LightNet: Generative Model for Enhancement of Low-Light Images

Chaitra Desai Nikhil Akalwadi Amogh Joshi Sampada Malagi
Chinmayee Mandi Ramesh Ashok Tabib Ujwala Patil Uma Mudenagudi
Center of Excellence in Visual Intelligence (CEVI), KLE Technological University
Hubballi, Karnataka, INDIA
{chaitra.desai, nikhil.akalwadi, ramesh_t, ujwalapatil, uma}@kletech.ac.in
{joshiamoghmukund, sampadamalagi12, chinmayeemandi2001}@gmail.com

Abstract

In this work, we propose a generative model for enhancement of images captured in low-light conditions. Sensor constraints and inappropriate lighting conditions are accountable for degradations introduced in the image. The degradations limit the visibility of the scene and impedes vision in applications like detection, tracking and surveillance. Recently, deep learning algorithms have taken a leap for enhancement of images captured in low-light conditions. However, these algorithms fail to capture information on fine grained local structures and limit the performance. Towards this, we propose a generative model for enhancement of low-lit images to exploit both local and global information, and term it as LightNet. In proposed architecture LightNet, we include a hierarchical generator encompassing encoder-decoder module to capture global information and a patch discriminator to capture fine grained local information. Typically, the encoder-decoder module downsamples the low-lit image into distinct scales. Learning at distinct scales helps to capture both local and global features thereby suppressing the unwanted features (noise, blur). With this motivation, we downsample the captured low-lit image into 3 distinct scales. The decoder upsamples the encoded features at respective scales to generate an enhanced image. We demonstrate the results of proposed methodology on custom and benchmark datasets in comparison with SOTA methods using appropriate quantitative metrics.

1. Introduction

In this paper, we propose a hierarchical generative model for enhancement of images captured in low-light conditions. We consider both global and local features to model the proposed framework. Enhancement of images captured in low-light conditions is the need of the hour as it contributes in expediting the vision for a wide range of applications. Applications like autonomous driving systems, traffic surveillance, wildlife photography, drone surveillance, underwater coral reef monitoring and protection demands the enhancement of low-light images for clear vision. However, the images captured in low-light conditions undergo several degradations limiting the visibility of the scene.

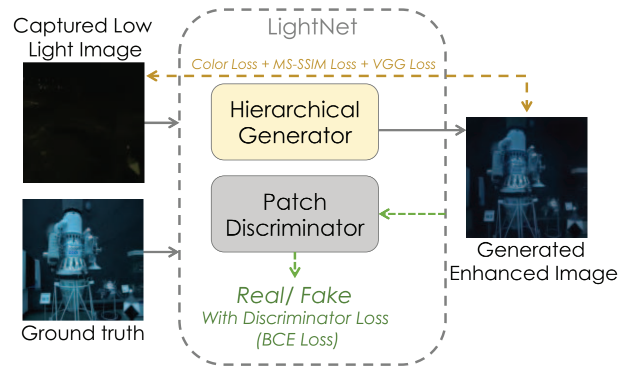


Figure 1: LightNet: Framework for enhancement of images captured in low-light conditions.

Most prominent degradations observed in low-light conditions [31] include loss of color, contrast, and implicit spatial noise. Spatial noise is introduced due to variation in each individual pixel or variation across the pixels suppressing the high-frequency components causing blur in the captured scene. Advancements in capturing technology facilitate to overcome these degradations, however deploying these sensors on edge devices is challenging due to its size and memory constraints. Towards this, we propose to enhance the quality of images captured in low-light conditions considering the variation per pixel (local information) and across the pixels (global information) with a combinational loss function.

Typically, the image captured is a function of its illuminance $Il(x, y)$ and reflectance $R(x, y)$ and is given by:

$$T(x, y) = f(Il(x, y), R(x, y)) \tag{1}$$

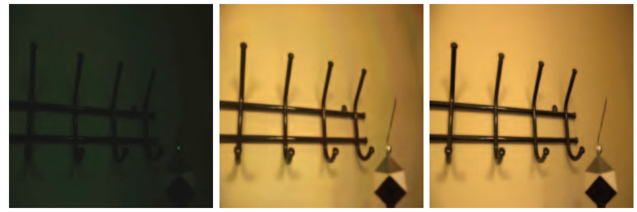
Here $T(x, y)$ is the true observation, f is a function of illuminance $Il(x, y)$ and reflectance $R(x, y)$ respectively.

In traditional methods, the incident light and reflectance is jointly estimated from a single true observation. Authors in literature propose Histogram Equalization [15] [1] and Retinex methods [27] [23] for enhancement of low-light images. However, Retinex methods use minimization framework to iteratively enhance illumination and reflectance component for each pixel in an image. Histogram Equalization methods focus on improving the overall contrast of an image resulting in unnatural colors. These methods possess limitations like assumptions on priors and are data specific. The methods do not explicitly consider noise introduced during the capture, and generally apply denoising as a post-processing module.

Authors in [24] perform simultaneous contrast enhancement with denoising using a stacked auto-encoder module. However, the method doesn't appreciate the true power of deep learning architectures. Authors in [10] propose three sub-modules namely feature extraction module, enhancement module, and fusion module (MBLLEN) and propose a variant of LLNet. The results of the feature extraction and enhancement module are fused towards enhancement of low-light images. [10] provides improved performance over LLEN with explicit feature extraction module necessary for enhancement of low-light images.

Authors of [10] propose another light-weight convolutional neural network to handle non-uniform illuminations [25]. [28] propose a hybrid network with content stream and edge stream to recover the content and high frequency components. [34] propose a fusion network to address high contrast and color biases introduced in low-light conditions. Authors in [2] propose a convolutional neural network for the enhancement of low-light images using raw images. Unlike others, we consider extremely dark images given by Night Rendering Photography, NTIRE 2022 challenge dataset towards training the proposed architecture as shown in Figure 2 and demonstrate the quality of enhancement with appropriate quantitative metrics.

The methods from literature, learn mapping directly between ground-truth data and corresponding low-light images. Alternatively, Retinex theory based methods enhance the illuminance and reflectance component independently with dedicated subnetworks. Authors in [32] extend the work of RetinexNet with new constraints and propose a novel architecture for enhancement of low-light images. Although most of the methods achieve considerable improvement in performance, the generalisation issue still persists due to the use of synthetically generated data. To address this, we propose a novel dataset consisting of real low-light



a) b) c)

Figure 2: Enhancement of images captured in low-light conditions. a) Input image from NTIRE 2022 Challenge Dataset [8]. b) Result of proposed methodology. c) ground-truth. (Note: For illustration purpose we have added a fixed constant to each pixel of input image.)

images captured with varying ISOs and exposures. We also release the corresponding ground-truth images. The captured low-light images coupled with ground-truth information facilitates to train deep learning frameworks towards enhancement of low-light images.

There is a paradigm shift with deep learning frameworks in terms of computational speed vs enhancement quality. However, trade-off between computational requirement and quality is the most crucial aspect of any algorithm design as it depends on the underlying application it is intended for. Authors in [18] propose a light weight model for the enhancement of low-light images. The main goal of our work is to improve the quality of enhancement irrespective of the computational time. Unlike deep Retinex methods, we intend to improve color and contrast of the low-light images without explicitly separating illuminance and reflectance parameters. Unlike authors in [20] we propose an architecture emphasizing on local and global features.

Contributions of the work include,

- We prepare customised low-light dataset, captured with varying ISOs and Exposures along with corresponding ground-truth information to train deep learning algorithms. (Section 2).
- We propose a hierarchical generative model with patch GAN to capture local information explicitly for low-light conditions. (Section 3).
- We propose a combinational loss function to exploit local illuminance keeping global features intact.
  - We consider VGG 19 loss to preserve the overall contextual features of the high lit image.
  - We consider MS-SSIM loss to preserve channel-wise structural information.
  - We consider Color loss to measure the color loss between the generated image and the ground-truth image.

- We demonstrate the results of proposed LightNet on NTIRE 2022 challenge dataset and our custom low-light dataset using appropriate quality metrics. (Section 4).t

2. Dataset Preparation: Captured Low-light Images

In this section, we discuss on the proposed dataset prepared for training the architecture towards enhancement of images captured in low-light conditions. From Retinex theory, we infer illuminance is the primary factor contributing for the visibility of the objects in the scene. However, the colors of each object in the scene vary under different illuminations leading to inconsistency of colors in the captured image. With this motivation, we vary the exposure settings in the camera to capture the change in color distribution of the scene as shown in Figure 6. We train the proposed architecture LightNet, with the proposed dataset to learn the color distribution under varying camera settings for consistent enhancement of images irrespective of illumination conditions.

2.1. Dataset Description

Images were captured with Samsung Galaxy Note 10+, for multiple scenes under varying camera settings. The camera settings for multiple scenes include keeping ISO constant under varying exposures. The ISO settings for the scene vary ranging from $ISO50$ to $ISO3200$ and exposure range from $1/2400s$ to $1/2s$. We capture the corresponding ground-truth information in auto settings of the camera. The dataset includes a total of 3500 raw image pairs and its corresponding $sRGB$ image pairs for varying scenes under varying camera settings. The few samples from the capture are shown in Figure 5.

3. LightNet: Generative Model for Enhancement of Low-Light Images

In this section, we provide the architectural details of the proposed methodology (LightNet) and we discuss in detail on the proposed combinational loss function.

3.1. Network Architecture

Encoder-decoder based architectures like UNET [29] are widely used for reconstruction of images. Encoder-decoder based architectures downsamples the input images into different scales and facilitate learning at different levels. Learning at different scales helps to capture the local and global variance of features thereby suppressing the unwanted features (noise, blur).

In the proposed architecture, we have hierarchical generator consisting of encoder and decoder block. At encoder we encode the input image in three distinct scales i.e., Encoding at Lower-Scale, Encoding at Medium-Scale, Encoding at Higher-Scale as shown in Figure 3. Encoder at Lower-Scale downsamples the input image by a factor of 2, Encoder at Medium-Scale downsamples the input image by a factor of 8, Encoder at Higher-Scale downsamples the input image by a factor of 32 as shown in Figure 3. We propose modified residual dense blocks (MRDB) as shown in Figure 4. The MRDB block facilitates learning of local features emphasising on fine grained structural information and minuet features. We include MRDB block in Encoder at Higher-Scale as it contains information of all the scales in the hierarchy. At decoder, the output of MRDBs is fused at two levels Decoding at Medium-Scale and Decoding at Lower-Scale to generate the enhanced low-light image.

Only encoder-decoder based architectures [18], [6], [5], [7], may not suffice the problem of low-light image enhancement as restoring true colors with such models is quite challenging. Towards this, we propose a new loss function to capture the lost colors as shown in Equation 5. The proposed encoder-decoder architecture performs downsampling and upsampling of the input image as shown in Figure 3. We include a patch-based discriminator to capture local color and contrast from each patch facilitating improved color and contrast reconstruction both locally and globally. Patch discriminator learns the local distribution of contextual and spatial information and resulting feedback is given to generator. Towards this, we propose a combinational loss function to capture local color, contrast, and content features.

In what follows, we discuss in detail the proposed combinational loss function. Unlike the authors in [1], the proposed methodology depicting a generator with a corresponding patch discriminator ensures the retention of local and global features is shown in Figure 3.

3.2. Proposed Combinational Loss Function

Images captured in low-light conditions are sensitive to spatial noise, color and contrast demanding modelling of a loss function to restore true colors, contrast, and noise. Towards this, we propose a combinational loss function to recover the lost color and contrast both locally and globally. Most commonly used loss functions for enhancement of images captured in low-light conditions include perceptual loss, smoothness loss, and reconstruction loss. However, these loss functions emphasize more on content features thereby ignoring the underlying color, and contrast component.

The proposed combinational loss function aims to recover lost colors, contrast, and structural information both locally and globally with the help of patch-based discriminator. The combinational loss function includes 3 components:

The first component (A) focuses on capturing local and

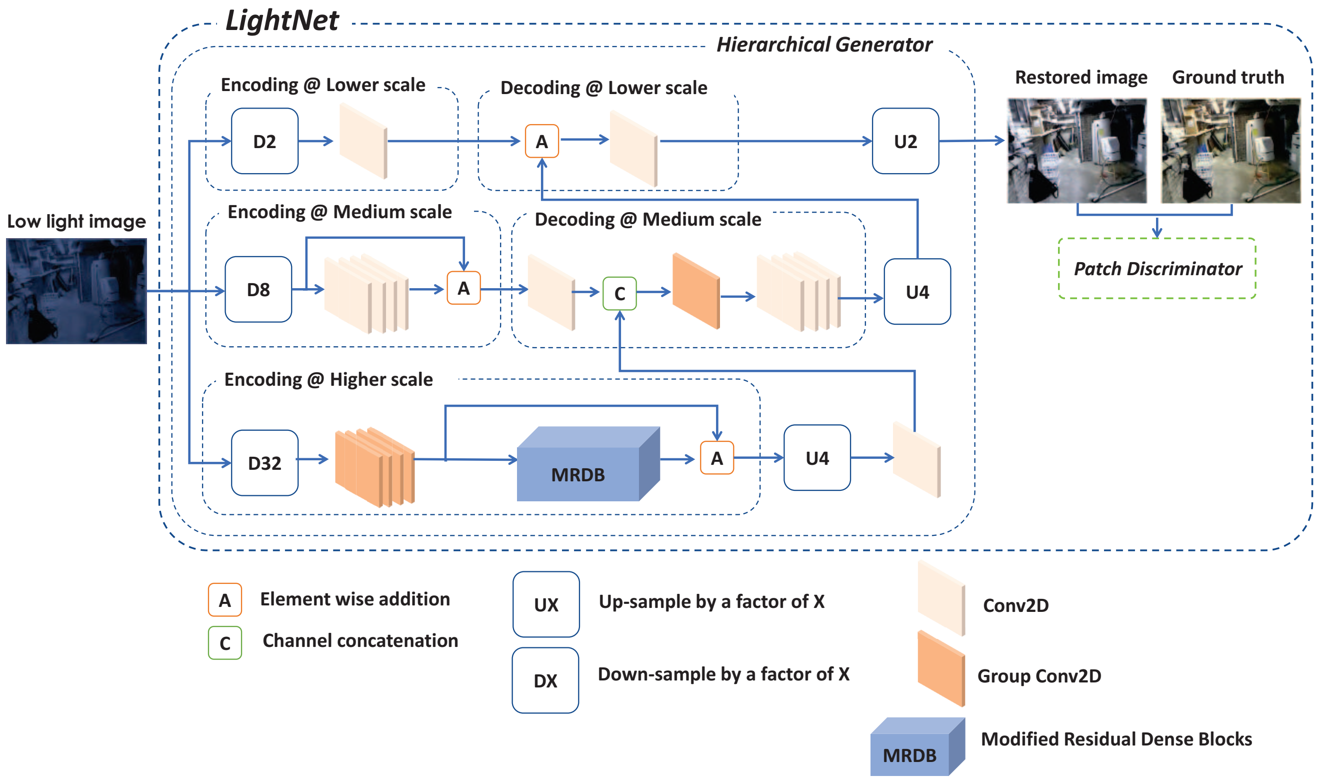


Figure 3: LightNet: Framework for enhancement of images captured in low-light conditions [9]. The hierarchical generator consists of two blocks encoder and decoder. At encoder we downsample the input image at 3 distinct scales namely $D2$, $D8$ and $D32$ respectively. At decoder we fuse the outputs in reverse order of encoder. The patch discriminator facilitates learning of local to global information and aids the learning of local and global features in generator.

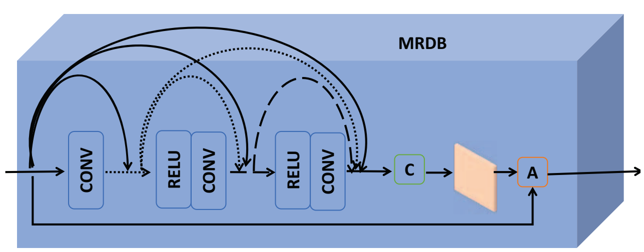


Figure 4: MRDB: Modified Residual Dense Blocks. MRDB for capturing local information across all scales.

global structural information as shown in Equation 2. The second component (B) focuses on recovering lost colors and contrast both locally and globally as shown in Equation 3. The third component (C) facilitates the reconstruction of channel-wise structural information as stated in [14] and is shown in Equation 4. Introducing color loss is the key towards enhancement of low-light images. During training, we use the combinational loss function and compute loss between generated RGB image and the corresponding ground-truth.

3.2.1 Perceptual Loss (A)

VGG Perceptual loss is based on the ReLU Activation layers of a pre-trained VGG 19 network and focus on pixel-wise loss computation. VGG 19 Perceptual loss is shown in Equation 2.

$$\mathrm{L}_{VGG/i,j} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\psi_{i,j}(\mathcal{I}^{GT})_{x,y} - \psi_{i,j}(\mathcal{G}_{\theta_G}(\mathcal{I}^{LL}))_{x,y})^2 \tag{2}$$

where, $\psi_j(x)$ be the activations of the j^{th} layer of the network ψ when processing the image x, I^{GT} is the ground-truth image, and I^{LL} is the enhanced image. By incorporating the VGG 19 loss function into our methodology, our objective is to comprehensively capture and preserve the discernible contextual features, and inherit the same in high-lit images.

Figure 5: Images captured in low-light conditions, 1^{st} column shows ground-truth image (captured in auto settings), 2^{nt} column shows images capture with $ISO50$, 3^{rd} column shows images captured with $ISO100$, 4^{th} column shows images capture with $ISO200$, 5^{th} column shows images captured with $ISO400$, 6^{th} column shows images captured with $ISO800$, 7^{th} column shows images captured with $ISO3200$. Exposure is fixed across the rows ($1^{st}row$: $1/24000s$, $2^{nd}row$: $1/4000s$, $3^{rd}row : 1/2000s$, $4^{th}row : 1/180s$, $5^{th}row : 1/20s$, $6^{th}row : 1/2s$

3.2.2 Color Loss (B)

To restore the lost colors in the generated image, we use Color loss as given by authors in [16]. The Color loss computes the Euclidean distance between generated image and the corresponding ground-truth. Initially, we apply gaussian blur, in order to eliminate sharp frequencies while computing the color difference as shown in Equation 3.

$$\mathcal{L}_{color} = ||GT_G - GN_G||_2^2 \tag{3}$$

where, GT_G and GN_G are the gaussian blurred image representations of ground-truth, and generated images respectively. Color loss is employed as a loss function or penalty term during the restoration process to mitigate these color discrepancies and enhance the color consistency.

3.2.3 MS-SSIM Loss (C)

To preserve the structure channel-wise, we use $\mathcal{MS}-\mathcal{SSIM}$ loss function as shown in Equation 4.

$$\mathcal{MS}-\mathcal{SSIM}(\S,\dagger) = L_m(x,y)^{\alpha_M} \cdot \prod_{j=1}^{M}[c_j(x,y)]^{\beta_j}[s_j(x,y)]^{\gamma_j} \tag{4}$$

where, L_m is luminance comparison at scale M, c_j, and s_j are contrast and structure comparison at j^{th} scale. By

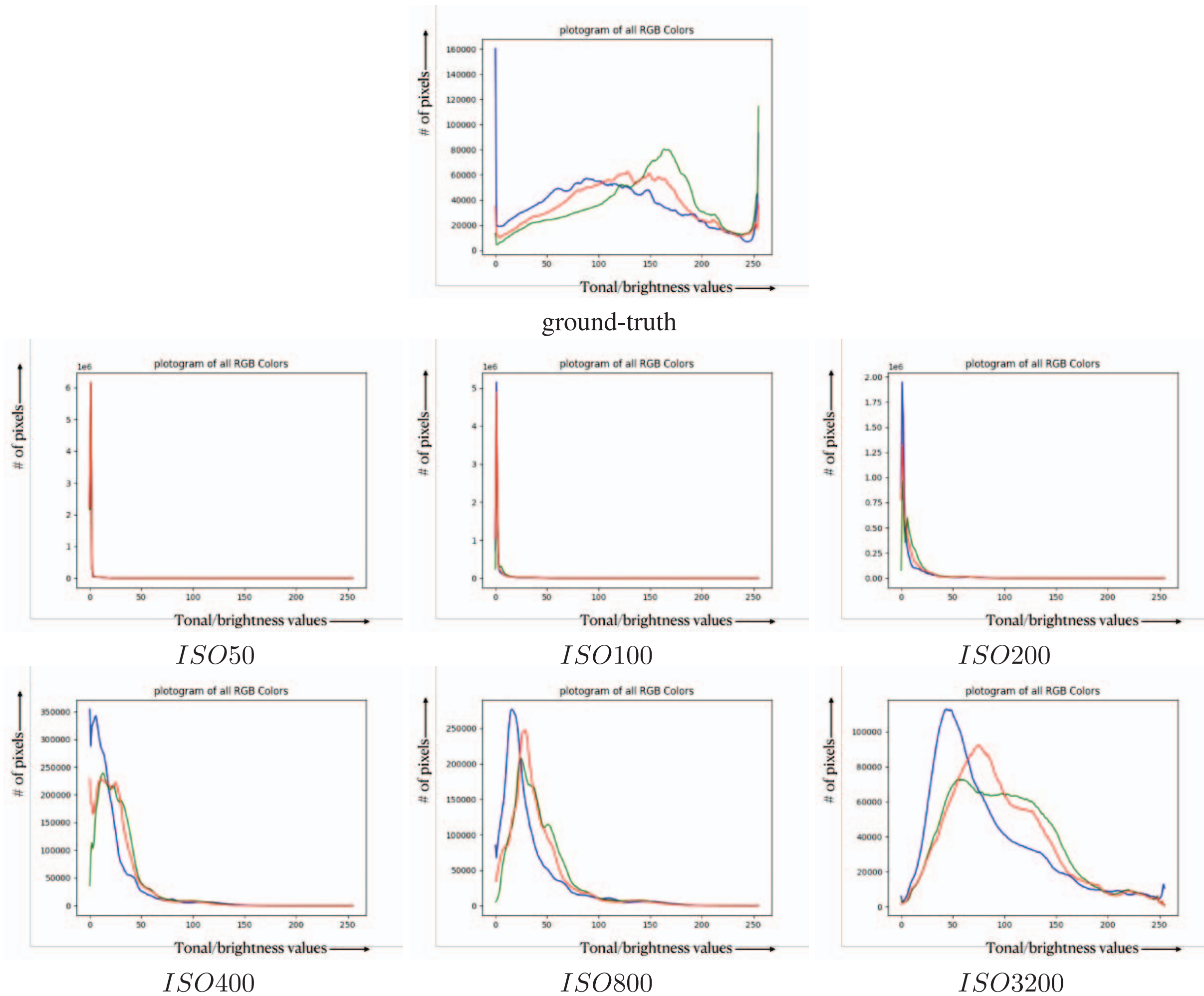


Figure 6: The color distribution of ground-truth image vs images captured under various ISO settings. We observe, with increase in ISO the span of color distribution increases. The proposed LightNet, aims to learn the difference in color distribution across ISO and exposures.

incorporating MS-SSIM loss into our methodology, our objective is to enable capturing of both fine and coarse structural information in the image. This is particularly beneficial for restoration tasks like denoising, deblurring, or super-resolution, where recovering details across different scales is necessary.

3.2.4 Total Loss

The combinational loss function is a combination of Perceptual loss, Color loss and MS-SSIM loss as shown in Equation 5.

$$Totalloss = \alpha * A + \beta * B + \gamma * C \tag{5}$$

where, $A = \mathcal{L}_{VGG/i,j}$, $B = L_{color}$, $C = MS - SSIM$. We set the values of $\alpha = 0.3$, $\beta = 0.4$, $\gamma = 0.3$ heuristically. We use the proposed combinational loss function as shown in Equation 5 to train the proposed generator module. We train the discriminator on $BCELoss$.

4. Results and Discussions

In this section, we present the results of the proposed methodology on proposed dataset and benchmark datasets using appropriate quantitative metrics. We also provide an overview of implementation settings considered for training the proposed LightNet.

4.1. Implementation Details

We use Python (v3.8) and PyTorch framework to develop the proposed architecture and train on Nvidia DGX Tesla V100. We use Adam optimizer with $lr = 0.0002$, $\beta_1 = 0.5$ and $\beta_2 = 0.99$ for both generator and discriminator. We train the model for 50k iterations, with Discriminator loss set to Binary crossentropy. We propose weighted combination of loss functions, more specifically we consider VGG 19 content loss to restore the content from ground-truth to generated image, MS-SSIM to restore structural similarity, Color loss to restore true colors and contrast levels as shown in Equation 5. We consider NTIRE 2022 [8] challenge dataset for training the proposed architecture consisting of 571 pairs along with our own dataset with a batch size of 16.

4.1.1 Results on NTIRE 2022 Challenge Dataset

We demonstrate the results of the proposed methodology on NTIRE 2022 challenge dataset [8]. The results of the proposed methodology are shown in Figure 7. The corresponding PSNR and SSIM scores of the same are shown in Figure 7 highlighted in **bold**.

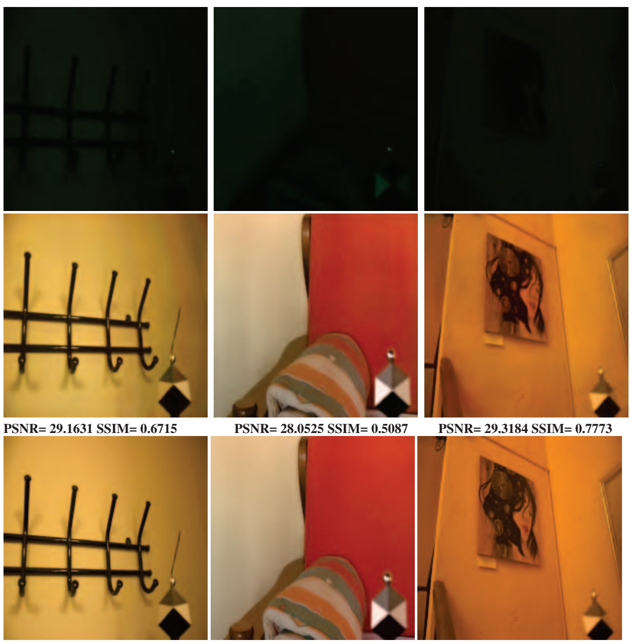


Figure 7: 1^{st} row represents input images (NTIRE 2022 Challenge Dataset [8]), 2^{nd} row represents results of proposed LightNet, and 3^{rd} row depicts the ground-truth images. We observe the proposed LightNet, restores local-global contextual and structural information consistently. (Note: For illustration purpose we have added a fixed constant to each pixel of input image.)

4.1.2 Comparison with SOTA methods on SID Dataset

We demonstrate the results of the proposed methodology on SID dataset [3]. The results of the proposed methodology are shown in Figure 8. The corresponding PSNR and SSIM scores in comparison with authors [18] are shown in Figure 8 highlighted in **bold**. Comparison with state-of-the-art methods on SID dataset using appropriate quantitative metric is shown in Table 1. We infer the proposed LightNet consistently enhances the images while retaining the color distribution both locally and globally. Figure 10, shows exemplar image from SID dataset [4]. We observe, the proposed LightNet consistently retains local contextual information (color distribution) in comparison with SOTA methods.

Figure 8: 1^{st} row represents input images (SID Dataset [3]), 2^{nd} row represents results of authors in [18], 3^{rd} row represents results of proposed method LightNet, and 4^{th} row represents the ground-truth images. We infer, the proposed LightNet restores color, contrast, and luminance consistently.

4.1.3 Comparison with SOTA methods on Our Dataset

We demonstrate the results of the proposed methodology on our proposed dataset. The results of the proposed methodology are shown in Figure 9. The corresponding PSNR and SSIM scores are shown in Figure 9 highlighted in **bold**. Table 2 shows quantitative comparison on state-of-the-art methods for custom dataset. To demonstrate the robustness of the proposed methodology, we consider the camera settings with ISO50 and ISO100 under Exposure 1/2s, 1/180s, and 1/24000s (Extreme low-light condition).

4.2. Ablation Study

LightNet Architecture. We trained our network on NTIRE 2022 Challenge, SID Sony Dataset, and custom dataset after including the proposed patch-discriminator with $BCELoss$. We observe a PSNR (in dB)/SSIM gain

Table 1: Results of proposed methodology in comparison with State-of-the-art-methods using PSNR and SSIM (Average across the selected data on SID dataset [3]) as a reference based quantitative metric. Last row corresponds to results of proposed LightNet (Represented in **bold**).

Methods	PSNR↑	SSIM ↑
SID [4] (2018)	28.80	0.787
DID [26] (2019)	28.41	0.780
SGN [11] (2019)	28.91	0.789
LLPackNet [17] (2020)	27.83	0.750
DCE [12] (2020)	26.53	0.730
LDC [30] (2020)	29.56	0.799
REDIRT [19] (2021)	28.66	0.790
LightNet (Ours)	**30.72**	**0.850**

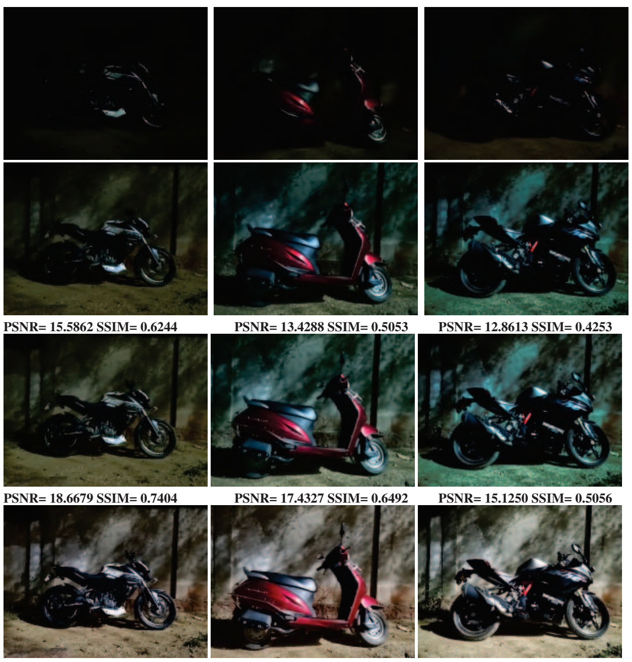


Figure 9: 1^{st} row represents input images (our dataset), 2^{nd} row represents results of authors in [18], 3^{rd} row represents results of proposed method (LightNet), and 4^{th} row represents ground-truth images.

from 28.66 dB/0.790 to 30.72 dB/0.850 indicating a 2.06 dB (7.15% ↑) gain.

5. Conclusions

In this work, we have proposed a generative model for enhancement of images captured in low-light conditions (LightNet). The proposed architecture includes a hierarchical encoder-decoder module along with a patch discriminator to capture local information as a key towards improving the quality of enhancement. We have proposed a combinational loss function to exploit lost color and contrast information both locally and globally. We have demonstrated the results of proposed method using NTIRE 2022 challenge dataset, SID dataset and our dataset. We have shown the results of low-light enhancement using different quality metrics.

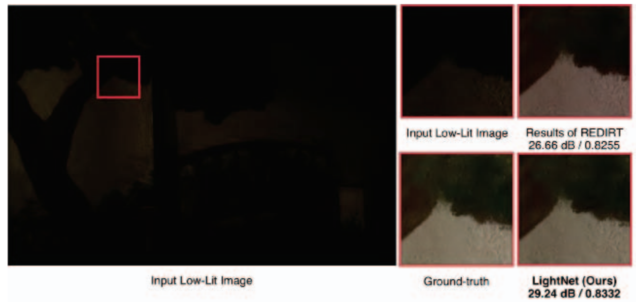


Figure 10: The zoomed in view of exemplar from SID Dataset [4]. 1^{st} row: shows the input low-lit image, results of authors in [19], 2^{nd} row shows ground-truth image and the results of the proposed LightNet. We observe, the proposed LightNet outperforms the SOTA methods.

Table 2: Results of proposed methodology on Custom dataset in comparison with state-of-the-art methods using PSNR (in dB) metric. To demonstrate the robustness of the proposed methodology, we consider the camera settings with ISO50 and ISO100 under Exposure 1/2s, 1/180s, and 1/24000s (Extreme low-light condition). The cells highlighted in "." represent the highest values, and the cells highlighted in "." represent the second best values.
Note: Enlighten Anything model [33] fails to segment the scene effectively under extreme low-light conditions (Exposure 1/240000s + ISO50) leading to failure in low-light image enhancement.

Exposure (in seconds)	1/2s		1/180		1/24000	
Quantitave Metrics	PSNR↑		PSNR↑		PSNR↑	
Methods	ISO 50	ISO 100	ISO 50	ISO 100	ISO 50	ISO 100
LIME [13](2016)	17.63	17.33	18.20	19.83	6.51	7.52
RetinexNet [23](2018)	12.71	10.89	16.24	14.13	11.09	12.22
Zero DCE[12](2020)	14.12	10.95	19.35	17.55	5.96	7.01
Zero DCE++ [21](2021)	13.10	10.49	20.86	17.14	5.99	7.14
Enlighten Anything[33](2023)	20.61	21.35	18.08	19.51	-	11.24
UHDFour [22](2023)	25.59	26.23	22.10	24.35	16.09	16.78
LightNet(Ours)	**31.17**	**31.98**	**30.03**	**30.34**	**25.61**	**26.11**

6. Acknowledgement

This project is partly carried out under Department of Science and Technology (DST) through ICPS programme-Indian Heritage in Digital Space for the project "Digital Poompuhar" (DST/ ICPS/ Digital Poompuhar/ 2017 (General)).

References

[1] M. Abdullah-Al-Wadud, Md. Hasanul Kabir, M. Ali Akber Dewan, and Oksam Chae. A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2):593–600, 2007.

[2] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark, 2018.

[3] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3291–3300, 2018.

[4] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3291–3300, 2018.

[5] Chaitra Desai, Sujay Benur, Ramesh Ashok Tabib, Ujwala Patil, and Uma Mudenagudi. Depthcue: Restoration of underwater images using monocular depth as a clue. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, pages 196–205, January 2023.

[6] Chaitra Desai, Badduri Sai Sudheer Reddy, Ramesh Ashok Tabib, Ujwala Patil, and Uma Mudenagudi. Aquagan: Restoration of underwater images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 296–304, June 2022.

[7] C S Dikshith, Sameeksha M Vernekar, Chaitra Desai, Ujwala Patil, and Uma Mudenagudi. Restoration of degraded underwater images with attenuation coefficients as a clue. *IOP Conference Series: Materials Science and Engineering*, 1187(1):012014, sep 2021.

[8] Egor Ershov, Alexey Savchik, Illya Semenkov, Nikola Banić, Alexander Belokopytov, Daria Senshina, Karlo Koščević, Marko Subašić, and Sven Lončarić. The cube++ illumination estimation dataset. *IEEE Access*, 8:227511–227527, 2020.

[9] 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.

[10] Jianhua Wu Chongsoon Lim Feifan Lv, Feng Lu. Mbllen: Low-light image/video enhancement using cnns. In *British Machine Vision Conference (BMVC)*, 2018.

[11] Shuhang Gu, Yawei Li, Luc Van Gool, and Radu Timofte. Self-guided network for fast image denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2511–2520, 2019.

[12] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020.

[13] 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, 2017.

[14] Dikshit Hegde, Tejas Anvekar, Ramesh Ashok Tabib, and Uma Mudengudi. Da-ae: Disparity-alleviation auto-encoder towards categorization of heritage images for aggrandized 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 5093–5100, June 2022.

[15] H. Ibrahim and N. S.P. Kong. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Trans. on Consum. Electron.*, 53(4):1752–1758, nov 2007.

[16] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3277–3285, 2017.

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

[18] Mohit Lamba and Kaushik Mitra. Restoring extremely dark images in real time. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3487–3497, 2021.

[19] Mohit Lamba and Kaushik Mitra. Restoring extremely dark images in real time. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3487–3497, 2021.

[20] E H Land. An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the National Academy of Sciences*, 83(10):3078–3080, 1986.

[21] Chongyi Li, Chunle Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[22] Chongyi Li, Chun-Le Guo, Man Zhou, Zhexin Liang, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Embedding fourier for ultra-high-definition low-light image enhancement. In *ICLR*, 2023.

[23] 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.

[24] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement, 2015.

[25] Feifan Lv, Bo Liu, and Feng Lu. Fast enhancement for non-uniform illumination images using light-weight cnns, 2020.

[26] Paras Maharjan, Li Li, Zhu Li, Ning Xu, Chongyang Ma, and Yue Li. Improving extreme low-light image denoising via residual learning. In *2019 IEEE International Conference on Multimedia and Expo (ICME)*, pages 916–921. IEEE, 2019.

[27] Seonhee Park, Byeongho Moon, Seungyong Ko, Soohwan Yu, and Joonki Paik. Low-light image enhancement using variational optimization-based retinex model. In *2017 IEEE International Conference on Consumer Electronics (ICCE)*, pages 70–71, 2017.

[28] Wenqi Ren, Sifei Liu, Lin Ma, Qianqian Xu, Xiangyu Xu, Xiaochun Cao, Junping Du, and Ming-Hsuan Yang. Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing*, 28(9):4364–4375, 2019.

[29] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.

[30] Ke Xu, Xin Yang, Baocai Yin, and Rynson WH Lau. Learning to restore low-light images via decomposition-and-enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2281–2290, 2020.

[31] 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 *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3060–3069, 2020.

[32] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021.

[33] Qihan Zhao, Xiaofeng Zhang, Hao Tang, Chaochen Gu, and Shanying Zhu. Enlighten anything: When segment anything model meets low-light image enhancement, 2023.

[34] Minfeng Zhu, Pingbo Pan, Wei Chen, and Yi Yang. Eemefn: Low-light image enhancement via edge-enhanced multi-exposure fusion network. In *AAAI*, 2020.