomly selected illuminant from our illumination set to prevent rendering the scenes with the same color vector of the light source. Lastly, we extract $32 \times 32 \times 3$ non-overlapping blocks from the images and remove blocks containing a color calibration object, consequently, we obtain a total of 1307943 blocks. Through utilizing image blocks rather than the whole image, we obtain a high number of inputs that satisfy the requirements of learning-based models [7].

Training Parameters. Our model learns its parameters by minimizing the Euclidean loss. We train our model using the Adam optimizer whose exponential decay rate for the first and second-moment estimates is assigned as 0.9 and 0.999, respectively. We set the number of epochs to 100 and the learning rate to $10e - 7$. We clip the gradient of each weight individually to avoid that its norm is higher than 1. Furthermore, we consider to add a weight decay of $10e - 8$ for better generalization. It is important to note that we did not fine-tune the model according to the angular error as in the prior work [7].

3 Experiments and Discussions

In this section, we detail our experimental setup, while briefly introducing the datasets and the evaluation strategy that is widely adopted in color constancy. To report the results of algorithms, we either run the original codes without any optimization, utilize the outcomes of the original works, or use the reported results from recent comprehensive publications [58].

3.1 Datasets

In our experiments, we utilize 2 publicly available, comprehensive single-illuminant benchmarks covering different lighting conditions and camera specifications. To evaluate our model on these datasets, if necessary, we perform preprocessing, i.e., masking the calibration objects, subtracting the black level, and clipping the under- and over-saturated pixels according to the masks provided in the benchmarks.

The ColorChecker RECommended dataset is the processed version of the Gehler-Shi dataset [27,32]. The benchmark contains a total of 568 distinct scenes: 254 indoor,

85 outdoor, and 229 close-up images [14], and their corresponding ground truths along with the masks to handle under- and over-saturated regions, and the position of the calibration object. Two cameras, i.e., Canon 1D and Canon 5D, are utilized to capture the scenes. The INTEL-TAU dataset is introduced by Lakoom *et al.* [38], and it includes a total of 7022 images: 1466 indoor, 2327 outdoor, and 3229 close-up scenes [14]. The images are captured with 3 different cameras, namely, Nikon D810, Canon 5DSR, and Sony IMX135. All the scenes in INTEL-TAU are preprocessed, i.e., images have a linear response, and their black level is calibrated. Moreover, all the sensitive data, i.e., faces, and license plates, is masked. During our experiments, we only utilized the scenes from the sets of "field1" and "field3" since some of the images in the other sets, i.e. "lab printouts", and "lab realscene", include an unmasked calibration object.

3.2 Evaluation Metric

To statistically investigate the model's performance, we adopt a widely used error metric in color constancy, namely, angular error [33]. We report the mean, median, and mean of the best and worst %25 of the angular errors.

The angle (ε) between the color vector of the estimated of the light source (L_{est}) and the ground truth illuminant (L_{gt}) can be calculated as follows:

$$\varepsilon(L_{est}, L_{gt}) = \cos^{-1} \left(\frac{L_{est} L_{gt}}{\|L_{est}\| \|L_{gt}\|} \right). \quad (5)$$

3.3 Discussion

We report our statistical analysis for both benchmarks in Table 1. The first noticeable result is that modifying the baseline model with our strategies improves the performance significantly. Each stage, i.e., guiding the network with information coming from the salient regions, and carrying the model into multiple scales, enhances the overall performance of the baseline method. Furthermore, the worst cases almost always improve significantly through the modifications which is a valuable outcome since in color constancy it is known that improving the algorithms' performance for the worst cases is important.

We obtain the best outcomes when we either modify the baseline model with the salient pixels or consider salient pixels with the scale-space computations. If we compare the contribution of the salient pixels and the scale-space computations, it is clear that considering the salient pixels benefits more in general. This result can be explained by the fact that not all regions are informative for color constancy, hence reducing the effect of the non-informative elements allows us to improve the overall performance. It is important to stress that while guiding the network to give more importance to the salient regions, our modification does not increase the number of trainable parameters. Hence, without increasing the complexity of the model, this modification allows us to enhance the performance significantly.

We provide visual investigations in Fig. 1 where we compare different modifications on the baseline model with other algorithms. We provide results chosen from the best and worst cases randomly. The baseline model modified with the salient pixels and/or

Table 1. Statistical results on 2 benchmarks. The top results are highlighted using color coding as follows, first: blue, second: cyan, third: green, forth: yellow, and fifth: orange. The abbreviations SP, MS and MSSP correspond to the salient pixels, multi-scale, and multi-scale with salient pixels, respectively.

		ColorChecker RECommended				INTEL-TAU			
Algorithms		B-25%	W-25%	Mean	Med.	B-25%	W-25%	Mean	Med.
Learning-free	max-RGB [40]	1.49	17.47	7.78	5.43	1.70	19.24	10.49	11.14
	GW [13]	0.93	10.44	4.71	3.54	0.93	10.59	4.90	3.85
	1 st - GE [60]	0.93	14.17	5.79	3.68	0.94	13.79	5.89	4.07
	2 nd - GE [60]	1.03	14.70	6.09	3.97	1.00	14.13	6.09	4.25
	wGE [31]	0.78	15.57	6.08	3.33	0.80	14.89	5.99	3.63
	DOCC [26]	0.79	18.04	7.23	4.26	0.80	16.97	7.18	4.66
	MSGP [46]	0.76	8.35	3.81	2.96	0.64	8.23	3.57	2.56
	GI [45]	0.44	8.02	3.19	1.90	0.56	8.03	3.32	2.18
	BIO-CC [54]	0.86	9.84	4.40	3.30	0.76	9.42	4.14	3.05
	BB-CC [55]	1.06	7.37	3.48	2.71	0.79	7.25	3.37	2.63
	MSSC [58]	0.62	7.32	3.16	2.16	0.59	7.47	3.23	2.23
Learning-based	Quasi-U CC [6]	-	-	3.46	2.23	0.60	7.28	3.12	2.19
	SIIE [2]	0.55	6.53	2.77	1.93	0.73	7.80	3.42	2.42
	C3AE [37]	0.80	4.00	2.10	1.90	0.90	7.00	3.40	2.70
	One-Net CCC [18]	-	-	-	-	1.10	5.90	3.30	3.20
	FFCC [5]	0.57	6.75	2.95	2.19	0.70	7.96	3.42	2.38
	FC4 [34]	0.34	4.29	1.77	1.11	0.70	5.50	2.60	2.00
	C5 [1]	0.53	5.46	2.50	1.99	0.52	5.96	2.52	1.70
	BoCF CC [36]	-	-	-	-	0.90	6.10	2.90	2.40
	Baseline	0.92	6.54	3.23	2.64	0.92	6.55	3.20	2.56
	Baseline w/ SP	0.62	5.98	2.70	2.00	0.79	6.11	2.90	2.27
Baseline w/ MS	0.75	6.32	3.00	2.34	0.84	6.56	3.16	2.51	
Baseline w/ MSSP	0.73	5.92	2.82	2.26	0.82	6.10	2.93	2.29	

the scale-space can output visually pleasing results with very low angular errors for the best cases while for the worst cases, an additional analysis should be conducted to enhance the outcomes even further.

As a final note, we would like to mention that in this study, we do not aim to design a state-of-the-art model that outperforms all other color constancy algorithms. We intend to present that with our simple yet effective modifications, we can improve the performance of a network, which we aimed as future direction in our previous works. As shown in Table 1, the baseline model alone cannot compete with the state-of-the-art models, but with our modifications, we can increase its performance so that this model performs well while providing competitive results. Thus, the effects of block-based strategies, salient pixels, and scale-space operations should be investigated thoroughly to improve the performance of learning-based color constancy algorithms.

4 Conclusion

Color is an important feature not only for us but also for a wide range of computer vision applications. Many image processing tasks, such as underwater image enhancement and image dehazing, use color constancy at the first stage of their pipelines. Com-

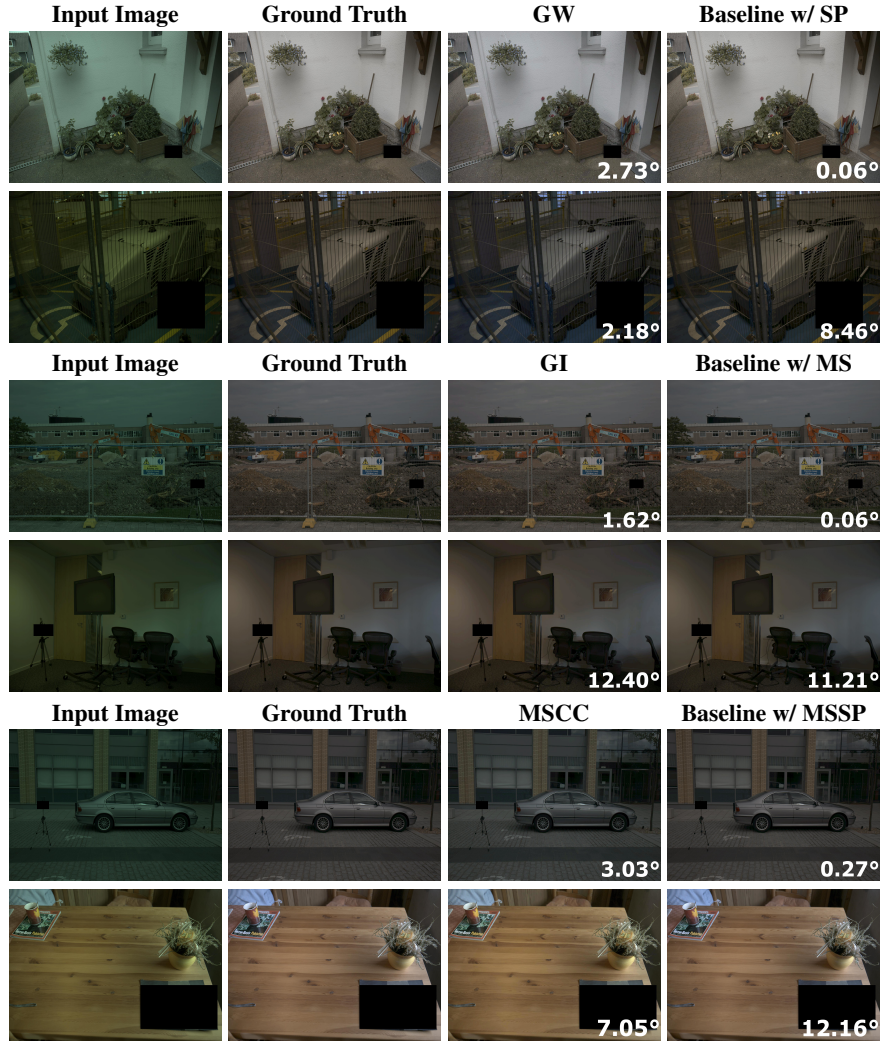


Fig. 1. Comparison of random results on best and worst cases for each stage of the modification. The angular errors are given on the bottom-right side of the images. Gamma is adjusted for better visualization.

putational color constancy is framed as a computational challenge that has been widely studied for more than five decades due to its significance. Researchers have developed numerous methods, and have aimed at improving the existing algorithms by exploiting the effectiveness of different approaches to provide simple yet effective solutions to the ill-posed nature of the problem. Motivated by this aim, alongside developing new methods, we also investigate how we can improve the performance of algorithms in estimating the illuminant. In our previous studies, we explicitly demonstrated that the

performance of learning-free methods can be significantly enhanced by using the following strategies: *(i)* considering varying local statistics of the scenes, i.e., block-based approaches, *(ii)* utilizing the salient pixels, i.e., the pixels having the highest luminance, and *(iii)* multi-scale operations. We mentioned that we will use our observations to investigate if they are beneficial for a learning-based model. Thereupon, in this study, we use our observations to modify a convolutional neural network designed for color constancy. According to our experiments, guiding the network to give more weight to the salient regions and carrying out scale-space computations improves the performance of the model, and allows it to compete with the state-of-the-art. Consequently, we can conclude that one can use the observations obtained from traditional algorithms to modify learning-based models with simple yet effective approaches to improve their effectiveness without significantly increasing their computational complexity. As future work, we will investigate the effects of our observations on different models, and further improve the algorithms' effectiveness in the worst cases.

References

1. Afifi, M., Barron, J.T., LeGendre, C., Tsai, Y.T., Bleibel, F.: Cross-camera convolutional color constancy. In: IEEE/CVF Int. Conf. Comput. Vision. pp. 1981–1990 (2021)
2. Afifi, M., Brown, M.S.: Sensor-independent illumination estimation for DNN models. In: Brit. Mach. Vision Conf. BMVA Press (2019)
3. Banić, N., Lončarić, S.: Unsupervised learning for color constancy. In: Int. Joint Conf. Comput. Vision Imag. Comput. Graph. Theory Appl. vol. 4, pp. 181–188. INSTICC (2018)
4. Barron, J.T.: Convolutional color constancy. In: IEEE/CVF Conf. Comput. Vis. Pattern Recog. pp. 379–387 (2015)
5. Barron, J.T., Tsai, Y.T.: Fast fourier color constancy. In: Conf. Comput. Vision Pattern Recognit. pp. 886–894. IEEE/CVF, Honolulu, HI, USA (2017)
6. Bianco, S., Cusano, C.: Quasi-unsupervised color constancy. In: IEEE/CVF Conf. Comput. Vis. Pattern Recog. pp. 12212–12221 (2019)
7. Bianco, S., Cusano, C., Schettini, R.: Color constancy using CNNs. In: IEEE Conf. Comput. Vis. Pattern Recog. Workshops. pp. 81–89 (2015)
8. Bianco, S., Cusano, C., Schettini, R.: Single and multiple illuminant estimation using convolutional neural networks. *IEEE Trans. Image Process.* **26**(9), 4347–4362 (2017)
9. Bianco, S., Schettini, R.: Color constancy using faces. In: IEEE Conf. Comput. Vis. Pattern Recog. pp. 65–72 (2012)
10. Bianco, S., Buzzelli, M.: Truncated edge-based color constancy. In: Int. Conf. Consum. Electronics. pp. 1–5. IEEE (2022)
11. Brainard, D.H., Freeman, W.T.: Bayesian method for recovering surface and illuminant properties from photosensor responses. In: Human Vision Vis. Process. Digit. Display V. vol. 2179, pp. 364–376. SPIE (1994)
12. Brainard, D.H., Freeman, W.T.: Bayesian color constancy. *J. Opt. Soc. America A* **14**(7), 1393–1411 (1997)
13. Buchsbaum, G.: A spatial processor model for object colour perception. *J. Franklin Inst.* **310**, 1–26 (1980)
14. Buzzelli, M., Zini, S., Bianco, S., Ciocca, G., Schettini, R., Tchobanou, M.K.: Analysis of biases in automatic white balance datasets and methods. *Color Res. Appl.* **48**(1), 40–62 (2023)
15. Buzzelli, M., Schettini, R., Bianco, S.: Learning color constancy: 30 years later. In: Color Imag. Conf. vol. 31, pp. 91–95. Society for Imaging Science and Technology (2023)

16. Cheng, D., Prasad, D.K., Brown, M.S.: Illuminant estimation for color constancy: Why spatial-domain methods work and the role of the color distribution. *J. Opt. Soc. America A* **31**, 1049–1058 (2014)
17. Das, P., Liu, Y., Karaoglu, S., Gevers, T.: Generative models for multi-illumination color constancy. In: *IEEE/CVF Conf. Comput. Vis. Pattern Recog.* pp. 1194–1203 (2021)
18. Domislović, I., Vršnak, D., Subašić, M., Lončarić, S.: One-net: Convolutional color constancy simplified. *Pattern Recognit. Letters* **159**, 31–37 (2022)
19. Drew, M.S., Joze, H.R.V., Finlayson, G.D.: Specularity, the zeta-image, and information-theoretic illuminant estimation. In: *Workshops Demonstrations: Eur. Conf. Comput. Vision.* pp. 411–420. Springer (2012)
20. Ebner, M.: A parallel algorithm for color constancy. *J. Parallel Distrib. Comput.* **64**, 79–88 (2004)
21. Ebner, M.: *Color Constancy*, 1st ed. Wiley Publishing, ISBN: 0470058299 (2007)
22. Finlayson, G., Hordley, S.: Improving gamut mapping color constancy. *IEEE Trans. Image Process.* **9**(10), 1774–1783 (2000)
23. Finlayson, G.D., Hordley, S.D., Tastl, I.: Gamut constrained illuminant estimation. *Int. J. Comput. Vision* **67**, 93–109 (2006)
24. Finlayson, G.D., Trezzi, E.: Shades of gray and colour constancy. In: *Color and Imag. Conf.* pp. 37–41. Society for Imaging Science and Technology (2004)
25. Forsyth, D.A.: A novel algorithm for color constancy. *Int. J. Comput. Vision* **5**(1), 5–35 (1990)
26. Gao, S.B., Yang, K.F., Li, C.Y., Li, Y.J.: Color constancy using double-opponency. *IEEE Transactions Pattern Anal. Mach. Intell.* **37**(10), 1973–1985 (2015)
27. Gehler, P.V., Rother, C., Blake, A., Minka, T., Sharp, T.: Bayesian color constancy revisited. In: *IEEE Conf. Comput. Vis. Pattern Recog.* pp. 1–8 (2008)
28. Geusebroek, J.M., Van Den Boomgaard, R., Smeulders, A.W., Dev, A.: Color and scale: The spatial structure of color images. In: *Eur. Conf. Comput. Vision.* pp. 331–341. Springer (2000)
29. Gijsenij, A., Gevers, T., Van De Weijer, J.: Physics-based edge evaluation for improved color constancy. In: *IEEE Conf. Comput. Vis. Pattern Recog.* pp. 581–588 (2009)
30. Gijsenij, A., Gevers, T., Van De Weijer, J.: Generalized gamut mapping using image derivative structures for color constancy. *Int. J. Comput. Vision* **86**, 127–139 (2010)
31. Gijsenij, A., Gevers, T., Van De Weijer, J.: Improving color constancy by photometric edge weighting. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 918–929 (2011)
32. Hemrit, G., Finlayson, G.D., Gijsenij, A., Gehler, P., Bianco, S., Funt, B., Drew, M., Shi, L.: Rehabilitating the colorchecker dataset for illuminant estimation. In: *Color Imag. Conf.* pp. 350–353. Society for Imaging Science and Technology (2018)
33. Hordley, S.D., Finlayson, G.D.: Reevaluation of color constancy algorithm performance. *JOSA A* **23**(5), 1008–1020 (2006)
34. Hu, Y., Wang, B., Lin, S.: Fc4: Fully convolutional color constancy with confidence-weighted pooling. In: *IEEE/CVF Conf. Comput. Vis. Pattern Recog.* (2017)
35. Joze, H.R.V., Drew, M.S., Finlayson, G.D., Rey, P.A.T.: The role of bright pixels in illumination estimation. In: *Color Imag. Conf.* pp. 41–46. Society for Imaging Science and Technology (2012)
36. Laakom, F., Passalis, N., Raitoharju, J., Nikkanen, J., Tefas, A., Iosifidis, A., Gabbouj, M.: Bag of color features for color constancy. *IEEE Trans. Image Process.* **29**, 7722–7734 (2020)
37. Laakom, F., Raitoharju, J., Iosifidis, A., Nikkanen, J., Gabbouj, M.: Color constancy convolutional autoencoder. In: *Symp. Ser. Comput. Intell.* pp. 1085–1090. IEEE (2019)
38. Laakom, F., Raitoharju, J., Nikkanen, J., Iosifidis, A., Gabbouj, M.: Intel-tau: A color constancy dataset. *IEEE Access* **9**, 39560–39567 (2021)

39. Land, E.H.: The retinex theory of color vision. *Scientific Amer.* **237**, 108–129 (1977)
40. Land, E.H., McCann, J.J.: Lightness and retinex theory. *J. Opt. Soc. America A* **61**(1), 1–11 (1971)
41. Li, B., Xu, D., Lee, M.H., Feng, S.H.: A multi-scale adaptive grey world algorithm. *IEICE Trans. Inf. Syst.* **90**(7), 1121–1124 (2007)
42. Linnell, K.J., Foster, D.H.: Space-average scene colour used to extract illuminant information. *John Dalton’s Colour Vision Legacy* pp. 501–509 (1997)
43. Morimoto, T., Kusuyama, T., Fukuda, K., Uchikawa, K.: Human color constancy based on the geometry of color distributions. *J. Vision* **21**(3), 7–7 (2021)
44. Ono, T., Kondo, Y., Sun, L., Kurita, T., Moriuchi, Y.: Degree-of-linear-polarization-based color constancy. In: *IEEE/CVF Conf. Comput. Vis. Pattern Recog.* pp. 19740–19749 (2022)
45. Qian, Y., Kamarainen, J.K., Nikkanen, J., Matas, J.: On finding gray pixels. In: *IEEE/CVF Conf. Comput. Vis. Pattern Recog.* pp. 8062–8070 (2019)
46. Qian, Y., Pertuz, S., Nikkanen, J., Kämäräinen, J.K., Matas, J.: Revisiting gray pixel for statistical illumination estimation. In: *Int. Joint Conf. Comput. Vision Imag. Comput. Graph. Theory Appl.* vol. 4, pp. 36–46. *INSTICC* (2019)
47. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18. pp. 234–241. Springer (2015)
48. Shapiro, A., Hedjar, L., Dixon, E., Kitaoka, A.: Kitaoka’s tomato: two simple explanations based on information in the stimulus. *i-Perception* **9**(1), 2041669517749601 (2018)
49. Shapiro, A., Lu, Z.L.: Relative brightness in natural images can be accounted for by removing blurry content. *Psychological Sci.* **22**(11), 1452–1459 (2011)
50. Uchikawa, K., Fukuda, K., Kitazawa, Y., MacLeod, D.I.: Estimating illuminant color based on luminance balance of surfaces. *J. Opt. Soc. America A* **29**(2), A133–A143 (2012)
51. Ulucan, D., Ulucan, O., Ebner, M.: Intrinsic image decomposition: Challenges and new perspectives. In: *Int. Conf. Image Process. Vision Eng.* pp. 57–64. *INSTICC*, Prague, Czech Republic (2023)
52. Ulucan, D., Ulucan, O., Ebner, M.: CC-NORD: A camera-invariant global color constancy dataset. In: *Eur. Signal Process. Conf.* pp. 541–545. *IEEE* (2023)
53. Ulucan, D., Ulucan, O., Ebner, M.: Multi-scale surface normal estimation from depth maps. In: *Int. Conf. Image Process. Vision Eng.* pp. 47–56 (2023)
54. Ulucan, O., Ulucan, D., Ebner, M.: BIO-CC: Biologically inspired color constancy. In: *Brit. Mach. Vision Conf.* *BMVA Press* (2022)
55. Ulucan, O., Ulucan, D., Ebner, M.: Block-based color constancy: The deviation of salient pixels. In: *IEEE Int. Conf. Acoust. Speech Signal Process.* pp. 1–5 (2023)
56. Ulucan, O., Ulucan, D., Ebner, M.: Color constancy beyond standard illuminants. In: *IEEE Int. Conf. Image Process.* pp. 2826–2830 (2022)
57. Ulucan, O., Ulucan, D., Ebner, M.: Multi-scale block-based color constancy. In: *Eur. Signal Process. Conf.* pp. 536–540. *IEEE* (2023)
58. Ulucan, O., Ulucan, D., Ebner, M.: Multi-scale color constancy based on salient varying local spatial statistics. *The Vis. Comput.* pp. 1–17 (2023)
59. Ulucan, O., Ulucan, D., Ebner, M.: Investigating color illusions from the perspective of computational color constancy. In: *Int. Joint Conf. Comput. Vision Imag. Comput. Graph. Theory Appl.* *INSTICC* (2024)
60. Van De Weijer, J., Gevers, T., Gijssenij, A.: Edge-based color constancy. *IEEE Trans. Image Process.* **16**, 2207–2214 (2007)
61. Zeki, S.: *A Vision of the Brain*. Blackwell Science, ISBN: 0632030545 (1993)