DS-LIF: NOVEL SPIKING NEURON FOR RELIABLE LOW-LIGHT IMAGE ENHANCEMENT

Anonymous authors
Paper under double-blind review

ABSTRACT

The event-triggered mechanism and rich dynamic characteristics of spiking neural networks (SNNs) are highly compatible with the dynamic adaptive repair requirements of low-level visual restoration tasks. However, SNNs adopted for low-light image enhancement (LLIE) fail to accurately distinguish variations in structural information within under-illuminated regions and produce ambiguous spiking responses triggered by image edges, textures, and noise. To address this issue, we propose a reliability-aware spiking neuron named Dynamic-Sleep Leaky Integrate-and-Fire (DS-LIF). It scales the input current prior to membrane potential integration according to local illumination and color contrast, and dynamically masks weakly effective feature channels for image restoration. Subsequently, it combines the mean spike response across all time steps with the local noise, chromatic stability, and underexposure level of the generated image to produce a reliability-guided map. Based on DS-LIF, we construct a dual-stream HVI enhancement framework, SpikeRelight, which can effectively improve the quality of LLIE. To better facilitate cross-stream information interaction, we design a Reliability-Guided Dual-Frequency Interaction (RGDFI) module in this architecture, guided by the reliability-guided map. It decomposes features in the illumination flow and chrominance flow into low-frequency and high-frequency parts, and performs asymmetric cross-stream interaction. Extensive experiments on both paired and unpaired benchmarks, along with ablation studies, demonstrate that SpikeRelight achieves competitive performance.

1 INTRODUCTION

Low-light image enhancement (LLIE) is important for camera imaging in dark environments (Jiang et al., 2025), nighttime autonomous-driving perception (Sakaridis et al., 2025), and surveillance-oriented visual perception under poor illumination (Jia et al., 2021). Reliable LLIE should improve image visibility while preserving scene structures and color fidelity under complex illumination conditions. To meet these requirements, recent studies have explored Retinex-based illumination modeling (Cai et al., 2023), color-space decoupling (Yan et al., 2025), illumination-invariant structural guidance (Dong et al., 2025), pixel-adaptive restoration (Jin et al., 2025), and lightweight restoration with multiple priors (Brateanu et al., 2026). Biologically inspired adaptive information allocation has also been investigated to make enhancement more selective (Liu et al., 2026). Nevertheless, these methods primarily optimize visibility, structural preservation, and color fidelity without explicitly estimating whether local responses are reliable enough to guide restoration. This limitation becomes especially critical in low-illumination areas, where recoverable structures and noise co-exist, making it necessary to balance illumination correction with the preservation of edge and texture details.

For high-quality LLIE tasks, this reliability gap manifests in two related ways. On the one hand, in low-light regions, responses from authentic edges and textures are hardly distinguishable from noise, forcing the model to either suppress weak but meaningful details or amplify noise and thus produce local artifacts. Existing methods alleviate this ambiguity through signal-to-noise modeling (Xu et al., 2022), intensity–chrominance decoupling (Yan et al., 2025), and structure-aware guidance (Dong et al., 2025). Although these strategies enhance feature representations for low-light degradations, they are typically integrated into end-to-end restoration networks and lack an explicit mechanism to judge whether local responses provides reliable information for restoration update. On the other

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

hand, the contributions of different feature channels to the restoration process vary, while existing methods mostly treat them in a relatively uniform manner. For example, RodNet (Liu et al., 2026) leveraged image-drived priors to direct channel-wise pruning in a global manner. PASD (Jin et al., 2025) applied a uniform pixel-adaptive scheme to all feature channels without differentiation. In these approaches, channels with limited contributions receive the same emphasis as informative ones, leading to redundant information and the propagation of unreliable information. Therefore, it is crucial to design a mechanism that both evaluates response reliability and adaptively adjust the participation of feature channels.

Recent studies have shown that spiking neural networks (SNNs) in low-level vision tasks can be adapted to continuous-valued image restoration and provide a possible adaptation solution for solving reliability problems. SpiDen (Castagnetti et al., 2023) investigates spiking computation for image denoising, while ESDNet (Song et al., 2024) and VLIF (Chen et al., 2026) explore spiking mechanisms for image deraining. Spike-RetinexFormer (Wang et al., 2026) further incorporates spiking operations into a Retinex-based low-light enhancement pipeline. These studies demonstrate that the membrane potential dynamics of spiking neurons can accumulate response information, and the event-driven activation provides a natural basis for selective feature processing. However, current spiking neuron firing strategies cannot explicitly distinguish whether the responses in dark areas originate from recoverable structures or noise, nor can they identify which feature channels should remain active during restoration. Therefore, how to enable spiking neurons to estimate the reliability of restoration information and selectively shield the channels with limited contributions, so as to enhance the overall restoration quality, is an important research topic.

To address the aforementioned issues, this paper first designs a new reliable perception-based spiking neuron named Dynamic-Sleep Leaky Integrate-and-Fire (DS-LIF), specifically for low light image enhancement, in order to obtain the reliability-guided map. Subsequently, based on this novel neuron, a HVI dual-stream enhancement architecture SpikeRelight is constructed. Within this architecture, the information interaction guided by reliability is achieved through the Reliability-Guided Dual-Frequency Interaction (RGDFI) module. Experimental results demonstrate that this architecture possesses high-efficient low-light image restoration capabilities.

The main contributions are summarized as follows:

- We propose a reliability-aware spiking neuron named DS-LIF. Unlike conventional spiking neurons with fixed membrane-integration and firing rules, DS-LIF dynamically adjusts the input current based on local illumination and local contrast to achieve spike adaptive firing rules. Then, it combines this with prior information related to image restoration to obtain a reliability guidance map for image restoration.
- We build an HVI-based dual-stream enhancement architecture SpikeRelight centered on DS-LIF. Within this architecture, we develop a RGDFI module for directional cross-stream interaction between intensity and chrominance features, which retains the full bandwidth cross-attention path and decomposes the features into low-frequency part and high-frequency part that are processed under the guidance of the reliability-guided map.
- Extensive experiments on paired and unpaired benchmarks demonstrate that SpikeRelight achieves highly competitive performance against representative state-of-the-art methods. Ablation studies verify the individual contributions of DS-LIF and RGDFI, while cross-architecture experiments on RetinexFormer further demonstrate the transferability of DS-LIF to other image-restoration architectures.

2 Related Work

2.1 Low-Light Image Enhancement

Early LLIE methods based on histogram equalization and hand-crafted Retinex priors (Li et al., 2021) often provide limited enhancement quality because their fixed assumptions are difficult to adapt to complex illumination and noise conditions. Deep-learning methods such as URetinex-Net (Wu et al., 2022) typically decompose an image into illumination and reflectance components in the RGB space, but decomposition errors and channel coupling may lead to color distortion issues. Transformer-based LLIE methods, such as Retinexformer (Cai et al., 2023), further model

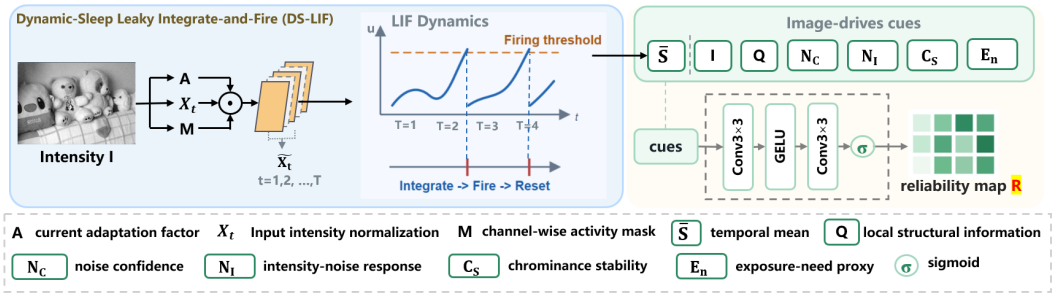

Figure 1: Dynamic-Sleep Leaky Integrate-and-Fire (DS-LIF). The input current $\widetilde{\mathbf{X}}_t$ is modulated by the illumination-aware adaptation factor A and the channel-wise activity mask M, while the resulting temporal spike statistics $\bar{S}$ and image-derived information are used to estimate the restoration reliability-guided map.

long-range dependencies during restoration. Recent studies have also explored SNR-aware processing (Xu et al., 2022), illumination decomposition, and paired or unpaired learning, while investigating more selective feature representations (Shi et al., 2025; Dong et al., 2025) and adaptive restoration strategies (Liu et al., 2026). However, none of these methods addressed the differences in the responses of local features.

2.2 Spiking Neural Networks for Image Restoration

SNNs have recently been extended from recognition to low-level image restoration by exploiting temporal neuronal dynamics and event-driven sparse computation. SpiDen (Castagnetti et al., 2023) investigates spiking computation for image denoising, while SpikingIR (Ouyang et al., 2026) generalizes spiking representations to image restoration. ESDNet (Song et al., 2024) introduces task-oriented spiking residual learning for image deraining, and VLIF (Chen et al., 2026) develops spatially adaptive neuronal dynamics based on the frequency response and saturation behavior of LIF neurons. Spike-RetinexFormer (Wang et al., 2026) further incorporates spiking operations into a Retinex-based low-light enhancement pipeline. These studies show that SNNs can support continuous-valued image restoration and may suit its adaptive processing demands. However, existing SNN-based LLIE methods generally retain fixed membrane-integration and thresholding rules, treating spikes mainly as reconstruction features. They neither estimate local response reliability nor suppress persistently weak channels or use temporal statistics to guide cross-stream information exchange. Therefore, Reliability-aware temporal processing remains underexplored in LLIE.

3 Method

How to achieve high-quality LLIE still poses significant challenges, as the same weak response could either correspond to the actual image structure or be derived from noise. Furthermore, the reliability of restoration information in different spatial regions, characteristic channels, and frequency components varies. Therefore, a restoration model should not only estimate an enhancement mapping, but also determine which responses are sufficiently reliable to participate in subsequent feature propagation. Based on this observation, we first introduce a reliability-aware spiking neuron DS-LIF (Fig. 1) to get a reliability-duided map of low-light features. We then embed the inferred reliability into an HVI-based dual-stream enhancement architecture SpikeRelight (Fig. 2). In this architeture, in order to convert the reliability state into asymmetric information routing between intensity and chrominance features, we develop a RGDFI (Fig. 2 (c)) module.

3.1 Dynamic-Sleep Leaky Integrate-and-Fire (DS-LIF)

In LLIE, a weak feature response may correspond either to a meaningful structure or to sensor noise. Conventional LIF neurons apply the same membrane dynamics and firing rules to all feature channels, which prevents the temporal processing from adapting to local illumination and structural conditions. In addition, channels with little restoration information may continue firing spikes across

all timesteps, introducing redundant activity. Therefore, we design DS-LIF as illustrated in Fig. 1 to adapt the input current and channel activity according to illumination and local structure.

Given an intensity map $\mathbf{I}$, we first compute its local mean using a 3×3 average pooling:

$$\bar{\mathbf{I}} = \mathrm{AvgPool}_{3\times3}(\mathbf{I}). \tag{1}$$

The local structural information $\mathbf{Q}$ is then estimated by

$$\mathbf{Q} = \mathrm{clip}\left(\frac{\left|\mathbf{I} - \bar{\mathbf{I}}\right|}{\bar{\mathbf{I}} + 0.03}, 0, 1\right), \tag{2}$$

where $\mathrm{clip}(\cdot)$ clamps values to $[0, 1]$. A larger $\mathbf{Q}$ means stronger local contrast or structural variation.

Let $\mathbf{D} = 1 - \mathrm{clip}(\mathbf{I}, 0, 1)$ denote the darkness map. DS-LIF constructs an illumination- and structure-dependent current adaptation factor:

$$\mathbf{A} = \frac{1 + \kappa\, \mathbf{D} \odot \mathbf{Q}}{1 + \eta\, \mathbf{D} \odot (1 - \mathbf{Q})}, \tag{3}$$

where κ and η control the enhancement of informative dark responses and the suppression of flat dark responses, respectively. Consequently, dark regions with clear structure are strengthened, while flat or ambiguous ones are dampened.

For an input feature $\mathbf{F}$, the projected current is first normalized and mapped to the spiking channels:

$$\mathbf{X} = \mathrm{Conv}_{3\times3}\left(\frac{\log(1 + \zeta\mathbf{I})}{T}\right), \qquad \mathbf{X}_t = \mathbf{X}, \quad t = 1, \ldots, T, \tag{4}$$

where T is the number of timesteps and ζ controls the intensity scaling. Before temporal integration, the current is modulated by $\mathbf{A}$ and a channel-wise activity mask $\mathbf{M}$:

$$\widetilde{\mathbf{X}}_t = \mathbf{X}_t \odot \mathbf{A} \odot \mathbf{M}. \tag{5}$$

The channel mask is determined from the normalized activity of each channel:

$$\begin{aligned} \mu_{b,c} &= \frac{1}{THW} \sum_{t=1}^{T} \sum_{h=1}^{H} \sum_{w=1}^{W} \left|\mathbf{X}_{t,b,c,h,w}\right|, \\ r_{b,c} &= \frac{\mu_{b,c}}{\max\left(\frac{1}{C}\sum_{c'=1}^{C} \mu_{b,c'}, \epsilon\right)}, \end{aligned} \tag{6}$$

where $\epsilon = 10^{-6}$, $\mu_{b,c}$ denotes the average activity magnitude of channel c in sample b, and $r_{b,c}$ denotes its activity normalized by the mean activity across all channels of same sample.

The soft keep probability is computed as:

$$p_{b,c} = \sigma\left(\frac{r_{b,c} - \left[1 + \eta\left(0.5 - \overline{Q}_b\right)\right]}{\tau_g}\right), \tag{7}$$

where $\overline{Q}_b$ is the spatial mean of the structural map for sample b, τ_g is the gate temperature, and $\sigma(\cdot)$ denotes the sigmoid function. The forward pass uses a hard binary mask, while the backward pass retains the gradient of the soft probability.

The membrane potential is then updated over T timesteps:

$$\begin{aligned} \lambda &= \exp(-1/\tau), \qquad \mathbf{s}_t = H\left(\mathbf{u}_t - \theta\right), \\ \mathbf{u}_t &= \lambda \mathbf{u}_{t-1} + \widetilde{\mathbf{X}}_t, \quad \mathbf{u}_t \leftarrow \mathbf{u}_t - \theta \mathbf{s}_t, \end{aligned} \tag{8}$$

where τ is the membrane time constant, θ is the firing threshold, $H(\cdot)$ is the Heaviside step function, and $\mathbf{s}_t$ is the binary spike output.

DS-LIF then summarizes the spike response by its temporal mean:

$$\overline{\mathbf{S}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{s}_t. \tag{9}$$

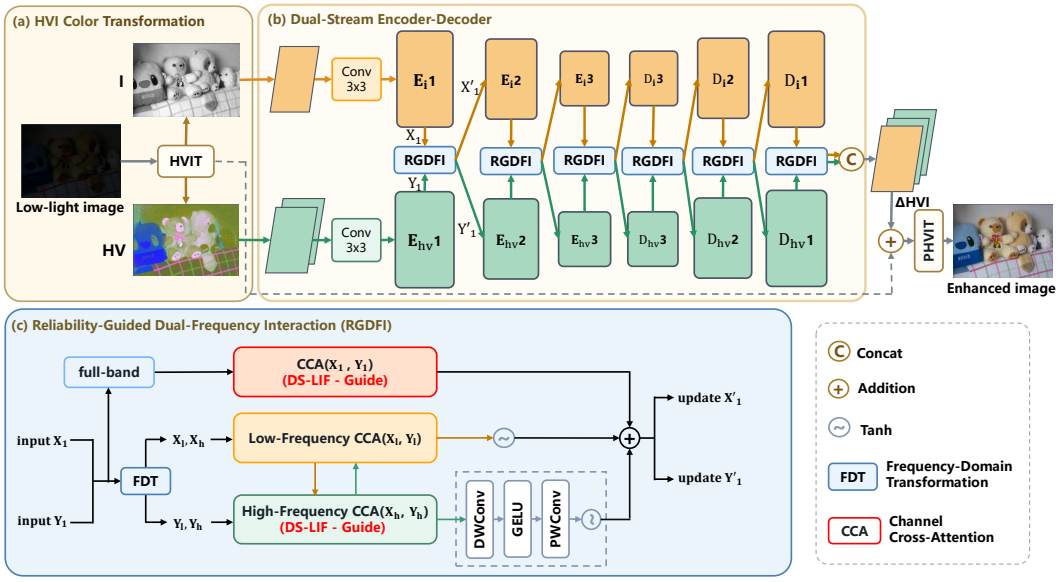

Figure 2: The overall framework of SpikeRelight. (a) HVI color transformation, (b) the dual-stream encoder–decoder backbone equipped with multi-scale RGDFI blocks, (c) detailed structure of RGDFI, where the DS-LIF reliability-guided map guides full-band and high-frequency cross-attention.

This statistic is combined with image-derived information, including intensity, local structure, noise confidence, noise response, chrominance stability, and exposure requirements. The restoration reliability is predicted by a lightweight two-layer 3×3 convolutional head:

$$\mathbf{R} = \sigma\left(\mathrm{Conv}_{3\times3}^{\mathrm{out}}\left[\mathrm{GELU}\left(\mathrm{Conv}_{3\times3}^{\mathrm{in}}\left([\overline{\mathbf{S}}, \mathbf{I}, \mathbf{Q}, \mathbf{N}_c, \mathbf{N}_I, \mathbf{C}_s, \mathbf{E}_n]\right)\right)\right]\right), \tag{10}$$

where $\mathbf{N}_c$, $\mathbf{N}_I$, $\mathbf{C}_s$, and $\mathbf{E}_n$ denote the noise confidence, normalized intensity-noise response, chrominance stability, and exposure-need proxy, respectively. The reliability-guided map $\mathbf{R}$ is passed to the subsequent recovery network to guide feature selection and cross-stream interaction.

3.2 SpikeRelight Framework

Since DS-LIF estimates restoration reliability rather than reconstructing image content, we develop an HVI-based dual-stream enhancement architecture SpikeRelight centered on it. Following the intensity-chrominance decoupling of HVI, SpikeRelight uses a dual-stream encoder–decoder to process intensity and chrominance information separately.

As illustrated in Fig. 2. Given a low-light RGB image $\mathbf{x}$, the HVI transformation first produces an intensity feature map and a chrominance feature map:

$$(\mathbf{I}, \mathbf{C}) = \mathrm{HVIT}(\mathbf{x}), \qquad \mathbf{C} = [\mathbf{H}, \mathbf{V}]. \tag{11}$$

The intensity feature map $\mathbf{I}$ encodes iuminance, whereas $\mathbf{C}$ encodes chrominance.

The transformed representation is fed into two feature streams:

$$\mathbf{F}_I^0 = \phi_I(\mathbf{I}), \qquad \mathbf{F}_C^0 = \phi_C(\mathbf{C}), \tag{12}$$

where ϕ_I and ϕ_C denote shallow feature projections. The two streams are processed by symmetric encoder–decoder paths, with downsampling enlarging the receptive field and skip connections restoring resolution while preserving fine details.

The intermediate features undergo directional information interaction through the RGDFI module. This process is guided by R from DS-LIF, then updated luminance information $\mathbf{F}_I^{k+1}$ and chrominance information $\mathbf{F}_C^{k+1}$ are obtained, which are respectively used as the input for the next stage:

$$\mathbf{F}_I^{k+1} = \mathrm{RGDFI}\left(\mathbf{F}_I^k, \mathbf{F}_C^k; \mathbf{R}_C^k\right), \qquad \mathbf{F}_C^{k+1} = \mathrm{RGDFI}\left(\mathbf{F}_C^k, \mathbf{F}_I^k; \mathbf{R}_I^k\right). \tag{13}$$

After decoding, the network predicts an HVI residual $\Delta\mathbf{HVI}$. Then the residual is added to the transformed input representation:

$$\widehat{\mathbf{hvi}} = \mathrm{HVIT}(\mathbf{x}) + \Delta\mathbf{HVI}. \tag{14}$$

The enhanced RGB image is finally obtained through the inverse HVI transformation:

$$\widehat{\mathbf{y}} = \mathrm{PHVIT}\left(\widehat{\mathbf{hvi}}\right). \tag{15}$$

3.3 Reliability-Guided Dual-Frequency Interaction (RGDFI)

3.3.1 Why reliability-guided frequency interaction is needed?

Low-frequency and high-frequency information play different roles in restoration: Low-frequency part encodes smooth illumination and chrominance variations, their exchange should remain spatially coherent. High-frequency part contains edges and textures but may also carry noise and unstable dark-region responses. Directly multiplying the reliability map with the feature map would treat each frequency part equally, which is inappropriate for LLIE. As a result, we design a reliability-guided for frequency-selective interaction.

3.3.2 Frequency decomposition

For an input feature $\mathbf{F}$, we use a low-pass operator $\mathcal{L}(\cdot)$ to obtain its low-frequency part:

$$\mathbf{F}^{\mathrm{low}} = \mathcal{L}(\mathbf{F}), \qquad \mathbf{F}^{\mathrm{high}} = \mathbf{F} - \mathbf{F}^{\mathrm{low}}. \tag{16}$$

The same decomposition is applied to both streams:

$$\mathbf{F}_I \rightarrow \left(\mathbf{F}_I^{\mathrm{low}}, \mathbf{F}_I^{\mathrm{high}}\right), \qquad \mathbf{F}_C \rightarrow \left(\mathbf{F}_C^{\mathrm{low}}, \mathbf{F}_C^{\mathrm{high}}\right). \tag{17}$$

3.3.3 Reliability-guided asymmetric interaction and aggregation

As illustrated in Fig. 2(c), RGDFI contains three complementary interaction paths. The full-band exchanges information between original features:

$$\mathbf{U}_I^{\mathrm{full}} = \mathcal{A}_{\mathrm{full}}\left(\mathbf{F}_I, \mathbf{F}_C\right), \qquad \mathbf{U}_C^{\mathrm{full}} = \mathcal{A}_{\mathrm{full}}\left(\mathbf{F}_C, \mathbf{F}_I\right), \tag{18}$$

where $\mathcal{A}_{\mathrm{full}}$ denotes full-band cross-attention, the corresponding reliability-guided map is used to regulate the transmitted update:

$$\widetilde{\mathbf{U}}_I^{\mathrm{full}} = \mathbf{R}_C \odot \mathbf{U}_I^{\mathrm{full}}, \qquad \widetilde{\mathbf{U}}_C^{\mathrm{full}} = \mathbf{R}_I \odot \mathbf{U}_C^{\mathrm{full}}, \tag{19}$$

where $\mathbf{R}_I$ and $\mathbf{R}_C$ denote the reliability-guided maps associated with the intensity and chrominance streams, respectively.

The low-frequency path performs cross-stream interaction on the smooth components:

$$\mathbf{U}_I^{\mathrm{low}} = \mathcal{A}_{\mathrm{low}}\left(\mathbf{F}_I^{\mathrm{low}}, \mathbf{F}_C^{\mathrm{low}}\right), \qquad \mathbf{U}_C^{\mathrm{low}} = \mathcal{A}_{\mathrm{low}}\left(\mathbf{F}_C^{\mathrm{low}}, \mathbf{F}_I^{\mathrm{low}}\right). \tag{20}$$

Because low-frequency parts mainly describe slowly varying illumination and chrominance, their interaction is not directly suppressed by the reliability gate. This preserves the spatial continuity required for global brightness and color correction.

The high-frequency path applies reliability-controlled interaction to the detail components:

$$\mathbf{U}_I^{\mathrm{high}} = \mathcal{A}_{\mathrm{high}}\left(\mathbf{F}_I^{\mathrm{high}}, \mathbf{F}_C^{\mathrm{high}}\right), \qquad \mathbf{U}_C^{\mathrm{high}} = \mathcal{A}_{\mathrm{high}}\left(\mathbf{F}_C^{\mathrm{high}}, \mathbf{F}_I^{\mathrm{high}}\right), \tag{21}$$

$$\widetilde{\mathbf{U}}_I^{\mathrm{high}} = \mathbf{R}_C \odot \mathbf{U}_I^{\mathrm{high}}, \qquad \widetilde{\mathbf{U}}_C^{\mathrm{high}} = \mathbf{R}_I \odot \mathbf{U}_C^{\mathrm{high}}. \tag{22}$$

The reliability gate attenuates high-frequency updates in regions where the corresponding source response is uncertain. Informative high-frequency responses remain available for detail reconstruction, whereas noise-sensitive updates receive a smaller contribution.

Then the three paths are aggregated with learnable branch scales:

$$\begin{aligned} \mathbf{F}_I^{\mathrm{out}} &= \mathbf{F}_I + \widetilde{\mathbf{U}}_I^{\mathrm{full}} + \tanh(\gamma_{\mathrm{low}})\mathbf{U}_I^{\mathrm{low}} + \tanh(\gamma_{\mathrm{high}})\widetilde{\mathbf{U}}_I^{\mathrm{high}}, \\ \mathbf{F}_C^{\mathrm{out}} &= \mathbf{F}_C + \widetilde{\mathbf{U}}_C^{\mathrm{full}} + \tanh(\gamma_{\mathrm{low}})\mathbf{U}_C^{\mathrm{low}} + \tanh(\gamma_{\mathrm{high}})\widetilde{\mathbf{U}}_C^{\mathrm{high}}, \end{aligned} \tag{23}$$

where γ_{low} and γ_{high} are learnable scalar parameters.

Table 1: Quantitative comparison with representative LLIE methods on LOLv1 and LOLv2-Real. Red and blue indicate the best and second-best results, respectively.

Method	Reference	LOLv1			LOLv2-Real		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
SNR-Aware	Xu et al. (2022)	22.91	0.83	0.14	20.93	0.85	0.15
Retinexformer	Cai et al. (2023)	**24.24**	0.83	0.13	22.32	0.84	0.17
Bread	Guo & Hu (2023)	20.18	0.81	0.27	20.87	0.82	0.22
PairLIE	Fu et al. (2023)	23.53	0.76	0.25	19.89	0.78	0.32
QuadPrior	Wang et al. (2024)	22.85	0.80	0.20	20.59	0.81	0.20
DarkIR	Feijoo et al. (2025)	23.22	0.84	0.18	23.74	0.86	0.19
MultiNex	Brateanu et al. (2026)	23.19	0.84	0.13	23.04	0.86	0.18
RodNet	Liu et al. (2026)	24.18	0.85	0.11	**24.24**	0.87	0.13
CIDNet	Yan et al. (2025)	24.17	**0.85**	**0.10**	23.50	0.86	0.13
SpikeRelight (Ours)	–	**24.93**	**0.86**	**0.09**	**23.73**	**0.87**	**0.13**

Table 2: Quantitative comparison on SID and SICE.

Method	Reference	SID		SICE	
		PSNR↑	SSIM↑	PSNR↑	SSIM↑
RetinexNet	Wei et al. (2018)	15.70	0.40	12.42	0.61
ZeroDCE	Guo et al. (2020)	14.09	0.09	12.45	0.64
URetinexNet	Wu et al. (2022)	15.52	0.32	10.90	0.61
RUAS	Liu et al. (2021)	12.62	0.08	8.66	0.49
LLFlow	Wang et al. (2022)	16.23	0.37	12.74	0.62
CIDNet	Yan et al. (2025)	22.55	0.68	13.52	0.61
SpikeRelight (Ours)	–	**23.00**	**0.69**	**14.39**	**0.65**

3.4 Loss Function

The loss function is defined as:

$$\mathcal{L}_{\text{total}} = \sum_{d\in\{\text{RGB,HVI}\}} \lambda_d \left(\mathcal{L}_1^d + \mathcal{L}_{\text{ssim}}^d + \mathcal{L}_{\text{edge}}^d + \lambda_{\text{p}}\mathcal{L}_{\text{perc}}^d\right) + \lambda_{\text{cal}}\mathcal{L}_{\text{cal}} + \lambda_{\text{rate}}\mathcal{L}_{\text{rate}}, \tag{24}$$

where $\lambda_{\text{RGB}} = 1$ and λ_{HVI} balances the reconstruction losses in the two representation domains. $\mathcal{L}_1$, $\mathcal{L}_{\text{ssim}}$, $\mathcal{L}_{\text{edge}}$, and $\mathcal{L}_{\text{perc}}$ denote the pixel, structural, edge, and perceptual losses, respectively. $\mathcal{L}_{\text{cal}}$ offers supervision for the reliability map and chrominance gate. Spike-rate term penalizes the mean firing rate outside a prescribed interval through $\mathcal{L}_{\text{rate}} = [r_{\min} - r_s]_+^2 + [r_s - r_{\max}]_+^2$, where r_s is average spike rate and $[\cdot]_+ = \max(\cdot, 0)$. The auxiliary terms are controlled by their weights.

4 Experiments

4.1 Implementation Details

We conduct experiments on LOLv1 Wei et al. (2018), LOLv2-Real Yang et al. (2021), SICE Cai et al. (2018), SID (Sony-Total-Dark) Chen et al. (2018), and FiveK Bychkovsky et al. (2011) using their official training and testing partitions, and further evaluate generalization on the unpaired datasets DICM Lee et al. (2013), LIME Guo et al. (2016), MEF Ma et al. (2015), NPE Wang et al. (2013), and VV Vonikakis et al. (2018). All experiments are conducted on a single NVIDIA GeForce RTX 4090 GPU. We use the Adam optimizer with an initial learning rate of 1×10^{-4}. Training images are randomly cropped into 256×256 patches and normalized to $[0, 1]$. Unless otherwise specified, the model is trained for 1000 epochs, with a three-epoch linear warm-up followed by cosine annealing. Further training and dataset details are provided in the supplementary material.

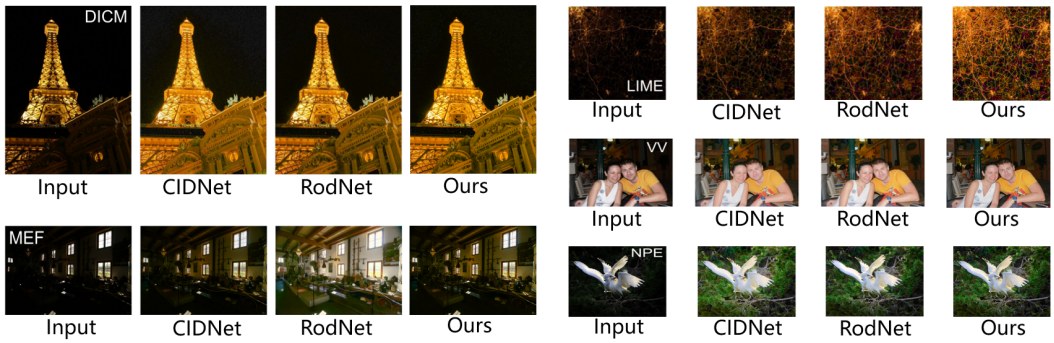

Figure 3: Visual comparison on the five unpaired datasets.

4.2 Main Experimental Results

Results on paired datasets. We use peak signal-to-noise ratio (PSNR) Hore & Ziou (2010) as the primary fidelity metric, with structural similarity (SSIM) Wang et al. (2004) and learned perceptual image patch similarity (LPIPS) Zhang et al. (2018) with AlexNet Krizhevsky et al. (2012) as complementary measures of structural consistency and perceptual similarity. Table 1 summarizes the comparison with representative low-light image enhancement methods on LOLv1 and LOLv2-Real. On LOLv1, SpikeRelight outperforms all compared methods across the three metrics, improving over the reproduced CIDNet baseline by 0.76 dB in PSNR, 0.01 in SSIM, and 0.01 in LPIPS. Averaged over LOLv1 and LOLv2-Real, SpikeRelight yields a 0.49 dB PSNR improvement over CIDNet. Results on SICE, SID, and FiveK, reported in Table 2 and Table 3, further show that SpikeRelight maintains competitive restoration quality.

Table 3: Quantitative comparison on FiveK.

Method	Reference	PSNR↑
DeepUPE	Guo et al. (2019)	23.04
MIRNet	Zamir et al. (2020)	23.73
SNR-Net	Xu et al. (2022)	23.81
Restormer	Zamir et al. (2022)	24.13
Retinexformer	Cai et al. (2023)	**24.94**
CIDNet	Yan et al. (2025)	24.34
SpikeRelight (Ours)	–	**24.43**

Results on unpaired datasets. Due to the lack of a properly aligned normal-light references, we use Naturalness Image Quality Evaluator (NIQE) Mittal et al. (2012b) and Blind/Referenceless Image Spatial Quality Evaluator (BRIS) Mittal et al. (2012a) metrics. Lower values indicate better perceptual quality. As reported in Table 5, the average BRIS score of SpikeRelight on the five unpaired datasets is lower than CIDNet (20.982 vs. 23.521), indicating that it produced fewer perceptual distortion under the no-reference evaluation protocol. The visual comparisons in Fig. 3 further show that SpikeRelight produces balanced illumination, avoids excessive highlight enhancement, and preserves local structures and color appearance in challenging low-light scenes.

4.3 Ablation Study

Effectiveness of DS-LIF. As shown in Table 4, the full-model uses DS-LIF with the default four timesteps. The *continuous activation* variant replaces binary spikes with sigmoid responses while keeping the temporal integration, input projection, and reliability heads unchanged. The *standard LIF* variant replaces DS-LIF, removing its input-current modulation and dynamic channel suppression, whereas the *single timestep* variant retains DS-LIF but sets $T = 1$. The *full model* outperforms *continuous activation* by 0.2191 dB, demonstrating the advantage of binary temporal responses. Replacing DS-LIF with standard LIF reduces PSNR by 0.3408 dB, validating the effectiveness of

Table 4: Ablation of DS-LIF on LOLv1.

Variant	PSNR↑	SSIM↑
full-model	**24.9253**	0.8594
continuous activation	24.7062	0.8547
standard LIF	24.5845	0.8598
single timestep ($T = 1$)	24.5494	**0.8623**

Table 5: Comparison of average metrics on five unpaired datasets (DICM, LIME, MEF, NPE, and VV).

Methods	Reference	NIQE↓	BRISQUE↓
RetinexNet	Wei et al. (2018)	4.558	23.286
ZeroDCE	Guo et al. (2020)	4.763	26.343
URetinexNet	Wu et al. (2022)	3.829	26.359
RUAS	Liu et al. (2021)	4.800	36.372
LLFlow	Wang et al. (2022)	4.221	28.087
CIDNet	Yan et al. (2025)	**3.523**	23.521
SpikeRelight (Ours)	–	3.878	20.982

input-current modulation and dynamic channel suppression. Setting $T = 1$ further reduces PSNR by 0.3759 dB relative to the full model, underscoring the necessity of multi-timestep response aggregation for reliability estimation.

Effectiveness of RGDFI. As shown in Table 6, the standard full-band cross-stream interaction serves as a baseline without reliability conditioning or frequency-specific branches. Compared with this baseline, *full-band only* introduces DS-LIF-guided routing and improves the results, confirming the benefit of reliability conditioning. Adding the low- and high-frequency branches further yields the complete RGDFI and brings further improvements, demonstrating the effectiveness of the dual-frequency design.

Table 6: Ablation of RGDFI on LOLv1.

Variant	PSNR↑	SSIM↑
full-model	**24.9253**	0.8594
standard cross-stream interaction	24.0990	0.8490
full-band only	24.8089	0.8540

Cross-architecture transferability of DS-LIF. As shown in Table 7, to evaluate cross-architecture transferability, we insert DS-LIF into RetinexFormer and compare it with *standard LIF* and *continuous activation* under identical surrounding architectures and training settings. DS-LIF outperforming the two alternatives by 0.2258 dB and 0.1978 dB, respectively. It indicates that the performance gain brought by DS-LIF is not dependent on the specific backbone network of SpikeRelight.

Table 7: DS-LIF transfer to RetinexFormer.

Variant	PSNR↑	SSIM↑
DS-LIF	**23.1879**	**0.8190**
continuous activation	22.9901	0.8190
standard LIF	22.9621	0.8100

5 Conclusion

This paper addresses the challenge of extracting reliable restoration information from low-light images, where genuine responses are often indistinguishable from noise, hindering the differentiation of dark regions. We develop a reliability-aware DS-LIF for LLIE. It scales the input current before membrane integration based on illumination and color contrast, while dynamically masking less contributive channels. It integrates the mean spike response with prior information from the input to generate a reliability map. Based on this neuron, we design a dual-stream enhancement architecture, SpikeRelight. Within this architecture, RGDFI uses the map to modulate asymmetric, frequency-dependent interactions between intensity and chrominance. The low-frequency branch captures smooth illumination and color variations, while the high-frequency branch applies reliability-based gating to retain details and suppress noise. Experiments across benchmarks show that SpikeRelight achieves competitive enhancement performance. Ablation studies confirm the effectiveness of DS-LIF and RGDFI, and transfer experiments on RetinexFormer further demonstrate the transferability of DS-LIF to other image restoration architectures. Although SpikeRelight is evaluated on diverse paired and unpaired low-light benchmarks, further validation under mixed degradations and highly uncontrolled illumination conditions remains an important direction for future work.

REFERENCES

Alexandru Brateanu, Tingting Mu, Codruta O Ancuti, and Cosmin Ancuti. Multinex: Lightweight low-light image enhancement via multi-prior retinex. In *2026 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 29887–29896, 2026.

Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *2011 Proceedings of the IEEE/CVF Confer- ence on Computer Vision and Pattern Recognition (CVPR)*, pp. 97–104. IEEE, 2011.

Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *2018 IEEE Transactions on Image Processing (TIP)*, 27(4):2049–2062, 2018.

Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *2023 IEEE/CVF international conference on computer vision (ICCV)*, pp. 12470–12479. IEEE, 2023.

Andrea Castagnetti, Alain Pegatoquet, and Benoît Miramond. Spiden: deep spiking neural networks for efficient image denoising. *Frontiers in Neuroscience*, 17:1224457, 2023.

Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3291–3300. Ieee, 2018.

Shuang Chen, Tom Krajnik, Farshad Arvin, and Amir Atapour-Abarghouei. Exploring the potentials of spiking neural networks for image deraining. In *2026 Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 3029–3037, 2026.

Wei Dong, Yan Min, Han Zhou, and Jun 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)*, pp. 1454–1461. IEEE, 2025.

Daniel Feijoo, Juan C Benito, Alvaro Garcia, and Marcos V Conde. Darkir: Robust low-light image restoration. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10879–10889. IEEE, 2025.

Zhenqi Fu, Yan Yang, Xiaotong Tu, Yue Huang, Xinghao Ding, and Kai-Kuang 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)*, pp. 22252–22261. IEEE, 2023.

C Guo et al. Deepupe: learning deep underexposed photo enhancement. In *2019 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1777–1786. Ieee, 2020.

Xiaojie Guo and Qiming Hu. Low-light image enhancement via breaking down the darkness. *2023 International Journal of Computer Vision (IJCV)*, 131(1):48–66, 2023.

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

Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th international conference on pattern recognition*, pp. 2366–2369. IEEE, 2010.

Xinyu Jia, Chuang Zhu, Minzhen Li, Wenqi Tang, and Wenli Zhou. Llvip: A visible-infrared paired dataset for low-light vision. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 3489–3497. IEEE, 2021.

Hai Jiang, Binhao Guan, Zhen Liu, Xiaohong Liu, Jian Yu, Zheng Liu, Songchen Han, and Shuaicheng Liu. Learning to see in the extremely dark. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7676–7685. IEEE, 2025.

Shuai Jin, Yuhua Qian, Feijiang Li, Guoqing Liu, and Xinyan Liang. Pasd: A pixel-adaptive swarm dynamics approach for unsupervised low-light image enhancement. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 9070–9079. IEEE, 2025.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *2012 Advances in Neural Information Processing Systems*, 25, 2012.

Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *2013 IEEE Transactions on Image Processing (TIP)*, 22(12): 5372–5384, 2013.

Chongyi Li, Chunle Guo, Linghao Han, Jun Jiang, Ming-Ming Cheng, Jinwei Gu, and Chen Change Loy. Low-light image and video enhancement using deep learning: A survey. *2021 IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 44(12):9396–9416, 2021.

Boheng Liu, Ziyu Li, Zhong Zhang, Mengrui Xu, Chenghua Duan, Dehao Liu, Qing Li, and Xia Wu. Rodnet: Visual pathway-inspired adaptive sparse network for efficient low-light image enhancement. In *2026 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4961–4970, 2026.

Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan 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)*, pp. 10556–10565. IEEE, 2021.

Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *2015 IEEE Transactions on Image Processing (TIP)*, 24(11):3345–3356, 2015.

Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *2012 IEEE Transactions on Image Processing (TIP)*, 21(12): 4695–4708, 2012a.

Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Signal processing letters*, 20(3):209–212, 2012b.

Yang Ouyang, Zihan Cheng, Xiaotong Luo, Guoqi Li, and Yanyun Qu. Spikingir: A novel converted spiking neural network for efficient image restoration. In *2026 Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 40, pp. 8215–8223, 2026.

Christos Sakaridis, Haoran Wang, Ke Li, Arpit Jadon, Wim Abbeloos, Daniel Olmeda Reino, Luc Van Gool, Dengxin Dai, et al. Acdc: The adverse conditions dataset with correspondences for robust semantic driving scene perception. *2025 IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2025.

Kangbiao Shi, Yixu Feng, Tao Hu, Yu Cao, Peng Wu, Yijin Liang, Yanning Zhang, and Qingsen Yan. Fusionnet: Multi-model linear fusion framework for low-light image enhancement. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1019–1028. IEEE, 2025.

Tianyu Song, Guiyue Jin, Pengpeng Li, Kui Jiang, Xiang Chen, and Jiyu Jin. Learning a spiking neural network for efficient image deraining. *arXiv preprint arXiv:2405.06277*, 2024.

Vassilios Vonikakis, Rigas Kouskouridas, and Antonios Gasteratos. On the evaluation of illumination compensation algorithms. *Multimedia Tools and Applications*, 77(8):9211–9231, 2018.

Hongzhi Wang, Xiubo Liang, Jinxing Han, and Weidong Geng. Spike-retinexformer: Rethinking low-light image enhancement with spiking neural networks. *2026 Advances in Neural Information Processing Systems (NIPS)*, 38:11425–11444, 2026.

Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *2013 IEEE Transactions on Image Processing (TIP)*, 22(9): 3538–3548, 2013.

Wenjing Wang, Huan Yang, Jianlong Fu, and Jiaying Liu. Zero-reference low-light enhancement via physical quadruple priors. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 26057–26066. IEEE, 2024.

Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *2022 Proceedings of the AAAI conference on artificial intelligence (AAAI)*, volume 36, pp. 2604–2612, 2022.

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

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

Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5891–5900. IEEE, 2022.

Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pp. 17693–17703. IEEE, 2022.

Qingsen Yan, Yixu Feng, Cheng Zhang, Guansong Pang, Kangbiao Shi, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. Hvi: A new color space for low-light image enhancement. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5678–5687. IEEE, 2025.

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

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *2020 European conference on computer vision (ECCV)*, pp. 492–511. Springer, 2020.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *2022 Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5728–5739, 2022.

Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 586–595. IEEE, 2018.

A Reproducibility Details - DS-LIF: Novel Spiking Neuron for Reliable Low-Light Image Enhancement

A.1 Overall Architecture

Given a low-light RGB input image $\mathbf{x} \in \mathbb{R}^{3\times H\times W}$, SpikeRelight first transforms it into an intensity component and a chrominance component using the HVI transformation:

$$(\mathbf{I}, \mathbf{C}) = \mathrm{HVIT}(\mathbf{x}), \qquad \mathbf{C} = [\mathbf{H}, \mathbf{V}]. \tag{25}$$

Here, $\mathbf{I}$ denotes the intensity information, whereas $\mathbf{C}$ denotes the chrominance information. This decomposition allows the network to model illumination and color variations separately while preserving their dependencies through cross-stream interaction.

The two HVI components are then projected into shallow features:

$$\mathbf{F}_I^0 = \phi_I(\mathbf{I}), \qquad \mathbf{F}_C^0 = \phi_C(\mathbf{C}), \tag{26}$$

where ϕ_I and ϕ_C denote the shallow feature projections of the two streams. The resulting intensity and chrominance streams are processed by separate encoder–decoder paths. The encoders progressively downsample the features to enlarge the receptive field, while the decoders restore the spatial resolution through upsampling and preserve fine-grained structures through skip connections.

At the k-th network stage, the two feature streams are denoted by $\mathbf{F}_I^k$ and $\mathbf{F}_C^k$. DS-LIF processes the intermediate features over multiple timesteps and estimates restoration reliability from their temporal spiking responses. For the two feature streams, DS-LIF produces the reliability maps $\mathbf{R}_I^k$ and $\mathbf{R}_C^k$. These reliability maps are then passed to RGDFI to control cross-stream feature updates.

RGDFI contains three interaction paths: a full-band path, a low-frequency path, and a high-frequency path. The full-band path first performs cross-stream attention directly on the undecomposed intensity and chrominance features. For the intensity stream, $\mathbf{F}_I^k$ serves as the query, while $\mathbf{F}_C^k$ provides the keys and values. The reverse direction is used to update the chrominance stream. This directional interaction enables information exchange between the two streams without reducing them to a simple symmetric fusion operation. Let the resulting updates be denoted by $\mathbf{U}_I^{\text{full}}$ and $\mathbf{U}_C^{\text{full}}$. They are regulated by the reliability maps of the corresponding source streams:

$$\widetilde{\mathbf{U}}_I^{\text{full}} = \mathbf{R}_C^k \odot \mathbf{U}_I^{\text{full}}, \qquad \widetilde{\mathbf{U}}_C^{\text{full}} = \mathbf{R}_I^k \odot \mathbf{U}_C^{\text{full}}.$$

To distinguish the roles of different frequency components, RGDFI decomposes the two feature streams using a low-pass operator $\mathcal{L}(\cdot)$:

$$\mathbf{F}^{\text{low}} = \mathcal{L}(\mathbf{F}), \qquad \mathbf{F}^{\text{high}} = \mathbf{F} - \mathbf{F}^{\text{low}}. \tag{27}$$

Accordingly, the two feature streams are decomposed as

$$\mathbf{F}_I^k \to \left(\mathbf{F}_I^{k,\text{low}}, \mathbf{F}_I^{k,\text{high}}\right), \qquad \mathbf{F}_C^k \to \left(\mathbf{F}_C^{k,\text{low}}, \mathbf{F}_C^{k,\text{high}}\right).$$

The low-frequency path performs cross-stream interaction on the smooth components of the two streams. The low-frequency intensity and chrominance features exchange information in both directions to model smooth illumination changes and large-scale chrominance variations. Because low-frequency information is spatially coherent, this path does not use explicit reliability gating, thereby preserving the continuous propagation required for global brightness and color correction.

The high-frequency path processes features containing edges, textures, and local structures. However, high-frequency responses may also contain sensor noise and unstable responses in dark regions. Therefore, the high-frequency interaction outputs are regulated by the reliability estimated by DS-LIF:

$$\widetilde{\mathbf{U}}_I^{\text{high}} = \mathbf{R}_C^k \odot \mathbf{U}_I^{\text{high}}, \qquad \widetilde{\mathbf{U}}_C^{\text{high}} = \mathbf{R}_I^k \odot \mathbf{U}_C^{\text{high}}.$$

Here, $\mathbf{U}_I^{\text{high}}$ and $\mathbf{U}_C^{\text{high}}$ denote the high-frequency cross-stream updates in the two directions. This operation preserves the propagation of reliable edges and textures while reducing the influence of noise-dominated or unstable responses on the other stream.

The outputs of the three paths are then aggregated to obtain the updated intensity and chrominance features:

$$\begin{aligned} \mathbf{F}_I^{\text{out}} &= \mathbf{F}_I^k + \widetilde{\mathbf{U}}_I^{\text{full}} + \tanh(\gamma_{\text{low}})\mathbf{U}_I^{\text{low}} + \tanh(\gamma_{\text{high}})\widetilde{\mathbf{U}}_I^{\text{high}}, \\ \mathbf{F}_C^{\text{out}} &= \mathbf{F}_C^k + \widetilde{\mathbf{U}}_C^{\text{full}} + \tanh(\gamma_{\text{low}})\mathbf{U}_C^{\text{low}} + \tanh(\gamma_{\text{high}})\widetilde{\mathbf{U}}_C^{\text{high}}, \end{aligned} \tag{28}$$

where γ_{low} and γ_{high} are learnable scalar parameters. The aggregated features are then passed to the next network stage or to the subsequent decoder stage.

After multiple stages of encoding, RGDFI interaction, and decoding, the network predicts an HVI residual $\Delta\mathbf{HVI}$, which is added to the HVI representation of the input image:

$$\widehat{\mathbf{hvi}} = \text{HVIT}(\mathbf{x}) + \Delta\mathbf{HVI}. \tag{29}$$

Finally, the enhanced RGB image is generated through the inverse HVI transformation:

$$\widehat{\mathbf{y}} = \text{PHVIT}\left(\widehat{\mathbf{hvi}}\right). \tag{30}$$

Table 8: Default training configuration.

Configuration	Setting
GPU	$1 \times$ NVIDIA GeForce RTX 4090
Framework	PyTorch
Optimizer	Adam
Initial learning rate	1×10^{-4}
Learning-rate schedule	3-epoch linear warm-up + cosine annealing
Batch size	8
Gradient accumulation	1
Training crop	256×256
Input range	$[0, 1]$
Data augmentation	random crop, horizontal flip, vertical flip
Data-loader workers	16
Random seed	2026
Deterministic execution	enabled when supported
Gradient clipping	enabled, maximum norm 0.01
Exponential moving average	not used
Validation frequency	every epoch

Overall, SpikeRelight first decomposes the input RGB image into intensity and chrominance streams. DS-LIF then processes the intermediate features over time and estimates restoration reliability. RGDFI converts this reliability into reliability-guided interaction for the full-band and high-frequency features, while its low-frequency path preserves coherent information exchange for smooth illumination and chrominance variations. After subsequent decoding and residual reconstruction, the enhanced HVI representation is transformed back into the final RGB image through the inverse HVI transformation.

B Reproducibility Details

B.1 Dataset Protocol

All experiments follow the official training and testing protocols of the corresponding datasets. For the unpaired evaluation, we use the following image directories:

$$\mathcal{D}_{\text{unpaired}} = \{\texttt{DICM}, \texttt{LIME}, \texttt{MEF}, \texttt{NPE}, \texttt{VV}\}.$$

These datasets do not provide aligned ground-truth images and are therefore evaluated using no-reference image-quality metrics and qualitative comparisons.

B.2 Training Configuration

The main training configuration is summarized in Table 8. Unless otherwise specified, the same configuration is used for all paired datasets.

The default launcher trains LOLv1, LOLv2-Real, LOLv2-Synthetic, SID, SICE, and FiveK for 1000 epochs. The validation batch size is one. Training and validation images are converted to tensors without additional dataset-specific normalization. For the LOLv2-Synthetic protocol, full-resolution training can be enabled through the corresponding launcher option; otherwise, the default implementation uses the same 256×256 random-crop protocol as the other paired datasets.

B.3 Loss Weights

The reconstruction objective is applied in both the RGB and HVI domains. The loss weights used by the implementation are listed in Table 9.

Table 9: Loss configuration used in the reported experiments.

Loss component	Weight
RGB L_1	1.0
HVI L_1	1.0
RGB/HVI SSIM loss	0.5
RGB/HVI edge loss	50.0
RGB/HVI perceptual loss	0.01
Reliability calibration on LOLv1/LOLv2-Real	0.10
Reliability calibration on other datasets	0.00
Calibration warm-up	10 epochs
Need-region weight	2.0
No-need-region weight	3.0
No-need suppression weight	1.0
No-need suppression margin	0.05
Reliability ranking weight	0.50
Reliability ranking margin	0.15
SNN-route calibration weight	1.0
Input-prior mixing ratio	0.70

The RGB and HVI reconstruction losses are computed with the same combination of L_1, SSIM, edge, and perceptual terms. Ground-truth reliability calibration is enabled for LOLv1 and LOLv2-Real, where real-camera noise is explicitly considered by the reliability front-end. It is disabled for LOLv2-Synthetic, SICE, SID, and FiveK unless explicitly overridden.

For completeness, the implementation monitors the DS-LIF spike rate within the interval $[0.05, 0.20]$. In the default reported launcher configuration, the spike-rate penalty weight is set to zero; therefore, this interval is used as a diagnostic operating contract rather than as an additional gradient-based optimization term. The measured spike rate and channel-sleep rate are reported separately in the efficiency analysis.

B.4 DS-LIF and RGDFI Configuration

The architectural settings of the proposed method are summarized in Table 10.

For LOLv1 and LOLv2-Real, the noise-aware reliability path is enabled and the effective chroma-gate floor is set to 0.20. For the synthetic and additional paired datasets, the default noise-aware path is disabled and the chroma-gate floor is set to zero. These settings are controlled through the dataset-specific launcher configuration.

B.5 Evaluation and Checkpoint Selection

During validation, the model is evaluated with a batch size of one and without gradient computation. PSNR and SSIM are computed after every training epoch using the generated RGB images and the corresponding reference images. No ground-truth mean correction is applied during evaluation. The checkpoint with the highest PSNR is selected as the primary checkpoint, while the checkpoint with the highest SSIM is retained as a secondary checkpoint.

LPIPS is computed only after training on the selected best-PSNR and best-SSIM checkpoints. This avoids using LPIPS as a frequent checkpoint selection criterion and prevents its relatively high computational cost from being introduced into every validation epoch.

For LOLv1, the standard HVI conversion path is used. For LOLv2-Real, the official gated HVI rendering path is used with an output scaling factor of $\alpha = 0.8$. SICE and FiveK preserve the original image dimensions during evaluation, and generated outputs are cropped back to the original spatial size after inference when necessary.

Table 10: Architectural configuration of DS-LIF and RGDFI.

Component	Setting
HVI representation	intensity + two chromatic components
DS-LIF internal steps T	4
Input current scale ζ	10.0
Membrane time constant τ	2.0
Firing threshold ϑ	1.0
Hidden channels	16
Maximum intensity residual	0.12
Sleep strength	0.75
Threshold-gain factor	1.0
Channel-gate temperature	0.1
Minimum kept-channel ratio	0.25
Reliability controller	Dynamic-Sleep LIF
Reliability scope	full
Chroma reliability gate	enabled
Noise-adaptive SNN route	enabled
SNN route floor	0.05
Dual-frequency interaction	enabled at stages 1–6
Anti-aliased reliability sampling	enabled
Sampling kernel	5×5
Detail-preservation floor	0.25
Adaptive chroma refinement	disabled

For the unpaired datasets DICM, LIME, MEF, NPE, and VV, no-reference evaluation is performed using NIQE and BRISQUE. Lower values are preferred for both metrics. Since no aligned target is available, these results are reported together with qualitative visual comparisons rather than PSNR, SSIM, or LPIPS.

B.6 Ablation Reproducibility

As shown in Table 11, all ablation experiments use the same dataset split, random seed, crop size, batch size, optimizer, learning-rate schedule, and number of training epochs as the full configuration. Only the component specified by the ablation name is changed.

The DS-LIF transfer experiment on RetinexFormer keeps the RetinexFormer backbone fixed and replaces only the corresponding reliability controller. The same temporal length and controller capacity are used for the DS-LIF, conventional-LIF, and matched-ANN variants. This design isolates the effect of the proposed neuron dynamics and tests whether DS-LIF is transferable beyond the SpikeRelight framework.

For all methods, the parameter count and computational complexity are reported for a 256×256 input. MACs are measured using the corresponding official or reproduced implementation, and FLOPs are estimated as twice the number of MACs. For SpikeRelight, the reported dense MACs/FLOPs represent an upper bound of the event-driven execution; the additional temporal behavior of DS-LIF is analyzed separately using spike-rate and channel-sleep statistics.

B.7 Results on Individual Unpaired Datasets

The main paper reports the average NIQE and BRISQUE scores of SpikeRelight over five unpaired low-light image enhancement datasets, including DICM, LIME, MEF, NPE, and VV. To provide a more detailed evaluation, Table 12 reports the dataset-specific results. SpikeRelight achieves consistently low no-reference image-quality scores across the five datasets. The last row gives the arithmetic mean of the five per-dataset results, which is consistent with the average values reported in the main paper, up to rounding.

Table 11: Parameter and computational complexity of the compared methods. MACs and FLOPs are reported in G for a 256×256 input.

Method	Params/M	MACs/FLOPs/G
Retinexformer	1.53	15.85
SNR-Aware	4.27	27.87
Bread	2.28	20.57
PairLIE	0.43	21.24
QuadPrior	1252.75	1103.20
DarkIR	3.51	7.84
MultiNex	0.0446	2.50
CIDNet	1.98	7.98/15.95
SpikeRelight (Ours)	2.58	9.78/19.56

Table 12: Dataset-specific evaluation of SpikeRelight on five unpaired low-light image enhancement datasets. Lower NIQE and BRISQUE scores indicate better perceptual quality. The last row reports the average across all five datasets.

Dataset	NIQE↓	BRISQUE↓
DICM	4.0138	20.9895
LIME	4.2245	19.9946
MEF	3.9966	23.0683
NPE	4.0298	16.6174
VV	3.1265	24.2399
Average	3.8782	20.9819