EnGMHE: Enhanced Geometric Mean Histogram Equalization for Low-Light Image Enhancement

Rawan Zaghloul[1], Hazem Hiary[2]
Department of Management Information Systems, Al-Balqa Applied University, Amman, 11934, Jordan[1]
Department of Computer Science, The University of Jordan, Amman, 11942, Jordan[2]

***Abstract*—Low-light image enhancement has been extensively studied, with numerous methods proposed to address this challenge. Among these, Geometric Mean Histogram Equalization (GMHE) emerged as a histogram-based technique specifically designed for enhancing low-light images. Despite its effectiveness, GMHE has notable limitations: it often oversaturates results under specific conditions and amplifies noise, limiting its practical applicability. These shortcomings become particularly pronounced in real-world scenarios where low-light conditions are frequently accompanied by significant noise artifacts. To address these shortcomings, this study introduces EnGMHE, an enhanced version of GMHE. The proposed method consists of three key steps: 1) introducing a novel Gaussian Histogram Equalization (GHE) to improve image contrast and brightness, 2) utilizing GMHE to enhance sharpness and detail clarity, and 3) denoising the enhanced image using a pretrained deep neural network model. Together, these steps offer a more robust solution for low-light image enhancement, balancing contrast improvement, detail preservation, and noise reduction. The experimental results reveal not only the efficiency but also the effectiveness of the proposed model when benchmarked against the state-of-the-art methods.**

***Keywords*—Histogram Equalization; image enhancement; low-light enhancement; denoising; deep learning**

I. Introduction

Enhancing low-light images is essential across various domains, including surveillance, medical imaging, and photography, especially within the context of computer vision and intelligent systems. Insufficient lighting often leads to images with low brightness, poor contrast, and noticeable noise, severely affecting visual quality and obscuring critical details necessary for effective automated processing.

In recent decades, researchers have extensively explored solutions for low-light image enhancement. One of the classic methods is Histogram Equalization (HE) [1]. While HE effectively enhances local contrast on dark images, it often oversaturates bright regions in the process [2]. Another prominent approach involves Retinex-based methods [2-8], which decompose low-light images into reflectance and illumination layers. While this decomposition can yield promising enhancements, many Retinex-based techniques require priors for effective layer separation [9].

Recent advancements in low-light image enhancement leverage dehaze-based methods [10] and illumination map estimation techniques [2], [4]. With the rise of deep learning, enhanced computational power and large datasets have driven progress. Supervised learning methods relying on paired datasets [11] are used to map low-light to enhanced images, effectively reducing noise and recovering details. However, their performance depends on high-quality datasets, which remain a significant challenge.

Unsupervised learning methods utilize unpaired datasets [12], which are more accessible, larger in scale, and diverse in content. These methods are adaptively trained to restore illumination, color, and contrast. However, the lack of paired supervision can limit their ability to recover fine details, often leaving noise in the enhanced results. The field of low-light image enhancement continues to evolve, with researchers exploring methodologies that balance effectiveness with inherent trade-offs [9].

While deep learning methods often deliver impressive results, their efficiency remains a concern, particularly in real-world scenarios where low-light conditions are frequently accompanied by noise artifacts. To address this gap, this study introduces the EnGMHE (Enhanced Geometric Mean Histogram Equalization) model, a novel approach for low-light image enhancement. Unlike classical histogram-based methods, EnGMHE combines Gaussian Histogram Equalization (GHE) for initial contrast enhancement, GMHE for structural preservation, and a pre-trained denoising network for noise suppression, providing a unified and robust enhancement pipeline.

EnGMHE is built on the foundation of GMHE (Geometric Mean Histogram Equalization) [13], a contrast enhancement technique inspired by traditional histogram equalization. A key innovation of EnGMHE is the introduction of Gaussian Histogram Equalization (GHE), proposed as the first step to enhance the input low-light image's contrast and luminosity. Following GHE, the traditional GMHE is applied to further refine the enhancement. The final step involves denoising using a pre-trained deep neural network, effectively mitigating noise introduced during low-light image capture. This novel model integrates the newly proposed GHE with traditional and advanced techniques, delivering a robust solution for improving the visual quality of images captured in low-light conditions.

The effectiveness of the proposed model is validated through extensive experiments, where its performance is benchmarked against state-of-the-art methods in low-light image enhancement.

The rest of the study is as follows: Section II reviews the literature and related work; Section III details the proposed

methodology; Section IV presents the experimental settings and results; and Section V concludes the study.

II. Related Work

The problem of low-light image enhancement has gained significant attention in recent years, leading to the development of diverse methodologies ranging from classical techniques to modern deep learning methods. These methods aim to improve the brightness, contrast, and overall visual quality of low-light images while addressing challenges such as noise and color distortion.

Traditional methods, such as HE [1], BPDHE [14], and GMHE [13], focus on enhancing image contrast by redistributing pixel intensity values. HE is computationally efficient and straightforward, but it often amplifies noise and over-saturates the image.

Retinex theory-based approaches, such as LIME [2], NPE [15], and SRIE [4], decompose an image into illumination and reflectance components. By enhancing the illumination while preserving the reflectance, these methods aim to balance brightness and detail clarity.

Supervised and semi-supervised deep learning methods, such as LLNET [16], DRBN [17], RetinexNet [3], MBLLEN [18], LLFlow [19], PairLIE [20], and MIRNet [21], leverage paired datasets to learn complex mappings from low-light images to enhanced images. These methods excel in preserving fine details and reducing noise. For instance, DRBN incorporates recursive band representation for effective enhancement, while MIRNet utilizes multi-scale feature extraction for robust performance. However, their reliance on high-quality paired datasets can limit scalability. In addition, LLFlow employs flow-based mechanisms for illumination correction.

Unsupervised deep learning methods, such as EnlightenGAN [12] and SCI [22], address the challenge of paired dataset availability by learning from unpaired data. EnlightenGAN uses adversarial learning to generate visually appealing results. Despite their adaptability, these methods may struggle with detail preservation and noise reduction.

Recently, zero-shot learning methods rely on models that are trained on a set of classes (seen classes) but can make predictions on new, unseen classes without direct examples of those unseen classes during training. These methods leverage transfer learning to generalize knowledge from seen to unseen classes. Examples of such methods include ZeroDCE [23], ZeroDCE++ [24], RUAS [25], and ExCNet [26].

Transformer-based architectures like Restormer [27] and U-EGformer [28] have also demonstrated significant potential in low-light enhancement. Restormer employs attention mechanisms to capture global context, while U-EGformer integrates exposure-guided mechanisms for enhanced adaptability. Although these methods offer high performance, they often come with increased computational complexity.

Emerging techniques like MTUR-Net [29] focus on multi-task learning for joint enhancement and noise reduction. Additionally, methods like ISSR [30] and LCDPNet [31] leverage advanced architectures to address specific challenges in low-light scenarios. For underwater and domain-specific scenarios, methods such as WaterNet [32], Ucolor [33] and PUGAN [34], address challenges like haze, color distortion, and poor visibility. These methods utilize specialized architectures and priors to handle unique environmental conditions effectively.

While the aforementioned methods have significantly advanced the field, challenges such as noise suppression, detail preservation, and computational efficiency persist. Many existing techniques either focus on enhancing brightness and contrast or on reducing noise, but few address these issues jointly. Furthermore, achieving a balance between enhancement quality and real-world applicability remains a critical research gap.

III. Methodology

The proposed methodology involves a sequence of preprocessing and noise-reduction phases. They include advanced histogram operations and a denoising approach using a pre-trained neural network, as shown in Fig. 1.

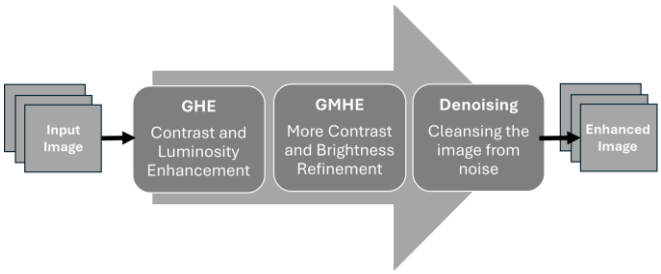

Fig. 1. The main phases of the proposed EnGMHE.

A. Gaussian Histogram Equalization (GHE)

This phase introduces the proposed GHE. It performs image enhancement through a Gaussian-smoothed histogram equalization technique. The following are the proposed steps for applying GHE to each of the image channels:

1) *Compute the histogram of the input image.*
2) *Apply Gaussian smoothing:*

A Gaussian filter, as defined in Eq. (1), is applied to the histogram to smooth intensity variations.

$$G(x,y) = \frac{1}{2\pi\tau^2} e^{-\frac{x^2+y^2}{\tau^2}} \quad (1)$$

where, *G(x,y)* is the value of the Gaussian function at pixel *(x,y)*. The parameter τ denotes the standard deviation of the Gaussian distribution and is empirically determined to control the degree of histogram smoothing. Pixel *(x,y)* represents the center of the filter. The filter size is determined using Eq. (2):

$$Filter_{Size} = 2 \times ceil(2 \times \tau) + 1 \quad (2)$$

τ is determined experimentally to control the degree of smoothing, as follows:

$\tau \in [0.5\text{-}1]$: Minimal smoothing.

$\tau \in [1.5\text{-}3]$: Moderate smoothing (balanced).

$\tau > 3.5$: Significant smoothing.

3) Normalize the smoothed histogram: The smoothed histogram h is normalized such that its values sum to 1, forming a probability distribution, as defined in Eq. (3):

$$Norm\_h_i = \frac{h_i}{\sum_i^p (h)}, i = \{1, \dots p\} \quad (3)$$

where, h_i is the value of the i^{th} bin in the histogram h, and p denotes the total number of bins.

4) Compute the cumulative distribution function (CDF) for the normalized histogram, as defined in Eq. (4):

$$CDF\_h_i = \sum_{j=0}^{i} Norm\_h_j, i = \{1, \dots p\} \quad (4)$$

5) Map pixel intensity values: Each pixel intensity i is transformed to a new value, as defined in Eq. (5):

$$New\_intensity\_level_i = CDF\_h_i * (L-1), i = \{1, \dots p\} \quad (5)$$

where, L is the maximum intensity value (265 in this study).

6) Output the enhanced image: Finally, the output image I' is generated by mapping the intensity of each pixel in the original image I according to the new intensity values stored in $New\_intensity\_level$. Note that steps (3) to (6) are the steps of the traditional HE method.

B. Geometric Mean Histogram Equalization (GMHE)

The key idea behind GMHE [13] is to preserve the overall structure of the image while enhancing contrast, particularly in dark regions. This is achieved by redistributing pixel intensities based on the geometric mean, which is less sensitive to extreme values than the arithmetic mean and therefore reduces the influence of outliers in the histogram.

In this phase, GMHE is applied to the intermediate enhanced image I'. The procedure is performed independently on each image channel as follows:

For each channel in I' apply the following steps:

1) Compute the histogram S of the input image.

2) Apply a geometric mean filter to the histogram to obtain a modified histogram (S') using Eq. (6):

$$S' = (\prod_{i=1}^{n} x_i)^{1/\gamma} \quad (6)$$

where, x_i represents the histogram values within the filter window, n denotes the filter length (set to 3 in this work), and γ is an experimentally determined constant that controls the strength of the geometric mean operation.

3) Apply steps (3–6) from the first phase (GHE) to the modified histogram S' to generate the enhanced image (I'').

C. Denoising

To suppress noise, while preserving image fidelity, a channel-wise image denoising is applied to the enhanced image I'' using a pre-trained deep learning model described in [35]. This network was selected due to its demonstrated effectiveness and stability in general-purpose image denoising tasks, particularly in scenarios involving low-light and noise-amplified images.

As shown in Fig. 2, the adopted model consists of 19 convolutional blocks, where each block is composed of a convolutional layer followed by batch normalization and a ReLU activation function. This relatively deep architecture enables the extraction of hierarchical features, ranging from low-level structures such as edges and intensity gradients to higher-level spatial patterns, which is essential for discriminating noise from meaningful image content. Padding is employed in all convolutional layers to preserve the spatial dimensions of the input image, thereby preventing the loss of structural details during the denoising process.

The use of a pre-trained denoising network is motivated by both technical and practical considerations. Leveraging a pre-trained model allows the proposed framework to benefit from robust, previously learned noise characteristics without requiring dataset-specific retraining, which enhances computational efficiency and generalizability. Moreover, the regression-based output layer is specifically designed to estimate clean image intensities, enabling effective noise suppression, while maintaining important visual and structural details. This balance between noise reduction and detail preservation makes the selected network particularly suitable for integration into the proposed EnGMHE-based enhancement pipeline.

IV. Experimental Results and Analysis

A. Datasets Used in the Comparisons

The datasets used in this study include both paired and unpaired collections. The ExDark-Bicycle [36], NPE [15], DICM [37], VV [38], MEF [39], LIME [2], and LoLi-Phone [40] datasets consist of 652, 84, 64, 24, 17, 10, and 600 real low-light images, respectively, with varying resolutions. These datasets are unpaired and cover a wide range of indoor and outdoor scenes, including natural landscapes, buildings, and indoor objects or decorations.

In contrast, the LOL [3], MIT-5K [41], SICE [42], and UIEB [32] datasets are paired datasets. The LOL dataset includes LOL-15, which comprises 15 test images, and LOL-V2, which contains two subsets: 100 synthetic and 100 real test images. The MIT-5K provides 5000 low-light images, with 500 designated for testing. The SICE dataset contains 4800 real and synthetic multi-exposure images of various resolutions, capturing diverse indoor and outdoor scenes under different exposure levels. In this study, two variations of the SICE dataset, namely SICE_Grad and SICE_Mix[1], were used, each consisting of 589 paired images. Finally, two subsets of the UIEB dataset were utilized: one comprising 890 paired images and another challenging subset containing 60 images.

B. Metrics

This section presents an overview of the evaluation metrics employed in this study to assess the performance of the proposed methods. Both reference-based metrics, which rely on a ground-

[1] SICE_Grad and SICE_Mix datasets are available at "https://github.com/ShenZheng2000/LLIE_Survey".

truth image for comparison, and no-reference metrics, which evaluate image quality without relying on a reference image, are considered. These metrics were selected to provide a comprehensive assessment of image quality and processing effectiveness across diverse experimental scenarios.

1	'InputLayer'	Image Input	50x50x1 images
2	'Conv1'	Convolution	64 3x3x1 convolutions with stride [1 1] and padding [1 1 1 1]
3	'ReLU1'	ReLU	ReLU
4	'Conv2'	Convolution	64 3x3x64 convolutions with stride [1 1] and padding [1 1 1 1]
5	'BNorm2'	Batch Normalization	Batch normalization with 64 channels
6	'ReLU2'	ReLU	ReLU
7	'Conv3'	Convolution	64 3x3x64 convolutions with stride [1 1] and padding [1 1 1 1]
8	'BNorm3'	Batch Normalization	Batch normalization with 64 channels
9	'ReLU3'	ReLU	ReLU
⋮	⋮	⋮	⋮
55	'Conv19'	Convolution	64 3x3x64 convolutions with stride [1 1] and padding [1 1 1 1]
56	'BNorm19'	Batch Normalization	Batch normalization with 64 channels
57	'ReLU19'	ReLU	ReLU
58	'Conv20'	Convolution	1 3x3x64 convolutions with stride [1 1] and padding [1 1 1 1]
59	'FinalRegressionLayer'	Regression Output	mean-squared-error with response 'Response'

Fig. 2. The structure of the adopted pre-trained network.

1) Reference-based measures

a) SSIM (Structural Similarity Index) is a perceptual metric that evaluates the similarity between two images by considering their luminance, contrast, and structural information. SSIM values range from -1 to 1, where a value of 1 indicates perfect similarity. It is widely used in image processing tasks to assess the quality of reconstructed images. It is defined in Eq. (7) [43]:

$$SSIM(a,b) = \frac{(2\mu_a\mu_b+(k_1L)^2)(2\sigma_{ab}+(k_2L)^2)}{(\mu_a^2+\mu_b^2+(k_1L)^2)(\sigma_a^2+\sigma_b^2+(k_2L)^2)} \quad (7)$$

where, a and b denote the input and the enhanced images, respectively. μ denotes the mean of the specified image. σ_a^2 and σ_b^2 denote the variances of images a and b, respectively, while σ_{ab} represents the covariance between a and b. The constants (K_1, K_2, L) are set to *0.01*, *0.03*, and *255* for 8-bit images, respectively.

b) PSNR (Peak Signal-to-Noise Ratio) is a quantitative metric used to evaluate the quality of an image by measuring the ratio between the maximum possible power of an image and the power of noise that affects the image's fidelity. Expressed in decibels (dB), PSNR quantifies image quality by comparing the original and distorted images, with higher values indicating better quality [44]. It is defined as follows in Eq. (8) and Eq. (9):

$$PSNR(a,b) = 10.\log_{10}(\frac{L^2}{MSE(a,b)}) \quad (8)$$

$$MSE(a,b) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[a_{i,j} - b_{i,j}]^2 \quad (9)$$

where, a and b are the input and the enhanced images, respectively. L is a constant referring to the maximum possible pixel value, which is equal to 255 for 8-bit images. MSE denotes the mean square error between two images. M and N are the dimensions of the image (width and height). $a_{i,j}$ denotes the pixel value at position (i,j) in image a.

c) LPIPS (Learned Perceptual Image Patch Similarity) is a recently developed metric designed to evaluate image quality from the perspective of human perception. Unlike traditional metrics, such as *PSNR* and *SSIM*, which focus on pixel-level comparisons, *LPIPS* leverages deep learning features to align more closely with human visual perception [45], [46]. It measures the similarity between an image and its corresponding ground truth by calculating the difference between their feature representations, extracted from a pre-trained deep neural network (e.g., *VGG* or *AlexNet*), as described in Eq. (10):

$$LPIPS\ (I_1,I_2) = \sum_l w_l \|f_l(I_1) - f_l(I_2)\|_2 \quad (10)$$

where, f_l denotes the feature maps from layer l of the network, w_l represents the learned weights for each layer l, and $\|.\|_2$ denotes the L2-Norm (Euclidean distance) of the deep features of a trained CNN.

d) Delta_E is a reference-based measure that quantifies the difference between two colors in the *CIELAB* color space. It is widely used to evaluate color accuracy, identify deviations, and assess color similarity. Lower Delta-E values indicate better color consistency and are therefore preferred [47]. Delta-E is defined as shown in Eq. (11):

$$Delta\_E = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\sqrt{(L1_{i,j} - L2_{i,j})^2 + (a1_{i,j} - a2_{i,j})^2 + (b1_{i,j} - b2_{i,j})^2} \quad (11)$$

where, ($L1_{i,j}$, $a1_{i,j}$, $b1_{i,j}$) are the LAB values for a pixel (i, j) in the input image, and ($L2_{i,j}$, $a2_{i,j}$, $b2_{i,j}$) correspond to those of the enhanced image. The overall score is obtained by averaging the pixel-wise Delta-E values, yielding a single scalar that reflects the overall perceptual color fidelity.

2) No reference-based measures

a) BRISQUE (Blind/Reference-less Image Spatial Quality Evaluator) is a no-reference image quality assessment model that operates in the spatial domain to evaluate image quality without relying on distortion-specific features or frequency-domain transformations, such as DCT or wavelets. It

uses natural scene statistics (NSS) of locally normalized luminance coefficients and their products to quantify the loss of image "naturalness" caused by distortions, offering a holistic and distortion-generic measure of quality. BRISQUE is a model-driven measure which does not have a single closed-form analytical equation. Instead, it employs a machine learning pipeline that estimates image quality using NSS features and a trained regressor, as shown in Eq. (12):

$$BRISQUE(I) = SVR(F(I)) \quad (12)$$

where, I denotes the input low-light image. F(.) represents a 36-dimensional NSS feature vector extracted from I. SVR(.) is the pre-trained support vector regressor. The feature vector F is derived from locally normalized luminance coefficients and is computed by fitting an Asymmetric Generalized Gaussian Distribution (AGGD) to the normalized image and its directional pairwise product maps at two image scales [48].

b) NIQE (The Natural Image Quality Evaluator) is a no-reference image quality assessment metric that measures the degree to which an image's features deviate from the statistical regularities commonly observed in natural images. It is built upon a robust spatial-domain NSS model designed to capture quality-aware statistical features learned from a dataset of natural, undistorted images. Lower NIQE scores indicate better perceptual quality. The quality score is computed as the Mahalanobis distance between the image's feature vector and a multivariate Gaussian (MVG) model fitted on high-quality images [49].

c) UCIQE (Underwater Color Image Quality Evaluation) is a no-reference image quality assessment metric specifically designed for underwater images. It evaluates visual quality in the CIELab color space by analyzing three key factors: Chroma, which measures the image's colorfulness and reflects the loss of color diversity due to light scattering and absorption in underwater environments; Saturation, which assesses the vividness or dullness of colors; and Contrast, which quantifies the visibility of image details often diminished in underwater scenes because of haze or turbidity. The UCIQE metric is computed using Eq. (13):

$$UCIQE = c_1 \times std_c + c_2 \times con_1 + c_3 \times Avg_s \quad (13)$$

where, std_c denotes the standard deviation of chroma, con_1 denotes the contrast of luminance, and Avg_s represents the average of saturation. The coefficients c1, c2, and c3 are empirically determined (c_1 =0.4680, c_2 =0.2745, c_3=0.2576). This metric provides a quantitative evaluation of underwater image quality, enabling the assessment and enhancement of visual clarity and color fidelity [50].

C. Experimental Environment and Parameter Settings

To ensure the reproducibility and reliability of the experiments, a controlled experimental setup was carefully established, and specific parameter configurations were defined. The following details provide an overview of the conditions under which the experiments were conducted:

- All experiments in this study were conducted in the MATLAB environment and were executed on an *Intel(R)* Core(TM) i7 CPU @ 2.60GHz device with *12 GB* of RAM.
- In the context of EnGMHE, the parameters τ and γ control histogram smoothing and geometric mean redistribution, respectively. Default values of τ = 2 and γ = 25 provide robust performance across most low-light datasets. For datasets with specific characteristics, such as SICE and ExDark, slightly different values (see Table I) were used to achieve optimal enhancement. These exceptions do not require exhaustive per-dataset tuning; they simply reflect natural variations in image statistics across diverse datasets, while the method remains effective with the general default settings.

TABLE I. PARAMETER SETTINGS FOR THE DATASETS OF INTEREST

Dataset	τ	γ
MEF | NPE | LOL | DICM | LIME | UIEB | VV | MIT5K	2	25
ExDark	10	15
SICE	2	50

D. Qualitative Results

Fig. 3 shows the performance of various low-light enhancement methods on an image from the LIME dataset. The input image is underexposed, with significant loss of visibility in shadowed regions. SCI [22] brightens the image but introduces a strong purplish tint, making the colors appear unnatural. URetinexNet [51] enhances the brightness and detail visibility but overexposes certain regions, resulting in a reddish hue. In contrast, EnGMHE achieves a balanced enhancement, maintaining natural color while preserving fine details.

Fig. 4 presents a comparison, where the input is an underwater image suffering from poor visibility and color distortion. Compared to state-of-the-art methods, EnGMHE delivers visually consistent results with enhanced clarity and more accurate color restoration for underwater images.

As shown in Fig. 5(a), the input image represents challenging low-light photography, with significant visibility issues in darker regions. While the illuminated monument retains some visibility, the surrounding darker areas lack detail, making the image a strong candidate for enhancement. EnGMHE effectively balances brightness enhancement with structural detail preservation, whereas FDMLNet [55] overexposes bright areas (see the monument's facade). In Fig. 5(b), the input image presents complex lighting conditions, with a mix of shadowed areas and bright overhead lights. ZeroDCE++ and RetinexNet enhance brightness, but introduce noticeable artifacts and color distortions. KinD++ [56] over-saturates certain regions, leading to unnatural visual results. LIVENet [57] produces more natural adjustments in terms of brightness but fails to recover fine details in shadowed regions [Fig. 5(c)]. EnGMHE, however, delivers the most balanced output, effectively preserving both color fidelity and structural details while enhancing brightness.

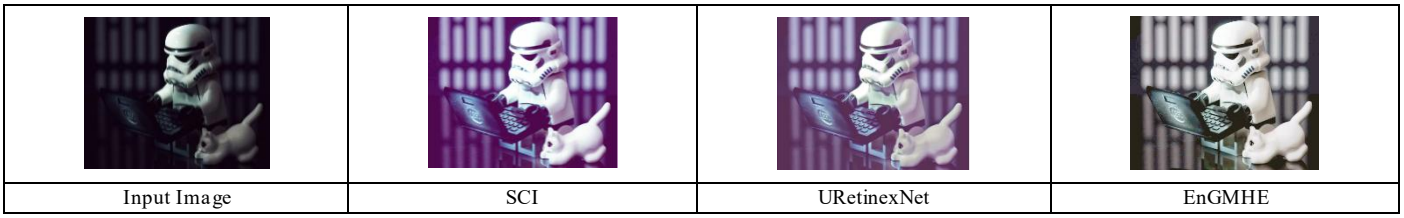

Fig. 3. Visual comparison with state-of-the-art methods on an input image from the LIME dataset.

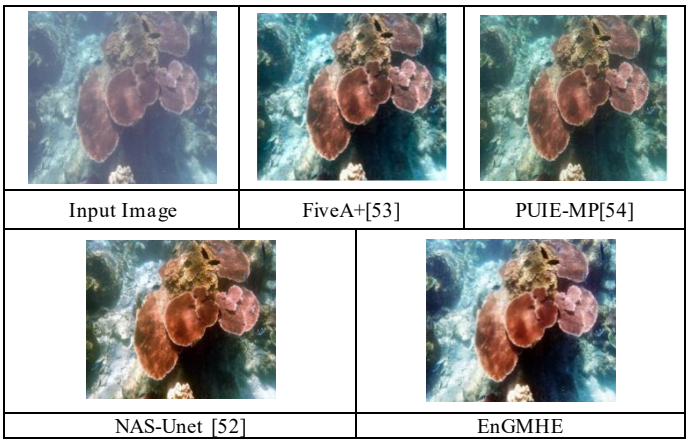

Fig. 4. Visual comparison with state-of-the-art methods on an input image from the UIEB-890 dataset.

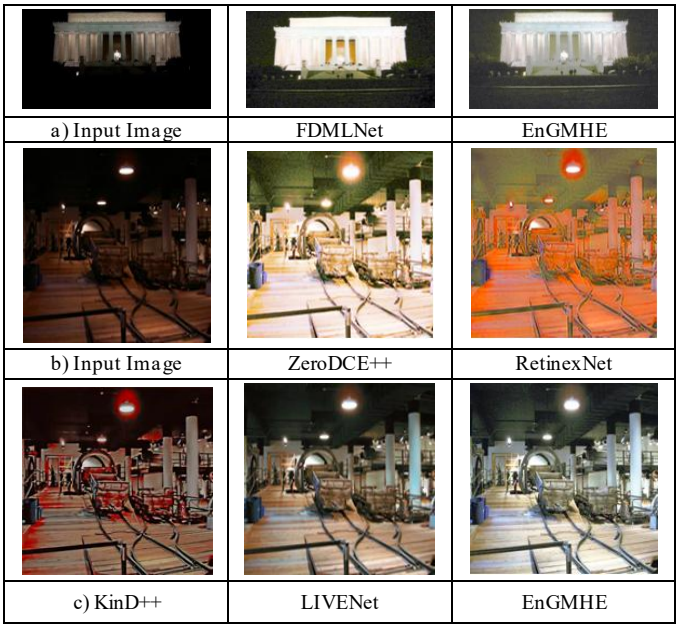

Fig. 5. Visual comparison with state-of-the-art methods on input images from the DICM dataset.

Fig. 6 shows a comparison with various state-of-the-art methods (KinD [58], SNR-Aware [59], and SPIC [60]) on challenging low-light and visually complex images from different datasets (NPE, VV, and ExDark). The overall results consistently demonstrate that EnGMHE excels in detail preservation, color accuracy, and balanced brightness enhancement across diverse datasets.

Fig. 7 evaluates the performance of LLIEDiff and EnGMHE against the ground truth for a low-light input image. The input image is significantly dark, with details obscured across various regions, while the ground truth serves as an ideal reference with vibrant colors, sharp textures, and excellent detail visibility. LLIEDiff improves brightness and restores some visibility in shadowed areas but struggles with color accuracy and fine detail preservation, particularly in the regions highlighted by the colored rectangles. In contrast, EnGMHE delivers a more balanced enhancement, achieving restored brightness, natural color tones, and excellent detail clarity that closely aligns with the ground truth. While EnGMHE may not perfectly match the ground truth in all regions, it demonstrates better color fidelity than LLIEDiff, producing vibrant and realistic colors and preserving natural tones, such as the red object and green grapes, with a closer resemblance to the ground truth in hue and saturation.

Fig. 8 highlights the text extraction results from a dark image and the EnGMHE-enhanced image. The text in this figure was recognized using "Image Recognize[2]" tool. As shown in Fig. 8, the EnGMHE-enhanced image not only retains high confidence for clearly visible text like "RESTAURANT" but also reveals additional details such as "Hpidavos" with significantly improved confidence (94%) compared to the dark image (44%). Furthermore, it introduces new textual elements such as "GROUP" and "be TEA", demonstrating its capability in enhancing low-light images for accurate text recognition.

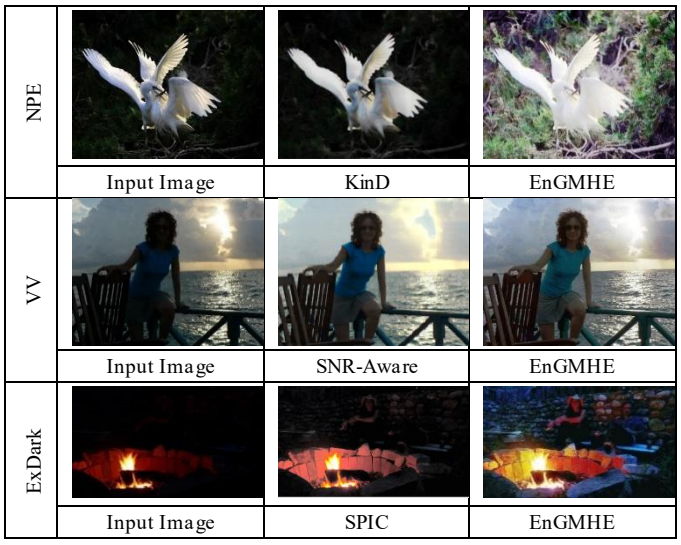

Fig. 6. Visual comparison with state-of-the-art methods on (NPE, VV, and Exdark) datasets.

[2] https://imagerecognize.com/text/#site-content

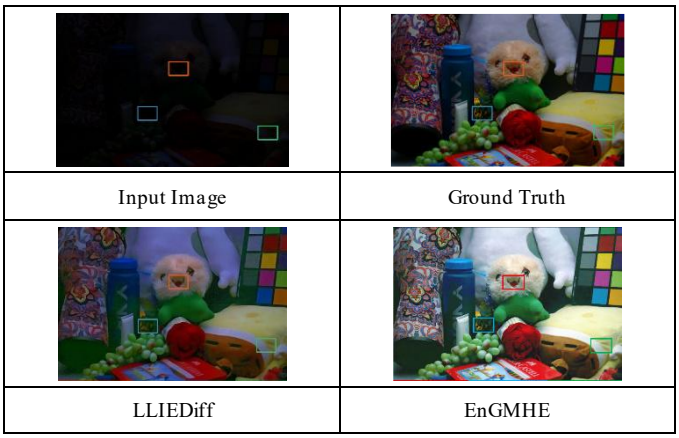

Fig. 7. Comparison with LLIEDiff against EnGMHE on the input image from the LOL-15 dataset.

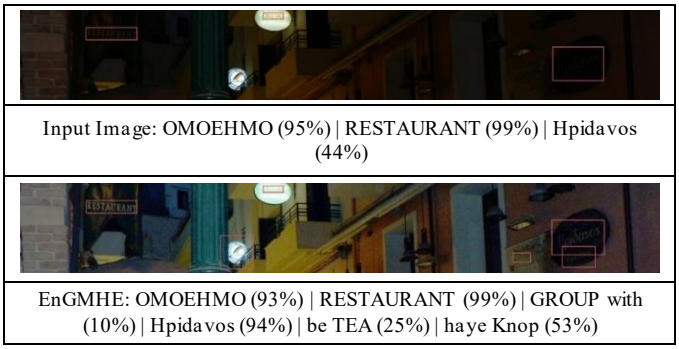

Fig. 8. Text extraction test. Red rectangles in the figure highlight the extracted text (confidence %) from the original dark image and the EnGMHE-enhanced image, respectively. The input image is a cropped portion of an image from the VV dataset.

E. Qualitative Results

This section quantitatively presents the comparative results obtained using both reference-based and non-reference-based methods across various datasets.

1) Comparisons using reference-based measures: The first comparisons were conducted to compare SSIM scores on different datasets. As shown in Table II, the proposed method achieved impressive SSIM scores of (0.72 and 0.78) on the LOL-V2 and MIT5K-500 datasets, respectively. On LOL-V2, it outperformed EnlightenGAN (0.68) and demonstrated a significant improvement over ZeroDCE [23] (0.58) and LIME (0.47), highlighting its ability to preserve structural details. On MIT5K-500, EnGMHE reported a score of 0.78, ranking second only to LIME (0.80). For SICE_GRAD, as presented in Table III, the proposed method achieved the highest score of (0.67), surpassing U-EGformer (0.64) and ZeroDCE (0.64). On SICE_MIX, it obtained a score of (0.64), outperforming URetinexNet (0.60) and RUAS (0.493).

On the LOL-15 dataset, as depicted in Table IV, the proposed method demonstrated a superior SSIM score of (0.734), significantly outperforming competitors like BIMEF [61] (0.595), EnlightenGAN (0.652), and ZeroDCE (0.559).

TABLE II. COMPARATIVE RESULTS ON SSIM ↑ ON (LOL-V2, MIT5K-500) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE, AND THE SECOND-RATED RESULTS ARE UNDERLINED

Method	*LOL-V2*	*MIT5K-500*
EnGMHE	**0.72**	0.78
LDR [37]	-	0.73
DRBN [17]	-	0.76
EnlightenGAN [12]	0.68	0.76
LIME [2]	0.47	**0.80**
ZeroDCE [23]	0.58	0.72

TABLE III. COMPARATIVE RESULTS ON SSIM ↑ ON (SICE GRAD AND SICE MIX) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE, AND THE SECOND-RATED RESULTS ARE UNDERLINED

Method	*SICE GRAD*	*SICE MIX*
EnGMHE	**0.67**	0.64
ZeroDCE [23]	0.64	0.64
RetinexNet [3]	0.60	0.61
URetinexNet [51]	0.61	0.60
RUAS [25]	0.49	0.49
SGZ [62]	0.60	0.62
LLFlow [19]	0.62	0.61
U-EGformer [28]	0.64	**0.65**

TABLE IV. COMPARATIVE RESULTS ON SSIM ↑ ON (LOL-15) DATASET, THE BEST RESULT IS DENOTED BY BOLDFACE AND THE SECOND-RATED RESULT IS UNDERLINED

EnGMHE	**EnlightenGAN**	**BIMEF [61]**	**LIME [2]**
0.734	0.652	0.595	0.484
STAR [64]	DRBN [17]	RRDNet [65]	RUAS [25]
0.518	0.551	0.457	0.500
ZeroDCE [23]	SDD [63]	SCI [22]	
0.559	0.637	0.522	

TABLE V. COMPARATIVE RESULTS ON PSNR ↑ ON (SICE_GRAD AND SICE_MIX) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

Dataset	**EnGMHE**	**Retinex Net [3]**	**URetinexNet [51]**	**KinD++ [56]**	**UEGformer [28]**
SICE_GRAD	**14.22**	12.45	10.89	13.24	13.27
SICE_MIX	13.66	12.39	10.90	13.19	**14.23**
Dataset	**RUAS [25]**	**ZeroDCE [23]**	**SCI [22]**	**LLFlow [19]**	**SGZ [62]**
SICE_GRAD	8.62	12.48	8.55	12.74	10.99
SICE_MIX	8.68	12.42	8.64	12.73	10.86

TABLE VI. COMPARATIVE RESULTS ON PSNR ↑ ON (LOL-15) DATASET, THE BEST RESULT IS DENOTED BY BOLDFACE AND THE SECOND-RATED RESULT IS UNDERLINED

EnGMHE	**RUAS [25]**	**DRBN [17]**	**SCI [22]**
16.29	**16.40**	16.29	14.78
SDD [63]	STAR [64]	ZeroDCE [23]	RRDNet [65]
13.34	12.91	14.86	11.40

The second set of comparisons was conducted to compare PSNR scores across various datasets. As shown in Table V, EnGMHE achieved a PSNR of 14.22 on the SICE_GRAD dataset, surpassing KinD++ (13.24) and U-EGformer (13.27). Similarly, on SICE_MIX, it recorded the second-highest PSNR with a score of 13.66, following the U-EGformer (14.23). On the LOL-15 dataset (Table VI), EnGMHE achieved a PSNR of 16.29, tying with DRBN for the top performance while outperforming other notable methods, such as ZeroDCE (14.86) and SCI (14.78). Additionally, on the MIT5K-500 dataset (Table VII), EnGMHE recorded the best PSNR score of 18.12, surpassing DRBN (16.37) and LIME (16.07), demonstrating its exceptional noise suppression and fidelity on this dataset.

The third set of comparisons evaluated the performance based on the LPIPS measure, where lower scores indicate better perceptual similarity. On the SICE_GRAD dataset (Table VIII), EnGMHE achieved a score of 0.33 as Kind++, securing the second rank, followed by U-EGformer with a score of 0.27, demonstrating strong perceptual similarity. On the LOL-15 dataset (Table IX), EnGMHE achieved the best score of 0.194, demonstrating superior perceptual quality, while methods like SDD (0.743) and STAR (0.366) performed significantly worse. For the LOL-V2 dataset (Table X), EnGMHE again achieved the best score of (0.204), slightly outperforming ISSR (0.206). Other competitors, such as ZeroDCE (0.313) and RetinexNet (0.365), showed comparatively weaker performance, further emphasizing EnGMHE's effectiveness in maintaining perceptual fidelity across various datasets.

TABLE VII. COMPARATIVE RESULTS ON PSNR ↑ ON (MIT5K-500) DATASET, THE BEST RESULT IS DENOTED BY BOLDFACE AND THE SECOND-RATED RESULT IS UNDERLINED

EnGMHE	EnlightenGAN	LIME [2]	DRBN	ZeroDCE
18.12	15.58	16.07	<u>16.37</u>	12.82

TABLE VIII. COMPARATIVE RESULTS ON LPIPS ↓ ON (SICE_GRAD) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

EnGMHE	RetinexNet [3]	URetinexNet	RUAS	SCI [22]
<u>0.33</u>	0.36	0.35	0.49	0.48
LLFlow [19]	SGZ [62]	KinD [58]	KinD++ [56]	U-EGformer [28]
0.38	0.36	0.34	<u>0.33</u>	**0.27**

TABLE IX. COMPARATIVE RESULTS ON LPIPS ↓ ON (LOL-15) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

EnGMHE	SDD [63]	STAR [64]	DRBN [17]
0.194	0.743	0.366	0.316
RRDNet [65]	RUAS [25]	SCI [22]	ExCNet [26]
0.362	0.270	0.339	0.373
MBLLEN [18]	RetinexNet [3]	GLADNet [66]	ZeroDCE [23]
<u>0.225</u>	0.379	0.321	0.335
PairLIE [20]	LLNET [16]	EnlightenGAN [12]	
0.248	0.360	0.322	

TABLE X. COMPARATIVE RESULTS ON LPIPS ↓ ON (LOL-V2) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

EnGMHE	RetinexNet [3]	ZeroDCE [23]	EnlightenGAN [12]
0.204	0.365	0.313	<u>0.309</u>
KinD [58]	ISSR [30]	MIRNet [21]	HWMNet [67]
0.375	0.206	0.317	0.359

TABLE XI. COMPARATIVE RESULTS ON DELTA-E ↓ ON (LOL-15) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

EnGMHE	SDD [63]	STAR [64]	MBLLEN [18]	Retinex Net [3]	KinD [58]
12.42	21.83	23.46	13.68	12.69	<u>12.49</u>
DRBN [17]	ZeroDCE [23]	RRDNet [65]	RUAS [25]	SCI [22]	EnlightenGAN [12]
13.44	18.81	26.43	16.85	19.52	14.5

The fourth comparison focused on evaluating Delta-E on the LOL-15 dataset (Table XI). EnGMHE achieved the lowest Delta-E score of (12.42), signifying minimal color distortion and high fidelity in color reproduction. The next closest competitor was KinD with a score of 12.49, while methods like ZeroDCE (18.81) and RRDNet (26.43) exhibited substantially higher errors, highlighting their relatively poorer performance in preserving color accuracy.

2) Comparisons using no reference-based measures: As depicted in Table XII, EnGMHE demonstrated excellent performance in minimizing NIQE scores across diverse datasets, underscoring its effectiveness in achieving high-quality, perceptually natural results. Its ability to outperform contemporary methods on DICM, NPE, and MEF highlighted its superiority in the domain of low-light image enhancement. Additionally, its performance on LIME and VV-24 was competitive with well-known methods such as EnlightenGAN and KinD++. Traditional methods like HE and KinD consistently recorded higher NIQE scores, reflecting their limitations in maintaining perceptual quality. Furthermore, as depicted in Table XIII and Table XIV, EnGMHE achieved the best performance on the MIT5K-500, LOL-V2, and LOL-15 datasets.

According to the BRISQUE measure (Table XV and Table XVI), EnGMHE led across all datasets, showing superior perceptual quality on VV-24 and LIME datasets and the second rank on DICM, NPE, and MEF datasets. This consistent performance across diverse datasets underscores EnGMHE's robustness and adaptability in various enhancement scenarios.

For the UCIQE measure (Table XVII), EnGMHE demonstrated a strong capability in underwater image enhancement, achieving the highest score on the UIEB-Ch60. It outperformed notable methods, including: MTUR-Net [29], WaterNet [32], Ucolor [33], IDMR [73], Restormer [27], FUnIEGAN [74], PUGAN [34], U-Shape [75], Semi-UIR [76], and UWE-Net [77], setting a new benchmark in this domain.

EnGMHE consistently outperforms or competes strongly with state-of-the-art techniques across a wide range of benchmark datasets and evaluation metrics. In reference-based

assessments, it achieved top or second-best SSIM and PSNR scores on datasets, including LOL-V2, MIT5K-500, SICE_GRAD, and LOL-15, indicating excellent structural preservation and noise suppression. It also demonstrated superior perceptual quality with low LPIPS values and achieved the best Delta-E score on LOL-15, highlighting its ability to maintain color accuracy. In no-reference evaluations, EnGMHE recorded the lowest or among the lowest NIQE and BRISQUE scores across diverse datasets such as VV-24, DICM, NPE, and MEF, highlighting its robustness in producing perceptually natural images. Additionally, EnGMHE outperformed other methods on the UCIQE metric in underwater image enhancement, setting a new benchmark in that domain.

TABLE XII. COMPARATIVE RESULTS ON NIQE ↓ ON (VV-24, LIME, DICM, NPE, MEF) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

Method	*VV-24*	*LIME*	*DICM*	*NPE*	*MEF*
Ours	2.50	3.92	**2.74**	**2.94**	**2.86**
+SFSA[68]	-	-	3.45	3.89	3.17
DRBN [17]	-	4.40	-	3.92	4.09
LIVENet[57]	3.44	4.21	3.60	3.89	3.89
PairLIE [20]	3.57	-	-	4.18	-
KinD [58]	4.30	4.64	-	4.69	4.13
RUAS [25]	4.29	-	-	-	-
LLFlow [19]	4.04	-	-	-	-
SKF [69]	-	3.98	3.53	3.82	3.63
HE [1]	3.20	4.12	3.64	4.28	3.47
KinD++ [56]	2.35	4.81	3.78	4.38	3.73
EnlightenGAN[12]	**2.25**	**3.67**	3.54	4.74	3.22
HWMNet [67]	-	4.35	3.92	4.06	4.21
Self-Supervised [70]	3.36	4.96	4.58	3.50	4.47
LCAE[71]	-	4.12	3.59	-	3.08

TABLE XIII. COMPARATIVE RESULTS ON NIQE ↓ ON (MIT5K-500) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE

EnGMHE	NPE [15]	SRIE [4]	LDR [37]	DRBN [17]	DSLR [72]	ZeroDCE [23]
2.18	3.92	3.93	3.67	4.38	4.20	3.84

TABLE XIV. COMPARATIVE RESULTS ON NIQE ↓ ON (LOL-15 AND LOL-V2) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE

Method	*LOL-15*	*LOL-V2*	*Method*	*LOL-15*	*LOL-V2*
EnGMHE	**2.82**	**3.11**	HE [1]	8.42	-
LLNET [16]	3.84	-	LIME [2]	5.32	5.37
BIMEF [61]	3.85	-	RetinexNet [3]	6.37	9.09
DRBN [17]	4.79	4.96	ExCNet [26]	3.05	-
+SFSA [68]	-	4.19	URetinexNet [51]	3.51	-
ZeroDCE [23]	7.76	8.05	KinD++ [56]	4.76	5.08
ZeroDCE++ [24]	7.86	8.05	EnlightenGAN [12]	3.26	-
PairLIE [20]	4.25	4.34	LCDPNet [31]	3.06	-
KinD [58]	5.18	6.84	HWMNet [67]	5.97	-
LLFlow [19]	5.41	-	SCI [22]	4.08	-
SKF [69]	4.14	3.96			

TABLE XV. COMPARATIVE RESULTS ON BRISQUE ↓ ON (VV-24) DATASET, THE BEST RESULT IS DENOTED BY BOLDFACE AND THE SECOND-RATED RESULT IS UNDERLINED

EnGMHE	KinD [58]	DCE [23]	RUAS [25]
24.93	50.56	34.66	38.37
SCI [22]	LLFLOW [19]	PairLIE [20]	SNR-Aware [59]
26.13	31.67	39.13	78.72

TABLE XVI. COMPARATIVE RESULTS ON BRISQUE ↓ ON (LIME, DICM, NPE, MEF) DATASETS, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

Method	*LIME*	*DICM*	*NPE*	*MEF*
KIND [58]	39.91	48.72	36.85	49.94
RUAS [25]	27.59	38.75	47.85	**23.68**
LLFLOW [19]	27.06	**26.36**	28.86	30.27
PairLIE [20]	25.23	33.31	28.27	27.53
SNR-Aware [59]	39.22	37.35	**26.65**	31.28
EnGMHE	**20.38**	26.78	26.85	24.13

TABLE XVII. COMPARATIVE RESULTS ON UCIQE ↑ ON (UIEB-CH60) DATASET, THE BEST RESULTS ARE DENOTED BY BOLDFACE AND THE SECOND-RATED RESULTS ARE UNDERLINED

EnGMHE	MTUR-Net [29]	WaterNet [32]	Ucolor [33]	IDMR [73]	Restormer [27]
0.64	0.5868	0.591	0.553	0.613	0.572
FUnIEGAN [74]	PUGAN [34]	U-Shape [75]	Semi-UIR [76]	UWE-Net[77]	
0.556	0.566	0.56	0.574	0.21	

F. Ablation Study

This section describes three tests, as they effectively address different aspects of the proposed model evaluation and improvement.

The first test evaluated the sensitivity of the model to the parameters γ and τ. As shown in Fig. 9, the model was applied to the dark image (Left). The test was first conducted by running EnGMHE multiple times on the input image while varying τ and keeping γ fixed at 25. Subsequently, the sensitivity to γ was examined by varying γ while keeping τ fixed at 2. As shown in Fig. 9, the sensitivity plot of τ indicates that τ significantly affects the model's performance and requires precise tuning to achieve optimal enhancement results. In comparison, the flat region in the γ sensitivity plot suggests that γ is less sensitive within its optimal range, making it slightly more robust than τ. To ensure better performance across datasets. Therefore, a sample image set from each dataset was initially selected to determine experimentally the most suitable parameter values for this model.

The second test (see Fig. 10) examined the incremental addition of the components of EnGMHE. This is to demonstrate how the following procedural steps enhance its performance.

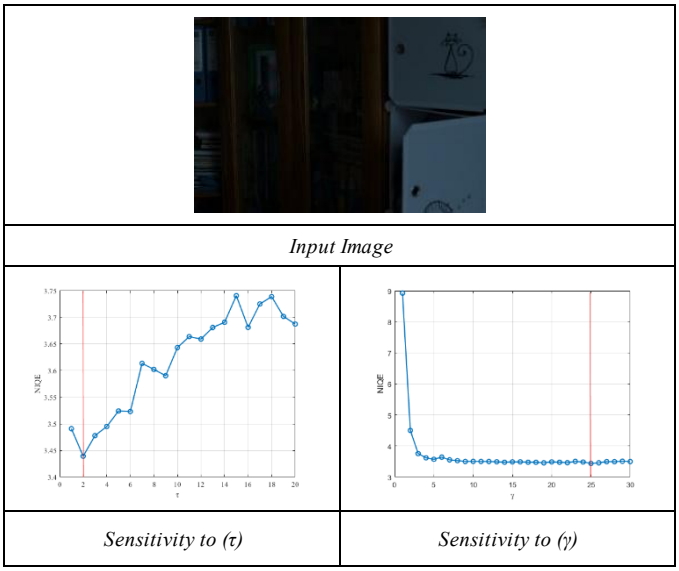

Fig. 9. Sensitivity to τ and γ parameters.

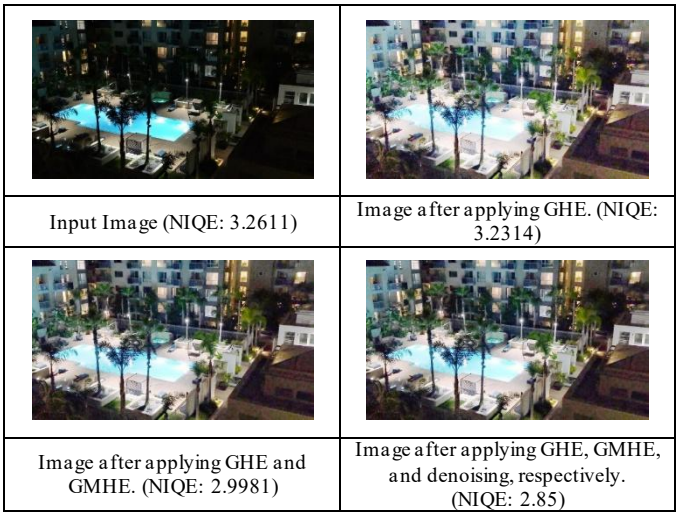

Fig. 10. The impact of the gradual addition of EnGMHE components. The input image extracted from the LOLi_Phone dataset [40].

The third test examined the sequence of enhancement steps and emphasized the importance of workflow design. It was applied across multiple datasets, and the proposed step sequence consistently produced the best results. For example, Fig. 11 illustrates how varying the order of steps in the enhancement workflow affected image quality, evaluated using NIQE and SSIM metrics. The input image exhibited low visibility and poor structural similarity, as reflected by its high NIQE and low SSIM scores. Among the tested configurations, the sequence GHE → GMHE → Denoise (EnGMHE) achieved the optimal balance between perceptual quality and structural similarity, recording the lowest NIQE (3.48) and the highest SSIM (0.84). This demonstrates that beginning with GHE, followed by GMHE, and concluding with denoising effectively preserved fine details and maintained natural image quality.

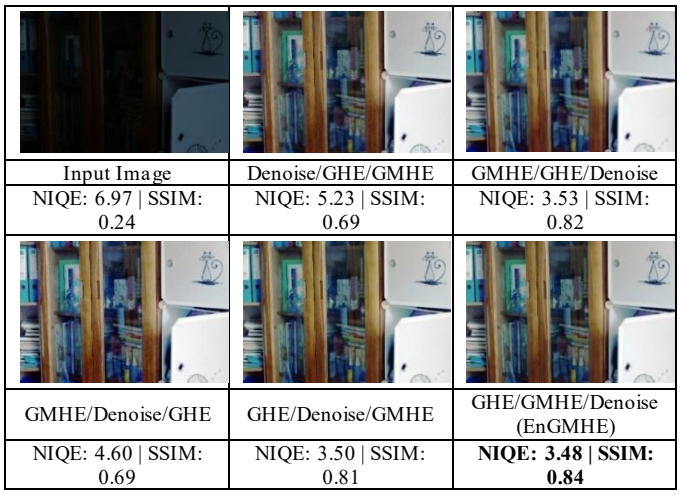

Fig. 11. Impact of procedural step order on image quality.

G. Limitations and Areas for Improvement

Fig. 12 highlights a potential limitation of EnGMHE in terms of color fidelity when compared to other enhancement methods. Although EnGMHE performs effectively in recovering fine details and improving text clarity, it may introduce slight oversaturation and color shifts when compared to the ground truth. Although this could be considered a weakness in terms of color fidelity, it may be advantageous in certain applications, such as text and object recognition, where clarity and detail preservation are often more important than strict closeness to a poor-quality ground truth.

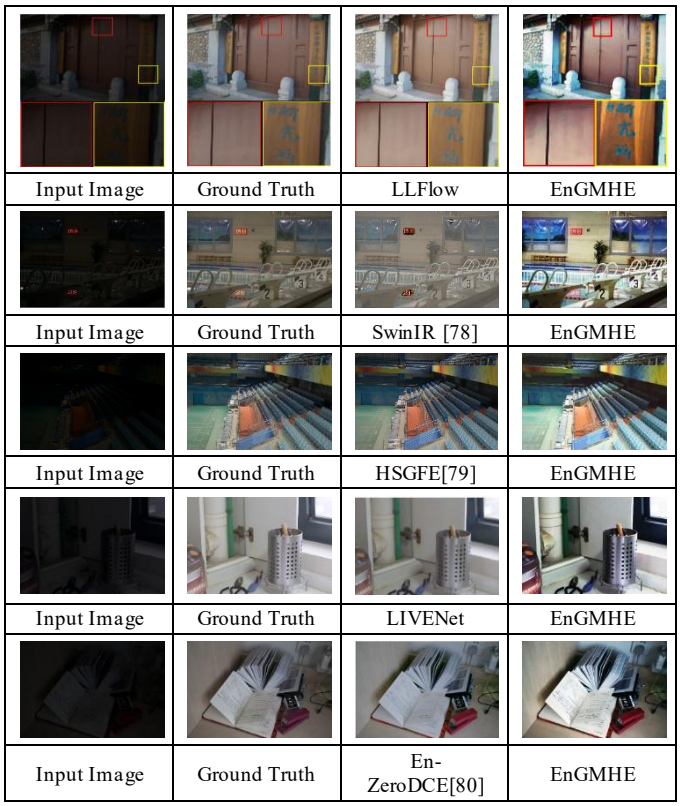

Fig. 12. Comparison of enhancement methods against the ground truth.

V. Conclusion

This study addresses the challenge of low-light image enhancement, considering the presence of noise in corrupted images. To tackle this problem, an effective joint low-light enhancement and denoising model, termed EnGMHE, is proposed. The model introduces a GHE-based technique combined with traditional GMHE to enhance image contrast and brightness. This enhancement stage is followed by a pretrained deep neural network for noise suppression, resulting in improved overall visual quality. EnGMHE effectively preserves fine details and produces visually appealing contrast and color distributions. Both qualitative and quantitative evaluations confirm its effectiveness, demonstrating superior performance compared to state-of-the-art methods.

References

[1] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: speed and effectiveness," in Proceedings of the First Conference on Visualization in Biomedical Computing, IEEE Comput. Soc. Press, 1990, pp. 337–345, doi: 10.1109/VBC.1990.109340.

[2] X. Guo, Y. Li, and H. Ling, "LIME: low-light image enhancement via illumination map estimation," IEEE Trans. Image Process., vol. 26, no. 2, pp. 982–993, February 2017, doi: 10.1109/TIP.2016.2639450.

[3] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in British Machine Vision Conference, British Machine Vision Association, 2018.

[4] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, June 2016, pp. 2782–2790, doi: 10.1109/CVPR.2016.304.

[5] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust retinex model," IEEE Trans. Image Process., vol. 27, no. 6, pp. 2828–2841, June 2018, doi: 10.1109/TIP.2018.2810539.

[6] X. Ren, M. Li, W.-H. Cheng, and J. Liu, "Joint enhancement and denoising method via sequential decomposition," in 2018 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, May 2018, pp. 1–5, doi: 10.1109/ISCAS.2018.8351427.

[7] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," IEEE Trans. Image Process., vol. 6, no. 7, pp. 965–976, July 1997, doi: 10.1109/83.597272..

[8] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," Signal Process., vol. 129, pp. 82–96, December 2016, doi: 10.1016/j.sigpro.2016.05.031.

[9] J. Tan, T. Zhang, L. Zhao, D. Huang, and Z. Zhang, "Low-light image enhancement with geometrical sparse representation," Appl. Intell., vol. 53, no. 9, pp. 11019–11033, May 2023, doi: 10.1007/s10489-022-04013-1.

[10] R. I. Zaghloul and H. Hiary, "A fast single image fog removal method using geometric mean histogram equalization," Int. J. Image Graph., vol. 21, no. 01, p. 2150001, January 2021, doi: 10.1142/S0219467821500017.

[11] J. Li, J. Li, F. Fang, F. Li, and G. Zhang, "Luminance-aware pyramid network for low-light image enhancement," IEEE Trans. Multimedia, vol. 23, pp. 3153–3165, 2021, doi: 10.1109/TMM.2020.3021243.

[12] Y. Jiang et al., "EnlightenGAN: deep light enhancement without paired supervision," IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349, 2021, doi: 10.1109/TIP.2021.3051462.

[13] H. Hiary, R. Zaghloul, A. Al-Adwan, and M. B. Al-Zoubi, "Image contrast enhancement using geometric mean filter," Signal Image Video Process., vol. 11, no. 5, pp. 833–840, July 2017, doi: 10.1007/s11760-016-1029-8.

[14] H. Ibrahim and N. Pik Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," IEEE Trans. Consumer Electron., vol. 53, no. 4, pp. 1752–1758, November 2007, doi: 10.1109/TCE.2007.4429280.

[15] S. Wang, J. Zheng, H.-M. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," IEEE Trans. Image Process., vol. 22, no. 9, pp. 3538–3548, September 2013, doi: 10.1109/TIP.2013.2261309.

[16] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: a deep autoencoder approach to natural low-light image enhancement," Pattern Recognit., vol. 61, pp. 650–662, 2017, doi: 10.1016/j.patcog.2016.06.008.

[17] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. 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), IEEE, June 2020, pp. 3060–3069, doi: 10.1109/CVPR42600.2020.00313.

[18] F. Lv, F. Lu, J. Wu, and C. Lim, "Mbllen: low-light image/video enhancement using cnns," in British Machine Vision Conference, British Machine Vision Association, 2018, pp. 1–13.

[19] Y. Wang, R. Wan, W. Yang, H. Li, L.-P. Chau, and A. Kot, "Low-light image enhancement with normalizing flow," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, no. 3, pp. 2604–2612, June 2022, doi: 10.1609/aaai.v36i3.20162.

[20] Z. Fu, Y. Yang, X. Tu, Y. Huang, X. Ding, and K.-K. Ma, "Learning a simple low-light image enhancer from paired low-light instances," in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, June 2023, pp. 22252–22261, doi: 10.1109/CVPR52729.2023.02131.

[21] S. W. Zamir, S. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Learning enriched features for fast image restoration and enhancement," IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 2, pp. 1934–1948, 2023, doi: 10.1109/TPAMI.2022.3167175.

[22] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2022, pp. 5627–5636, doi: 10.1109/CVPR52688.2022.00555.

[23] C. Guo et al., "Zero-reference deep curve estimation for low-light image enhancement," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Jun. 2020, pp. 1777–1786.

[24] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 8, pp. 4225–4238, 2022.

[25] R. Liu, L. Ma, J. Zhang, X. Fan, and Z. Luo, "Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 10556–10565.

[26] L. Zhang, L. Zhang, X. Liu, Y. Shen, S. Zhang, and S. Zhao, "Zero-shot restoration of back-lit images using deep internal learning," in Proceedings of the 27th ACM International Conference on Multimedia, New York, NY, USA: ACM, October. 2019, pp. 1623–1631.

[27] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: efficient transformer for high-resolution image restoration," in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, June. 2022, pp. 5718–5729.

[28] E. Adhikarla et al., "Unified-EGformer: exposure guided lightweight transformer for mixed-exposure image enhancement," in Pattern Recognition, A. Antonacopoulos, S. Chaudhuri, R. Chellappa, CL. Liu, S. Bhattacharya, and U. Pal, Eds., ICPR 2024, Lecture Notes in Computer Science, vol. 15329. Springer, Cham, 2025, pp. 260–275. doi: 10.1007/978-3-031-78110-0_17.

[29] K. Yan, L. Liang, Z. Zheng, G. Wang, and Y. Yang, "Medium transmission map matters for learning to restore real-world underwater images," Applied Sciences, vol. 12, no. 11, p. 5420, May 2022.

[30] M. Fan, W. Wang, W. Yang, and J. Liu, "Integrating semantic segmentation and retinex model for low-light image enhancement," in Proceedings of the 28th ACM International Conference on Multimedia, New York, NY, USA: ACM, Oct. 2020, pp. 2317–2325.

[31] H. Wang, K. Xu, and R. W. H. Lau, "Local color distributions prior for image enhancement," in A. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds., Computer Vision – ECCV 2022, Lecture

Notes in Computer Science, vol. 13678. Cham: Springer, 2022, pp. 343–359. doi: 10.1007/978-3-031-19797-0_20.

[32] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," IEEE Trans. Image Process., vol. 29, pp. 4376–4389, 2019, doi: 10.1109/TIP.2019.2955241.

[33] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," IEEE Trans. Image Process., vol. 30, pp. 4985–5000, 2021, doi: 10.1109/TIP.2021.3076367.

[34] R. Cong et al., "PUGAN: physical model-guided underwater image enhancement using GAN with dual-discriminators," IEEE Trans. Image Process., vol. 32, pp. 4472–4485, 2023, doi: 10.1109/TIP.2023.3286263.

[35] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: residual learning of deep cnn for image denoising," IEEE Trans. Image Process., vol. 26, no. 7, pp. 3142–3155, 2017, doi: 10.1109/TIP.2017.2662206.

[36] Y. P. Loh and C. S. Chan, "Getting to know low-light images with the exclusively dark dataset," Comput. Vis. Image Underst., vol. 178, pp. 30–42, 2019, doi: 10.1016/j.cviu.2018.10.010.

[37] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2d histograms," IEEE Trans. Image Process., vol. 22, no. 12, pp. 5372–5384, 2013, doi: 10.1109/TIP.2013.2284059.

[38] V. Vonikakis, R. Kouskouridas, and A. Gasteratos, "On the evaluation of illumination compensation algorithms," Multimed. Tools Appl., vol. 77, no. 8, pp. 9211–9231, 2018, doi: 10.1007/s11042-017-4783-x.

[39] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," IEEE Trans. Image Process., vol. 24, no. 11, pp. 3345–3356, 2015, doi: 10.1109/TIP.2015.2442920.

[40] C. Li et al., "Low-light image and video enhancement using deep learning: a survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 12, pp. 9396–9416, 2022, doi: 10.1109/TPAMI.2021.3126387.

[41] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image pairs," in CVPR, IEEE, 2011, pp. 97–104, doi: 10.1109/CVPR.2011.5995413.

[42] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," IEEE Trans. Image Process., vol. 27, no. 4, pp. 2049–2062, 2018, doi: 10.1109/TIP.2018.2794218.

[43] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.

[44] E. Cuevas and A. N. Rodríguez, Image processing and machine learning, Volume 2: Advanced topics in image analysis and machine learning, Chapman & Hall, 2024.

[45] S. Ghazanfari, S. Garg, P. Krishnamurthy, F. Khorrami, and A. Araujo, "R-LPIPS: an adversarially robust perceptual similarity metric," in ICML 2023, Workshop on New Frontiers in Adversarial Machine Learning, unpublished.

[46] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018, pp. 586–595. doi: 10.1109/CVPR.2018.00068.

[47] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations," Color Res. Appl., vol. 30, no. 1, pp. 21–30, 2005. doi: 10.1002/col.20070.

[48] A. Mittal, A. K. Moorthy, and A. C. Bovik, "Blind/referenceless image spatial quality evaluator," in 2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR), IEEE, 2011, pp. 723–727. doi: 10.1109/ACSSC.2011.6190099.

[49] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," IEEE Signal Process. Lett., vol. 20, no. 3, pp. 209–212, 2013. doi: 10.1109/LSP.2012.2227726.

[50] M. Yang and A. Sowmya, "An underwater color image quality evaluation metric," IEEE Trans. Image Process., vol. 24, no. 12, pp. 6062–6071, 2015. doi: 10.1109/TIP.2015.2491020.

[51] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang, and J. Jiang, "URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement," in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, pp. 5891–5900. doi: 10.1109/CVPR52688.2022.00581.

[52] Y. Weng, T. Zhou, Y. Li, and X. Qiu, "NAS-Unet: neural architecture search for medical image segmentation," *IEEE Access*, vol. 7, pp. 44247–44257, 2019, doi: 10.1109/ACCESS.2019.2908991.

[53] J. Jiang, T. Ye, S. Chen, E. Chen, Y. Liu, S. Jun, J. Bai, and W. Chai, "Five A+ network: you only need 9k parameters for underwater image enhancement," in 34th British Machine Vision Conference 2023, Aberdeen, UK: BMVA, pp. 1–12.

[54] Z. Fu, W. Wang, Y. Huang, X. Ding, and K.-K. Ma, "Uncertainty inspired underwater image enhancement," in Computer Vision – ECCV 2022, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds., Lecture Notes in Computer Science, vol. 13678, Springer, Cham, pp. 465–482. doi: 10.1007/978-3-031-19797-0_27.

[55] H. Lu, J. Gong, Z. Liu, R. Lan, and X. Pan, "FDMLNet: a frequency-division and multiscale learning network for enhancing low-light image," Sensors, vol. 22, no. 21, p. 8244. doi: 10.3390/s22218244.

[56] Y. Zhang, X. Guo, J. Ma, W. Liu, and J. Zhang, "Beyond brightening low-light images," Int. J. Comput. Vis., vol. 129, no. 4, pp. 1013–1037. doi: 10.1007/s11263-020-01407-x.

[57] D. Makwana, G. Deshmukh, O. Susladkar, S. Mittal, and S. C. T. R, "LIVENet: a novel network for real-world low-light image denoising and enhancement," in 2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), IEEE, pp. 5844–5853. doi: 10.1109/WACV57701.2024.00575.

[58] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness," in Proceedings of the 27th ACM International Conference on Multimedia, New York, NY, USA: ACM, pp. 1632–1640. doi: 10.1145/3343031.3350926.

[59] X. Xu, R. Wang, C.-W. Fu, and J. Jia, "SNR-aware low-light image enhancement," in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 17693–17703. doi: 10.1109/CVPR52688.2022.01719.

[60] Y. P. Loh, X. Liang, and C. S. Chan, "Low-light image enhancement using Gaussian process for features retrieval," Signal Process. Image Commun., vol. 74, pp. 175–190, 2019. doi: 10.1016/j.image.2019.02.001.

[61] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in 34th Conference on Neural Information Processing Systems (NeurIPS 2020), Vancouver, Canada.

[62] S. Zheng and G. Gupta, "Semantic-guided zero-shot learning for low-light image/video enhancement," in 2022 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW), Waikoloa, HI, USA, pp. 581–590, 2022, doi: 10.1109/WACVW54805.2022.00064.

[63] S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," IEEE Trans. Multimedia, vol. 22, no. 12, pp. 3025–3038, 2020, doi: 10.1109/TMM.2020.2969790.

[64] J. Xu et al., "STAR: a structure and texture aware retinex model," IEEE Trans. Image Process., vol. 29, pp. 5022–5037, 2020, doi: 10.1109/TIP.2020.2974060.

[65] A. Zhu, L. Zhang, Y. Shen, Y. Ma, S. Zhao, and Y. Zhou, "Zero-shot restoration of underexposed images via robust retinex decomposition," in 2020 IEEE International Conference on Multimedia and Expo (ICME), IEEE, pp. 1–6, 2020, doi: 10.1109/ICME46284.2020.9102962.

[66] W. Wang, C. Wei, W. Yang, and J. Liu, "GLADNet: low-light enhancement network with global awareness," in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), IEEE, pp. 751–755, 2018, doi: 10.1109/FG.2018.00118.

[67] C.-M. Fan, T.-J. Liu, and K.-H. Liu, "Half wavelet attention on M-Net+ for low-light image enhancement," in 2022 IEEE International Conference on Image Processing (ICIP), IEEE, Oct. 2022, pp. 3878–3882, doi: 10.1109/ICIP46576.2022.9897503.

[68] X. Wang and Q. Zheng, "Revitalize supervised low-light image enhancer: learning source-free fast scene adaptation," in Advanced Intelligent Computing Technology and Applications, D. S. Huang, H. Chen, B. Li, and Q. Zhang, Eds., Lecture Notes in Computer Science, vol. 15843, Springer, Singapore, 2025, pp. 86–98, doi: 10.1007/978-981-96-9866-0_8.

[69] Y. Wu et al., "Learning semantic-aware knowledge guidance for low-light image enhancement," in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 2023, pp. 1662–1671, doi: 10.1109/CVPR52729.2023.00166.

[70] J. Liang, Y. Xu, Y. Quan, B. Shi, and H. Ji, "Self-supervised low-light image enhancement using discrepant untrained network priors," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 11, pp. 7332–7345, Nov. 2022, doi: 10.1109/TCSVT.2022.3181781.

[71] Q. Mu, Y. Guo, X. Ge, X. Wang, and Z. Li, "Local content-aware enhancement for low-light images with non-uniform illumination," Computers, Materials & Continua, vol. 82, no. 3, pp. 4669–4690, 2025, doi: 10.32604/cmc.2025.058495.

[72] S. Lim and W. Kim, "DSLR: deep stacked Laplacian restorer for low-light image enhancement," IEEE Trans. Multimedia, vol. 23, pp. 4272–4284, 2021, doi: 10.1109/TMM.2020.3039361.

[73] A. Sridhar, A. B. N., T. Radhakrishnan, and S. Tripathi, "Underwater image enhancement using image dehazer and multiscale retinex," in 2024 6th International Conference on Image, Video and Signal Processing, New York, NY, USA: ACM, Mar. 2024, pp. 86–94, doi: 10.1145/3655755.3655767.

[74] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," IEEE Robot. Autom. Lett., vol. 5, no. 2, pp. 3227–3234, Apr. 2020, doi: 10.1109/LRA.2020.2974710.

[75] L. Peng, C. Zhu, and L. Bian, "U-shape transformer for underwater image enhancement," IEEE Trans. Image Process., vol. 32, pp. 3066–3079, 2023, doi: 10.1109/TIP.2023.3276332.

[76] S. Huang, K. Wang, H. Liu, J. Chen, and Y. Li, "Contrastive semi-supervised learning for underwater image restoration via reliable bank," in 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, June. 2023, pp. 18145–18155, doi: 10.1109/CVPR52729.2023.01740.

[77] S. Jayasurya, S. Geetha, A. S. Abdullah, and U. Mishra, "UWE-Net: deep learning framework for underwater image enhancement integrating CBAM and Charbonnier loss," Procedia Comput. Sci., vol. 258, pp. 689–698, 2025, doi: 10.1016/j.procs.2025.04.302.

[78] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: image restoration using Swin transformer," in 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021, pp. 1833–1844, doi: 10.1109/ICCVW54120.2021.00210.

[79] W. Dong, Y. Min, H. Zhou, and J. Chen, "Towards scale-aware low-light enhancement via structure-guided transformer design," in 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE, 2025, pp. 1454–1461, doi: 10.1109/CVPRW67362.2025.00135.

[80] J. Premasagar and S. Pelluri, "Low-light image enhancement for video object detection using modified Zero DCE deep learning model," SSRG Int. J. Electron. Commun. Eng., vol. 11, no. 9, pp. 290–304, 2024, doi: 10.14445/23488549/IJECE-V11I9P125.