ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 3 - 2026

HVI-CoLIE: Color-Aware Context-Based Low-Light Image Enhancement via Neural Implicit Representations and Illumination-Guided Restoration

Chaoqun Wang[1,*] *and Jiazheng Hou*[2]

[1] School of Science, Hebei University of Technology, Tianjin, 300401 China
[2] School of Mechanical Engineering, Hebei University of Technology, Tianjin, 300401, China

SUMMARY: *Current low-light image enhancement methods still suffer from poor visual perception, color bias, and over-exposure caused by aggressive brightening when dealing with unknown illumination conditions. In this paper, we propose a zero-shot method termed HVI-CoLIE, which redefines the enhancement process through mapping the 2D coordinates of an underexposed image to its HVI intensity component, while using the restored intensity to guide chromatic restoration and adaptive gamma correction to control over-exposure. We transform the image to HVI space and propose an implicit neural function combined with an illumination-guided Chromatic Restoration Module. Moveover, We introduce an adaptive decision criterion to effectively alleviate the over-exposure problem commonly observed in current zero-shot low-light enhancement methods. Through comprehensive evaluations, we investigate the characteristics of our proposed framework and demonstrate its advantages in image quality and scene adaptability. Moreover, we further evaluate its performance on downstream tasks under low-light scenarios, highlighting the practical utility of HVI-CoLIE.*

KEYWORDS: *Low-light image enhancement; HVI color space; neural implicit representation; illumination-guided Transformer; exposure-adaptive restoration.*

1 Introduction

In the real world under uncontrolled lighting conditions, the taken images usually show an uneven distribution of exposure, including large dark areas, unstable color tones and sometimes over-exposed local parts. Such degradation weakens perceptual clarity, obscures structural cues needed for human observation, and reduces the reliability of downstream vision algorithms under poor-visibility conditions [1–3].

Motivated by these problems, low-light image enhancement has received more and more attention in computer vision. Contrast-limited adaptive histogram equalization improves local visibility by redistributing contrast under certain constraints [4], whereas Retinex-based enhancement and illumination-map estimation offer reasonable exposure correction under uneven lighting conditions [5-7]. Some early deep autoencoder models have indicated that data-driven mappings can also restore natural low-light images without depending entirely on hand-made priors [8]. Most deep learning methods try to obtain high-quality images by using supervised, semi-supervised, unsupervised and zero-shot restoration approaches. Supervised and semi-supervised models acquire low-to-normal-light mappings or illumination-reflectance decomposition from paired or partially paired data [9-13]. Zero-reference, unpaired,

*chaoqun825@163.com
https://doi.org/10.65102/is20261288
Published: 10 June, 2026

unfolding-based, self-calibration and paired-instance learning strategies decrease the need for paired datasets, but they still require learned priors or training based on a specific dataset [14-19]. CoLIE does not need to train with prior datasets, but it optimizes a coordinate-dependent neural implicit representation on a single input image [20]. In addition, the hidden variable image (HVI) color space model suggests that a more stable chromatic-intensity representation can decrease the sensitivity to color under weak illumination [21]. Fig. 1 shows a visual comparison of the degradation and enhancement behaviors of low-light images.

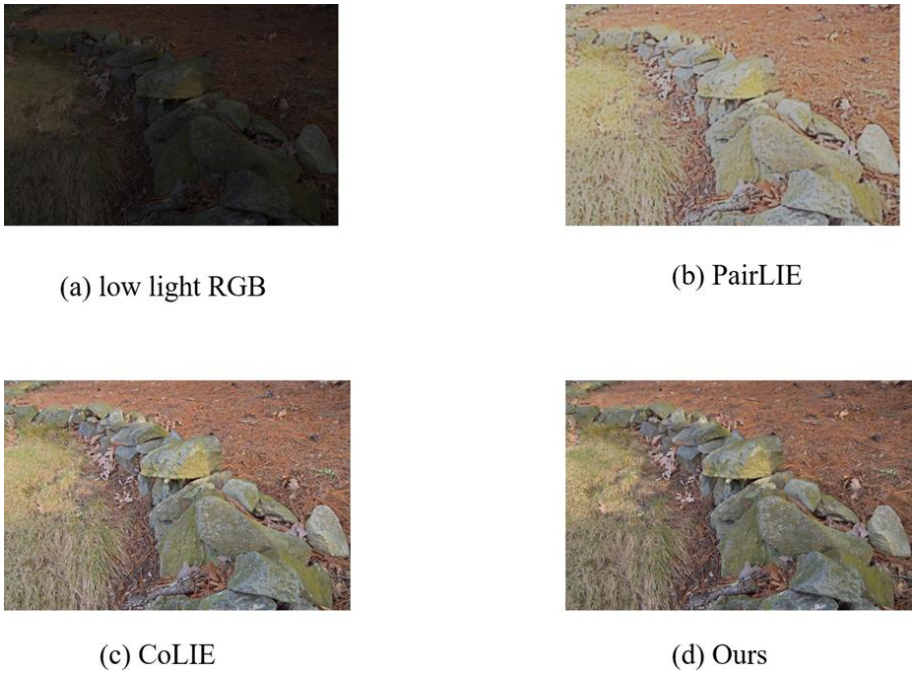

Figure 1: Visual comparison on a low-light image from the MIT dataset. Low-light conditions obscure structural details and degrade color perception, which limits both human observation and downstream vision tasks. Compared with existing methods, our approach enhances under-exposed regions while better preserving natural color and saturation, highlighting the importance of low-light image enhancement.

However, these zero-shot methods still provide insufficient restoration of chromatic components and struggle with exposure control in challenging low-light scenes. Since it reconstructs images in the HSV space by enhancing only the Value component, the degraded Hue and Saturation information is directly inherited, which may introduce color-space noise and chromatic distortion in dark regions. In addition, its aggressive brightening strategy can easily cause local over-exposure. These issues moti-vate a more color-aware and exposure-adaptive enhancement framework.

In this paper, we propose a new method for enhancing low-light images robustly by using an adaptive decision criterion and an Illumination-guided Chromatic Restoration Module. Recent studies have shown that the HVI color space is more robust and stable against interference, We build our zero-shot network in the HVI space of the input image. Con-sidering the over-exposure tendency of zero-shot methods, we introduce a gated gamma module for adaptive exposure-aware restoration. Beyond restoring the I component, the recovered I is further used to guide chromatic restoration. Our model enables the zero-shot framework to reduce over-exposure, color bias, and chromatic distortion. The contributions of this paper are as follows:

- We reformulate the network in the HVI space by re-placing HSV-based Value

restoration with HVI-based Intensity restoration, thereby reducing the noise effects introduced by the original color space.

- We propose a gated gamma correction strategy with an adaptive decision criterion to alleviate the over-exposure problem caused by aggressive brightening in zero-shot low-light enhancement.
- We introduce an illumination-guided HV chromatic restoration module, which uses the restored Intensity component to guide the recovery of degraded chromatic information and improve color fidelity in dark regions.
- We conduct extensive evaluations to analyze the prop-erties of our proposed framework, demonstrating its ad-vantages in image quality, color consistency, and scene adaptability under unknown low-light conditions.

2 Related Work

2.1 Low-Light Image Enhancement

Early low-light image enhancement methods mainly rely on Retinex-based enhancement, where illumination maps provide direct evidence for spatial brightness distribution.

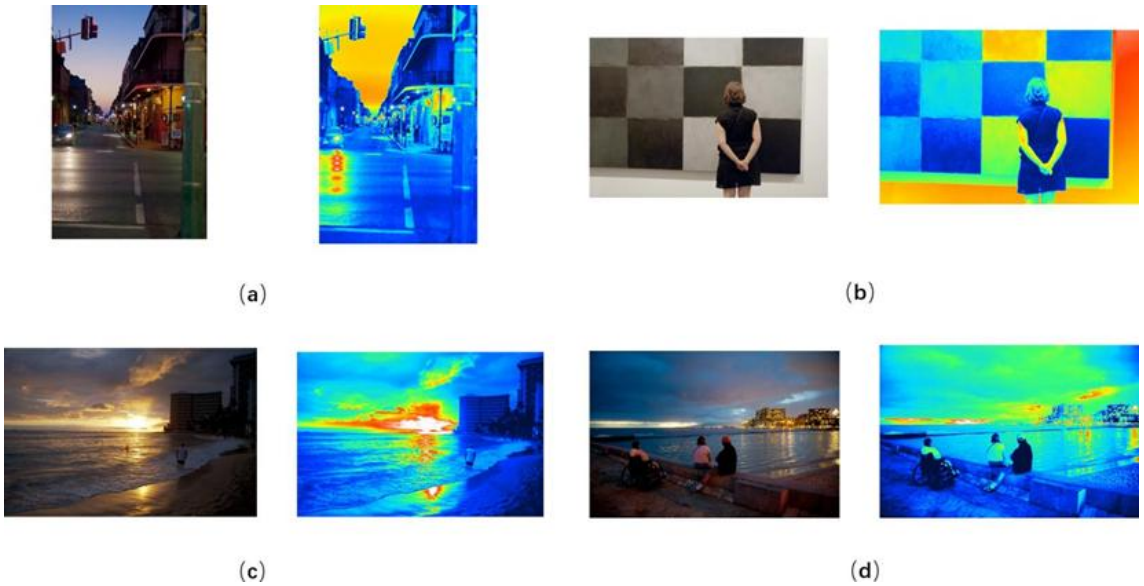

Figure 2: Representative illumination maps are visualized as heatmaps using images from the MIT dataset. These maps describe the spatial distribution of scene brightness and provide explicit guidance for the subsequent low-light enhancement process. Following the Retinex assumption that illumination and reflectance jointly determine image appearance, the recovered illumination information helps the model enhance under-exposed regions and guide the HV component to restore.

As shown in Fig. 2, such maps help identify under-exposed regions that require stronger restoration and brighter observed low-light image is modeled as the product of illumination and reflectance components. Although these Retinex-based methods provide a physically interpretable formulation for LLIE, their performance is often sensitive to manually designed priors and multiple hyper-parameters. As a result, they may introduce color distortion or fail to achieve a proper balance between severely under-exposed and over-exposed regions, which restricts their robustness in real-world low-light scenarios.

With the advance of deep learning, LLIE has achieved substantial progress owing to the stronger representation ca-pability of modern architectures. In particular, Retinex-based methods have been successively combined with CNNs. Existing approaches can be roughly divided into two categories. Some methods embed the Retinex theory into deep networks to explicitly model illumination and reflectance components, while others directly learn the mapping from low-light inputs to enhanced images in a supervised manner. Nevertheless, supervised learning usually requires paired low-light and normal-light images, which are expensive and difficult to collect in real-world scenarios. To reduce this dependence, recent studies attempt to train enhancement models using only low-light images. For example, ZeroDCE estimates image-specific enhancement curves with a deep network. Liu et al. proposed a Retinex-based unrolling framework with architecture search, and Ma et al. designed a self-calibrating module for image brightening through cascaded illumination learning. Fu et al. further introduced adaptive priors learned from low-light image pairs to guide Retinex decomposition. Al-though these methods are often categorized as zero-reference or self-supervised approaches, they still require prior training on low-light data. Therefore, their performance may become unstable when the under-exposure level of testing images differs from that of the training data.

2.2 Neural Implicit Representations

Neural implicit representations (NIRs) have recently shown strong potential in image processing by representing visual signals as continuous coordinate-based functions. Local implicit image functions support continuous image representation for resolution-adaptive reconstruction [22], and local texture estimators improve the ability of implicit functions to represent fine-grained spatial details [23]. while the enhancement process

However, the enhancement process can still struggle with adequate exposure correction in extremely dark regions [23]. In contrast, our method formulates the zero-shot enhancement framework in the HVI color space, where an extremely-dark gated gamma correction is employed for adaptive exposure correction, while the recovered Intensity further guides the restoration of HV chromatic components. This design extends NIR-based LLIE from illumination-only enhancement to color-aware and exposure-adaptive restoration.

3 METHOD

Fig. 3 shows the overall structure of our approach. The framework is based on CoLIE and transforms the low-light image enhancement into the HVI space. It includes a gamma-gated adaptive implicit intensity estimator and an illumination-directed HV restorer. The intensity estimator directly associates the 2D coordinates and local context with the illumination part of the I channel in a zero-shot way, whereas the HV restorer utilizes the recovered intensity to influence the chromatic restoration.

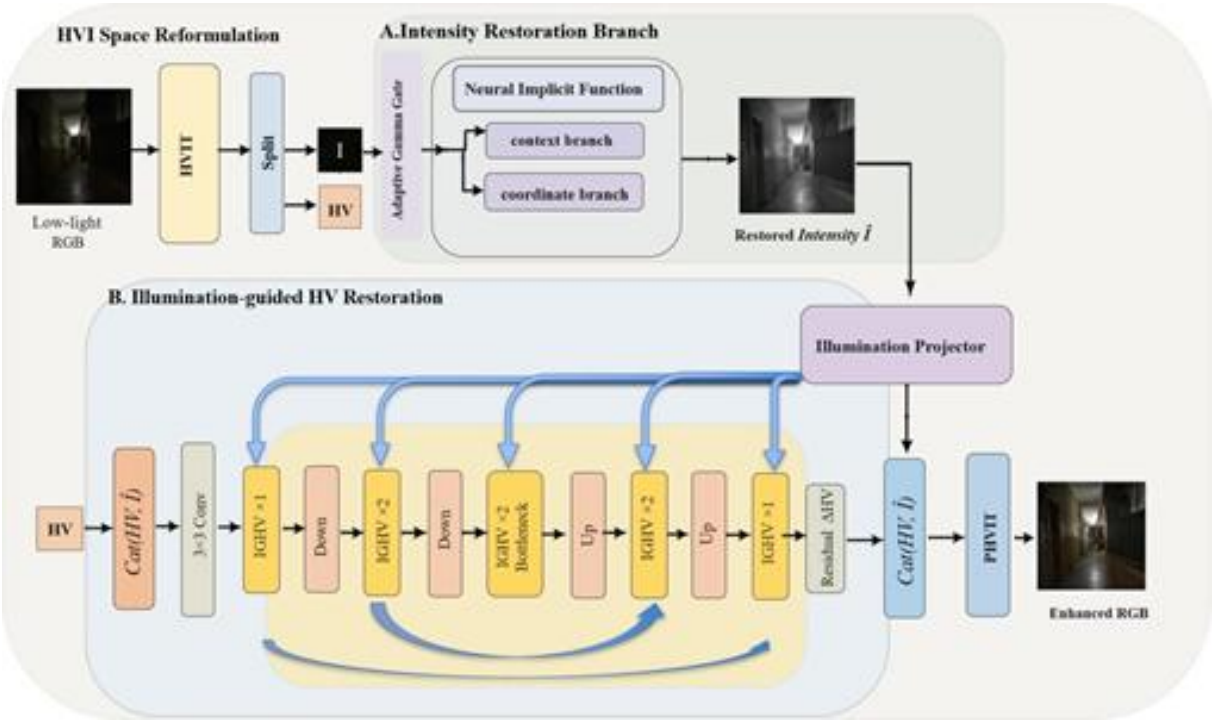

Figure 3: The overview of our method.

3.1 HVI-based Implicit Enhancement Framework

Following CoLIE, low-light image enhancement can be viewed as illumination learning in a zero-shot coordinate-conditioned manner. Different from the original CoLIE formulation that operates in the HSV space and reconstructs only the Value component, our model reformulates the enhancement process in the HVI space and explicitly considers both intensity recovery and chromatic restoration.

Given a low-light RGB image Y ∈ R^{H×W×3}, we first transform it into the HVI space as

$$(\mathbf{Y}_{HV}, \mathbf{Y}_I) = \mathbf{T}_{HVI}(\mathbf{Y}) \tag{1}$$

where YHV ∈ R^{H×W×2} denotes the HV color map and YI ∈ R^{H×W} denotes the intensity map.

Compared with the original CoLIE assumption that the chromatic components can be directly reused, we observe that the HV plane in low-light images still contains noticeable chromatic corruption, especially in extremely dark regions and after the light-up process. Therefore, we model the low-light HVI observation as

$$\mathbf{Y}_I = \mathbf{X}_I \odot \mathbf{Z}_I + \mathbf{C}_I, \ \mathbf{Y}_{HV} = \mathbf{Z}_{HV} + \mathbf{C}_{HV} \tag{2}$$

where ZI and ZHV denote the desired normal-light intensity and HV components, XI is the illumination component, and CI and CHV represent the residual degradations in the intensity and chromatic domains, respectively.

To better address extremely under-exposed regions, we introduce an adaptive gamma-gating strategy before implicit illumination learning. Specifically, a gating map is predicted from the input intensity map to determine whether local regions require additional gamma-based compensation:

$$\mathbf{M}_\gamma = \sigma(\mathrm{G}(\mathbf{Y}_I)) \tag{3}$$

$$\mathbf{Y}_I^g = (1 - \mathbf{M}_\gamma) \odot \mathbf{Y}_I + \mathbf{M}_\gamma \odot \Gamma_\gamma(\mathbf{Y}_I) \tag{4}$$

where G(·) denotes the gating predictor, σ(·) is the sigmoid function, and Γγ(·) denotes the gamma correction operator.

We then follow the neural implicit representation paradigm of CoLIE to estimate the illumination component from the gamma-gated intensity map in a coordinate-conditioned manner. Let pi = (ai, bi) be the 2D coordinate of pixel i, and N(pi) be the local context centered at pi. The implicit mapping is defined as

$$f\theta\colon (\mathbf{p}_i, \mathrm{N}(\mathbf{p}_i)) \rightarrow \mathbf{r}_i \tag{5}$$

where ri is the corresponding residual illumination value. Based on this mapping, the estimated illumination component and the restored intensity map are obtained as

$$\widehat{\mathbf{X}}_I = \mathbf{Y}_I^g + f\theta(\mathbf{p}, \mathrm{N}(\mathbf{p})) \tag{6}$$

$$\hat{\mathbf{Z}}_I = \mathbf{Y}_I^g \oslash \widehat{\mathbf{X}}_I \tag{7}$$

where X̂I denotes the estimated illumination component and ẐI denotes the restored intensity component.

After the light-up process, we obtain a coarse HVI representation

$$Y_{lu}^{HVI} = \mathrm{Cat}\big(Y_{HV}, \hat{Z}_I\big) = \mathrm{Cat}\big(Z_{HV} + C_{HV}, Z_I + \tilde{C}_I\big) \tag{8}$$

where C̃I denotes the residual error remaining in the restored intensity branch. Similar to the corruption-restoration formulation in Retinexformer, our second stage aims to suppress the residual chromatic corruption in the HV plane under the guidance of the restored intensity representation. Accordingly, we formulate our one-stage framework as

$$(\hat{\mathbf{Z}}_I, \mathbf{F}_I) = \mathrm{E}(\mathbf{Y}_I^g, \mathbf{p}, \mathrm{N}(\mathbf{p})) \tag{9}$$

$$\hat{\mathbf{Z}}_{HV} = \mathrm{R}(\mathrm{Cat}(\mathbf{Y}_{HV}, \hat{\mathbf{Z}}_I), \mathbf{F}_I) \tag{10}$$

$$\widehat{\mathbf{Y}} = \Phi_{HVI}^{-1}(\mathrm{Cat}(\hat{\mathbf{Z}}_{HV}, \hat{\mathbf{Z}}_I)) \tag{11}$$

where E denotes the adaptive gamma-gated implicit intensity estimator, FI is the restored intensity feature used as illumination guidance, R denotes the HV corruption restorer implemented by the illumination-guided transformer, and Φ−1 denotes the perceptual-inverse HVI Transformation. The final enhanced RGB image Ŷ is thus obtained by first restoring the intensity component with coordinate-conditioned implicit illumination learning and then recovering the HV chromatic components under the guidance of the restored intensity.

Different from previous CoLIE-style formulations that only reconstruct the brightness component and directly reuse the original chromatic channels, our framework explicitly models the chromatic corruption in the HVI space. Moreover, rather than performing enhancement directly in the RGB space, our method follows a one-stage coarse-to-refined strategy: it first restores the intensity component through implicit illumination learning, and then leverages the restored intensity to guide HV restoration, leading to a more stable and photometrically consistent enhancement process.

3.2 Illumination-Guided HV Restorer

Although the restored intensity component improves the global visibility of a low-light image, the chromatic information in extremely dark regions may still be degraded. Directly reusing the original chromatic components, as in the baseline pipeline, can therefore leave color inconsistency or weak local details in the final result. To address this issue, we introduce an illumination-guided HV restorer in the HVI space. The restored intensity map is used as a guidance cue to refine the HV components before converting the image back to RGB.

Given an input image transformed into the HVI space, we denote its chromatic components and intensity component as YHV and YI. After the CoLIE-based implicit branch restores the intensity component $\hat{\mathrm{I}}$, we first construct a preliminary HVI representation:

$$\widetilde{\mathbf{Y}}_{HVI} = \mathrm{Cat}(\mathbf{Y}_{HV}, \hat{\mathbf{I}}) \tag{12}$$

This representation provides the image content to be refined, while $\hat{\mathrm{I}}$ provides illumination guidance for chromatic restoration.

The proposed HV restorer follows an encoder-decoder structure. The preliminary HVI image is first embedded by a 3 × 3 convolution. In parallel, the restored intensity map is projected into an illumination feature through a lightweight projector consisting of a 1 × 1 convolution and a 5 × 5 depth-wise convolution:

$$\mathbf{F}_0 = \mathrm{E}(\widetilde{\mathbf{Y}}_{HVI}),\ \mathbf{G}_0 = \mathrm{P}(\hat{\mathbf{I}}) \tag{13}$$

where F0 and G0 denote the image feature and illumination feature, respectively.

The restorer contains two downsampling stages, a bottleneck, and two upsampling stages. At each scale, the image feature is refined by an illumination-guided attention block, while the illumination feature is downsampled accordingly to keep the two feature streams spatially aligned. Skip connections are used between the encoder and decoder to preserve local structures. In each attention block, the illumination feature modulates the value branch of self-attention, allowing the restored intensity map to guide the aggregation of chromatic features. This design encourages regions with different illumination levels to be processed adaptively, which is especially useful for dark areas where color information is weak or unstable.

After multi-scale refinement, a 3 × 3 mapping layer predicts an HV residual, which is added to the original HV components:

$$\widehat{\mathbf{Y}}_{HV} = \mathbf{Y}_{HV} + \mathrm{M}(\mathbf{F}_{out}) \tag{14}$$

Finally, the restored HV components and the restored intensity component are concatenated to form the enhanced HVI representation:

$$\widehat{\mathbf{Y}}_{HVI} = \mathrm{Cat}(\widehat{\mathbf{Y}}_{HV}, \hat{\mathbf{I}}) \tag{15}$$

which is transformed back to RGB using the inverse HVI transform.

In implementation, the HV restorer uses a feature width of 40, two encoder-decoder levels, and the block configuration [1, 2, 2]. To handle images with arbitrary spatial sizes, the input is padded to a multiple of 22 before restoration and cropped back to its original size after prediction.

4 Experiments

4.1 Implementation Details.

Implementation Details During training, we use the Adam optimizer with parameters $\beta 1 = 0.9$ and $\beta 2 = 0.999$. The learning rate was set to $10{-}5$ to initialize the optimization for 100 epochs. The regularization parameters were set as $\alpha = 1$, $\beta = 20$, $\gamma = 8$, and $\delta = 5$, while the exposure-control target was fixed to $L = 0.5$. Unlike the original CoLIE, the input RGB image was first transformed into the HVI color space, and the I component was used in place of the V component for zero-shot implicit restoration. In the current implementation, the size of the context window was set to 1×1.

The neural implicit function was implemented using a SIREN-based MLP, where both the context branch and the coordinate branch comprised two layers, and the fused output branch further contained two layers for the final prediction. Prior to optimization, the input I component was downscaled to the resolution of 256×256 pixels to improve computational efficiency. After restoration, the outputs were upscaled to the original image size using a guided filter with radius $r = 1$ and $\varepsilon = 10{-}8$.

To utilize the local context for pixels near the image boundary, reflection padding was employed during patch extraction. For chromatic restoration, we adopted an illumination-guided transformer adapted from the Corruption Restorer of Retinexformer, where the input and output channel numbers were set to 3 and 2, respectively, the feature width was set to 40, the network depth level was set to 2, and the numbers of attention blocks at different stages were set to [1, 2, 2]. In the HVI transform, the density parameter k was initialized to 0.2.

4.2 Compared Methods.

We evaluate eleven advanced techniques with various pre-trained parameters in this benchmarking research. The tested methods comprise supervised, unsupervised, optimization-based and zero-shot conditions. Besides the mentioned LLIE methods, the benchmark also includes DeepLPF, HWMNet and the recurrent attentional underexposure-enhancement network introduced by Zhao et al. [24-26]. The performance is assessed using MIT-Adobe FiveK and SICE as typical photographic and multi-exposure inputs, UHD-LL for high-definition reference test and DarkFace for low-light detection evaluation [27-29].

4.3 Performance Evaluation.

We conduct a reference-based quantitative analysis between our method and the state-of-the-art methods on the UHD-LL and MIT datasets, as reported in Table 1 and Table 2, respectively. Our model achieves the best PSNR and SSIM performance on the MIT dataset. On the UHD-LL dataset, it obtains the highest PSNR and the third-best SSIM, while still outperforming existing zero-shot methods in terms of SSIM. In addition, we report a no-reference evaluation on a subset of the DarkFace dataset using the EMEE metric, as shown in Table 3. Our method obtains the highest EMEE score of 7.0122, compared with 6.7473 achieved by CoLIE. This suggests that the enhanced images contain stronger measurable contrast and information entropy, which may be beneficial for subsequent high-level vision tasks under low-light conditions.

Table 1: Benchmarking study on the testing set of the UHD-LL dataset. We highlight the best and the second and third best results.

Method	Training Setting	PSNR↑	SSIM↑	Training Set
DeepLPF [34]	Supervised	16.3087	0.5322	MIT-Adobe FiveK
HWMNet [35]	Supervised	14.3063	0.5441	MIT-Adobe FiveK
RetinexNet [27]	Supervised	16.0355	0.5796	LOL
Zhao *et al.* [36]	Supervised	15.1646	0.5465	MIT-Adobe FiveK
ZeroDCE [20]	Unsupervised	12.6484	0.5961	DarkFace
SCI-easy [22]	Unsupervised	15.5433	0.6078	MIT-Adobe FiveK
SCI-medium [22]	Unsupervised	15.4666	0.6218	LOL+LSRW
SCI-difficult [22]	Unsupervised	17.8476	0.5757	DarkFace
RUAS-LOL [28]	Unsupervised	11.6932	0.6970	LOL-v2
RUAS-MIT5K [28]	Unsupervised	14.2146	0.5894	MIT-Adobe FiveK
RUAS-DARK [28]	Unsupervised	11.2424	0.5737	DarkFace
EnGAN [21]	Unsupervised	13.1644	0.5964	Assembled
PairLIE [29]	Unsupervised	17.5349	0.6847	LOL+SICE
LIME [9]	Zero-shot	17.3496	0.5418	–
CoLIE [23]	Zero-shot	17.8927	0.6252	–
Ours	Zero-shot	18.2020	0.6455	–

Table 2: Benchmarking evaluation on a subset of the MIT dataset. We highlight the best and the second and third best results.

Method	Training Setting	PSNR↑	SSIM↑	Training Set
HWMNet [35]	Supervised	17.0284	0.8190	LOL
RetinexNet [27]	Supervised	12.0806	0.6653	LOL
ZeroDCE [20]	Unsupervised	15.2448	0.7509	SICE
SCI-medium [22]	Unsupervised	12.2299	0.7385	LOL+LSRW
SCI-difficult [22]	Unsupervised	15.0544	0.7941	DarkFace
RUAS-LOL [28]	Unsupervised	7.3771	0.4903	LOL-v2
RUAS-DARK [28]	Unsupervised	7.0460	0.4787	DarkFace
EnGAN [21]	Unsupervised	12.5576	0.7662	Assembled
PairLIE [29]	Unsupervised	12.8361	0.7677	LOL+SICE
LIME [9]	Zero-shot	14.9763	0.5418	–
CoLIE [23]	Zero-shot	18.6703	0.8317	–
Ours	Zero-shot	19.3930	0.8340	–

Table 3: Benchmarking evaluation on a subset of images from the DarkFace dataset using a no-reference metric.

Method	Training Setting	EMEE↑
Input	–	2.6286
DeepLPF [34]	Supervised	4.0142
HWMNet [35]	Supervised	4.3744
ZeroDCE [20]	Unsupervised	5.5686
SCI-difficult [22]	Unsupervised	6.3856
RUAS-DARK [28]	Unsupervised	6.2555
EnGAN [21]	Unsupervised	2.9794
PairLIE [29]	Unsupervised	1.7985
CoLIE [23]	Zero-shot	6.7473
Ours	Zero-shot	7.0122

Figure 4: Visual perception comparison with SOTA methods on a real-world low-light image from the MIT dataset. Our method not only suppressing over-exposed regions and improving under-exposed areas, but also better restores color appearance and saturation.

Additionally, the qualitative comparison in Fig. 4 shows that the proposed method suppresses over-exposed regions, improves under-exposed areas, and better restores color appearance and saturation, producing results that are more natural and closer to real-world normal-light images.

4.4 DarkFace Detection

We also assess our approach using Grounding DINO in the DarkFace detection task [30]. The improved images are inputted into the detector with "person" as the prompt. The results indicate that our method can enhance the perceptual clarity of dark images; nevertheless, this improvement does not necessarily result in higher detection accuracy.

5 Ablation Study

We conduct an ablation study on the UHD-LL dataset to investigate the effectiveness of the three key components in our framework, including the Gated Gamma Correction, the illumination-guided HV restoration module, and the HVI Transform. The quantitative results are reported in Table 4.

Table 4: Ablation study of different moudle based on the UHD-LL DATASET.

PSNR↑	SSIM↑	Gamma	IGHV	HVI-Transform
17.7936	0.6260		✓	✓
18.1758	0.6454	✓		✓
18.0641	0.6298	✓	✓	
18.2020	0.6455	✓	✓	✓

5.1 Gated Gamma Correction

By comparing the model without the gamma-gated module and the full model, we observe that introducing gamma gating improves the PSNR from 17.7936 dB to 18.2020 dB and the SSIM from 0.6260 to 0.6455. This indicates that the adaptive gamma decision is beneficial for handling extremely under-exposed regions. Instead of applying a fixed enhancement strength to all inputs, the gamma gate allows the model to adjust the enhancement according to the illumination condition, thereby alleviating insufficient restoration in very dark areas while maintaining stable exposure.

5.2 HVI transform

We investigate the effect of substituting the original HSV representation with the HVI color space. If the HVI Transformation is omitted, the model has 18.0641 dB PSNR and 0.6298 SSIM. However, after incorporating the HVI Transformation, the complete model reaches 18.2020 dB PSNR and 0.6455 SSIM. The improvement is especially clear in SSIM, suggesting that the HVI representation provides a more suitable decomposition for low-light enhancement. Compared with directly relying on the HSV Value component, the HVI intensity component offers a more stable basis for illumination restoration, while the continuous HV chromatic plane helps reduce the color instability caused by hue and saturation degradation in dark regions.

5.3 Illumination-Guided HV Restorer

We also evaluate the contribution of the illumination-guided HV restoration module. Removing this module leads to 18.1758 dB PSNR and 0.6454 SSIM, while the full model obtains 18.2020 dB PSNR and 0.6455 SSIM. Although the numerical improvement in PSNR and SSIM is relatively moderate, this module directly targets the HV chromatic components rather than only the intensity component. Therefore, it is complementary to the implicit intensity restoration branch: the restored I component provides illumination guidance for recovering degraded color and saturation information in the HV plane. This design is important for improving chromatic fidelity, since PSNR and SSIM mainly reflect pixel-level and structural consistency and may not fully capture perceptual improvements in color and saturation.

6 Conclusion

In this paper, we introduce HVI-CoLIE, a novel zero-shot low-light image enhancement model utilizing a neural implicit representation function (NIR) in the HVI color space. Our approach focuses on improving exposure correction and chromatic restoration, aiming to enhance image quality and perceptual fidelity under unknown illumination conditions. Particularly, we incorporate gated gamma correction for adaptive exposure control and use the recovered Intensity component to guide the restoration of HV chromatic components, beyond the illumination-only restoration strategy of conventional zero-shot methods. Experimental results demonstrate superior performance on real-world night and under-exposed images, showing that our method effectively reduces over-exposure, color bias, and chromatic distortion while preserving the flexibility of zero-shot enhancement.

References

[1] Yang, W., Yuan, Y., Ren, W., Liu, J., Scheirer, W. J., Wang, Z., Zhang, T., Zhong, Q., Xie, D., Pu, S., Zheng, Y., Qu, Y., Xie, Y., Chen, L., Li, Z., Hong, C., Jiang, H., Yang, S., & Liu, Y. (2020). Advancing image understanding in poor visibility environments: A collective benchmark study. IEEE Transactions on Image Processing, 29, 5737-5752.

[2] Huang, Y., Zha, Z.-J., Fu, X., Hong, R., & Li, L. (2020). Real-world person re-identification via degradation invariance learning. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 14084-14094.

[3] Xu, X., Wang, S., Wang, Z., Zhang, X., & Hu, R. (2021). Exploring image enhancement for salient object detection in low light images. ACM Transactions on Multimedia Computing, Communications, and Applications, 17(1s), 1-19.

[4] Pizer, S. M., Johnston, R. E., Ericksen, J. P., Yankaskas, B. C., & Muller, K. E. (1990). Contrast-limited adaptive histogram equalization: Speed and effectiveness. Proceedings of the First Conference on Visualization in Biomedical Computing, 337, 2.

[5] Jobson, D. J., Rahman, Z.-u., & Woodell, G. A. (1997). A multiscale Retinex for bridging the gap between color images and the human observation of scenes. IEEE Transactions on Image Processing, 6(7), 965-976.

[6] Guo, X., Li, Y., & Ling, H. (2017). LIME: Low-light image enhancement via illumination map estimation. IEEE Transactions on Image Processing, 26(2), 982-993.

[7] Ren, X., Yang, W., Cheng, W.-H., & Liu, J. (2020). LR3M: Robust low-light enhancement via low-rank regularized Retinex model. IEEE Transactions on Image Processing, 29, 5862-5876.

[8] Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. Pattern Recognition, 61, 650-662.

[9] Chen, C., Chen, Q., Xu, J., & Koltun, V. (2018). Learning to see in the dark. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3291-3300.

[10] Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep Retinex decomposition for low-light enhancement. Proceedings of the British Machine Vision Conference.

[11] Zhang, Y., Zhang, J., & Guo, X. (2019). Kindling the darkness: A practical low-light image enhancer. Proceedings of the 27th ACM International Conference on Multimedia, 1632-1640.

[12] Zhang, Y., Guo, X., Ma, J., Liu, W., & Zhang, J. (2021). Beyond brightening low-light images. International Journal of Computer Vision, 129(4), 1013-1037.

[13] Yang, W., Wang, S., Fang, Y., Wang, Y., & Liu, J. (2020). From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3063-3072.

[14] Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., & Cong, R. (2020). Zero-reference deep curve estimation for low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1780-1789.

[15] Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., & Wang, Z. (2021). EnlightenGAN: Deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 30, 2340-2349.

[16] Liu, R., Ma, L., Zhang, J., Fan, X., & Luo, Z. (2021). Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 10561-10570.

[17] Wu, W., Weng, J., Zhang, P., Wang, X., Yang, W., & Jiang, J. (2022). URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5901-5910.

[18] Ma, L., Ma, T., Liu, R., Fan, X., & Luo, Z. (2022). Toward fast, flexible, and robust low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5637-5646.

[19] Fu, Z., Yang, Y., Tu, X., Huang, Y., Ding, X., & Ma, K.-K. (2023). Learning a simple low-light image enhancer from paired low-light instances. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 22252-22261.

[20] Chobola, T., Liu, Y., Zhang, H., Schnabel, J. A., & Peng, T. (2024). Fast context-based low-light image enhancement via neural implicit representations. European Conference on Computer Vision, 413-430.

[21] Yan, Q., Feng, Y., Zhang, C., Pang, G., Shi, K., Wu, P., Dong, W., Sun, J., & Zhang, Y. (2025). HVI: A new color space for low-light image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5678-5687.

[22] Chen, Y., Liu, S., & Wang, X. (2021). Learning continuous image representation with local implicit image function. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 8628-8638.

[23] Lee, J., & Jin, K. H. (2022). Local texture estimator for implicit representation function. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1929-1938.

[24] Moran, S., Marza, P., McDonagh, S., Parisot, S., & Slabaugh, G. (2020). DeepLPF: Deep local parametric filters for image enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 12826-12835.

[25] Fan, C.-M., Liu, T.-J., & Liu, K.-H. (2022). Half wavelet attention on M-Net+ for low-light image enhancement. IEEE International Conference on Image Processing, 3878-3882.

[26] Zhao, L., Lu, S.-P., Chen, T., Yang, Z., & Shamir, A. (2021). Deep symmetric network for underexposed image enhancement with recurrent attentional learning. Proceedings of the IEEE/CVF International Conference on Computer Vision, 12075-12084.

[27] Bychkovsky, V., Paris, S., Chan, E., & Durand, F. (2011). Learning photographic global tonal adjustment with a database of input/output image pairs. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 97-104.

[28] Cai, J., Gu, S., & Zhang, L. (2018). Learning a deep single image contrast enhancer from multi-exposure images. IEEE Transactions on Image Processing, 27(4), 2049-2062.

[29] Li, C., Guo, C.-L., Zhou, M., Liang, Z., Zhou, S., Feng, R., & Loy, C. C. (2023). Embedding Fourier for ultra-high-definition low-light image enhancement. arXiv preprint arXiv:2302.11831.

[30] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., Jiang, Q., Li, C., Yang, J., Su, H., Zhu, J., & Zhang, L. (2024). Grounding DINO: Marrying DINO with grounded pre-training for open-set object detection. European Conference on Computer Vision, 38-55.