Learning Color Representations for Low-Light Image Enhancement

Bomi Kim* Sunhyeok Lee* Nahyun Kim Donggon Jang
Dae-Shik Kim
Korea Advanced Institute of Science and Technology
{by5747, sunhyeok.lee, nhkim21, jdg900, daeshik}@kaist.ac.kr

Abstract

Color conveys important information about the visible world. However, under low-light conditions, both pixel intensity, as well as true color distribution, can be significantly shifted. Moreover, most of such distortions are non-recoverable due to inverse problems. In the present study, we utilized recent advancements in learning-based methods for low-light image enhancement. However, while most "deep learning" methods aim to restore high-level and object-oriented visual information, we hypothesized that learning-based methods can also be used for restoring color-based information. To address this question, we propose a novel color representation learning method for low-light image enhancement. More specifically, we used a channel-aware residual network and a differentiable intensity histogram to capture color features. Experimental results using synthetic and natural datasets suggest that the proposed learning scheme achieves state-of-the-art performance. We conclude from our study that inter-channel dependency and color distribution matching are crucial factors for learning color representations under low-light conditions.

1. Introdcution

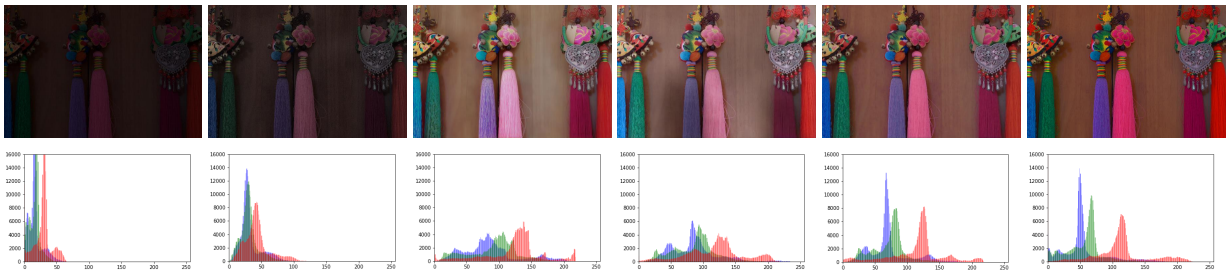
At night or in the shade, we often obtain dark results due to some physical constraints such as slow shutter speed, low aperture brightness, and image sensor sensitivity. When we need these low-light images for a security system or various computer vision systems, we have to repair the dark pictures while maintaining the original content. However, the restoration is a highly ill-posed problem and requires a detailed understanding of underlying image statistics. Many researchers attempt to recover adequate brightness by manipulating the image statistics. The Histogram Equalization (HE) methods [37, 11, 18] increase the contrast by expanding the dynamic range for better visibility. Retinex-based approaches [17, 7, 25] decompose the dark image into reflectance and illumination channels and recover each channel.

Recent learning-based methods attain promising results on high-level visual recognition, verifying the capability of learning visual representations. Along with the recent development of deep learning, many studies suggest deep models for low-light image enhancement [36, 30, 13, 15, 16, 29]. While the low-light degradation causes disruptive changes in pixel intensity and color distribution of the captured photos, most deep learning methods focus on visual features except color representations to densely predict the most probable pixel values. There remains a question for learning-based approaches: Whether the high-order color representations are learnable under low-light conditions?

To address this question, we propose a new color representation learning method for low-light image enhancement, coined *Channel-aware Color histogram Matching method for low-light image enhancement (CCM)*. We aim to capture the inter-channel dependencies and informative channels for learning color representations for the darkness restoration. To this end, we present the layer normalization integrated channel-attention block (γ-Channel-attention Residual Block). The γCRB consists of convolutional layers, layer normalization layers, activation functions, and following channel attention block. We use the re-scaling factor of the layer normalization, γ, as a scaling factor for the global average pooled vectors of the channel attention block.

To consider the massive changes in color distribution, we propose a new color distribution loss by utilizing a differentiable color histogram. We apply the Gaussian filters to segregate the high and low-frequency parts of the estimated and ground-truth images, minimizing the interference of edge and color representations. For recovering the original color distribution, we utilize the Kernel Density Estimation to approximate the differentiable color histogram and measure the absolute distance between the estimated color histograms of low-frequency segments. To restore the edge details, we measure the Structural Similarity [28] between

*These authors contributed equally.



(a) Input (b) CLAHE (c) KinD (d) DRBN (e) **CCM** (f) Ground-truth

Figure 1. A visual illustration of previous approaches and our method for the low-light enhancement on real-world dataset [29]. The upper row shows the RGB images, and the lower row demonstrates the corresponding (discrete) color histograms. The prediction result of our method successfully recovers the original color histogram while other methods generate incomplete restoration results.

the high-frequency areas.

Experimental results validate that our method effectively learns color representations under low-light conditions. Subsequently, our model consistently outperforms the comparison methods on benchmark datasets regardless of synthetic domain or real-world domain. The significant performance improvements verify that channel level dependency and constraint on color histogram, previously ignored, are pivotal factors for low-light image enhancement.

The contributions of the proposed CCM are as follows:

- We observe that CCM learns high-order color representations, and those are crucial factors for low-light image enhancement.
- We figure out that the inter-channel dependency is decisive information for color representation learning under low-light conditions of substantial changes in color distributions.
- We propose a new color distribution loss for color representation learning and suggest a specialized kernel function for improved differential histogram estimation.

2. Related Work

Low-light image enhancement The classical methods perform low-light enhancement via Histogram Equalization (HE) or Retinex-based approaches. HE methods [37, 11, 18] increase the contrast of images by extending the dynamic range at both global and local levels. The limitation of HE-based methods is evident in that they ignore the information about the image structure. Retinex-based approaches [17, 7, 25] decompose the images into reflectance and illumination maps and adjust the illumination maps. Our method segregates high and low-frequency parts of RGB images and utilizes each component to restore all the visual features and color distributions.

Recent deep learning-based methods show promising results, yet they disregard color distributions of the images and dependency between channels of the feature maps. KinD [36] and Retinex-Net [29] estimate reflectance and illumination maps using convolution layers for low-light enhancement. To overcome the cost of requiring a large amount of paired data, EnlightenGAN [13] proposes GAN-based architecture to utilize unpaired datasets. DRBN [30] suggests a two-staged learning method to enhance perceptual quality by utilizing additional unpaired high-quality data.

We propose a new color representation learning method under low-light conditions, suggesting a channel attention block combined with normalization and color distribution loss.

Channel dependency and sparsity The normalization methods of deep neural networks assist better convergence by mitigating the internal covariate shift and acting as a regularizer. The main differences among the proposed normalization methods are the dimension of the normalization. For example, Batch Normalization [12] normalizes each channel of a mini-batch, Layer Normalization [3] applies to every channel of an instance, and Instance Normalization [26] normalizes each channel of one instance.

Attention mechanism helps the network preserve informative features and obtain more precise results. Recently, many deep models apply attention mechanisms to various tasks, including image generation [33], image classification [9], and image restoration [34]. SENet [9] and ECA-Net [27] propose a method for modeling the relationship between channels in image classification. In the image restoration task, RCAN [34], HAN [23], and RESCAN [20] improve the performance by focusing on more relevant channels of feature maps.

Our method exploits layer normalization and a channel attention mechanism together for learning color representations.

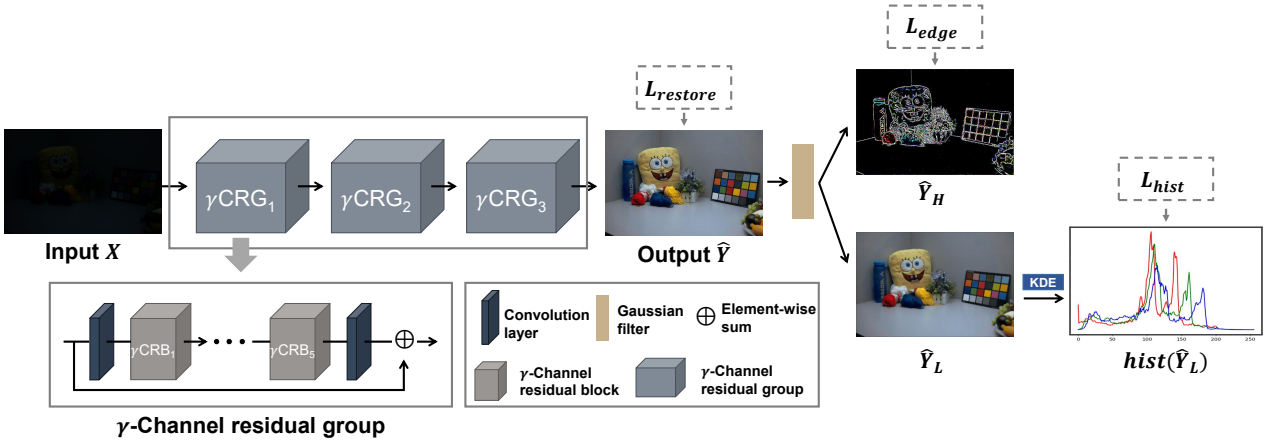


Figure 2. An overview of the proposed CCM. Our method utilizes the inter-channel dependency through the LN integrated channel attention structure. We estimate the differentiable color histogram of the low-frequency component and use the histogram distance from the smoothed ground-truth image as a color histogram loss. The edge recovering loss and the color distribution loss also assist in learning color and other visual representations together.

3. Method

In this section, we describe our methodology for color representation learning. Figure 2 represents the overall construction of the proposed method. We use a deep residual network as a base structure, well known that learns visual features for image restoration. We integrate the a layer normalization technique and a channel attention block to manipulate the channel dependency, a viable factor for color representation. Also, we employ low-pass filters and differentiable histogram estimation for learning color and other visual attributes at once.

3.1. Network architecture

Residual network as a base network Deep residual networks yield promising image restoration results by learning visual representations [21, 35, 24]. We also use a residual block with a skip connection as our basic unit. The Figure 3 represents our specialized residual block, γ Channel Residual Block (γCRB). We build our residual group with the γCRB, convolution layers, and a prolonged skip connection. The first layer of our residual group converts the input three-channel RGB low-light image into a 48-channel feature map using a convolutional layer. After that, the feature map goes through five consecutive residual blocks (all the feature map dimensions kept consistent as $48 \times 100 \times 100$). The last layer of the residual group generates an RGB-channel image. We stack three residual groups to construct our network.

Channel dependency for color representation The normalization methods of deep neural networks assist better convergence by mitigating the internal covariate shift and acting as a regularizer [12, 3]. Table 2 shows the comparison results of various normalization methods. Batch Normalization (BN) shows robust performance in high-level vision tasks such as classification, but we figure out that BN is fruitless in low-light enhancement tasks. The probable reason is that BN performs normalization independent of each channel, and we assume that channel dependency is a pivotal factor in learning the color representations. The Layer Normalization (LN), on the other hand, computes over all channels within an instance (eq. 1). Thus the channel dependency can propagate through deeper layers.

$$h^l = \gamma(\frac{a^l - \mu^l}{\sigma^l}) + \beta \tag{1}$$

where $\mu^l = \frac{1}{H}\sum_{t=1}^{H} a_t^l$ and $\sigma^l = \sqrt{\frac{1}{H}\sum_{t=1}^{H}(a_t^l - \mu^l)^2}$. H is the number of hidden unit in the layer, γ and β denote the scale and shift parameter. LN learns the γ and β which are the same dimension as the channel.

Using LN prevents covariate shift while maintaining channel dependency, viable in the reverse mapping of low to normal color distribution. We attach the LN to every convolutional layer in the RB to manipulate the channel dependency for learning color representations. We integrate the LN and the Channel Attention Block [35] to consider channel dependency and boost channel sparsity for enhanced color representations. Figure 3 depicts our γCRB. We multiply the re-scaling factor γ obtained from the LN layer to the average activation vector computed by the global average pooling. With this step, our residual block can consider both dependency and saliency of all channels. We empirically decide to use the γ of the second layer among the two LNs of the RB.

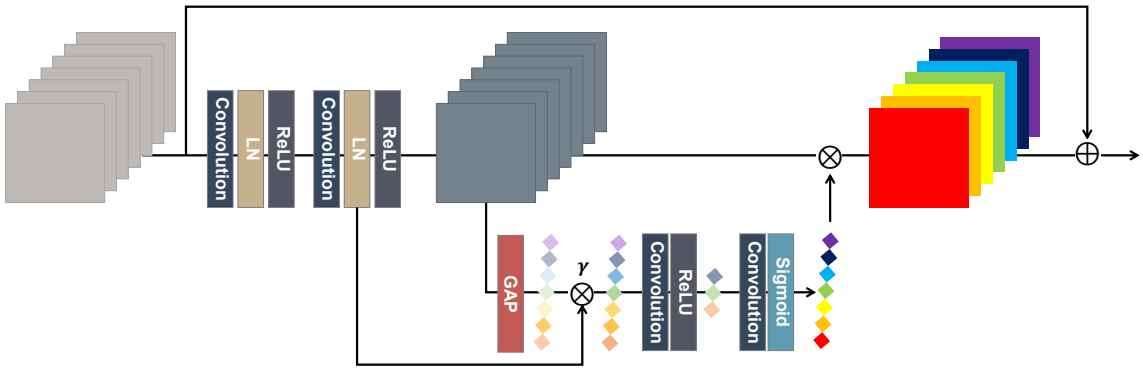


Figure 3. An illustration of the proposed γCRB. We figure out that channel dependency is a pivotal factor in low-light image enhancement. To consider the channel dependency, we use the residual block augmented with LN and channel attention. The re-scaling factor γ from the second LN multiplies to the globally averaged activation vector.

3.2. Learning color and other visual representations

In Figure 1, the color histograms of the restored results demonstrate that the previous deep learning methods disregard color representations. We assume this color histogram carries crucial underlying image statistics and is profitable for learning color representations. To use the histogram for parameter optimization, we approximate the differentiable histogram through the Kernel Density Estimation (KDE).

Differentiable histogram estimation Differentiable estimation is required to use the color histogram for model parameter training, and we utilize KDE for the estimation. According to Avi-Ahron *et al*. [2], KDE produces adequate differentiable histograms for model optimization.

Considering RGB image space consists of continuous intensity values in [0,1], we define the pixel intensity of an image pixel $x \in P$ as $I(x) \in [0, 1]$. We use the KDE for estimating the density f_I for each RGB channel I as follows:

$$\hat{f}_I(i) = \frac{1}{nh} \sum_{x \in \Pi} \kappa\left(\frac{I(x) - i}{h}\right) \tag{2}$$

where $i \in [0, 1]$, $\kappa(\cdot)$ is the kernel, h is the bandwidth and $n = |\Pi|$ is the number of pixels. Since the RGB color histograms are discrete and spiky, we assume that using the derivative of a function close to the unit step function as a kernel will be more suitable for estimating the color histograms.

$$\kappa(z) = \frac{d}{dz} s(z), \tag{3}$$

where $s(z) = \frac{1}{1+\exp(-\alpha \cdot z)}$. As the estimation error converges when α exceeds 1000, we set it as 1000. We split the interval $[0, 1]$ into K sub-intervals h_k. The length of each interval L is $\frac{1}{K}$ and μ_k is $-1 + L(k + \frac{1}{2})$, and h_k is

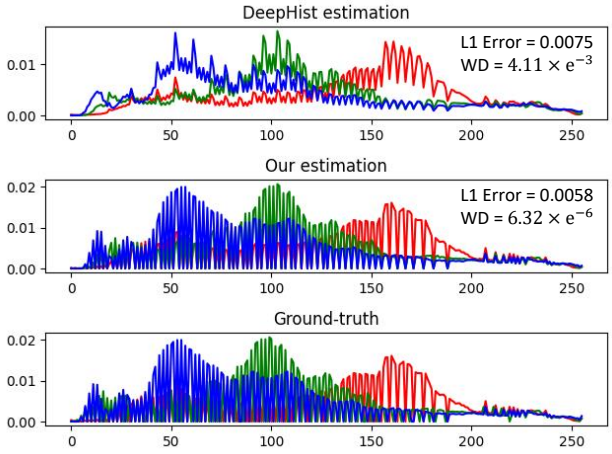


Figure 4. Comparison of differentiable color histogram estimation. The upper graph shows the estimated histogram by DeepHist [2] and the second graph is the approximation of our method. The last graph is the target (discrete) color histogram. Using the derivative of the step-shaped function as the kernel, we achieve more precise estimations, still differentiable.

$[\mu_k - \frac{L}{2},\ \mu_k + \frac{L}{2}]$. We define the probability $P_I(K)$ that a pixel in the image belongs to a particular bin.

$$P_I(k) = Pr(i \in h_k) = \int_{h_k} \hat{f}_I(i) di \tag{4}$$

Since the kernel function is the derivative of $s(z)$, eq. 4 can be easily integrated. Finally we calculate the function $P_I(k)$ which is the value of the k^{th} bin. We define the differentiable histogram $\mathbf{h}$ as follows:

$$\mathbf{h} = \{\mu_k, P_I(k)\}_{k=0}^{K-1} \tag{5}$$

Figure 4 illustrates the estimated histograms and the actual histogram. Our estimation has significantly lowered estima-

tion errors, verifying our assumption on the kernel function. We propose a method for learning the color representation by combining this estimated histogram vector to the loss function.

Low-pass filtering and histogram matching First, we separate the low-frequency and high-frequency regions through a Gaussian filter to segregate the color and shape representation. We empirically decided the Gaussian kernel size as three. After estimating the differentiable color histogram of the low-frequency regions, we measure the L1 distance between the differentiable histogram vectors of the restored results and ground-truth images.

$$\begin{aligned} Y_L &= GaussianBlur(Y) \\ Y_H &= Y - Y_L \end{aligned} \tag{6}$$

We also focus on other vital visual representations of shape. We calculate the Structural Similarity [28] between the images in the high-frequency region and use it as a loss function. As with the previous approaches, we use the L1 distance in the RGB pixel space as a loss function to learn overall spatial representations for the restoration task. The loss functions for training the proposed model are as follows:

$$\begin{aligned} L &= L_{restore} + \lambda_1 \cdot L_{hist} + \lambda_2 \cdot L_{edge} \\ &= \|Y - \hat{Y}\|_1 + \lambda_1 \cdot \|hist(Y_L) - hist(\hat{Y}_L)\|_1 \\ &\quad - \lambda_2 \cdot SSIM(Y_H, \hat{Y}_H) \end{aligned} \tag{7}$$

where Y is ground-truth, $\hat{Y}$ is the restored result, λ_1 and λ_2 are the loss weights and both are empirically set as 1. $hist(\cdot)$ is the estimated color histogram matrix of $[3 \times 256]$.

4. Experiments

To empirically validate that the proposed method can learn color representations while restoration, we compare CCM with previous approaches for low-light image enhancement, including CLAHE [37], BPDHE [11], Dong [5], SRIE [7], DHECE [22], MF [6], EFF [31], CRM [32], LIME [8], JED [25], RRM [19], Retinex-Net [29], KinD [36], EnlightenGAN (EG) [13], and DRBN [30]. We reproduce all the other methods with their original codes and settings to compare. We report the best performance for the comparison table.

Our method consistently outperforms all the other approaches on various datasets (Table 1) in terms of PSNR, SSIM of the images, and Wasserstein Distance (WD) [1] of color histograms. The considerable margin between CCM and comparison methods suggests that color representation could be a credible factor in improving the overall performance of deep learning methods for image restoration.

4.1. Comparison results

Datasets We compare our model with other state-of-the-art methods on synthetic and real-world paired datasets [29]. Wei *et al*. [29] collect 1000 raw images from RAISE [4] and generate the synthetic dataset by adjusting the histogram of the Y channel. We split the 1000 image pairs of the synthetic dataset into 900 training and 100 testing pairs. The original real-world dataset [29] consists of 485 image pairs for training and 15 images for testing. Since the test dataset contains only a few image pairs, we randomly choose 85 more image pairs from the training dataset and enlarge the test dataset to become 100 pairs. Comparison results for the original train/test split(485/15) are given in Appendix C.

Metrics We evaluate PSNR [10] and SSIM [28] on the Y channel of transformed YCbCr space for quantitative evaluation. In addition, Wasserstein Distance (WD) [1] is used to measure the quantitative value of the color histogram between the output image and the ground truth.

Other details For detailed experimental settings, please refer to Appendix A.

Main results In Table 1, we provide the quantitative evaluation results of our model and other comparison methods. CCM achieves the best performance; in both synthetic and real-world domains, CCM consistently outperforms other comparison methods with considerable margins in terms of PSNR, SSIM, and WD. In particular, CCM gains at least *+3.15dB* and *+1.12dB* of PSNR margins on the synthetic and real-world domains, respectively. Furthermore, CCM effectively recovers the structural complexity and achieves at least *+2.3%* and *+4.0%* improvements in terms of SSIM on the synthetic and real-world datasets, respectively. The closest WD confirms that CCM learns color representations and effectively recovers the original color distribution.

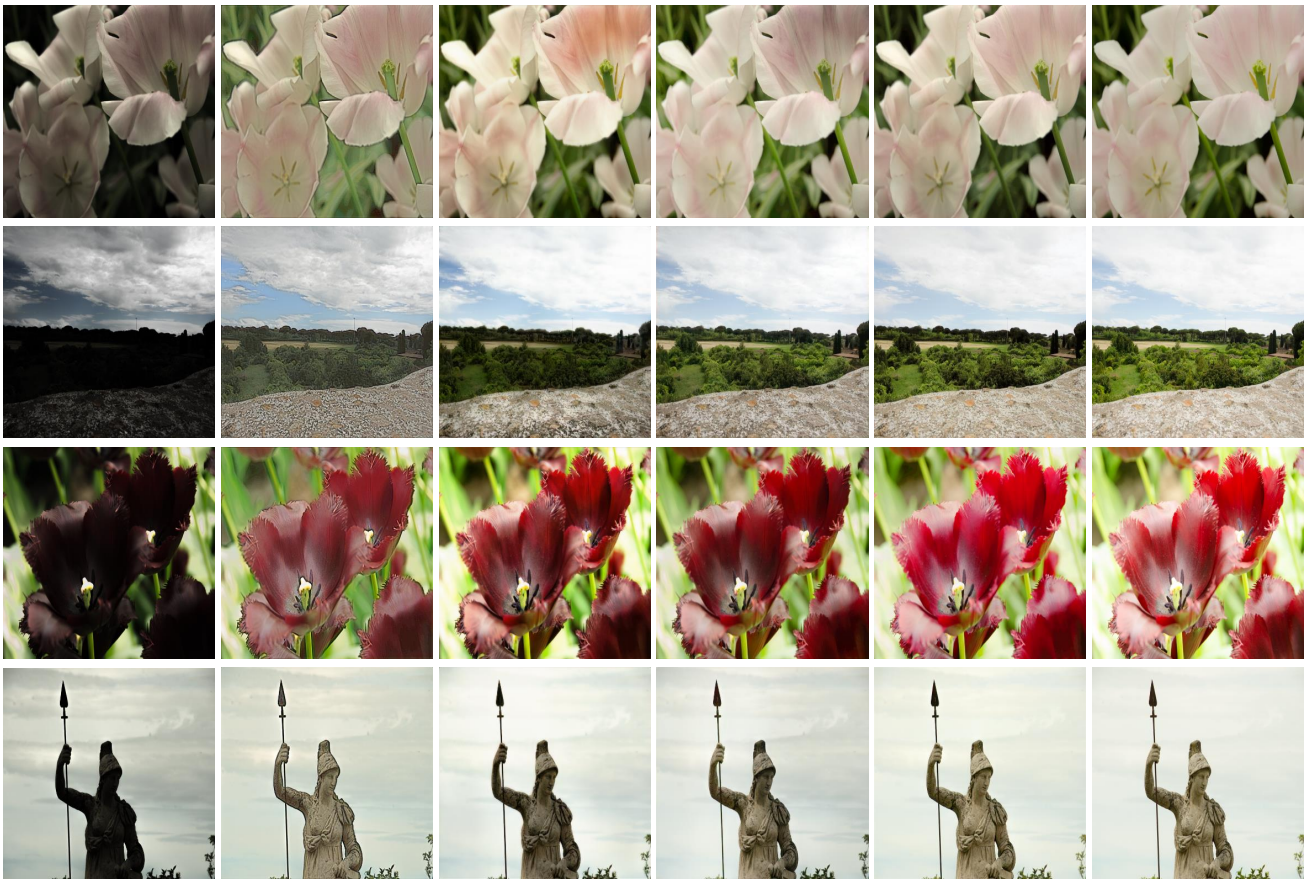
Figure 5 and 6 show the qualitative comparison results on synthetic datasets and real datasets. Previous methods tend to underexpose the images and fail to capture the color distribution of the input images. In Figure 5, other approaches have apparent differences between restoration outputs and ground-truth, especially in the petals, ground, and statue, while CCM restores the closest color distribution, shape, texture to the original without artifacts. In Figure 6, similar to Figure 5, our method generates the most natural restoration results. For more qualitative results, please refer to Appendix D.

4.2. Ablation study

We evaluate our base network and each component of CCM (Table 2 and Figure 7). When we use LN with base

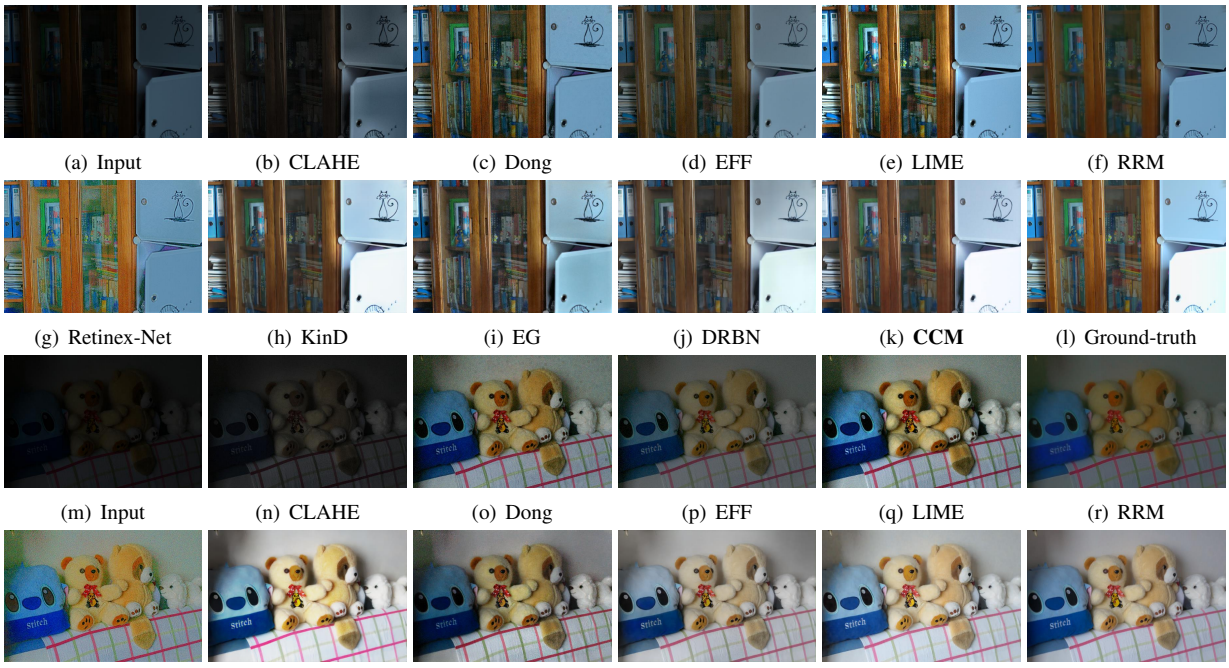
		CLAHE	BPDHE	Dong	SRIE	DHECE	MF	EFF	CRM
Synthetic	PSNR	12.58	12.50	17.02	14.54	18.14	17.75	17.9	19.83
	SSIM	0.5604	0.5771	0.7539	0.6107	0.8157	0.7916	0.8096	0.8733
	WD	971.0	1073.0	585.6	760.6	931.9	560.6	691.4	1126.7
		LIME	JED	RRM	Retinex-Net	KinD	EG	DRBN	**CCM**
	PSNR	17.67	17.05	17.31	18.50	22.34	17.84	<u>23.52</u>	**26.67**
	SSIM	0.7935	0.7507	0.7471	0.8274	0.9203	0.8192	<u>0.946</u>	**0.9673**
	WD	693.0	655.0	639.0	541.4	413.9	479.9	<u>328.2</u>	**291.5**
		CLAHE	BPDHE	Dong	SRIE	DHECE	MF	EFF	CRM
Real-world	PSNR	8.56	11.60	15.91	11.28	17.11	16.79	13.63	16.94
	SSIM	0.3244	0.3601	0.5503	0.5299	0.4902	0.5876	0.6450	0.6968
	WD	2287.1	1692.0	935.1	1204.0	1386.6	957.9	1470.7	3053.2
		LIME	JED	RRM	Retinex-Net	KinD	EG	DRBN	**CCM**
	PSNR	16.89	13.12	13.15	17.28	<u>23.00</u>	17.11	21.27	**24.12**
	SSIM	0.5620	0.6778	0.6791	0.5109	<u>0.8993</u>	0.7164	0.8983	**0.9351**
	WD	1173.0	1304.8	1295.1	961.2	880.5	974.1	<u>779.8</u>	**760.7**

Table 1. The quantitative comparison results with other state-of-the-art methods on the synthetic and real-world datasets [29]. The proposed method achieves the new state-of-the-art performance on both synthetic and natural datasets. **Bold** and <u>underline</u> indicate the best and the second best score, respectively.



(a) Input (b) Retinex-Net (c) KinD (d) DRBN (e) **CCM** (f) Ground-truth

Figure 5. Qualitative evaluation results on the synthetic dataset [29].



(a) Input (b) CLAHE (c) Dong (d) EFF (e) LIME (f) RRM

(g) Retinex-Net (h) KinD (i) EG (j) DRBN (k) **CCM** (l) Ground-truth

(m) Input (n) CLAHE (o) Dong (p) EFF (q) LIME (r) RRM

(s) Retinex-Net (t) KinD (u) EG (v) DRBN (w) **CCM** (x) Ground-truth

Figure 6. Qualitative evaluation results on the real-world dataset [29].

	Norm.	γCAB	L_{edge}&L_{hist}	PSNR	SSIM	WD
Base RN	-	-	-	19.21	0.8582	903.55
RN + BN [12]	BN	-	-	19.28	0.8803	851.76
RN + IN [26]	IN	-	-	18.69	0.7672	1290.83
RN + LN [3]	LN	-	-	22.94	0.9171	769.79
RN + L_{edge}+L_{hist}	-	-	+	19.75	0.9017	878.99
RN + γCRB	LN	+	-	23.31	0.9212	753.60
CCM	LN	+	+	**24.12**	**0.9351**	**745.68**

Table 2. Ablation study on real-world domain dataset [29].

RN, we achieve considerable performance gain. As we described in Section 3, LN supports learning color representations through channel-dependent normalization. On the other hand, using BN is fruitless, and Instance Normalization harms the performance. These results necessitate channel dependency for color representation learning and low-light image enhancement.

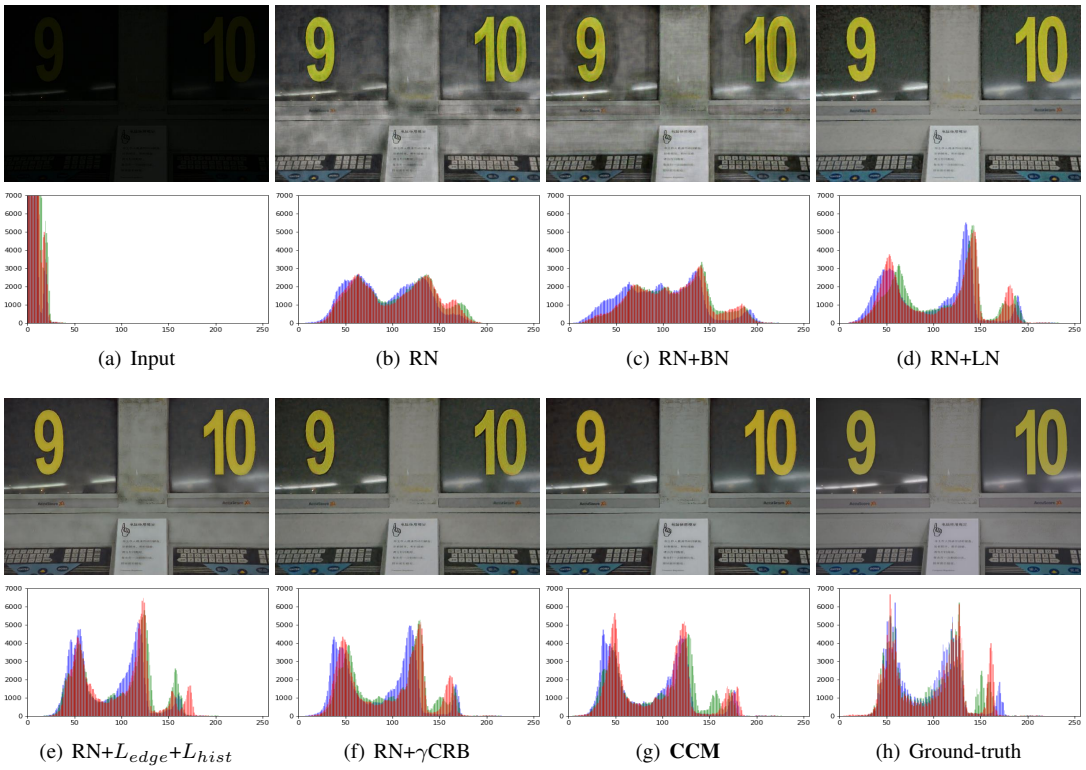
When we use LN and γCAB together, the model gains further performance gain because it manipulates the channel dependency of the intermediate feature maps and helps to improve the channel sparsity.

L_{edge} and L_{hist} enhance the performance of our base RN. In particular, the model attains remarkable improvements in SSIM, and the color histogram of the estimated image gets way closer to the ground-truth image. This result validates the suitability of the training scheme for learning color and shape representations together.

Our final model achieves the best performance of low-light image enhancement, and all the components contribute in a complementary way to assist in learning color representation and other visual representations together.

5. Conclusion

In this study, we verified that high-order color representation under low-light conditions is learnable through deep parametrization. Experimental results validated that learning color representation is a linchpin of the low-light image enhancement. We also figured out the importance of channel dependency throughout the normalization of feature maps and proposed to use LN with channel attention weights. We suggested constructing a histogram matching loss by leveraging differential histogram estimation and low-pass filters. Our CCM acquired the new state-of-the-art performance while effectively learning color representa-



(a) Input (b) RN (c) RN+BN (d) RN+LN

(e) RN+L_{edge}+L_{hist} (f) RN+γCRB (g) **CCM** (h) Ground-truth

Figure 7. Sample results of the ablation study on the real-world dataset [29]. Each component of our method assists in learning color representations under low-light conditions.

tions by utilizing the proposed channel-aware structure and color distribution loss.

Future work Our approach aims to learn color representations in a supervised manner, yet the supervised learning methods have the inherent problem of requiring paired datasets of high annotation costs. The subsequent issue is that insufficient data highly likely generate the bias of the training. Thus we plan to extend our work on semi-supervised and unsupervised learning schemes with the out-of-distribution problems.

6. Acknowledgement

This research has been supported by the SHINSEGAE I&C Corporation. (Project No. G01210148).

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International conference on machine learning*, pages 214–223. PMLR, 2017.

[2] Mor Avi-Aharon, Assaf Arbelle, and Tammy Riklin Raviv. Deephist: Differentiable joint and color histogram layers for image-to-image translation. *arXiv preprint arXiv:2005.03995*, 2020.

[3] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016.

[4] Duc-Tien Dang-Nguyen, Cecilia Pasquini, Valentina Conotter, and Giulia Boato. Raise: A raw images dataset for digital image forensics. In *Proceedings of the 6th ACM Multimedia Systems Conference*, pages 219–224, 2015.

[5] Xuan Dong, Guan Wang, Yi Pang, Weixin Li, Jiangtao Wen, Wei Meng, and Yao Lu. Fast efficient algorithm for enhancement of low lighting video. In *2011 IEEE International Conference on Multimedia and Expo*, pages 1–6. IEEE, 2011.

[6] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, 129:82–96, 2016.

[7] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2782–2790, 2016.

[8] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26(2):982–993, 2016.

[9] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.

[10] Quan Huynh-Thu and Mohammed Ghanbari. Scope of validity of psnr in image/video quality assessment. *Electronics letters*, 44(13):800–801, 2008.

[11] Haidi Ibrahim and Nicholas Sia Pik Kong. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(4):1752–1758, 2007.

[12] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015.

[13] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30:2340–2349, 2021.

[14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[15] Dokyeong Kwon, Guisik Kim, and Junseok Kwon. Dale: Dark region-aware low-light image enhancement. *arXiv preprint arXiv:2008.12493*, 2020.

[16] Mohit Lamba, Atul Balaji, and Kaushik Mitra. Towards fast and light-weight restoration of dark images. *arXiv preprint arXiv:2011.14133*, 2020.

[17] Edwin H Land. The retinex theory of color vision. *Scientific american*, 237(6):108–129, 1977.

[18] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE transactions on image processing*, 22(12):5372–5384, 2013.

[19] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018.

[20] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 254–269, 2018.

[21] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 136–144, 2017.

[22] Keita Nakai, Yoshikatsu Hoshi, and Akira Taguchi. Color image contrast enhacement method based on differential intensity/saturation gray-levels histograms. In *2013 International Symposium on Intelligent Signal Processing and Communication Systems*, pages 445–449. IEEE, 2013.

[23] Ben Niu, Weilei Wen, Wenqi Ren, Xiangde Zhang, Lianping Yang, Shuzhen Wang, Kaihao Zhang, Xiaochun Cao, and Haifeng Shen. Single image super-resolution via a holistic attention network. In *European Conference on Computer Vision*, pages 191–207. Springer, 2020.

[24] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3937–3946, 2019.

[25] Xutong Ren, Mading Li, Wen-Huang Cheng, and Jiaying Liu. Joint enhancement and denoising method via sequential decomposition. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 1–5. IEEE, 2018.

[26] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016.

[27] Q Wang, B Wu, P Zhu, P Li, W Zuo, and Q Hu. Eca-net: Efficient channel attention for deep convolutional neural networks, 2020 ieee. In *CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE*, 2020.

[28] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[29] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018.

[30] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3063–3072, 2020.

[31] Zhenqiang Ying, Ge Li, Yurui Ren, Ronggang Wang, and Wenmin Wang. A new image contrast enhancement algorithm using exposure fusion framework. In *International Conference on Computer Analysis of Images and Patterns*, pages 36–46. Springer, 2017.

[32] Zhenqiang Ying, Ge Li, Yurui Ren, Ronggang Wang, and Wenmin Wang. A new low-light image enhancement algorithm using camera response model. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 3015–3022, 2017.

[33] Han Zhang, Ian Goodfellow, Dimitris Metaxas, and Augustus Odena. Self-attention generative adversarial networks. In *International Conference on Machine Learning*, pages 7354–7363. PMLR, 2019.

[34] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 286–301, 2018.

[35] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)*, pages 286–301, 2018.

[36] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM International Conference on Multimedia*, pages 1632–1640, 2019.

[37] Karel Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems*, pages 474–485, 1994.