

Pixel-Wise Polynomial Estimation Model for Low-Light Image Enhancement

Muhammad Tahir Rasheed¹, and Daming Shi^{1*}

¹ College of Computer Science and Software Engineering, Shenzhen University
Shenzhen, 518060 China
[e-mail: tahir@email.szu.edu.cn, dshi@szu.edu.cn]

*Corresponding author: Daming Shi

*Received June 4, 2023; revised August 4, 2023; accepted August 7, 2023;
published September 30, 2023*

Abstract

Most existing low-light enhancement algorithms either use a large number of training parameters or lack generalization to real-world scenarios. This paper presents a novel lightweight and robust pixel-wise polynomial approximation-based deep network for low-light image enhancement. For mapping the low-light image to the enhanced image, pixel-wise higher-order polynomials are employed. A deep convolution network is used to estimate the coefficients of these higher-order polynomials. The proposed network uses multiple branches to estimate pixel values based on different receptive fields. With a smaller receptive field, the first branch enhanced local features, the second and third branches focused on medium-level features, and the last branch enhanced global features. The low-light image is downsampled by the factor of 2^{b-1} (b is the branch number) and fed as input to each branch. After combining the outputs of each branch, the final enhanced image is obtained. A comprehensive evaluation of our proposed network on six publicly available no-reference test datasets shows that it outperforms state-of-the-art methods on both quantitative and qualitative measures.

Keywords: Deep learning, polynomial estimation, low-light image enhancement, multi-branch.

This work is supported by Ministry of Science and Technology China (MOST) Major Program on New Generation of Artificial Intelligence 2030 No. 2018AAA0102200. It is also supported by Natural Science Foundation China (NSFC) Major Project No. 61827814 and Shenzhen Science and Technology Innovation Commission (SZSTI) project No. JCYJ20190808153619413.

1. Introduction

The quality of images strongly affects computer vision tasks. Low-light conditions lead to poor contrast, poor visibility, and substantial noise in an image [24, 19, 55]. When such degraded images are used for computer vision tasks, obtaining accurate features is challenging, and as a consequence, computer vision tasks perform decrease [39]. Histogram equalization-based methods [6, 33] use different approaches to balance the histogram of the intensities of pixels. It is possible that the balancing histogram can saturate certain regions of the image and introduce strange artifacts in it. Another non-linear classical method, Gamma correction [43, 35], suppresses bright regions while increasing the brightness of dark areas. Aside from suppressing the bright region, it treats each pixel as a separate entity and does not consider the entire image structure, which reduces the quality of the image.

To produce an improved image in RGB space from the scene radiance information, the camera manufacturers apply several nonlinear functions on the captured light, such as white balance, demosaick and color correction. Over the years, several approaches have been proposed to estimate the best Camera Response Function (CRF). Grossberg and Nayar [10] devised an Empirical Model of Response (EMoR) that relates scene radiance to image brightness. Later on, they collected a Database of real-world camera Response Functions (DoRF) and used this database to design an Empirical Model of Response (EMoR) [10]. Ng and Tsui [31] estimated the CRF using geometry invariants. In [29], the radiometric response function is designed for imaging systems. A radiometric response function can be used to estimate the CRF based on a rough estimate of the exposures used. A study conducted by Ying et al. [47] examined the relationship between two images with different exposures and devised a CRF based on this relationship. A designed CRF estimates the exposure ratio map for low-light images and adjusts pixel-by-pixel exposure accordingly. The majority of classical camera response function estimation algorithms lack robustness.

In the past, different functions have been used to estimate the camera response model and traditional optimization criteria have been employed to find the best model parameters. The main disadvantage of these models is their lack of robustness as a result of inaccurate model parameters. It is the primary objective of this paper to develop a lightweight model with improved generalization capabilities. To accomplish this goal, we make use of deep learning and higher-order polynomials to create a lightweight and better low-light enhancement network. In this paper, we use deep learning to fit pixel-wise polynomial approximations for low-light enhancement. For each pixel, the proposed module uses a Deep Curve Estimation (DCE) like module of a Zero-Deep Curve Estimation (Zero-DCE) [14] network to estimate the polynomial's coefficients. In order to enhance local, medium, and global features, multiple branches are used. By combining the outputs of these multiple branches, a final image is produced. Using data-driven multi-branch polynomial approximation based on deep learning, we obtain images that look more natural, have better contrast, and are less noisy. The flowchart of the proposed method can be seen in Fig. 1. Experiments on no-reference enhancement test datasets with various image quality assessment metrics demonstrate our approach's generalization ability. In summary, this paper makes the following contributions:

- A pixel-wise higher-order polynomial approximation module is proposed in order to map low-light inputs to high-quality images. To estimate the coefficients of the polynomials, a deep convolution network is employed.
- A multi-branch network is designed using the proposed higher-order polynomial approximation module to enhance the local, medium, and global features of low-light

images.

- Comprehensive experimental results on six publicly available test datasets using four image quality assessment metrics demonstrate that our proposed method has better generalization ability as compared to traditional and advanced state-of-the-art methods.

The remainder of the paper is organized as follows: Section 2 provides a brief overview of methodologies for low-light enhancement, Section 3 describes the proposed approach in detail, and Section 4 presents detailed information about the implementation of the proposed approach, training data, objective function, ablation study, qualitative, quantitative, and computational analysis. Finally, a conclusion and future work is presented in Section 5.

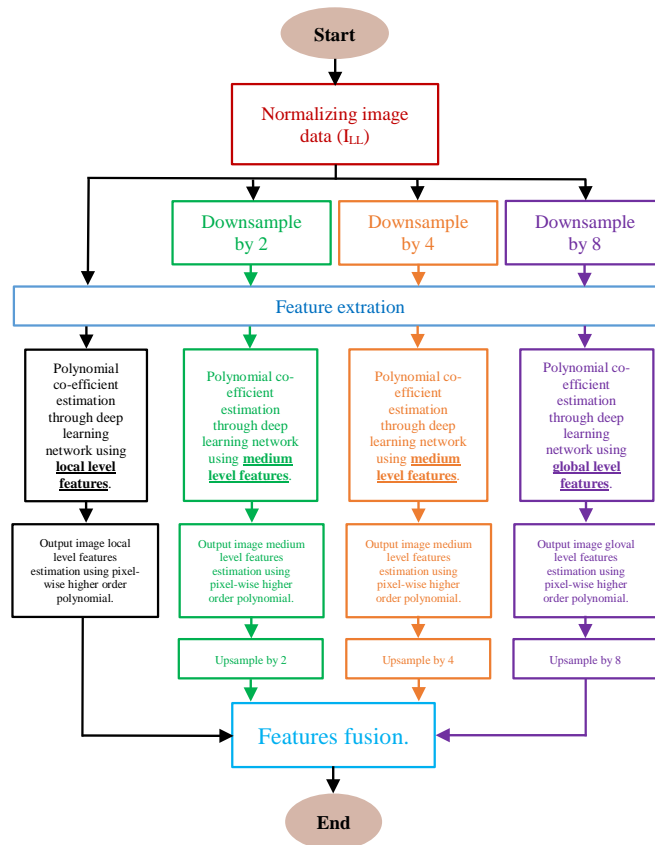


Fig. 1. Flowchart of the proposed system.

2. Literature Review

In the past, Histogram Equalization (HE) based methods such as Weighted Approximated Histogram Equalization (WAHE) [2], Bi-Histogram Equalization (BiHE) [22] are employed to enhance the contrast of low-light images by balancing pixel values distribution. Balancing pixel values based on probability introduces annoying artifacts and may increase the noise in the image. Gamma correction-based methods [17, 3] apply the same non-linear function on each pixel value individually without considering their mutual dependencies. As a result of

Gamma correction, some regions of images are over-enhanced and others under-enhanced. In addition, there is no mechanism to reduce the noise in the output image.

Deep learning has found application in supply chain management [57], health care systems [58], IoT [59], and missing data imputation [11, 54, 56]. In recent years, deep learning-based methods [5, 18, 44, 37, 38] have outperformed classical image enhancement techniques. Lore et al. [25] are the first ones to propose a deep learning-based sparse auto-encoder for simultaneous low-light enhancement and noise reduction using synthetic dataset. Training networks on synthetic datasets may limit their application to real-world images. Wei et al. [42] combined the idea of Retinex theory and deep learning to develop Retinex-Net. Retinex-Net is composed of Decom-Net for decomposition and an Enhance-Net for illumination adjustment. During the training, there is no ground truth of illumination and reflectance for Decom-Net. In GLObal illumination-Aware and Detail-preserving Network (GLADNet) [40], an encoder-decoder network is used to estimate global illumination. A convolution network is used for detail reconstruction using global estimation knowledge to produce enhanced images. Multi-branch fusion module is used for combining the output of multiple subnets to produce the final result in Multi-Branch Low-Light Enhancement Network (MBLLEN) [27]. The major drawback of GLADNET and MBLLEN is the use of Synthetic datasets for their training. Chen et al. [5] proposed an Unet-based pipeline for enhancing extremely low-light images using RAW image format. However, in reality, a pipeline based on RAW data cannot be used for sRGB images. Xu et al. [44] observed that noise detection is easier in the low-frequency layer and proposed a decomposition and enhancement network. A decomposition and enhancement network is used to recover low-frequency objects, which are used to enhance high-frequency details.

Several approaches based on Generative Adversarial Networks (GANs) have also been found to be effective in translating low-light images into enhanced images. Jiang et al. [19] introduced a global-local discriminator and Unet-based generator to enhance images using unpaired data for training. Jung et al. [21] introduced the concept of Multi-Frame GAN (MFGAN) to translate bad illumination images into good illumination for stereo visual odometry in low-light conditions. Training a GAN is difficult because of its instability. In an attempt to enhance low-light images, Guo et al. [14] proposed a Zero-reference Deep Curve Estimation (Zero-DCE) method. Zero-DCE adjusts the dynamic range of a given image by estimating the higher-order curve. Based on the idea of deep image priors, Zhao et al. [52] proposed the generative Retinex Deep Image Prior (RetinexDIP) network for Retinex decomposition. Robust Retinex Decomposition Network (RRDNet) [53] is another zero-shot learning-based network based on Retinex decomposition. Both RetinexDIP and RRDNet methods minimize non-reference loss iteratively to produce enhanced images. Iteratively minimizing errors increases computation costs and makes them less efficient. Rasheed et al. [36] have presented a comprehensive experimental review of various enhancement methods on nine publicly available datasets to assess their generalization ability. In this experimental review, supervised learning-based methods outperform not only classical methods but also self-supervised and zero-shot learning-based approaches.

Even though several deep learning-based methodologies have been developed to date, their network architectures are either computationally inefficient or have limited generalization capabilities. The primary objective of this study is to develop a network that is not only efficient but also capable of producing more natural results in a variety of situations. The classical methods lack the model capacity to remove noise, enhance contrast, sharpen details and improve the aesthetic of low-light images. Furthermore, the majority of classical methods try to enhance the local region and ignore the global aspect (i.e., Gamma correction)

while the others enhance the global and ignore the local information (i.e., HE). Whereas Zero-shot methods (i.e., RetinexDIP and RRDNet) iteratively enhance the image and take several seconds to process a single image. Based on these drawbacks and Rasheed et al. [36] experimental analysis of low-light enhancement, an efficient and lightweight method is required that enhances both local and global details, removes noise, and improves an image's aesthetics. Moreover, it should be more generalizable in real-world scenarios.

3. Proposed Architecture

In this section, the basic building blocks as well as the overall architecture of the proposed network are discussed in detail. The proposed architecture can be seen in Fig. 2.

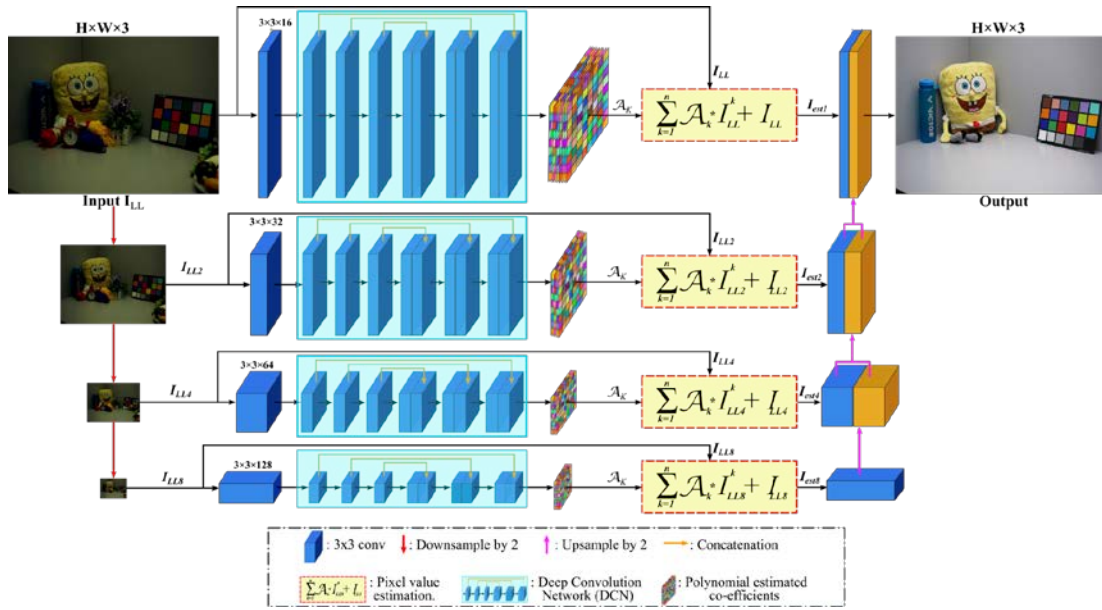


Fig. 2. Proposed pixel-wise polynomial estimation based deep convolution network architecture. The network is composed of four parallel branches. The first branch estimates the pixel value based on local regions, the second and third branches estimates based on medium regions and the last branch estimates based on global region.

3.1 Polynomial-based Pixel Value Estimation Unit

In this subsection, we describe the working of the polynomial-based pixel value estimation unit. This unit consists of two parts; 1) Deep Convolution Network (DCN), and 2) pixel-wise higher-order polynomial fitting. The DCN can be seen in Fig. 2. It is composed of six convolutional layers with skip connections among them. This deep convolution network first calculates the coefficient of the higher-order polynomial (i.e., \mathcal{A}_k). Afterwards, the higher-order polynomial uses the low-light pixel value as input for estimating the pixel value of the enhanced output image. The following equation can be used to describe the operation of this unit,

$$I_{est}(x, y) = \sum_{k=1}^n \mathcal{A}_k(x, y) * I_{LL}(x, y)^k + I_{LL}(x, y), \quad (1)$$

where $I_{est}(x, y)$ represents the estimated output pixel-wise. \mathcal{A}_k and n represent the matrix containing the coefficient and degree of the polynomial, respectively. The x and y denote the pixel coordinates. The coefficient of the polynomial is estimated using a convolution neural network (CNN). The architecture of the CNN used is similar to the one described in [14], except that each convolution layer is followed by the ReLU activation function. The CNN output totals nine feature maps, which are used in the above Equation (1) to calculate the enhanced image. The degree (i.e., n) is selected to be 3 for the polynomial equation. In order to learn the non-linear relationship between low-light pixel values and enhanced pixel values, higher-order polynomials are used to map the low-light pixels to enhanced pixels.

3.2 Multi-branch Network Architecture

This subsection discusses the overall structure of the multi-branch network. The proposed network consists of four branches. Each branch is composed of a polynomial-based pixel value estimation unit. The main difference between each branch is that the pixel value is learned based on a different receptive field. The first branch learns pixel values based on local features, while the second and third branches learn pixel values based on medium-level features, and the last branch learns pixel values based on global features. The input to the first branch is the low-light image, and the input to the second, third, and fourth branches is the downsampled version of the low-light by a factor of two, four, and eight, respectively. Learning the pixel values based on different receptive fields helps to enhance the local texture of the image as well as the global details. Each branch of the network is denoted by the following set of equations,

$$I_{est1}(x, y) = \sum_{k=1}^n \mathcal{A}_k(x, y) * I_{LL}(x, y)^k + I_{LL}(x, y), \quad (2)$$

$$I_{est2}(x, y) = \sum_{k=1}^n \mathcal{A}_k(x, y) * I_{LL2}(x, y)^k + I_{LL2}(x, y), \quad (3)$$

$$I_{est4}(x, y) = \sum_{k=1}^n \mathcal{A}_k(x, y) * I_{LL4}(x, y)^k + I_{LL4}(x, y), \quad (4)$$

$$I_{est8}(x, y) = \sum_{k=1}^n \mathcal{A}_k(x, y) * I_{LL8}(x, y)^k + I_{LL8}(x, y), \quad (5)$$

where I_{est1} , I_{est2} , I_{est4} and I_{est8} are the first, second, third, and fourth branches' outputs, respectively. The low-light input image is denoted by I_{LL} . The I_{LL2} , I_{LL4} , and I_{LL8} are the downsampled versions of I_{LL} by a factor of two, four and eight, respectively. The final enhanced image is produced by fusing the estimated output of multi-branches. The I_{est8} is upsampled by a factor of 2 and concatenated with the feature of I_{est4} and a convolution filter is applied to them. The same process is repeated to produce the final enhanced image I_{en} .

4. Experimental Results

4.1 Implementation

The proposed pixel-wise polynomial estimation-based network is implemented with the Tensorflow 2.4 framework in Python 3.6. To train the network, a workstation equipped with an NVIDIA Titan Xp GPU is used. Adam optimizer with an initial learning rate of 10^{-4} is

applied to optimize the objective function. As part of each training cycle, patches of size $256 \times 256 \times 3$ are fed to the network. A batch size of 16 is used for training. Training this network for 560 epochs takes approximately three and a half days. A flip, resize, and rotation based augmentation of the data is also applied during the training process to avoid overfitting of the model.

4.2 Training Loss Function

We need to enhance the aesthetics, brightness, and contrast of the image. The generated image and the ground truth image must be structurally similar. The structural similarity index (SSIM) based objective function is used for training networks. The objective function can be expressed as follows:

$$\mathcal{L}_{\text{SSIM}} = 1 - \text{SSIM}(I_{en}, I), \quad (6)$$

Where I_{en} and I represent the predicted and ground truth images, respectively. In general, SSIM values fall between 0 and 1. The value of $\mathcal{L}_{\text{SSIM}}$ can lie between 1 and 0. The value of $\mathcal{L}_{\text{SSIM}}$ closer to 1 indicates that the structure of the predicted image is far from the ground truth image and vice versa.

4.3 Training Dataset

For the purpose of training an enhancement network, we have collected low-light images from the LOW-Light (LOL), Single Image Contrast Enhancement (SICE), GLADNet, Large-Scale Real-World (LSRW), and Vassilios Vonikakis (VV) datasets. The LOL contains 485 images taken in low light, most of which are taken indoors. On the other hand, the SICE dataset contains approximately 589 images, the majority of which are taken outside. As a means of balancing indoor and outdoor images and expanding the training dataset, low-light images from other datasets, such as LSRW and VV, are also included. We randomly cropped 1455 patches from the LOL dataset, 1783 patches from SICE, 3491 patches from the LSRW dataset, and 5000 patches from the GLADNet dataset. In total, 11829 patches are randomly cropped from the above-mentioned datasets, and each patch measured $256 \times 256 \times 3$. We trained the network with 10646 patches and evaluated it with 1183 patches. The primary reason for collecting patches from different training datasets is to increase the size and diversity of the training dataset. Larger and more diverse datasets have a higher entropy, which enhances the learning ability of a neural network. In summary, the larger the dataset, the better the network will be able to generalize in real-life scenarios.

4.4 Ablation Study

The purpose of this subsection is to demonstrate the effectiveness of each component of the proposed method. Controlled experiments are conducted on 1183 patches with a size of $256 \times 256 \times 3$ and the results are shown in [Table 1](#) and [Fig. 3](#). The default settings use each convolution filter of size 3×3 followed by a ReLU activation function. There are four parallel branches in the network, and the degree of polynomial is chosen to be 3. The network is optimized using the SSIM loss function presented in subsection 4.2. Fig. 4b and 4a show the results of the evaluation of the models based on Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) metrics. The higher the PSNR and SSIM values, the better the image quality. We have changed the loss function, filter size, number of parallel branches, and activation

function of the proposed network one by one, and the best results for each setting have been presented in **Table 1**. The default setting produces the best results, with a PSNR of 19.463 and an SSIM of 0.767. When the filter size is changed from 3×3 to 6×6 , the PSNR and SSIM values are reduced. Applying batch normalization or layer normalization after the convolution filter also reduces the value of these metrics. When the network is optimized using the mean square error (MSE) loss function instead of the SSIM loss function, there is a significant difference between PSNR and SSIM metrics.

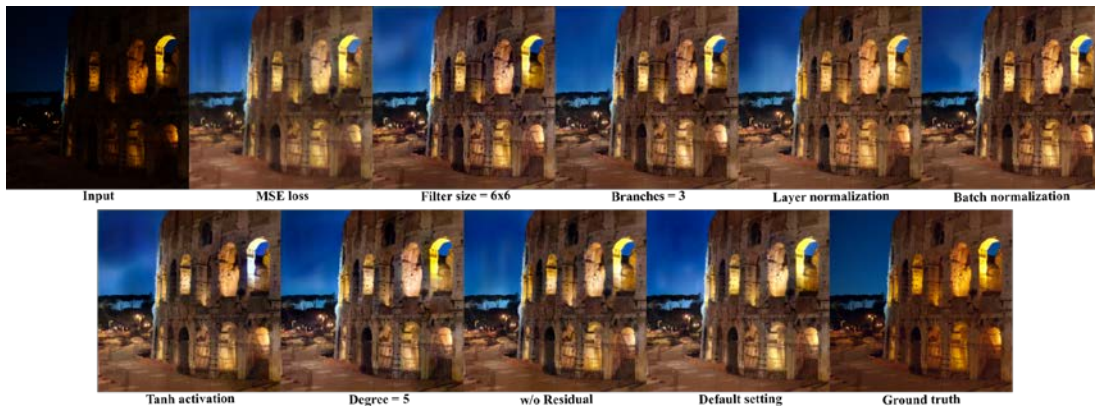


Fig. 3. A visual comparison of the results produced by the proposed method with different settings.

Table 1. Several experiments have been conducted to evaluate the performance of the proposed architecture.

Settings	PSNR \uparrow	SSIM [41] \uparrow
Default	19.463	0.767
SSIM \rightarrow MSE	19.03	0.72
3×3 fil \rightarrow 6×6 fil	19.352	0.759
4 branches \rightarrow 3 branches	19.098	0.761
with layer normalization	19.344	0.758
with batch normalization	19.334	0.756
ReLU \rightarrow tanh	19.327	0.765
degree 3 \rightarrow 5	19.355	0.766
w/o Residual	19.311	0.733

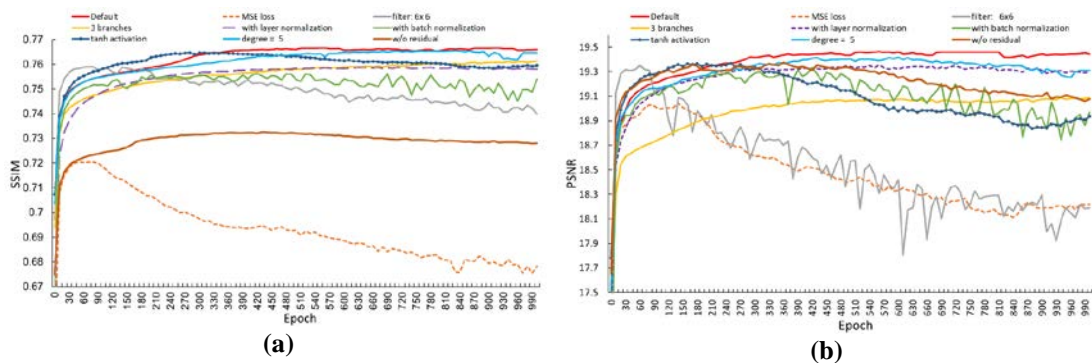


Fig. 4. The evaluation of the different models for the ablation study. (a) shows the PSNR vs the number of epochs and (b) shows the SSIM vs the number of epochs for different variants of the proposed methods on the evaluation data.

4.5 Quantitative Comparison

The proposed model is quantitatively compared with other methods using six publicly available test datasets: Low-light Image Enhancement via Illumination Map Estimation (LIME), LOL (15 images paired), Multi-Exposure Image Fusion (MEF). (17 images unpaired), DICM (69 images unpaired), Extremely Dark (ExDark) (150 images unpaired), and Synthetic Low-Light (SLL) (154 images paired).

Table 2. A quantitative comparison of the proposed methodology is conducted with state-of-the-art methods on six non-reference low-light test datasets using BIQME [13] metric (A higher value of BIQME indicates a higher quality image). Red represents the best result, while blue and green represent the second and third best results, respectively.

Methods	LIME [15]	LOL[42]	MEF [28]	DICM [23]	ExDark [24]	SLL [26]	Average
Input	0.366	0.194	0.377	0.505	0.376	0.311	0.355
CVC [4]	0.517	0.452	0.485	0.568	0.469	0.464	0.492
CLAHE [32]	0.458	0.249	0.428	0.547	0.417	0.423	0.420
WAHE [2]	0.511	0.427	0.474	0.562	0.461	0.450	0.481
LDR [23]	0.514	0.531	0.469	0.560	0.471	0.524	0.511
BiHE [22]	0.471	0.483	0.416	0.504	0.442	0.486	0.467
AGCWD [17]	0.493	0.485	0.487	0.554	0.465	0.482	0.495
IAGC [3]	0.505	0.367	0.476	0.557	0.453	0.453	0.468
BIMEF [45]	0.561	0.430	0.544	0.556	0.510	0.520	0.520
MF [8]	0.562	0.484	0.541	0.555	0.504	0.480	0.521
CRF [47]	0.590	0.476	0.565	0.543	0.529	0.590	0.549
EFF [46]	0.561	0.430	0.544	0.556	0.510	0.520	0.520
PM-SIRE [7]	0.497	0.384	0.488	0.539	0.456	0.435	0.467
MSRCR [20]	0.375	0.408	0.358	0.409	0.340	0.255	0.357
pmea [34]	0.571	0.469	0.562	0.565	0.514	0.533	0.536
Retinex Net [42]	0.572	0.475	0.548	0.532	0.531	0.557	0.536
KinD [51]	0.567	0.514	0.558	0.548	0.525	0.569	0.547
DPED [18]	0.439	0.465	0.453	0.535	0.410	0.420	0.453
KinD++ [50]	0.597	0.525	0.576	0.555	0.545	0.598	0.566
White-box [16]	0.476	0.497	0.460	0.520	0.457	0.497	0.485
Zero-DCE [14]	0.564	0.441	0.547	0.528	0.523	0.509	0.519
RRDNet [53]	0.499	0.364	0.500	0.559	0.453	0.428	0.467
IBA [1]	0.588	0.480	0.561	0.513	0.543	0.563	0.541
StableLLVE [48]	0.530	0.449	0.543	0.496	0.523	0.562	0.517
RetinexDIP [52]	0.587	0.282	0.571	0.555	0.525	0.472	0.499
SS Network [49]	0.606	0.534	0.578	0.506	0.572	0.618	0.569
Ours	0.609	0.546	0.599	0.588	0.589	0.600	0.589

It is the primary objective of this comparison to determine whether these methods can be generalized. Five publicly available test datasets are considered real low-light datasets, while one dataset is considered synthetic. These test datasets include a variety of images, including dynamic and static objects, as well as nature scenes and indoor photographs. A wide range of lighting conditions is also covered by these test datasets. Testing the proposed method against conventional and state-of-the-art methods on this selection of test datasets is appropriate for evaluating its generalization ability. The results of all the methods are evaluated using four metrics, namely Blind Image Quality Measure of Enhanced Images (BIQME), No-Reference Image Quality Metric for Contrast Distortion (NIQMC), Natural Image Quality Evaluator (NIQE), and Information Entropy (IE). An image with a low NIQE value indicates better image quality, while an image with a high IE, NIQMC, and BIQME value demonstrates better image quality. The NIQE measures the quality of distorted images by measuring the distance between the Natural Scene Statistical (NSS) and MultiVariate Gaussian (MVG) feature models. NIQMC measures the contrast quality of the enhanced image. NIQMC works on an information maximization approach. BIQME is a machine learning-based technique for

assessing the quality of enhanced images. BIQME uses 17 features, including brightness, contrast, color, sharpness, and naturalness, to blindly predict the quality of an image. The Information Entropy (IE) of an image is a no-reference metric used to quantify how much information it contains.

Table 3. A quantitative comparison of low-light enhancement algorithms on six test datasets using the NIQE [30] metric. Lower NIQE values indicate better performance. The last column shows the average result on all test datasets by the method. Red represents the best result, while blue and green represent the second and third best results, respectively.

Methods	LIME [15]	LOL [42]	MEF [28]	DICM [23]	ExDark [24]	SLL [26]	Average
Input	4.357	6.748	4.263	4.274	5.128	5.358	5.021
CVC [4]	4.029	8.014	3.636	3.823	4.662	5.828	4.999
CLAHE [32]	3.907	7.268	3.606	3.792	4.734	5.756	4.844
WAHE [2]	3.921	7.802	3.510	3.858	4.569	5.649	4.885
LDR [23]	4.162	8.336	3.632	3.937	4.596	5.745	5.068
BiHE [22]	6.009	8.138	5.362	4.629	7.077	7.782	6.500
AGCWD [17]	4.032	7.528	3.868	3.629	4.582	5.660	4.883
IAGC [3]	3.951	7.418	3.652	4.015	4.557	5.740	4.889
BIMEF [45]	3.859	7.515	3.329	3.845	4.514	5.747	4.802
MF [8]	4.067	8.877	3.492	3.844	4.531	6.039	5.142
CRF [47]	3.854	7.686	3.264	3.801	4.525	6.008	4.856
EFF [46]	3.859	7.515	3.329	3.845	4.514	5.747	4.802
PM-SIRE [7]	4.050	7.506	3.450	3.978	4.383	5.435	4.800
MSRCR [20]	3.939	8.006	3.688	3.948	4.904	5.574	5.010
pmea [34]	3.843	8.281	3.431	3.836	4.296	6.237	4.987
Retinex Net [42]	4.597	8.879	4.410	4.415	4.551	7.573	5.738
KinD [51]	4.763	4.709	3.876	4.150	4.340	4.450	4.381
DPED [18]	7.995	5.785	8.713	8.088	6.143	5.836	7.093
KinD++ [50]	4.385	4.616	3.738	3.804	4.343	5.090	4.329
White-box [16]	4.598	7.819	4.622	4.630	5.534	7.138	5.724
Zero-DCE [14]	3.769	7.767	3.283	3.567	3.917	5.998	4.717
RRDNet [53]	3.936	7.436	3.508	3.637	4.010	5.524	4.675
IBA [1]	4.062	7.884	3.536	3.723	4.273	5.837	4.886
StableLLVE [48]	4.234	4.372	3.924	4.061	4.053	4.185	4.138
RetinexDIP [52]	3.735	7.096	3.245	3.705	4.234	5.883	4.650
SS Network [49]	4.819	3.753	4.351	4.717	4.048	5.400	4.514
Ours	3.838	4.304	3.174	3.457	4.294	4.961	4.005

The results based on the above-mentioned metrics are presented in Table 2, 3, 4, and 5. The best results, the second best results, and the third best results have been colored in red, blue, and green, respectively. A summary of the average results for each method on all six test datasets is provided in the last column of each table. We can gain a better understanding of the overall performance of each method by analyzing this average result. Based on a BIQME analysis, we have determined that the proposed method performs best on five test datasets and second best on SLL test datasets. As indicated by the BIQME measurement, our method produces improved brightness, contrast, sharpness, and color. Based on the average score, self-supervised networks and Kindling the Darkness (KinD++) rank second and third, respectively. Both self-supervised network and KinD++ are deep learning methods based on Retinex theory. As a result of the NIQMC metric, our method got the best results on four and the second best on two test datasets, as shown in Table 4. NIQMC indicates that IBA and LDR had the second and third best performances, respectively. The NIQE metric results in Table 3 show that the proposed method produced more natural and less distorted results on MEF and DICM test datasets, while it has the second best performance on LOL and third best performance on LIME and SLL test datasets. But it still has the best overall performance on all test datasets. Our proposed method has the highest amount of information entropy on five test datasets and

the second highest amount of information on LOL test data according to our information entropy.

Table 4. A quantitative comparison of the proposed methodology is performed with state-of-the-art methods on six non-reference low-light test datasets using NIQMC [12] (A higher value of NIQMC indicates better image quality). Red represents the best result, while blue and green represent the second and third best results, respectively.

Methods	LIME [15]	LOL[42]	MEF [28]	DICM [23]	ExDark [24]	SLL [26]	Average
Input	3.918	1.602	4.197	4.641	3.982	3.655	3.666
CVC [4]	4.997	4.244	5.066	5.294	4.945	4.841	4.898
CLAHE [32]	4.735	2.697	4.756	5.171	4.588	4.685	4.438
WAHE [2]	4.812	3.847	4.935	5.213	4.812	4.638	4.710
LDR [23]	5.024	4.654	5.048	5.247	4.975	5.163	5.019
BiHE [22]	3.987	4.639	3.877	4.446	4.040	4.322	4.218
AGCWD [17]	4.832	4.947	5.027	5.198	4.870	4.957	4.972
IAGC [3]	4.703	3.191	4.823	5.005	4.702	4.487	4.485
BIMEF [45]	4.721	3.711	4.879	5.047	4.743	4.933	4.672
MF [8]	4.869	4.502	5.042	5.116	4.834	4.761	4.854
CRF [47]	4.974	4.277	5.033	5.040	4.917	5.348	4.932
EFF [46]	4.721	3.711	4.879	5.047	4.743	4.933	4.672
PM-SIRE [7]	4.592	3.532	4.803	5.008	4.603	4.540	4.513
MSRCR [20]	3.877	4.176	3.565	4.040	3.405	2.642	3.617
pmea [34]	4.883	4.285	4.990	5.128	4.794	4.934	4.836
Retinex Net [42]	4.697	4.225	4.747	4.763	4.671	4.989	4.682
KinD [51]	4.943	4.504	5.093	5.009	4.860	5.324	4.955
DPED [18]	4.287	4.076	4.579	4.707	4.355	4.548	4.425
KinD++ [50]	4.903	4.588	5.060	4.986	4.864	5.344	4.957
White-box [16]	4.503	4.683	4.541	4.666	4.534	4.708	4.606
Zero-DCE [14]	4.839	4.015	4.944	4.927	4.843	4.864	4.738
RRDNet [53]	4.738	3.536	4.969	5.159	4.627	4.551	4.597
IBA [1]	5.245	4.488	5.206	4.874	5.124	5.298	5.039
StableLLVE [48]	4.644	3.691	4.849	4.644	4.755	5.044	4.604
RetinexDIP [52]	4.875	2.726	4.968	5.008	4.682	4.478	4.456
SS Network [49]	5.189	4.764	5.174	4.541	4.949	5.385	5.001
Ours	5.232	4.781	5.273	5.295	5.256	5.483	5.220

Table 5. A quantitative comparison of low-light enhancement algorithms on six test datasets using the Information Entropy [9] metric. Higher information entropy values indicate better performance. The last column shows the average result on all test datasets by the method. Red represents the best result, while blue and green represent the second and third best results, respectively.

Methods	LIME [15]	LOL[42]	MEF [28]	DICM [23]	ExDark [24]	SLL [26]	Average
Input	6.148	4.915	6.075	6.686	5.744	5.616	5.864
CVC [4]	6.875	6.409	6.755	7.055	6.465	6.549	6.685
CLAHE [32]	6.764	5.679	6.583	7.088	6.302	6.591	6.501
WAHE [2]	6.924	6.570	6.732	7.065	6.457	6.644	6.732
LDR [23]	6.845	7.044	6.641	6.969	6.443	6.822	6.794
BiHE [22]	5.452	4.395	5.434	5.911	5.010	4.948	5.192
AGCWD [17]	6.792	6.415	6.648	6.925	6.248	6.278	6.551
IAGC [3]	6.991	6.247	6.878	7.015	6.554	6.554	6.707
BIMEF [45]	7.006	6.145	6.898	7.029	6.464	6.464	6.668
MF [8]	7.238	6.881	7.193	7.203	6.686	6.686	6.981
CRF [47]	6.487	4.971	6.203	6.640	5.921	5.921	6.024
EFF [46]	7.006	6.145	6.898	7.029	6.464	6.464	6.668
PM-SIRE [7]	7.006	6.322	6.894	7.084	6.441	6.325	6.679
MSRCR [20]	6.563	6.841	6.455	6.677	6.319	5.936	6.465
pmea [34]	7.284	6.824	7.273	7.220	6.725	6.638	6.994
Retinex Net [42]	7.489	7.233	7.448	7.413	7.273	7.385	7.374
KinD [51]	7.388	7.017	7.328	7.211	6.905	7.408	7.210
DPED [18]	6.486	6.804	6.563	7.099	6.387	6.830	6.695

KinD++ [50]	7.486	7.065	7.463	7.332	7.034	7.452	7.305
White-box [16]	5.984	5.925	5.391	6.051	5.275	5.460	5.681
Zero-DCE [14]	7.166	6.531	7.093	7.224	6.932	6.545	6.915
RRDNet [53]	6.646	5.457	6.453	7.142	6.426	6.077	6.367
IBA [1]	5.905	4.913	5.749	6.826	6.971	5.465	5.972
StableLLVE [48]	7.227	6.625	7.241	7.010	7.129	7.439	7.112
RetinexDIP [52]	6.974	5.375	6.661	7.214	6.668	6.213	6.518
SS Network [49]	7.497	7.404	7.469	6.675	7.296	7.484	7.304
Ours	7.604	7.266	7.577	7.559	7.464	7.696	7.528

The high average score indicates that our method generalizes well on all test datasets. The average score of Retinex Net and KinD++ indicates they have the second and third-best performance, respectively. For a better understanding of the quantitative analysis, the results of [Table 3](#) to [Table 4](#) are summarized in [Fig. 5](#). For a better comparison of the methods, we have used the ranking scheme proposed in [\[38\]](#). The methods are ranked based on the average value of each metric and shown by the graph in [Fig. 5](#). The methods with the best performance got the rank 1 and poor performance got the highest ranking (i.e., 27). Our method got ranked one on all four metrics, whereas KinD++ got 3rd rank based on NIQE, BIQME and entropy, and got 6th rank based on NIQMC. The average rank is also calculated to compare the overall performance of the methods measured on all four metrics and shown in red color in the graph. The red color graph summarized the results of all four metrics and allowed a comparison of the overall performance of all methods on six publicly available test datasets. The average graph (in red color) in [Fig. 5](#) provides further insight; for example, each of the top five methods belongs to the field of deep learning. This graph also includes metrics scores for the low-light test datasets.

According to the average ranking as shown in [Fig. 5](#), learning-based methods outperformed classical methods. Classical methods are limited in their ability to generalize due to their lack of model capacity. Furthermore, classical methods do not take into account the noise factor when enhancing an image. Both KinD++ and KinD are supervised learning methods that use only the LOL dataset for training. There are only 485 paired training images in the LOL dataset, while the proposed method is trained on a larger dataset. The size of training data plays a crucial role in the network's performance. Zero-shot learning methods (such as Self-supervised Nets and Zero-DCE) have also shown superior performance. In real-world situations, zero-shot learning-based methods can perform better since they utilize the image's internal information. Methods based on zero-shot learning require a limited number of parameters. However, they require the development of an appropriate loss function to have robust performance.

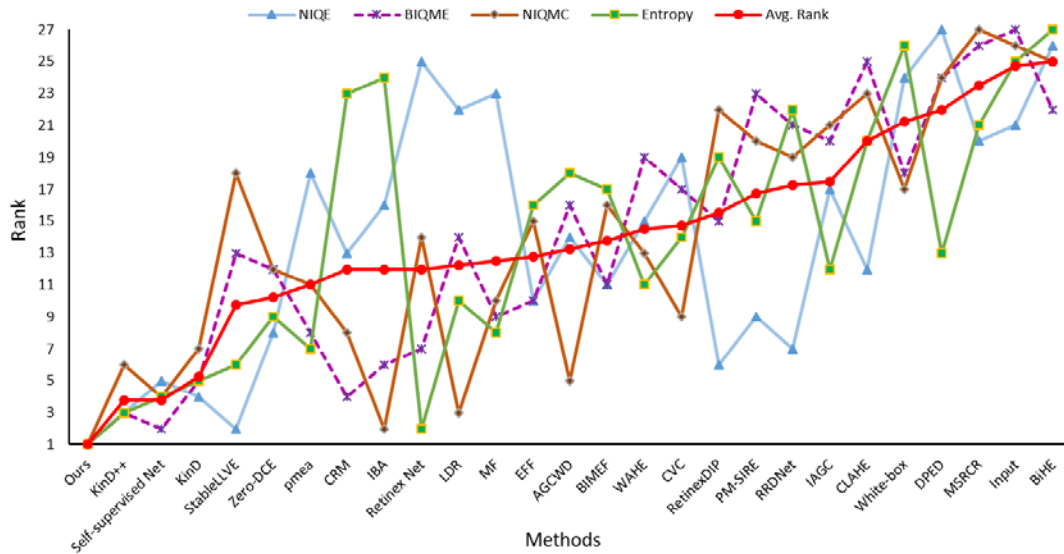


Fig. 5. Based on the average values calculated in [Tables 2 to 5](#), the overall rank of the different enhancement methods is shown. Rank values range from 1 to 27. According to a specific IQA method, the first rank indicates the best performance and the highest rank indicates the worst performance. Red represents the average of the assigned rank based on IQA metrics.

4.6 Qualitative Comparison

A visual comparison of the images of LIME [15], MEF [28], DICM [23] and synthetic low-light (SLL) [26] test datasets are shown in [Fig. 6, 7, 8, and 9](#), respectively. In light of the visual comparison of the figures mentioned above, we make the following observations: It has been demonstrated that among classical methods, the results of the majority of Retinex theory-based methods (e.g., Bio-Inspired Multi-Exposure Fusion framework (BIMEF) [45], Camera Response Function (CRF) [47], EFF [46], PMEA [34], Multi-Fusion (MF) [28], Probabilistic Method with Simultaneous Illumination and Reflectance Estimation (PM-SIRE)[7]) are better than those of histogram and gamma correction-based methods. These Retinex-based methods produce images that appear more natural. Although MultiScale Retinex with Color Restoration (MSRCR) is a Retinex Theory-based method, it suffers greatly from global lightness distortion, cannot correctly recover colors, and a whitewash effect can be observed in all results. Generally, the results of histogram based methods (such as Contextual and Variational Contrast enhancement (CVC) [4], Contrast-limited adaptive histogram equalization (CLAHE) [32], Weighted Approximated Histogram Equalization (WAHE) [2], Layered Difference Representation (LDR) [23]) have low brightness, low contrast, and underexposed regions. The image has not been uniformly enhanced by these methods. As a result of the underexposure problem, these methods could not reveal hidden details. The Improved Adaptive Gamma Correction (IAGC) results are superior to the Adaptive Gamma Correction with Weighting Distribution (AGCWD) [17] results. The robustness of the Retinex theory makes it suitable for use even in learning-based methods. RetinexNet [42] is a deep learning network based on Retinex theory. The visual results of RetinexNet are not only subject to severe noise and contrast distortion but also appear unnatural. KinD is another Retinex-based method, displaying an underexposure issue in the DICM test data image. Eventually, an advanced version of KinD [51] was developed and named KinD++ [50]. As a result of KinD++, this

problem of underexposure is eliminated. StableLLVE suffers from a smoothness around the edges as well as from a global distortion of lightness. On the DICM test image, the smoothness is more apparent. When compared to other learning-based methods, the results of DPED [18], White-box, and Robust Retinex Decomposition Network (RRDNet) [53] are relatively less bright. White-box suffers greatly from underexposure on LIME, DICM, and MEF test images, whereas it performs relatively well on synthetic images. Kind++ and self-supervised networks (SS Net) revealed pale yellow hues in the test images of MEF and DICM. In RetinexDIP [52] results on MEF test images, this pale yellow color is also visible, but it is more prominent in Self-supervised Net results. On the DICM test image, the illumination boost algorithm (IBA) suffers from overexposure, which leads to the loss of texture details. In comparison to the traditional methods, advanced deep learning-based methods produce more natural results. The method we have proposed not only enhances local regions but also enhances global regions. It has a high contrast ratio, less noise, better sharpness, and better brightness. In addition, it solves the underexposure issue and reveals the hidden details. Details and contrast are richer in our method. Using these visual comparisons, we can demonstrate the robustness and generalizability of our proposed method.

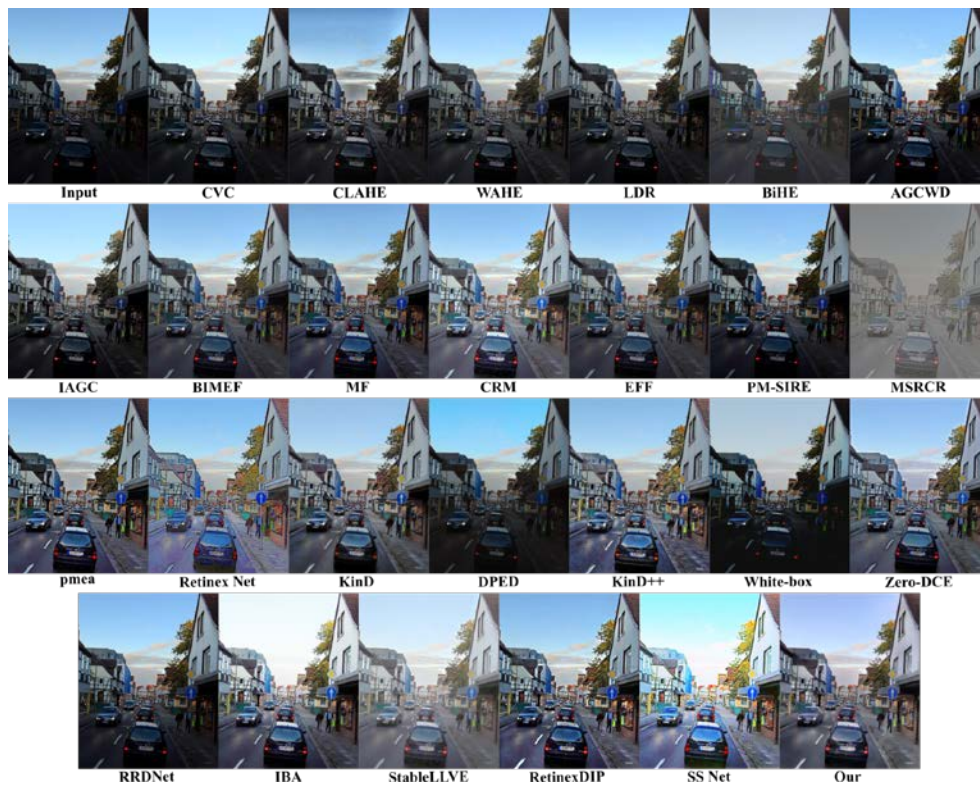


Fig. 6. Visual comparison with state-of-the-art methods on the input image from the LIME test dataset [15]. Please zoom in for a more detailed comparison.

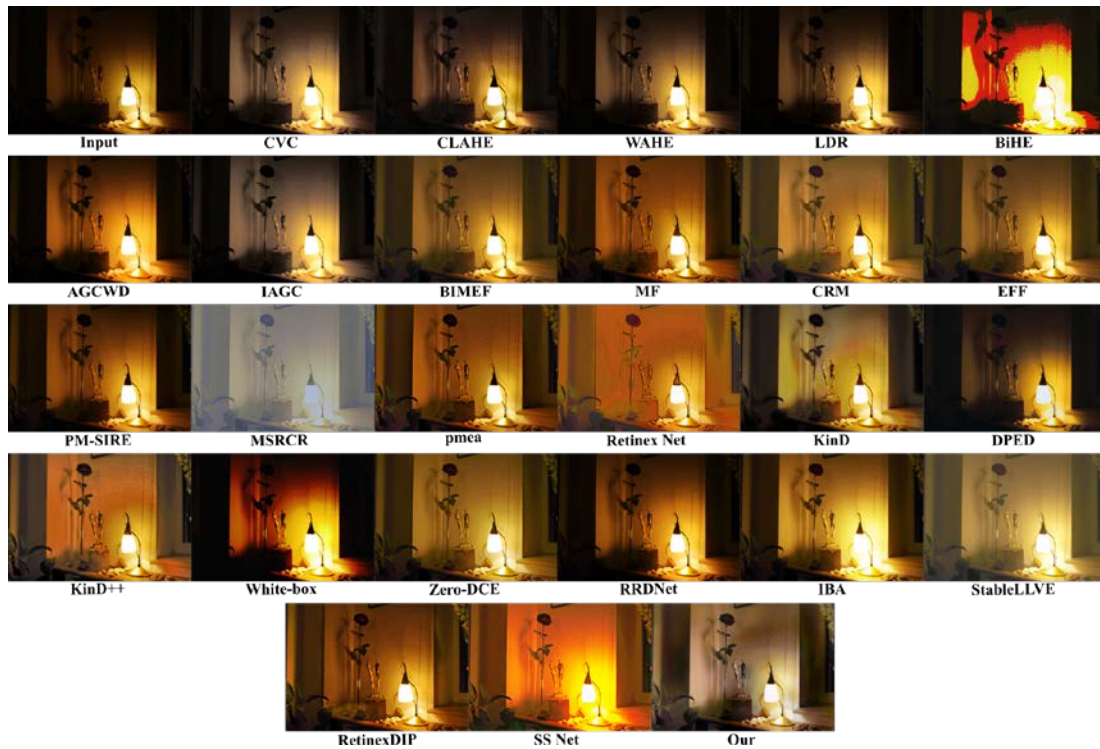


Fig. 7. Visual comparison with state-of-the-art methods on the input image from the MEF test dataset [28]. Please zoom in for a more detailed comparison.

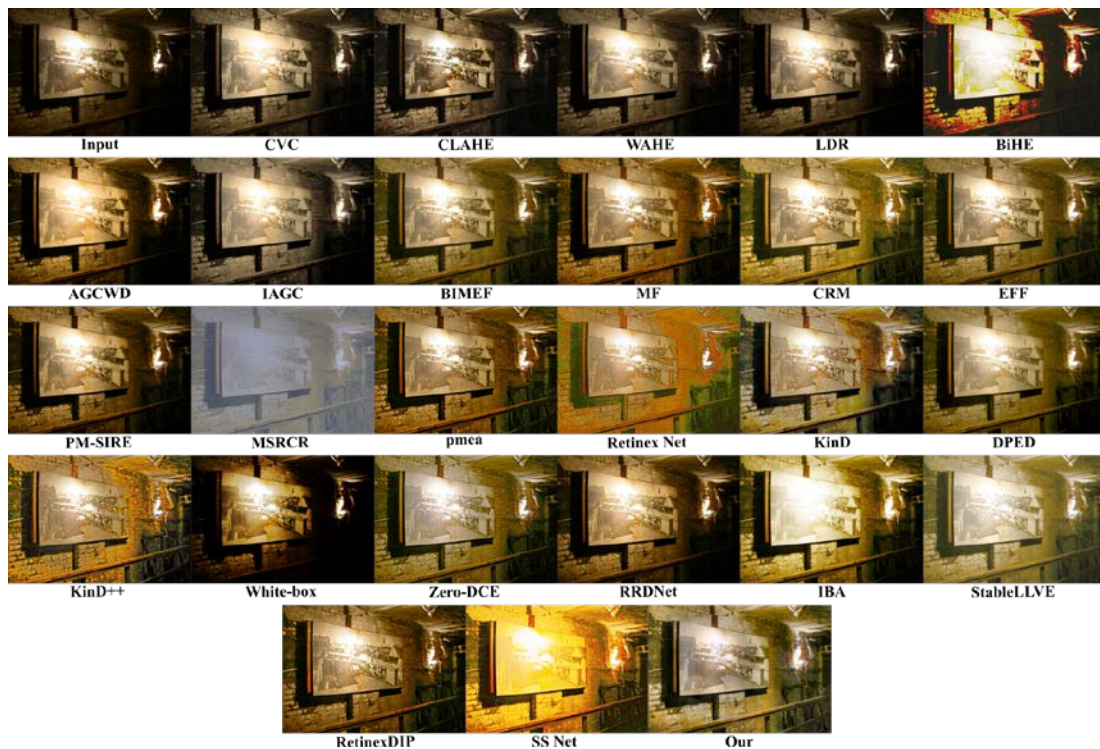


Fig. 8. Visual comparison with state-of-the-art methods on the input image from the DICM test dataset [23]. Please zoom in for a more detailed comparison.

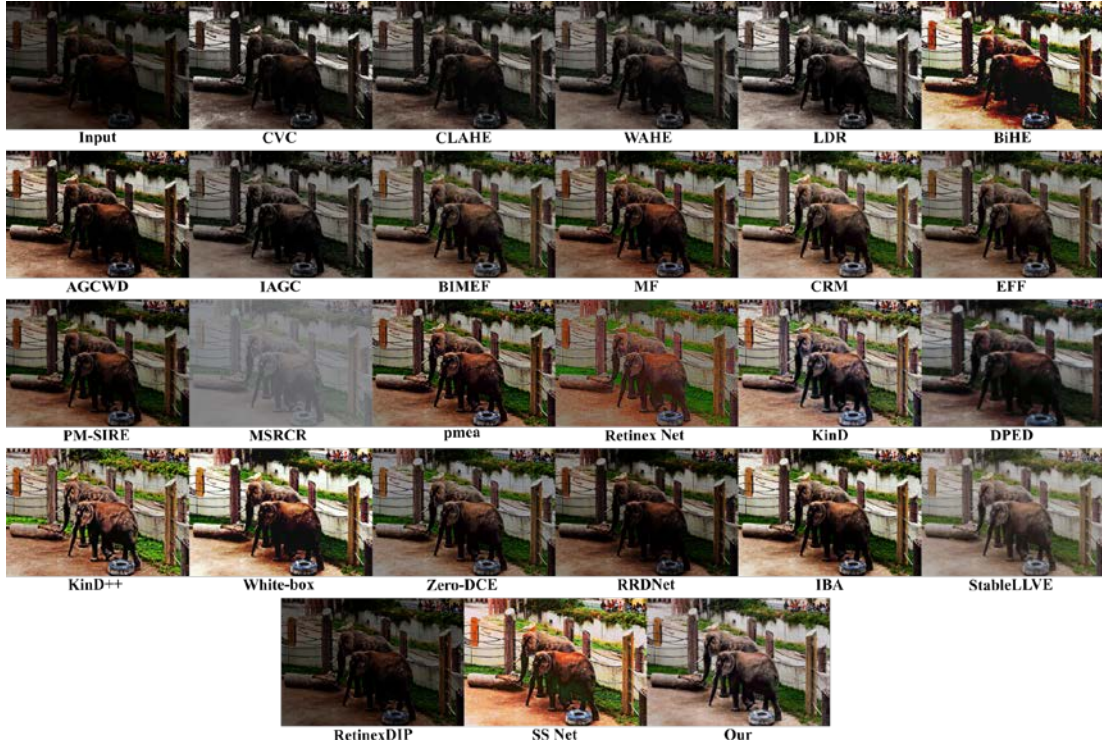


Fig. 9. Visual comparison with state-of-the-art methods on the input image from the Synthetic Low-Light (SLL) test dataset [26]. Please zoom in for a more detailed comparison.

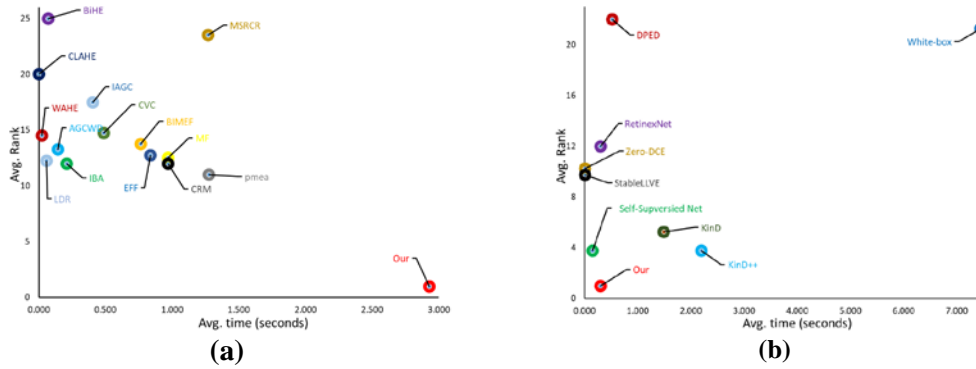


Fig. 10. For a better comparison, average time versus average rank is shown. (a) shows the average time versus average rank for classical methods and (b) shows the average time versus average rank for learning-based methods.

Table 6. Computing time comparison of deep learning-based methods on different resolution images (in seconds) on GPU (NVIDIA TITAN Xp).

Methods	600 × 400 × 3	960 × 640 × 3	2304 × 1728 × 3	Parameter	Avg.
Retinex Net [42]	0.155	0.162	0.591	0.440M	0.303
KinD [51]	0.334	0.604	3.539	0.255M	1.492
KinD++ [50]	0.337	0.857	5.408	8.275M	2.201
DPED [18]	0.247	0.282	1.040	7.760M	0.523
White-box [16]	6.040	6.483	9.833	8.560M	7.452
Zero-DCE [14]	0.0025	0.0026	0.021	0.079M	0.009
RRDNet [53]	59.479	128.217	893.0	0.128M	360.232
StableLLVE [48]	0.0047	0.005	0.0076	4.310M	0.006
RetinexDIP [52]	33.924	37.015	63.443	0.707M	44.794
Self-supervised Network [49]	0.022	0.054	0.366	0.485M	0.147
Ours	0.106	0.168	0.625	0.507M	0.300

Table 7. Computing time comparison of classical methods on different resolution images (in seconds) on CPU.

Methods	600 × 400×3	960 × 640×3	2304 × 1728×3	Avg.
CVC [4]	0.086	0.230	1.150	0.489
CLAHE [32]	0.00033	0.00099	0.0058	0.002
WAHE [2]	0.003	0.009	0.055	0.022
LDR [23]	0.013	0.027	0.136	0.059
BiHE [22]	0.016	0.036	0.161	0.071
AGCWD [17]	0.032	0.053	0.344	0.143
IAGC [3]	0.038	0.155	1.025	0.406
BIMEF [45]	0.123	0.359	1.811	0.764
MF [8]	0.215	0.422	2.275	0.971
CRF [48]	0.215	0.422	2.275	0.971
EFF [47]	0.136	0.407	1.973	0.839
PM-SIRE [7]	0.402	1.340	28.948	10.230
MSRCR [20]	0.322	0.704	2.787	1.271
pmea [34]	0.646	0.874	2.307	1.276
IBA [1]	0.0311	0.0829	0.512	0.209
Ours	0.618	1.080	7.093	2.930

4.7 Computational Comparison

We have compared the computing time (in seconds) of the proposed method with learning based method on GPU (NVIDIA TITAN Xp) in **Table 6** and classical ones on CPU (Intel(R) Core(TM) i7-6700 CPU @ 3.40GHz) in **Table 7**. All methods are computed using three different resolution images. The number of training parameters (in Million) is also reported in the second last column of **Table 6** for learning-based methods. The scatter plots of average time vs average rank for classical methods and learning-based methods are plotted in **Fig. 10a** and **Fig. 10b**, respectively. Only methods having a duration of less than 10 seconds are reported in these plots. The method closer to the origin of these graphs is more efficient and has better generalization ability. It can be seen from this that our method is least efficient as compared to classical methods, but it has better generalization ability. The simple nature of classical methods makes them faster, whereas they lack the model capacity to generalize well in real-world images. Furthermore, the Retinex theory-based classical methods (i.e., MSRCR, pmea, EFF, BIMEF) are less efficient as compared to Histogram Equalization (i.e., Bi-Histogram Equalization (BiHE), CVC, CLAHE, WAHE, LDR, CVC) and Gamma correction (i.e., IAGC, AGCWE) based methods. The main reason for the less efficient performance of the Retinex-based method is the need to decompose an image into illumination and reflectance maps first. The runtime comparison of the proposed method against learning-based methods is shown in **Fig. 10b**. The closeness of our method to the origin reveals it is efficient as

compared to the majority of methods. The zero-DCE, StableLLVE, and self-supervised networks are faster than the proposed methods due to their simpler network architecture. Due to the longer computation times, RetinexDIP and RRDNet are not included in the scatter plot. In both methods, errors are minimized iteratively. For RetinexDIP, 300 iterations are required to process a single image, whereas RRDNet does not have a fixed number of iterations. This makes these two methods more time-consuming.

5. Conclusion

The paper presents a lightweight and robust deep network based on pixel-wise polynomial estimation that can enhance low-light images. A deep convolution network is used to estimate the coefficient of higher-order polynomials in a supervised learning way. In addition, the proposed network uses four parallel branches to estimate pixel values based on different receptive fields, which enhances both local and global details. The combination of CNNs with higher-order polynomials facilitates the development of the lightweight model. The proposed network has demonstrated well-adjusted contrast, better color details, sharper details, and better generalization ability on six publicly available datasets, however, noise and color shift can be observed in the visual results on the DICM dataset. A denoising module will be designed as part of our future work to mitigate these issues. A SSIM-based objective function is used for the network's training. The SSIM uses the specific structure as a basis for assessing a prediction. It is, therefore, necessary to design a more rational loss function that is capable of evaluating not only any possible structure but also details (e.g., color, contrast, dynamic range, etc.). With our proposed network, we are able to improve an image for a resolution of 960 x 640 x 3 in just 0.168 seconds on the GPU. Despite this speed being adequate for real-time enhancement of images, it is not quite sufficient for low-light video enhancement. An additional research direction can be the optimization and improvement of the presented pipeline for video enhancement.

Acknowledgement

This work is supported by Ministry of Science and Technology China (MOST) Major Program on New Generation of Artificial Intelligence 2030 No. 2018AAA0102200. It is also supported by Natural Science Foundation China (NSFC) Major Projects No. U22A2097 and No. 61827814, as well as Shenzhen Science and Technology Innovation Commission (SZSTI) project No. JCYJ20190808153619413.

References

- [1] Al-Ameen, Z, "Nighttime Image Enhancement Using a New Illumination Boost Algorithm," *IET Image Processing*, vol. 13, no. 8, pp. 1314-1320, June, 2019. [Article \(CrossRef Link\)](#)
- [2] Arici T, Dikbas S, and Altunbasak Y, "A Histogram Modification Framework and Its Application for Image Contrast Enhancement," *IEEE Transactions on Image Processing*, vol. 18, no. 9, pp. 1921-1935, September, 2009. [Article \(CrossRef Link\)](#)
- [3] Cao G, Huang L, Tian H, Huang X, Wang Y, and Zhi R, "Contrast Enhancement of Brightness-Distorted Images by Improved Adaptive Gamma Correction," *Computers & Electrical Engineering*, vol. 66, pp. 569-582, April, 2018. [Article \(CrossRef Link\)](#)
- [4] Celik T, and Tjahjadi T, "Contextual and Variational Contrast Enhancement," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp. 3431-3441, December, 2011. [Article \(CrossRef Link\)](#)

- [5] Chen C, Chen Q, Xu J, & Koltun V, "Learning to see in the dark," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3291-3300, 2018. [Article \(CrossRefLink\)](#)
- [6] Cheng H D, and Shi X J, "A Simple and Effective Histogram Equalization Approach to Image Enhancement," *Digital Signal Processing*, vol. 14, no. 2, pp. 158-170, March, 2004. [Article \(CrossRefLink\)](#)
- [7] Fu X, Liao Y, Zeng D, Huang Y, Zhang X P, and Ding X, "A Probabilistic Method for Image Enhancement With Simultaneous Illumination and Reflectance Estimation," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4965-4977, December, 2015. [Article \(CrossRefLink\)](#)
- [8] Fu X, Zeng D, Huang Y, Liao Y, Ding X, and Paisley J, "A Fusion-Based Enhancing Method for Weakly Illuminated Images," *Signal Processing*, vol. 129, pp. 82-96, December, 2016. [Article \(CrossRefLink\)](#)
- [9] Gonzalez R C, Woods R E, and Masters B R, "Digital Image Processing, Third Edition," *Journal of Biomedical Optics*, vol. 14, no. 2, pp. 029901-1 - 209901-2, 2009. [Article \(CrossRefLink\)](#)
- [10] Grossberg M D, and Nayar S K, "Modeling the Space of Camera Response Functions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 10, pp. 1272-1282, October, 2004. [Article \(CrossRefLink\)](#)
- [11] Khan H, Wang X, Liu H, "Missing value imputation through shorter interval selection driven by Fuzzy C-Means clustering," *Computers & Electrical Engineering*, vol. 93, July, 2021. [Article \(CrossRefLink\)](#)
- [12] Gu K, Lin W, Zhai G, Yang X, Zhang W, and Chen C W, "No-Reference Quality Metric of Contrast-Distorted Images Based on Information Maximization," *IEEE Transactions on Cybernetics*, vol. 47, no. 12, pp. 4559-4565, December, 2017. [Article \(CrossRefLink\)](#)
- [13] Gu K, Tao D, Qiao J F, and Lin W, "Learning a No-Reference Quality Assessment Model of Enhanced Images With Big Data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 4, pp. 1301-1313, April, 2018. [Article \(CrossRefLink\)](#)
- [14] Guo C, Li C, Guo J, Loy C C, Hou J, Kwong S, & Cong R, "Zero-reference deep curve estimation for low-light image enhancement," in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1777-1786, 2020. [Article \(CrossRefLink\)](#)
- [15] Guo X, Li Y, and Ling H, "LIME: Low-Light Image Enhancement via Illumination Map Estimation," *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 982-993, February, 2017. [Article \(CrossRefLink\)](#)
- [16] Hu Y, He H, Xu C, Wang B, and Lin S, "Exposure: A white-box photo post-processing framework," *ACM Transactions on Graphics*, vol. 37, no. 2, pp. 1-17, July, 2018. [Article \(CrossRefLink\)](#)
- [17] Huang S C, Cheng F C, and Chiu Y S, "Efficient Contrast Enhancement Using Adaptive Gamma Correction With Weighting Distribution," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 1032-1041, March, 2013. [Article \(CrossRefLink\)](#)
- [18] Ignatov A, Kobyshev N, Timofte R, Vanhoey K, & Van G L, "Dslr-quality photos on mobile devices with deep convolutional networks," in *Proc. of the IEEE International Conference on Computer Vision*, pp. 3297-3305, 2017. [Article \(CrossRefLink\)](#)
- [19] Jiang Y, Gong X, Liu D, Cheng Y, Fang C, Shen X, Yang J, Zhou P, and Wang Z, "EnlightenGAN: Deep Light Enhancement Without Paired Supervision," *IEEE Transactions on Image Processing*, vol. 30, pp. 2340-2349, 2021. [Article \(CrossRefLink\)](#)
- [20] Jobson DJ, Rahman Z, and Woodell G A, "A Multiscale Retinex for Bridging the Gap between Color Images and the Human Observation of Scenes," *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 965-976, July, 1997. [Article \(CrossRefLink\)](#)
- [21] Jung E, Yang N, & Cremers D, "Multi-frame GAN: Image enhancement for stereo visual odometry in low light," in *Proc. of Conference on Robot Learning*, pp. 651-660, 2020. [Article \(CrossRefLink\)](#)
- [22] Yeong Taeg Kim, "Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization," *IEEE Transactions on Consumer Electronics*, vol. 43, no. 1, pp. 1-8, February, 1997. [Article \(CrossRefLink\)](#)

- [23] Lee C, Lee C, and Kim C S, “Contrast Enhancement Based on Layered Difference Representation of 2D Histograms,” *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5372-5384, December, 2013. [Article \(CrossRefLink\)](#)
- [24] Loh Y P, and Chan C S, “Getting to Know Low-Light Images with the Exclusively Dark Dataset,” *Computer Vision and Image Understanding*, vol. 178, pp. 30-42, January, 2019. [Article \(CrossRefLink\)](#)
- [25] Lore K G, Akintayo A, and Sarkar S, “LLNet: A Deep Autoencoder Approach to Natural Low-Light Image Enhancement,” *Pattern Recognition*, vol. 61, pp. 650-662, January, 2017. [Article \(CrossRefLink\)](#)
- [26] Lv F, Li Y, and Lu F, “Attention Guided Low-Light Image Enhancement with a Large Scale Low-Light Simulation Dataset,” *International Journal of Computer Vision*, vol. 129, no. 7, pp. 2175-2193, July, 2021. [Article \(CrossRefLink\)](#)
- [27] Lv F, Lu, F, Wu J, Lim C, “Mblen: Low-light image/video enhancement using cnns,” in *Proc. of British Machine Vision Conference*, 2018. [Article \(CrossRefLink\)](#)
- [28] Ma K, Kai Zeng, and Zhou Wang, “Perceptual Quality Assessment for Multi-Exposure Image Fusion,” *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345-3356, November, 2015. [Article \(CrossRefLink\)](#)
- [29] Mitsunaga T, & Nayar S K, “Radiometric self-calibration,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 374-380, 1999. [Article \(CrossRefLink\)](#)
- [30] Mittal A, Soundararajan R, and Bovik A C, “Making a ‘Completely Blind’ Image Quality Analyzer,” *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, March, 2013. [Article \(CrossRefLink\)](#)
- [31] Ng T T, Chang S F, & Tsui M P, “Using geometry invariants for camera response function estimation,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-8, 2007. [Article \(CrossRefLink\)](#)
- [32] Pizer S M, Johnston R E, Ericksen J P, Yankaskas B C, Muller K E, “Contrast-limited adaptive histogram equalization: Speed and effectiveness,” in *Proc. of the First Conference on Visualization in Biomedical Computing*, pp. 337-345, May, 1990. [Article \(CrossRefLink\)](#)
- [33] Pizer S M, Amburn E P, Austin J D, Cromartie R, Geselowitz A, Greer T, & Zuiderveld K, “Adaptive Histogram Equalization and Its Variations,” *Computer Vision, Graphics, and Image Processing*, vol. 39, no. 3, pp. 355-368, September, 1987. [Article \(CrossRefLink\)](#)
- [34] Pu T, and Wang S, “Perceptually Motivated Enhancement Method for Non-uniformly Illuminated Images,” *IET Computer Vision*, vol. 12, no. 4, pp. 424-433, June, 2018. [Article \(CrossRefLink\)](#)
- [35] Rahman S, Rahman M M, Abdullah-Al-Wadud, M, Al-Quaderi G D, and Shoyaib M, “An Adaptive Gamma Correction for Image Enhancement,” *EURASIP Journal on Image and Video Processing*, October, 2016. [Article \(CrossRefLink\)](#)
- [36] Rasheed M T, Guo G, Shi D, Khan H, and Cheng X, “An Empirical Study on Retinex Methods for Low-Light Image Enhancement,” *Remote Sensing*, vol. 14, no. 18, September, 2022. [Article \(CrossRefLink\)](#)
- [37] Rasheed M T, and Shi D, “LSR: Lightening Super-Resolution Deep Network for Low-Light Image Enhancement,” *Neurocomputing*, vol. 505, pp. 263-275, September, 2022. [Article \(CrossRefLink\)](#)
- [38] Rasheed M T, Shi D, and Khan H, “A Comprehensive Experiment-Based Review of Low-Light Image Enhancement Methods and Benchmarking Low-Light Image Quality Assessment,” *Signal Processing*, vol. 204, March, 2023. [Article \(CrossRefLink\)](#)
- [39] Wang J, Wang W, Wang R, and Gao W, “CSPS: An Adaptive Pooling Method for Image Classification,” *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1000–1010, March, 2016. [Article \(CrossRefLink\)](#)
- [40] Wang W, Wei C, Yang W, and Liu J, “Gladnet: Low-light enhancement network with global awareness,” in *Proc. of IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 751-755, May, 2018. [Article \(CrossRefLink\)](#)
- [41] Wang Z, Bovik A C, Sheikh H R, and Simoncelli E P, “Image Quality Assessment: From Error Visibility to Structural Similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April, 2004. [Article \(CrossRefLink\)](#)

- [42] Wei C, Wang W, Yang W, Liu J, “Deep retinex decomposition for low-light enhancement,” *arXiv preprint arXiv:1808.04560*, 2018. [Article \(CrossRefLink\)](#)
- [43] Xiao Z, Zhang X, Zhang F, Geng L, Wu J, Su L, and Chen L, “Diabetic Retinopathy Retinal Image Enhancement Based on Gamma Correction,” *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 1, pp. 149-154, February, 2017. [Article \(CrossRefLink\)](#)
- [44] Xu K, Yang X, Yin B, and Lau R W, “Learning to restore low-light images via decomposition-and-enhancement,” in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2278-2287, 2020. [Article \(CrossRefLink\)](#)
- [45] Ying Z, Li G, and Gao W, “A bio-inspired multi-exposure fusion framework for low-light image enhancement,” *arXiv preprint arXiv:1711.00591*, 2017. [Article \(CrossRefLink\)](#)
- [46] Ying Z, Li G, Ren Y, Wang R, and Wang W, “A new image contrast enhancement algorithm using exposure fusion framework,” in *Proc. of Computer Analysis of Images and Patterns, 17th International Conference*, pp. 36-46, August, 2017. [Article \(CrossRefLink\)](#)
- [47] Ying Z, Li G, Ren Y, Wang R, & Wang W, “A new low-light image enhancement algorithm using camera response model,” in *Proc. of the IEEE International Conference on Computer Vision Workshops*, pp. 3015-3022, 2017. [Article \(CrossRefLink\)](#)
- [48] Zhang F, Li Y, You S, & Fu Y, “Learning temporal consistency for low light video enhancement from single images,” in *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4965-4974, 2021. [Article \(CrossRefLink\)](#)
- [49] Zhang Y, Di X, Zhang B, Li Q, Yan S, Wang C, “Self-supervised low light image enhancement and denoising,” *arXiv preprint arXiv:2103.00832*, 2021. [Article \(CrossRefLink\)](#)
- [50] Zhang Y, Guo X, Ma J, Liu W, and Zhang J, “Beyond Brightening Low-Light Images,” *International Journal of Computer Vision*, January, vol. 129, pp. 1013-1037, 2021. [Article \(CrossRefLink\)](#)
- [51] Zhang Y, Zhang J, & Guo X, “Kindling the darkness: A practical low-light image enhancer,” in *Proc. of the 27th ACM International Conference on Multimedia*, pp. 1632-1640, October, 2019. [Article \(CrossRefLink\)](#)
- [52] Zhao Z, Xiong B, Wang L, Ou Q, Yu L, and Kuang F, “RetinexDIP: A Unified Deep Framework for Low-Light Image Enhancement,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 3, pp. 1076 - 1088, March, 2022. [Article \(CrossRefLink\)](#)
- [53] Zhu A, Zhang L, Shen Y, Ma Y, Zhao S, & Zhou Y, “Zero-shot restoration of underexposed images via robust retinex decomposition,” in *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 1-6, July, 2020. [Article \(CrossRefLink\)](#)
- [54] Khan H, Wang X, Liu H, “Handling missing data through deep convolutional neural network,” *Information Sciences*, vol. 595, pp. 278-293, May, 2022. [Article \(CrossRefLink\)](#)
- [55] Khan H, Wang X, Liu H, “A study on relationship between prediction uncertainty and robustness to noisy data,” *International Journal of Systems Science*, vol. 54, no. 6, pp. 1243-1258, January, 2023. [Article \(CrossRefLink\)](#)
- [56] Khan, H, Liu H, & Liu C, “Missing label imputation through inception-based semi-supervised ensemble learning,” *Advances in Computational Intelligence*, December, 2021. [Article \(CrossRefLink\)](#).
- [57] Ashraf S, Saleem S, Chohan A H, Aslam Z, Raza A, “Challenging strategic trends in green supply chain management,” *International Journal of Research in Engineering and Applied Sciences*, vol. 5, no. 2, pp. 71-74, March, 2020. [Article \(CrossRefLink\)](#)
- [58] Ahmad A, Ullah A, Feng C, Khan M, Ashraf S, Adnan M, Nazir S, Khan HU, “Towards an improved energy efficient and end-to-end secure protocol for iot healthcare applications,” *Security and Communication Networks*, vol. 2020, September, 2020. [Article \(CrossRefLink\)](#)
- [59] Ashraf S, Ahmed T, Aslam Z, Muhammad D, Yahya A, Shuaeeb M, “Depuration based Efficient Coverage Mechanism for Wireless Sensor Network,” *Journal of Electrical and Computer Engineering Innovations (JECEI)*, vol. 8, pp. 145-160, July, 2020. [Article \(CrossRefLink\)](#)



Muhammad Tahir Rasheed received the bachelor's degree from National University of Science & Technology, Pakistan in 2013 and master's degree from North China Electric Power University, China in 2016. He was a Lecturer at National University of Science & Technology, Pakistan from 2016 to 2019. He is currently pursuing his Ph.D. degree in computer science from Shenzhen University under the supervision of Prof. Daming Shi. His research interests include deep learning, low-light image enhancement, and super-resolution.



Daming Shi received the first Ph.D. degree in mechanical control from Harbin Institute of Technology, China, and the second Ph.D. degree in computer science from the University of Southampton, U.K. He was an Assistant Professor at Nanyang Technological University, Singapore, a Reader at Middlesex University, U.K, a Distinguished Professor at Harbin Institute of Technology. He is currently a Distinguished Professor at Shenzhen University, China. His current research interests include machine learning, medical image processing, pattern recognition, and neural networks.