



OPEN ACCESS

EDITED BY
Manning Wang,
Fudan University, China

REVIEWED BY
Jinxing Liang,
Wuhan Textile University, China
Kexue Fu,
Shandong Academy of Sciences, China

*CORRESPONDENCE
Qiang Wang
✉ caption_wang@21cn.com

RECEIVED 20 September 2023
ACCEPTED 02 February 2024
PUBLISHED 05 March 2024

CITATION
Wang X, Wang Q, Zhang L, Qu Y, Yi F, Yu J,
Liu Q, Xia R, Xu Z and Tong S (2024) DCENet-
based low-light image enhancement
improved by spiking encoding and convLSTM.
Front. Neurosci. 18:1297671.
doi: 10.3389/fnins.2024.1297671

COPYRIGHT
© 2024 Wang, Wang, Zhang, Qu, Yi, Yu, Liu,
Xia, Xu and Tong. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

DCENet-based low-light image enhancement improved by spiking encoding and convLSTM

Xinghao Wang, Qiang Wang*, Lei Zhang, Yi Qu, Fan Yi,
Jiayang Yu, Qiuhan Liu, Ruicong Xia, Ziling Xu and Sirong Tong

Equipment Management and Unmanned Aerial Vehicle Engineering School, Air Force Engineering University, Xi'an, China

The direct utilization of low-light images hinders downstream visual tasks. Traditional low-light image enhancement (LLIE) methods, such as Retinex-based networks, require image pairs. A spiking-coding methodology called intensity-to-latency has been used to gradually acquire the structural characteristics of an image. convLSTM has been used to connect the features. This study introduces a simplified DCENet to achieve unsupervised LLIE as well as the spiking coding mode of a spiking neural network. It also applies the comprehensive coding features of convLSTM to improve the subjective and objective effects of LLIE. In the ablation experiment for the proposed structure, the convLSTM structure was replaced by a convolutional neural network, and the classical CBAM attention was introduced for comparison. Five objective evaluation metrics were compared with nine LLIE methods that currently exhibit strong comprehensive performance, with PSNR, SSIM, MSE, UQI, and VIFP exceeding the second place at 4.4% (0.8%), 3.9% (17.2%), 0% (15%), 0.1% (0.2%), and 4.3% (0.9%) on the LOL and SCIE datasets. Further experiments of the user study in five non-reference datasets were conducted to subjectively evaluate the effects depicted in the images. These experiments verified the remarkable performance of the proposed method.

KEYWORDS

intensity-to-latency, spiking encoding, low-light enhancement, unpaired image, deep learning

1 Introduction

The lack of illumination leads to the loss of image information, which severely affects the execution of visual tasks, e.g., face recognition, object detection, dataset preparation, and autonomous driving (Li J. et al., 2021; Liu et al., 2021; Tang et al., 2022; Guo et al., 2023). Capturing images in low-light conditions poses a challenge owing to the limited aperture size, demand for instantaneous processing, and limited memory resources. To mitigate the issues of structuring and the high expense of research and development associated with hardware, refining images in low-light settings through minimalistic software algorithms aligns better with predictable requirements.

In low-light image enhancement (LLIE), the first effective methods were based on histogram equalization, the Retinex model, gamma transform, and fusion. Fusion-based methods achieve better performance in terms of image indicators, such as brightness and color, through exposure-splicing fusion methods. This method is typically synthesized by collecting images under different exposure conditions (Wang et al., 2016). Another method

fuses the illumination map of night and day to enhance the image (Rao et al., 2010); however, such processing generally renders a poor visual effect.

The method based on the Retinex model divides the low illumination image into reflection and illumination components or adds a noise component by constructing a suboptimal problem. The estimated reflection component is considered the result of enhancement. Previous attempts to improve Retinex replaced the logarithmic solution with a typical enhanced Lagrange solver to enhance the image with a long image processing time. However, the variational optimization algorithm has a high computational cost. Moreover, it introduces unnecessary pseudo-details in the image.

The adaptive GAMMA transform can improve an image's contrast; however, most algorithms of this class still cause local overexposure or underexposure in the enhanced result. As most images are captured in non-uniform lighting conditions, Chen et al. (2022) proposed a naturalness- and information-preserving method for processing them. The MEMBHE algorithm (Dar and Mittal, 2020) improved the functionality of the transform through histogram equalization after multiple exposure smoothing. Nevertheless, it overconsumes memory and requires arduous incremental updates.

Several methods for achieving LLIE with deep learning (DL) have been researched. Among them, supervised learning, a mature and informative DL method typically constructed by an end-to-end network, was the first to be applied to an LLIE field. Low-light net (LLNet) (Lore et al., 2017) was the first end-to-end LLIE network established by constructing a deep auto-encoder structure. MBLLEN (Lv et al., 2018) uses three subnetworks to extract rich image features of different levels and introduces a regional loss function into the network loss function to employ different loss weights for high- and low-light regions. In the same vein, Li et al. (2021) determined that enhancing the low-frequency layer of a low-light image with noise was easier than directly enhancing the whole image. Progressive recursive networks (Cai et al., 2018) were used to perform staging, which is a more efficient method for preserving image details and removing noise. In that method, each subnetwork could better achieve its own function, which was eventually enhanced by gradually improving the quality of the image.

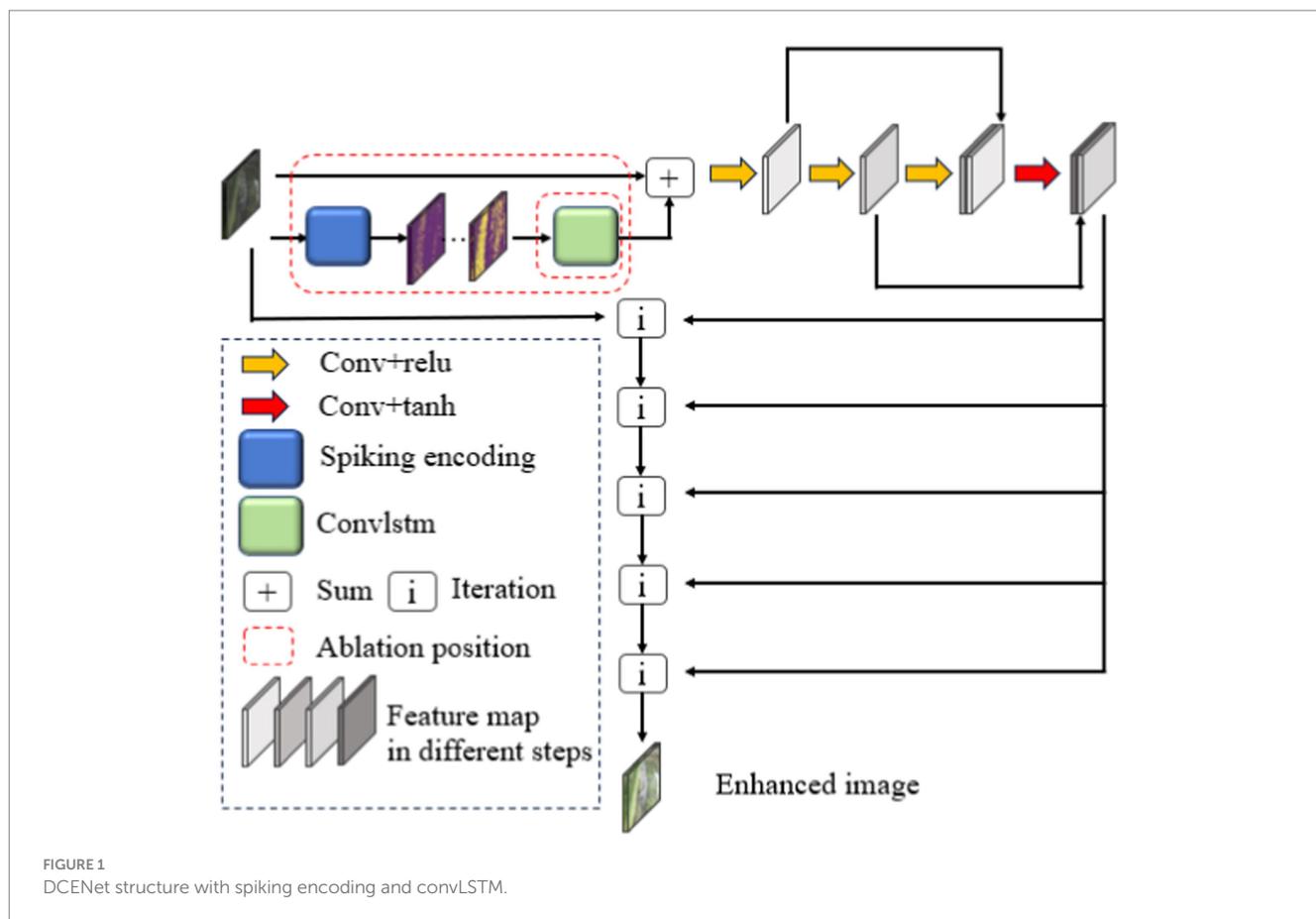
Ke et al. (2020) established an SCIE multiexposure dataset (Ke et al., 2020) consisting of low-contrast images with different exposure levels and their corresponding high-quality reference images. Furthermore, they introduced the high- and low-frequency components of images as prediction targets. A double-exposure fusion algorithm (Ying et al., 2017) was proposed to design the weight matrix of image fusion using an illuminance-estimation technique. Then, a camera response model was introduced to synthesize the multiexposure images. Low- and high-exposure images can also be used to estimate the perceptual gain, signal strength, signal structure, and mean intensity. Perceptual gain suits an underexposed image. The feature fusion and recalibration module (FFRM) (Singh et al., 2024) was proposed to recalibrate and merge the features to provide an enhanced output image. Intrinsic image decomposition (Zhang and Ma, 2023) can be applied to the fusion of multiexposure to generate HDR images.

Retinex was combined with DL for enhanced performance (Chen et al., 2018; Zhang Y. et al., 2019; Tang et al., 2023). The attention mechanism was combined with the Retinex model to construct DL

networks for enhancement (Chen et al., 2022). A decomposition network (Liu et al., 2023) was developed with a self-supervised fine-tuning strategy that achieved promising performance without manual hyperparameter tuning. Different sensitivities relate to different regions. The low-rank regularized Retinex model (Bao et al., 2022) can represent the image as low-rank decomposition, preserve the image details and high-frequency information, and improve the visual quality of the image. A plug-and-play framework for image enhancement and noise removal based on the Retinex theory (Wu et al., 2023) was introduced. Inspired by guided filtering and using synthetic data for network training, Li et al. (2018) designed a lightweight network architecture based on the Retinex theory. By including the unsettling V channel image component in the HSV color space, the component was converted to a reflection component using a DL network (Jiang Z. et al., 2021). Owing to their significant worth, their Retinex and DL-based methods were applied in image dehazing and underwater image enhancement (Xu et al., 2022; Shen et al., 2023).

The development of LLIE in DL is not limited. Creative thinking models, such as those based on unsupervised learning, represented by the unsupervised learning method (Zhu et al., 2020; Li et al., 2021), generative network architecture (Jiang Z. et al., 2021), and normalizing flow (Wang et al., 2022), show the immense research potential of LLIE. The strategy network learns the local exposure sequentially using reinforcement learning for a segmented subimage (Rong et al., 2018). In the generated adversarial network architecture, global-local discriminators (Jiang Z. et al., 2021) were used to ensure that the enhanced results resemble real normal light images. With the strong capability of image generation, diffusion models were applied to LLIE. For example, the pyramid diffusion model (Zhou et al., 2023) was constructed to solve the RGB shift. Moreover, the inference speed of the diffusion model was accelerated. As a scientific structure for image feature extraction, transformers have become some of the most prevalent network structures in vision processing. The regional distributions have been effectively managed, and the histogram loss has been designed in a stage transformer-guided network (Jiang et al., 2023). Half-wavelet attention block and hierarchical M-Net were utilized to improve computation consumption and reserve context information, aided by the DAU block and discrete wavelet transformation (Fan et al., 2022).

Spiking neural networks (SNNs) are frequently employed in numerous pixel-level classification tasks (Martinez-Seras et al., 2023), such as object detection (Zhang et al., 2023b), image segmentation (Zhang et al., 2023a), and anomaly detection (Yusob et al., 2018). Research centered on SNNs includes methods for neural network learning, data coding, and hardware platforms. The learning approaches for SNNs can be divided into supervised and unsupervised learning, which are represented by spike-timing-dependent plasticity (STDP). Spiking encoding, which involves utilizing discrete pulsed signals to convey information, is a method of signal transmission. Neuroscience computing has access to specialized offline or online application-specific integrated circuit platforms, as well as neuromorphic computing cores that can support various learning rules and neuronal models. Nonetheless, spiking neural network research continues to confront significant barriers. The training of the transformed SNN still relies on the backpropagation algorithm of artificial neural networks (ANNs). As



the performance difference between the SNN and the core ANN is small, the former cannot provide significant advantages. Moreover, generative tasks, such as LLIE, image patching, multimodal image generation, and network deployment, present significant challenges. As a new neural network structure, the SNN's internal algorithm can be implemented in LLIE.

The main contributions of this study are as follows: (1) According to the progressive output results with the specified number concluding the embodiment of the image structure characteristics, the application of the SNN in a spiking encoding method for LLIE has distinct advantages in extracting structural features from images (the intensity-to-latency encoding outputs multiple feature maps with structure and specified steps); and (2) a convLSTM structure that can better absorb the features from multiple feature maps. Based on unlabeled, unsupervised, and unpaired image training via simplified DCENet, the proposed structure is improved by spiking encoding and the convLSTM module. The research introduces spiking encoding, which concludes the image's backbone information to describe the hierarchical information. The rest of the paper follows this structure: Section 2 describes the proposed enhancement method. Section 3 describes the user study and ablation experiment carried out in the study and compares the performance of the proposed method with the state-of-the-art network structure based on seven objective indicators. Section 4 concludes the study and discusses the potential applications.

2 Proposed method

2.1 DCENet structure

The DCENet structure, as the primary structure used for unsupervised enhancement, divides the LLIE into a high-order iterative process, i.e., the input dark light image is finally enhanced through several iterations of the same operation. Figure 1 depicts the overall enhancement process and part of the ablation study, which can also have a description in literal form. The input passes through the spiking encoding module and ConvLSTM described in subheadings 2.2 and 2.3, respectively, and then through the convolution module containing skip links. The sum module in Figure 1 means a direct overlay between the ConvLSTM's final output and the input dark light images. The resulting features select the feature graph of a certain channel in order and combine the matrix of the same size in the length and width scale of the output and input images with the initial input tensor according to Equation (1). The matrix is used as the input for the next iteration, and the feature graph of the next channel is selected as needed for the next iteration operation.

Compared with the mathematical relationship represented by the previous gamma transform, the DCENet structure changes the training coefficient of the second term of the right-hand side of Equation (1) into a training coefficient matrix with the same dimensions as those of the

input image. This can restrain the problem of over-enhancement or under-enhancement of the image to a certain extent. Finally, the normal brightness area in the image is maintained, and the low illumination area is restored. A is the output of the network, which can be divided into several pieces denoted by A_n . Based on the number of iterations n , the final output enhancement result is x_n .

$$x_n = x_{n-1} + A_n(x_{n-1}^2 - x_{n-1}) \tag{1}$$

2.2 Spiking encoding method

In this study, spiking encoding from the overall DCENet structure equals the intensity-to-latency transform (Mozafari and Ganjtabesh, 2019), as illustrated in Figure 2. First, the intensity-to-latency transform requires an initial parameter, i.e., time step S . Then, the grayscale image, which corresponds to a matrix with shape (H, W) , is reshaped to a vector with $H \times W$ dimensions by $R(\cdot)$ as illustrated in Equation (2). We named this original vector V . For the next step, the vector was arranged in descending order. This procedure generated two vectors with the same dimensions: the first vector is the descending order vector V_d , while the second one is the index vector V_i , corresponding to the index in V and this relation is represented by Equation (3).

$$V[1, H * W] = R(I[H, W]) \tag{2}$$

$$V_d, V_i = D(V) \tag{3}$$

where K is the number of non-zero elements in an original vector V . The split parameter θ is set in Equation (4). V_d and V_i are split into small vector pieces; θ decides the shape of these pieces. The small vector piece returns to the dimension $H \times W$, which is called the spiking encoding vector T_m in Equation (5), with the complementary element filled with 0. The small label m ranges from 0 to the time step S . The start time step T_0 is composed of the value in the first split piece, and the value in T_0 is rearranged to the original position in V according to V_i . The second time step T_1 , which is based on T_0 , adds the second small piece, and the value in the second piece is adjusted to the original position in the same way. Thus, the intensity-to-latency transformation is complete. The sequence of outputs T_m is reshaped to similar dimensions as those of the input image, which are denoted by E_m , with the dimensions of (S, H, W) . This procedure is formulated as Equation (6). $R^*(\cdot)$ means the reverse calculation manipulation of $R(\cdot)$. (Considering the length of the paper, its time step in the figure is set to 6.)

$$\theta = \begin{cases} \frac{K}{S}, \frac{K}{S} > 2 \\ 2, \frac{K}{S} \leq 2 \end{cases} \tag{4}$$

$$T_m = split(V_d, \theta) \tag{5}$$

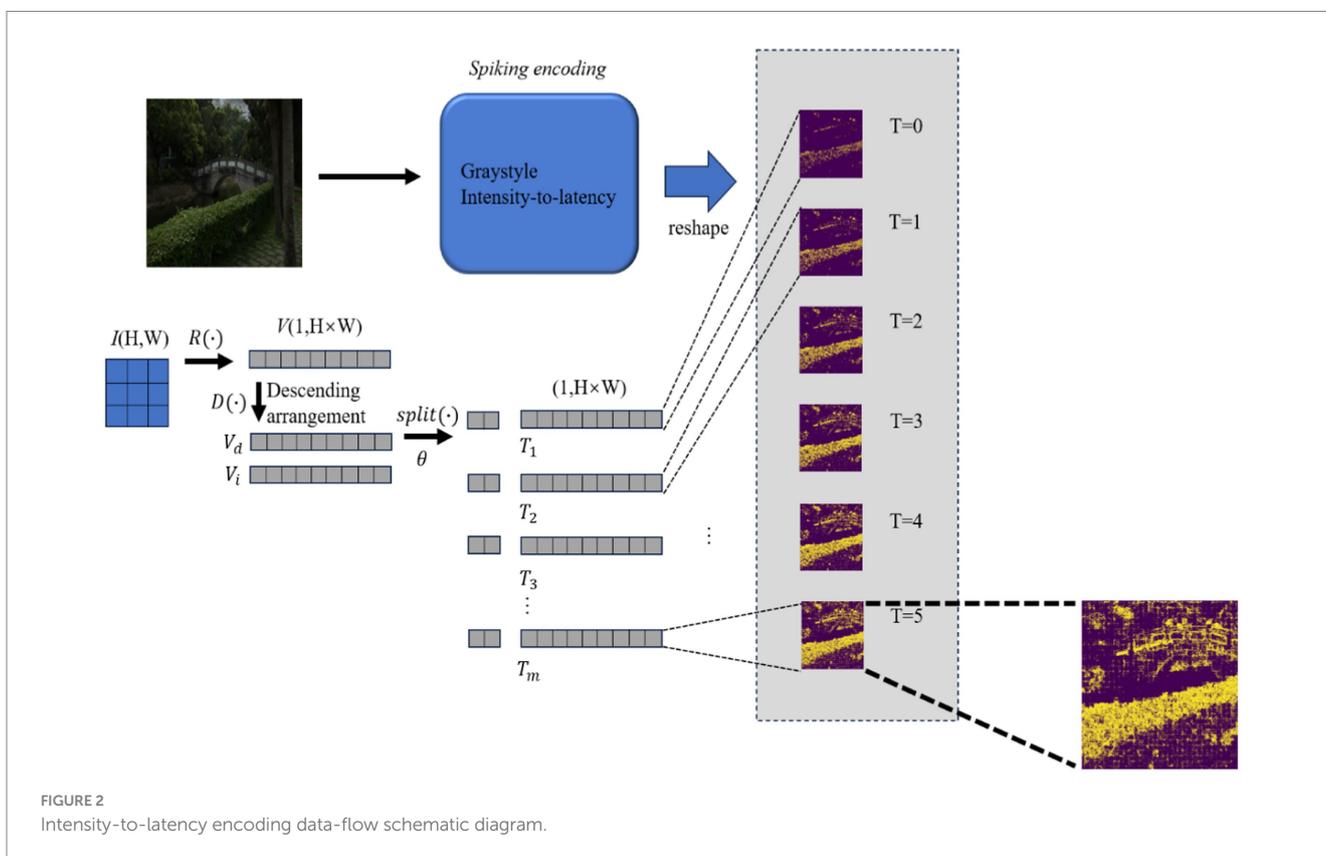


FIGURE 2 Intensity-to-latency encoding data-flow schematic diagram.

$$E_m = R^*(T_m, V_i) \tag{5}$$

2.3 ConvLSTM

The features extracted by the intensity-to-latency transform have certain similarities and differences. These features will constitute an image sequence with fluent features. The convLSTM structure is applicable in this scenario. ConvLSTM is proposed for precipitation nowcasting (Shi et al., 2015), the backbone of which is the recurrent neural network (RNN) for spatiotemporal prediction with convolutional structures. This design is convenient for video and image sequence-related tasks. ConvLSTM is similar to LSTM, which is also called FC-LSTM, and its block structure is illustrated in Figure 3.

The convLSTM computation method is based on LSTM's gate relationship. The distinction between its different layers is the input and output dimensions. The core of convLSTM is the convLSTM cell, which represents one convLSTM layer. ConvLSTM cell is an RNN-like structure; therefore, a specific hidden-layer parameter called *hidden state* is required. In every convLSTM layer, the hidden state is initialized with a zero element of dimensions (C1,H,W). One of the input sequence time step tensors *I*, which is another input of convLSTM and the outputs from spiking encoding, with dimensions (C,H,W), were concatenated with δ . It outputs a combined tensor with dimensions (C + C1,H,W), corresponding to the concatenate calculation represented by *concat* in Equation (6), which needs two different

variables. The convLSTM cell accepts this combined tensor and outputs the tensor with dimensions (4 × C1,H,W). The outputs were divided into four tensors with dimensions (C1,H,W) for the outputs of different gates: input, forget, and output gates, and a new δ for the subsequent layer and input time step. This divided single step is represented by the *split*. The calculation procedure is summarized in Equations 6–10 and Figure 2.

$$mid_1 = conv(concat(\delta_{t-1}, I_{t-1}), dim = 1) \tag{6}$$

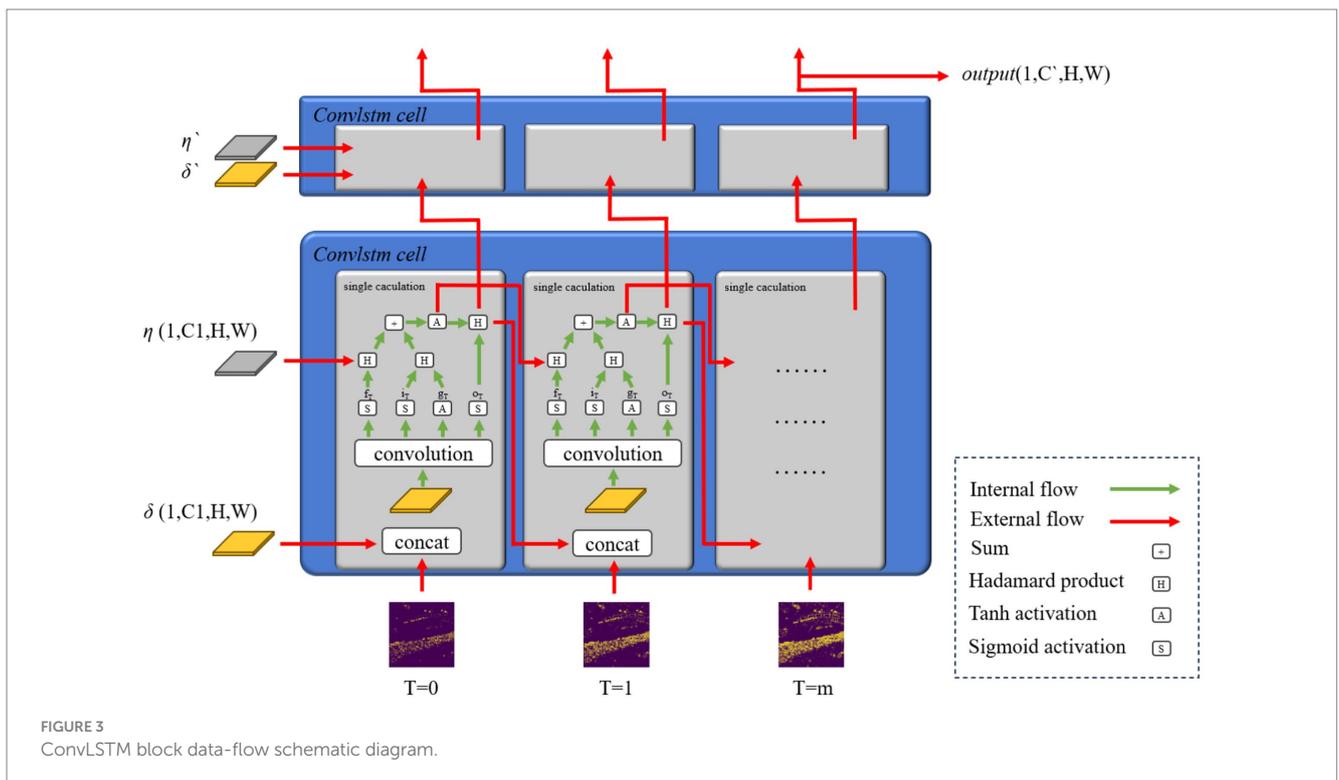
$$c_i, c_f, c_o, c_g = split(mid_1, hidden, dim = 1) \tag{7}$$

$$a_t = S(c_a), a = i, f, o; g_t = A(c_g) \tag{8}$$

$$\eta_t = f_t \circ \eta_{t-1} + i_t \circ g_t \tag{9}$$

$$\delta_t = o_t \circ A(\eta_t) \tag{10}$$

In the convLSTM structure, δ and η , which are output by one convLSTM cell, pass to the next cell at a certain time. This time



corresponds to the next time step in the same layer. This RNN-like network structure will preserve the main features from the previous time step image feature. The δ is also output to the convLSTM cell, combined with the new hidden states δ' and η' in the next layer in the same time step. The overall output of the convLSTM module is the tensor with dimensions $(1, C, H, W)$, which is labeled *output* in Figure 2. The time step dimension is eliminated with the convLSTM module and $S(\cdot)$, sigmoid activation function, and $A(\cdot)$, tanh activation function.

2.4 The loss items of the DCENet structure

Four loss items, namely spatial consistency loss, color constancy loss, exposure control loss, and illumination smoothness loss, were considered for the convergence of the network. The loss function used by the network is represented by Equation (12). The spatial consistency loss item was calculated by Equation (13). The purpose of setting the spatial consistency loss item was to maintain the difference between the original image and the adjacent area of a pixel in the enhanced image as small as possible. The \bar{X} represents the tensor X after channel averaging and average pooling for every 4×4 area. K is the number of pixels after average pooling in one feature map channel. These pixels are separated by a distance of 1, which corresponds to a point assemble called $R(i)$. This difference logic will enhance the pixel neighborhood within the same spatial structure. By introducing the sum item, the pixel neighborhood consistency can be promoted to the spatial position consistency of the whole image. The setting of this loss item will maintain the spatial consistency of the image before and after enhancement.

$$L = L_{spa} + L_{exp} + W_{col}L_{col} + W_{tv}L_{tv} \tag{12}$$

$$L_{spa} = \frac{1}{K} \sum_{i=1}^K \sum_{j \in R(i)} (|\bar{E}_i - \bar{E}_j| - |\bar{O}_i - \bar{O}_j|)^2 \tag{13}$$

To ensure the overall improvement in brightness, the exposure loss was established as Equation (14). The average value of pixels in the pixel block corresponding to the gray-level image of the output-enhanced image should meet certain size requirements, and the reference average value was set to 0.7. \bar{E}_m' represents the m^{th} pixel value after image channel mean processing and pooling for the enhancement of the final result. The pooling operation may have different parameters. Hence, quotes were added to distinguish it from the spatial consistency loss term. The number of pixels after pooling was set to M .

$$L_{exp} = \frac{1}{M} \sum_{m=1}^M (\bar{E}_m' - 0.7)^2 \tag{14}$$

The value of one color channel of the image should not significantly exceed that of the other channels. Hence, the loss of color was set to a constant value represented by L_{col} represented by Equation (15). This loss should go through all pairings in the three color channels. To better satisfy this condition, the spatial average of

the enhanced image is calculated, and a three-channel difference loss term was constructed to satisfy this conclusion. $(c1, c2)$ traverses all pairwise combinations in the three RGB color channels. \bar{E}^{c1} and \bar{E}^{c2} represent the enhancement result's mean value of one RGB channel.

$$L_{col} = \sum_{\forall (c1, c2) \in c} (\bar{E}^{c1} - \bar{E}^{c2})^2, c = \{(R, G), (G, B), (B, R)\} \tag{15}$$

Different from the final enhanced image result, A is the network output. In Equation (16), N , which equals to $H \times W$ dimensioned by A_n , represents the shape of the input. d represents the gradient of A ; for instance, A_{iy}^d relates to the longitudinal gradient of A in the i^{th} iteration. The illumination smooth loss L_{illu} was established here.

$$L_{illu} = \frac{1}{N} \sum_{i=0}^n \sum_d (A_{ix}^d + A_{iy}^d) \tag{16}$$

Considering that the brightness change between adjacent pixels is not significant, the gradient term was introduced to the network output to ensure a monotonic relationship between adjacent pixels. No texture was introduced in the network output. Instead, it was introduced from the original image through the relationship. As a common loss term for LLIE, the estimation of the illumination smooth loss term is similar to the calculation of light smoothness loss in Zhang Y. et al. (2019).

3 Experiments and evaluation

3.1 Experimental setup

The hardware part adopts an 11 GB GTX 1080 Ti. The software is PyTorch framework 1.10.0 v. The spiking encoding convLSTM-augmented LLIE model was constructed using the Python 3.7 library of PyTorch and trained using datasets consisting of unpaired images. The optimization process of the proposed network employed the ADAM optimizer with default parameters and a fixed learning rate of 1×10^{-4} . The weights W_{col} and W_{tv} were set to 0.5 and 20, respectively. These parameters remained constant in all experiments.

The datasets, i.e., LLIE fields, were divided into referenced and unreferenced image datasets. Typical referenced image datasets include LOL, SCIE, and MIT-Adobe FiveK, while unreferenced datasets include VV, NPE, and LIME. The LOL dataset has a considerably different degree of underexposure from the rest, which is suitable for the comparison of the overall performance of LLIE algorithms. The SCIE dataset is a multiexposure image sequence dataset with rich illumination information, which is highly suitable for algorithm debugging. Hence, we selected the LOL and SCIE datasets for the experiments. We retained the original training and test dataset distributions for the LOL dataset. In each image sequence of the SCIE dataset, the first image was chosen as the low-light image to be enhanced, whereas the most suitable one was chosen as the high-light reference image among the third, fourth, and fifth images. We used a user study to evaluate five common unreferenced datasets, namely VV, NPE, LIME, DICM, and MEF. We hypothesized that the key performance of LLIE should lie in the size of the space occupied by its running process, which can influence the integration of related

TABLE 1 A calculation of the objective image evaluation index.

Image evaluation index	Mathematical expression	Range	Trend for better
PSNR	$10\log_{10} \frac{MAX(H)^2}{MSE}$	[0,+∞)	To positive infinity
SSIM	$\frac{1}{3} \sum_{channel} \frac{[2\bar{R}\bar{H} + K1 \times MAX(H)^2] + [2\hat{R}\hat{H} + K2 \times MAX(H)^2]}{[\bar{R}^2 + \bar{H}^2 + K1 \times MAX(H)^2][\hat{R}^2 + \hat{H}^2 + K2 \times MAX(H)^2]}$	[0,1]	Closer to 1
MSE	$\frac{1}{H \times W} \sum_{(a,b) \in I} (R_{ab} - H_{ab})$	[0,+∞)	Smaller
UQI	$\frac{4\bar{R}\bar{H}\hat{R}\hat{H}}{(\bar{R}^2 + \bar{H}^2)(\hat{R}^2 + \hat{H}^2)}$	[-1,1]	Closer to 1
VIF	$\sum_{gauss(a,b) \in I} \log_{10} \left[1 + g^2 \frac{H^2 - \bar{H}^2}{svsq + 2} \right]$	[0, +∞]	Closer to 1

TABLE 2 The performance comparison of the ablation study for different substructures (red bold for the best, black bold for the second best).

Detection methods	datasets	PSNR	SSIM	MSE	UQI	VIF
SimpleDCE	LOL	17.1200	0.5969	0.0309	0.7694	0.8564
	SCIE	15.4942	0.6036	0.0337	0.8047	0.5617
Spiking + CNN	LOL	17.2140	0.5999	0.0291	0.7689	0.9394
	SCIE	15.1792	0.6007	0.0353	0.7945	0.5856
Proposed (Spiking+convLSTM)	LOL	18.3374	0.5974	0.0227	0.8369	1.1019
	SCIE	16.7519	0.7289	0.0221	0.8,413	0.5358
Proposed + CBAM	LOL	16.9693	0.5968	0.0315	0.7653	0.8539
	SCIE	15.3300	0.6026	0.0347	0.8010	0.5635

tiny systems. This feature represents the application’s ability to integrate with other functions and algorithms of the testing process and of the model itself.

There are five assessment indices for image objective evaluation, namely peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM), mean square error (MSE), universal image quality index (UQI) (Wang and Bovik, 2002), and visual information fidelity (VIF) (Sheikh and Bovik, 2006). The calculated evaluation indices are listed in Table 1. In this table, h, r corresponds to H s and R ’s results of the Laplace filter. The $nonzero$ (*) function realizes 0 to 1. \sum_{gauss} represents the summation of the results of different Gaussian filter parameters. Gaussian filtering was used for H and R . The number n

depicts the Gaussian filtering times. For instance, \widehat{R}_{n+1}^2 indicates that the square calculation was performed first, followed by Gaussian filtering. \widetilde{R}_{n+1}^2 indicates that Gaussian filtering was performed first, followed by square calculation. \bar{x} represents the uniform filter for x . PSNR and MSE are non-negative. Test images with reference images were calculated to get the PSNR value. The larger the PSNR, the less the image noise and the better the image quality, and SSIM reflects structural similarity. It is typically used to measure whether the image backbone of the image recovered by the LLIE has also been restored.

The SSIM ranges from 0 to 1; only when two sets of identical image data converge will the SSIM reach 1. The indicator, UQI, reflects the measure of the degree of linear correlation, the closeness of the mean luminance, and the similarity of contrast between the enhanced result and the reference image. VIF combines a natural image statistical model, an image distortion model, and a human vision system model. Compared to the PSNR, SSIM, and other indicators, because the numerator of the VIF index calculation formula is the information fidelity criterion (IFC), VIF has a higher consistency with subjective vision. The higher its value, the better the image quality.

3.2 Ablation study

As the proposed method is based on the DCENet structure, the change in the enhancement properties after introducing the spiking+convLSTM structure must be considered. The study demonstrates the influence of each loss term of the loss function on the enhancement results under different loss combinations. In the ablation experiment, different loss combinations were used for retraining. The necessity of each loss item was retested using the proposed DCENet-based method to prevent the negative effects of spiking encoding and convLSTM.

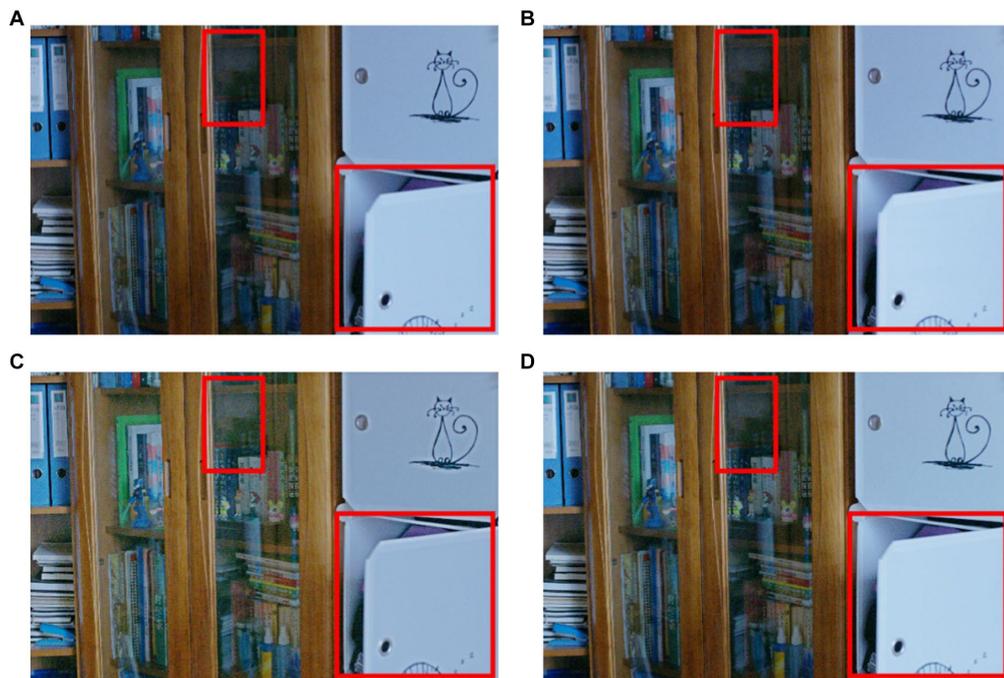


FIGURE 4 Ablation study by three substructures (obvious areas for specific differences) (A) CBAM (B) spikingCNN result (C) the simple dce structure (D) the result obtained by proposed structure.

TABLE 3 Different loss function assemblies of ablation study in the LOL and SCIE datasets (red bold for the best, black bold for the second best).

Lcol	Ltv	Lspa	Lexp	Datasets	PSNR	SSIM	MSE	UQI	VIF
☑	☑			LOL	10.7278	0.3607	0.0940	0.3944	0.2599
				SCIE	9.7752	0.3467	0.1326	0.2811	0.2252
☑		☑		LOL	10.8758	0.3745	0.0908	0.4175	0.2800
				SCIE	10.2279	0.3660	0.1241	0.3243	0.2342
☑			☑	LOL	12.4312	0.4727	0.0670	0.7956	0.5626
				SCIE	13.1966	0.4463	0.0602	0.7650	0.7316
		☑	☑	LOL	11.0469	0.4633	0.0844	0.7478	0.4725
				SCIE	11.3607	0.4384	0.0793	0.7144	0.6979
	☑		☑	LOL	8.0262	0.3444	0.1657	0.5044	0.4472
				SCIE	8.6462	0.3565	0.1435	0.4935	0.8410
	☑	☑		LOL	10.9274	0.3682	0.0909	0.4029	0.2717
				SCIE	9.8145	0.3491	0.1317	0.2844	0.2273
☑		☑	☑	LOL	14.2867	0.5006	0.0470	0.8284	0.4368
				SCIE	14.3752	0.4610	0.0437	0.7912	0.8286
	☑	☑	☑	LOL	8.1473	0.3512	0.1607	0.5078	0.4504
				SCIE	8.6934	0.3591	0.1417	0.4963	0.7412
☑	☑	☑		LOL	10.7671	0.3631	0.0932	0.3984	0.2608
				SCIE	9.8145	0.3491	0.1317	0.2844	0.2273
☑	☑		☑	LOL	13.0447	0.5859	0.0693	0.8047	0.1875
				SCIE	14.9476	0.5691	0.0335	0.8125	0.7470
☑	☑	☑	☑	LOL	18.3374	0.5974	0.0227	0.8369	1.1019
				SCIE	16.7519	0.7289	0.0221	0.8413	0.5358

Another ablation experiment should also be considered, which focuses on spiking encoding and convLSTM itself. Thus, three ablation study experiments, whose network is made up of the only light DCENet structure, the structure with the CBAM attention mechanism, or the CNN structure that replaces the convLSTM, have been considered for comprehensively verifying the proposed structure's necessity. In the two ablation studies, the training parameter did not change. The SCIE dataset was applied for specific calculations.

Only light DCENet structure: Without the proposed spiking encoding and convLSTM structure, the enhancement is only realized by DCENet.

Structure with CBAM attention mechanism: Based on the only-light DCENet structure, the CBAM attention mechanism is set after the first layer.

CNN structure that replaces convLSTM: The enhancement was running using a CNN structure instead of convLSTM. The dimensions of the spiking encoding image sequence were trimmed, and the image sequence was superimposed to form a feature map.

The ablation study about the importance of spiking encoding and convLSTM is summarized in Table 2. Compared with the basic DCENet structure, spiking combined with the CNN structure revealed that the integration of spiking encoding alone improved the



FIGURE 5 (Continued)

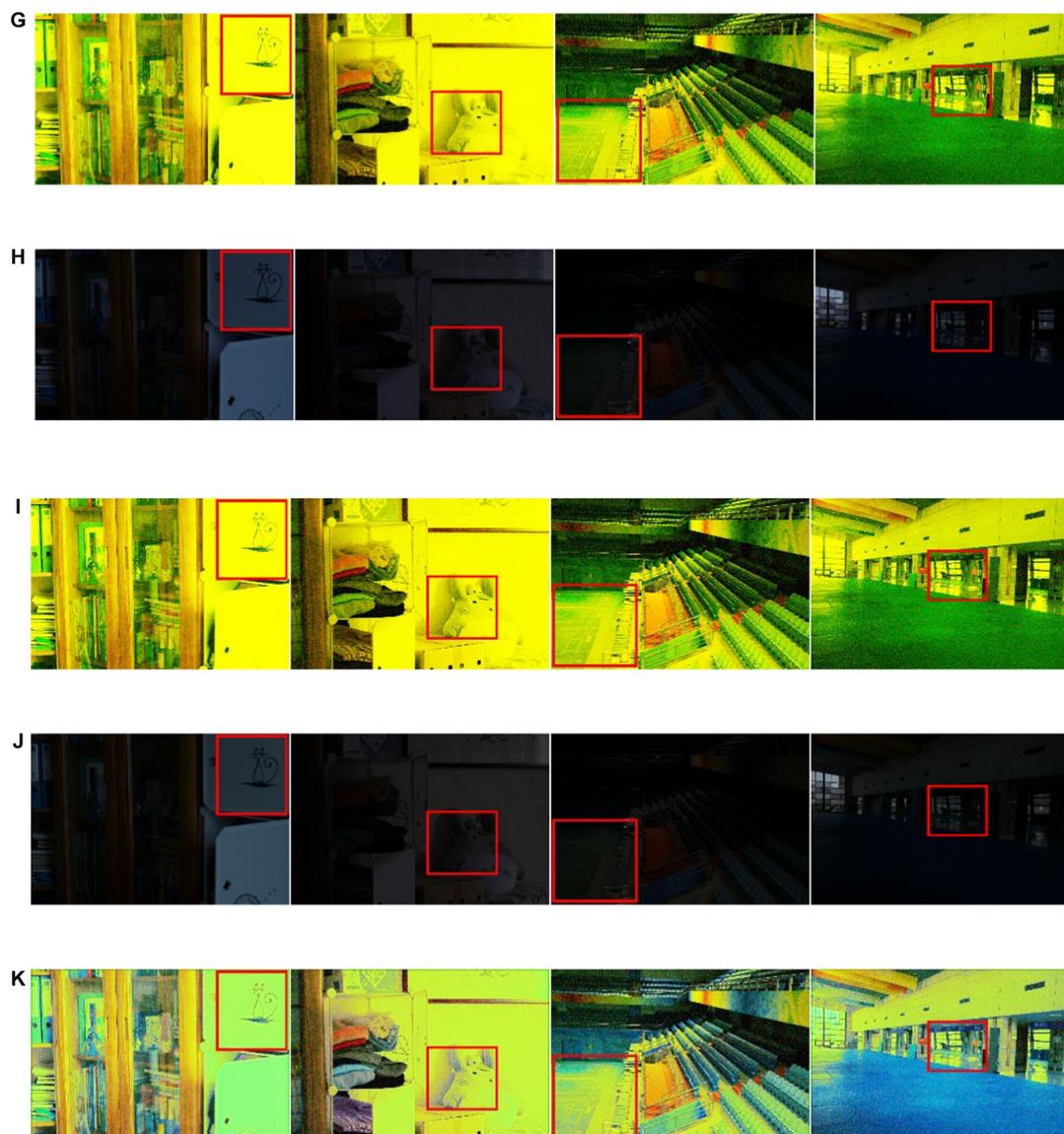


FIGURE 5

The loss function assembly in LOLdataset (A) original low-light image (B) color constancy loss and exposure loss (C) illumination smooth loss, exposure loss (D) color constancy loss, spaital consistency loss, (E) spaital consistency loss, exposure loss, (F) illumination smooth loss, color constancy loss, (G) illumination smooth loss, spaital consistency loss (H) illumination loss, exposure loss and color constancy loss (I) color constancy loss, exposure loss and spaital consistency loss (J) illumination loss, color constancy loss and spaital consistency loss (K) illumination loss, exposure loss and spaital consistency loss.

performance. Specifically, the VIF parameter showed significant increments of 9.7 and 4.3% on the LOL and SCIE datasets, respectively. However, when considering other indicators, the objective evaluations of the LOL and SCIE datasets demonstrated a contrary trend. This suggests that the combination of spiking and CNN methods may not be beneficial for enhancing model generalization stability. To improve upon this, the study adopted the classic CBAM attention mechanism as a representative approach for introducing attention mechanisms. Data suggest that incorporating attention mechanisms alone reduced the number of essential evaluation criteria, such as PSNR, SSIM, and UQI. Additionally, the combination of convLSTM and spiking encoding not only elevated the evaluation index on the SCIE dataset but also surpassed the effect of the convolutional network

combination. In addition, we identified only minor differences in the subjective effects of the methods under the ablation experiments. These effects are presented in Figure 4.

Ablation experiments assess the impact of different loss function terms on the image enhancement quality. The proposed approach employed four loss function terms. Their pairwise and three-way combinations and the corresponding image evaluation index parameters are listed in Table 3. Of the six paired combination parameters, color constant loss and exposure loss substantially enhanced image quality, followed by spatial consistency loss and exposure loss. Consequently, we infer that exposure loss is the most crucial loss item, followed by color constant loss and spatial consistency loss, which exert the least impact on light smoothness loss.

TABLE 4 The performance comparison of the ablation study in the SICE dataset (red bold for the best, black bold for the second best).

Detection methods	PSNR	SSIM	MSE	UQI	VIF
RRDNet	12.4675	0.5469	0.0594	0.5713	0.3487
zero-DCE	16.0794	0.6618	0.0365	0.8208	0.6829
DRBN	15.8745	0.4667	0.0290	0.8008	0.2575
EXCNet	16.0427	0.6006	0.0334	0.7626	0.3675
LightenNet	11.2150	0.3216	0.0810	0.7228	0.4736
DSLRL	15.1002	0.5996	0.0318	0.7839	0.3348
BIMEF	15.7917	0.5904	0.0326	0.7936	0.3678
LLFLOW	15.0300	0.5830	0.0364	0.7918	0.4452
Enlighten anything	14.9228	0.6193	0.0443	0.8393	0.4487
Proposed (Spiking+convLSTM)	16.7519	0.7289	0.0221	0.8,413	0.5358
EnlightenGAN	16.6135	0.6219	0.0260	0.8306	0.5310
BFSA	12.0203	0.4419	0.0721	0.5401	0.5002
Bread	16.0787	0.6209	0.0292	0.7962	0.6295

TABLE 5 The performance comparison of the ablation study in the LOL dataset (red bold for the best, black bold for the second best).

Detection methods	PSNR	SSIM	MSE	UQI	VIF
RRDNet	13.1360	0.5598	0.0695	0.5286	0.4951
zero-DCE	17.5592	0.5750	0.0228	0.8355	1.0560
DRBN	17.2850	0.5174	0.0238	0.7222	0.5668
EXCNet	14.8137	0.5539	0.0643	0.7010	0.4894
LightenNet	10.5513	0.1243	0.1183	0.6142	0.1467
DSLRL	16.1505	0.6273	0.0389	0.6932	0.5073
BIMEF	17.0586	0.5565	0.0254	0.7792	0.4692
LLFLOW	15.5407	0.5625	0.3552	0.7353	0.5331
Enlighten Anything	16.8056	0.5646	0.0280	0.8678	0.5288
Proposed (Spiking+convLSTM)	18.3374	0.5974	0.0227	0.8369	1.1019
EnlightenGAN	17.2322	0.6945	0.0287	0.8189	0.7589
BFSA	11.0324	0.4429	0.1031	0.3894	0.3573
Bread	17.6990	0.6530	0.0273	0.7888	0.8823

The study revealed a consistent trend among the four pairs of three-way combination parameters. The method that incorporated exposure loss, color constant loss, and spatial consistency loss outperformed all others in the overall index. However, in terms of UQI, the method combined with exposure loss, color constant loss, and illumination smooth loss performed similarly to the rest. Notably, all four loss functions operated simultaneously. In other words, the index value corresponding to the method proposed in Table 1 is still the best. However, in both the LOL and SCIE datasets, UQI and VIF were marginally inferior to the composite approach of exposure loss, color constant loss, and spatial consistency loss. This highlights the indispensability of using four loss functions. Figures 5, 6 illustrate the influence of each loss function on the image enhancement effect. As

observed, exposure loss directly controls image enhancement, while color constant loss mainly controls image distortion after enhancement.

3.3 Performance comparison

After gaining an understanding of the proposed LLIE method, we conclude that the LLFLOW (Wang et al., 2022), BIMEF (Ying et al., 2017), RRDNet (Zhu et al., 2020), zero-DCE (Guo et al., 2020a), DRBN (Yang et al., 2020), EXCNet (Zhang Y. et al., 2019), Lightnetnet (Zhang et al., 2019), Enlighten Anything (Zhou et al., 2023), EnlightenGAN, DSLR (Ignatov et al., 2017), BREAD (Hu and Guo,

2022), and BFSFA (Long et al., 2023) algorithms have strong robustness and potential applications. The proposed method was compared with two referenced datasets in the LLIE field based on five performance indicators. Figure 7 directly demonstrates the enhancement effect. Tables 4, 5 list the performance index values of the proposed method and several of the most popular enhancement methods in the two reference image datasets. The proposed method yielded the best values, with the PSNR, SSIM, MSE, UQI, and VIFP exceeding the second place at 4.4% (0.8%), 3.9% (17.2%), 0% (15%), 0.1% (0.2%), and 4.3% (0.9%), respectively. The numbers inside parentheses

represent the increase in the SCIE dataset. In addition, we also measured the parameters related to the actual application characteristics of the resulting algorithm. In the actual application of the image algorithm, the hardware space occupied by the model and the space occupied by the test process warrant attention.

Enlighten Anything performs similarly to the EXCNet method. However, it has a good research starting point, which is combined with the large segmentation pretrained model algorithm (Kirillov et al., 2023). Although the LightenNet method meets the characteristics of lightweight, it yields several poor indices. The



FIGURE 6 (Continued)

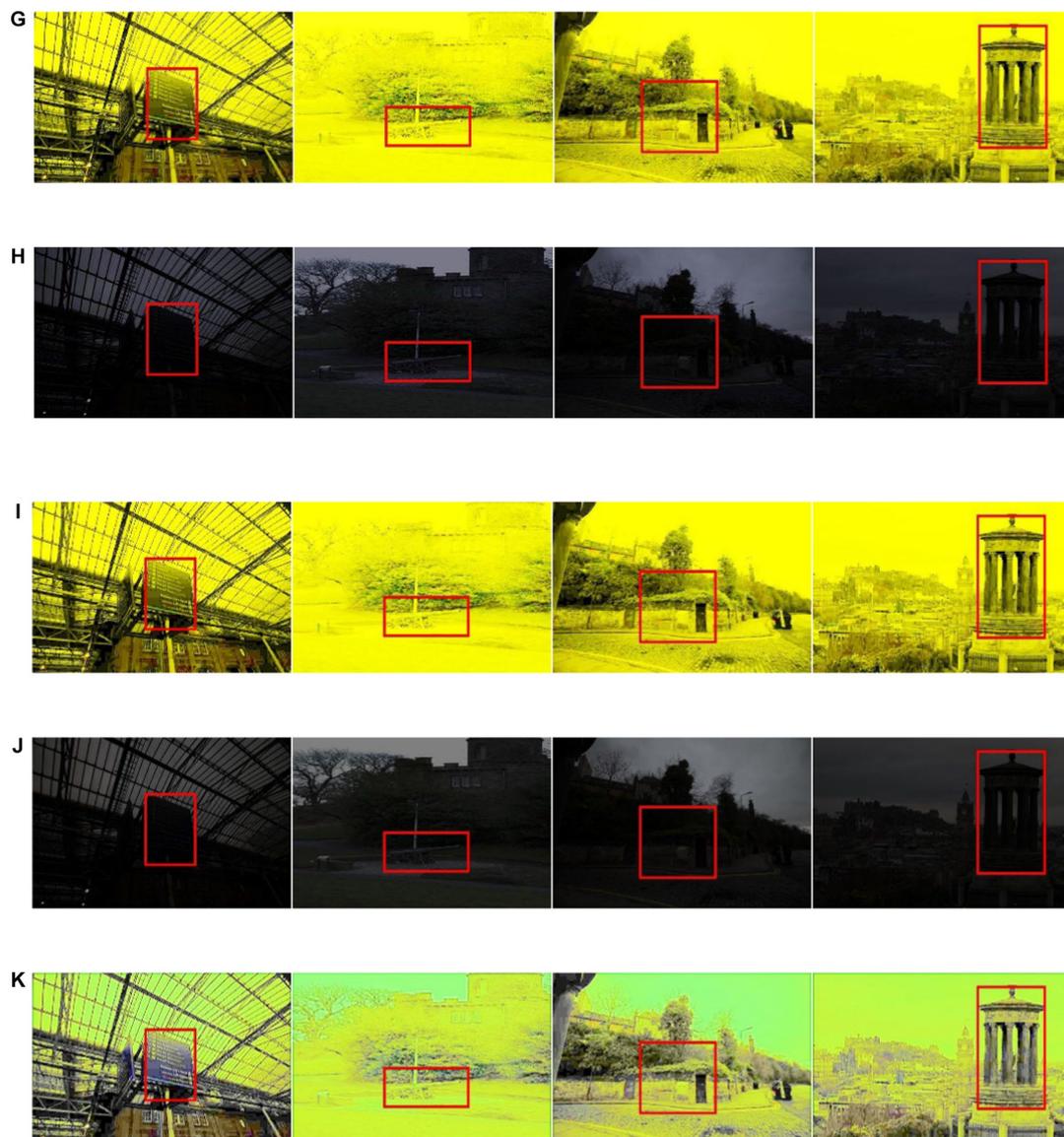


FIGURE 6

The loss function assembly in SCIE dataset (A) original low-light image (B) color constancy loss and exposure loss (C) illumination smooth loss, exposure loss (D) color constancy loss, spatial consistency loss, (E) spatial consistency loss, exposure loss, (F) illumination smooth loss, color constancy loss, (G) illumination smooth loss, spatial consistency loss (H) illumination loss, exposure loss and color constancy loss (I) color constancy loss, exposure loss and spatial consistency loss (J) illumination loss, color constancy loss and spatial consistency loss (K) illumination loss, exposure loss and spatial consistency loss.

performance of EXCNet methods is relatively moderate compared to other state-of-the-art methods. The model occupies a relatively large space. In the performance evaluation, the proposed methods, zero-DCE and EnlightenGAN, ranked the highest. The primary role of LLIE methods is to assist with enhancing the realization of other algorithmic functions. Generally, the model and testing process should occupy less space for better integration with other product features. As indicated in Table 6, the space occupied by the proposed algorithm in the test process ranks second, which is only larger than the poorly performing RRDNet, while the space occupied by the model itself reaches 151 KB, which is more than half of the space occupied by the second place.

3.4 User study

Certain LLIE-related datasets have no reference images corresponding to normal light, only images under dark lighting conditions, and therefore it was difficult to use objective evaluation indicators, such as PSNR, to evaluate image quality. To make the performance comparison clearer, more intuitive, and more efficient for these non-reference image datasets, a user study was performed to assess the human perception of the proposed method. The images tested by the user study included various image contents in different environments, including animals, exterior scenes, and buildings. Based on the user

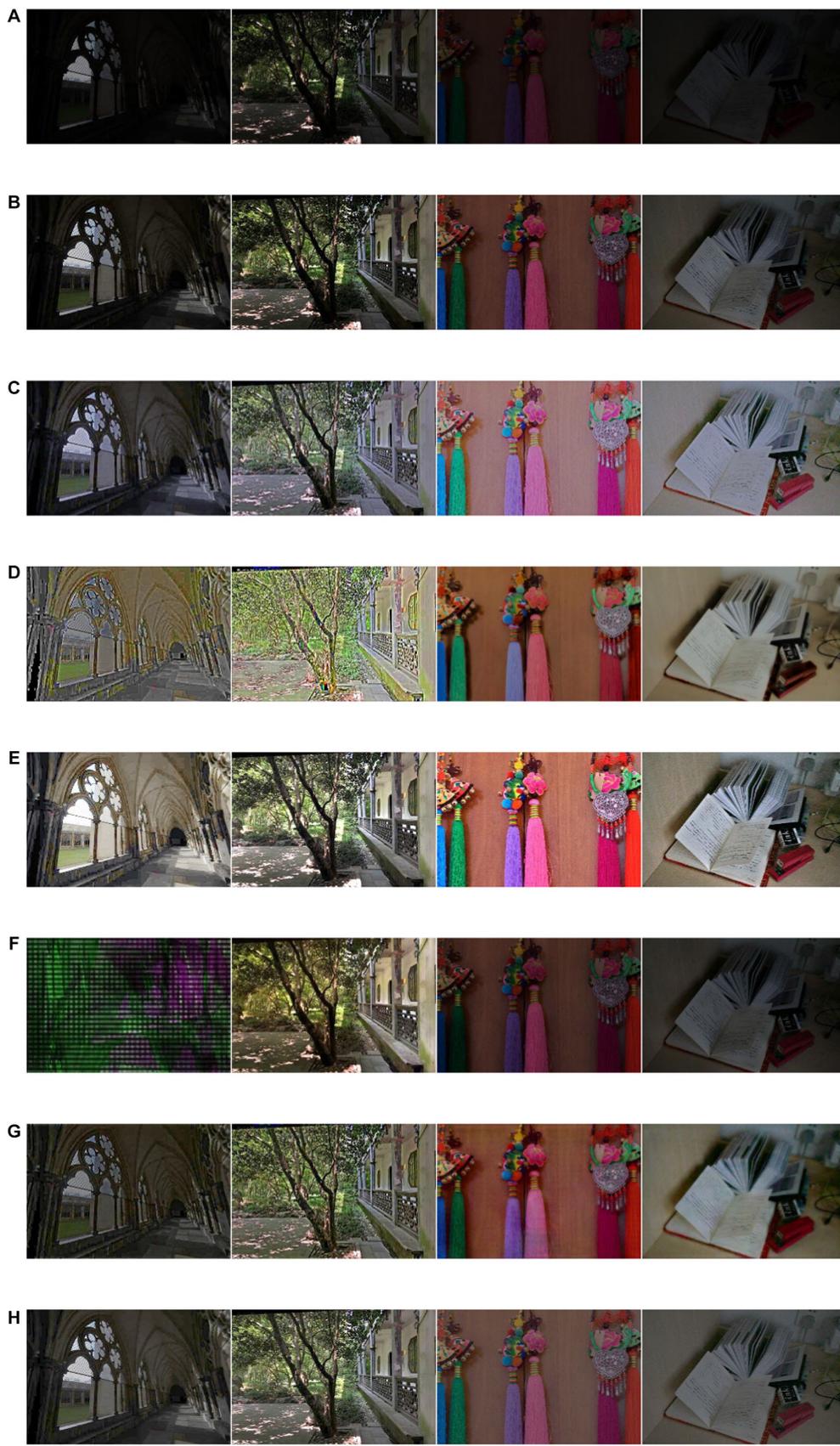


FIGURE 7 (Continued)

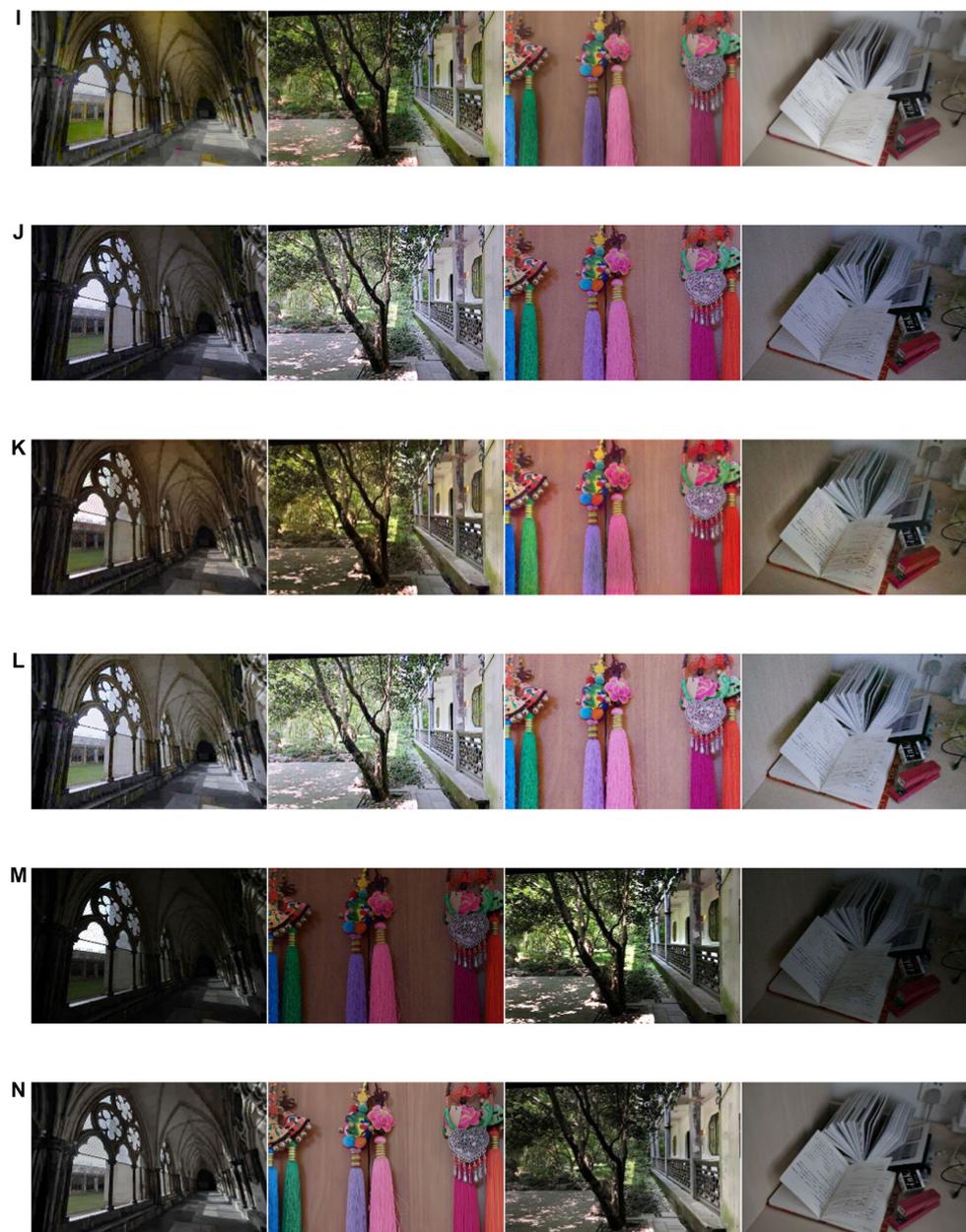


FIGURE 7

The performance comparisons for different combinations of loss function in SCIE dataset (A) original low-light image (B) RRDNet (C) zerodce (D) DRBN, (E) EXCNet (F) Lightnnet, (G) DSLR, (H) BIMEF (I) LLLFLOW (J) Enlighten Anything (K) EnlightenGAN (L) Proposed method (M) Bread (N) BFSa.

feedback data, we constructed a radar map with a maximum of 100 score points for each index, which answered the following five questions:

- Are the details noticeable?
- Are the colors vivid?
- Is the result visually realistic?
- Do the results contain overexposed/underexposed artifacts or over-enhanced/under-enhanced regions?
- Do the results have unnatural texture and noticeable noise?

A single radar map can clearly compare the performance of different methods in various aspects of an unreferenceed image dataset. The larger the area of the radar map, the better the subjective comprehensive evaluation of the method. Each angular direction, which ranges from 70 to 100 on the radar map, represents the user rating score for a specific problem. The five radar plots in Figure 8 illustrate the distributions of scores evaluated on different questions for different LLIE methods, where the bright red lines in the radar map represent the proposed method. We compared the results of the proposed method for the user study with those of the other LLIE

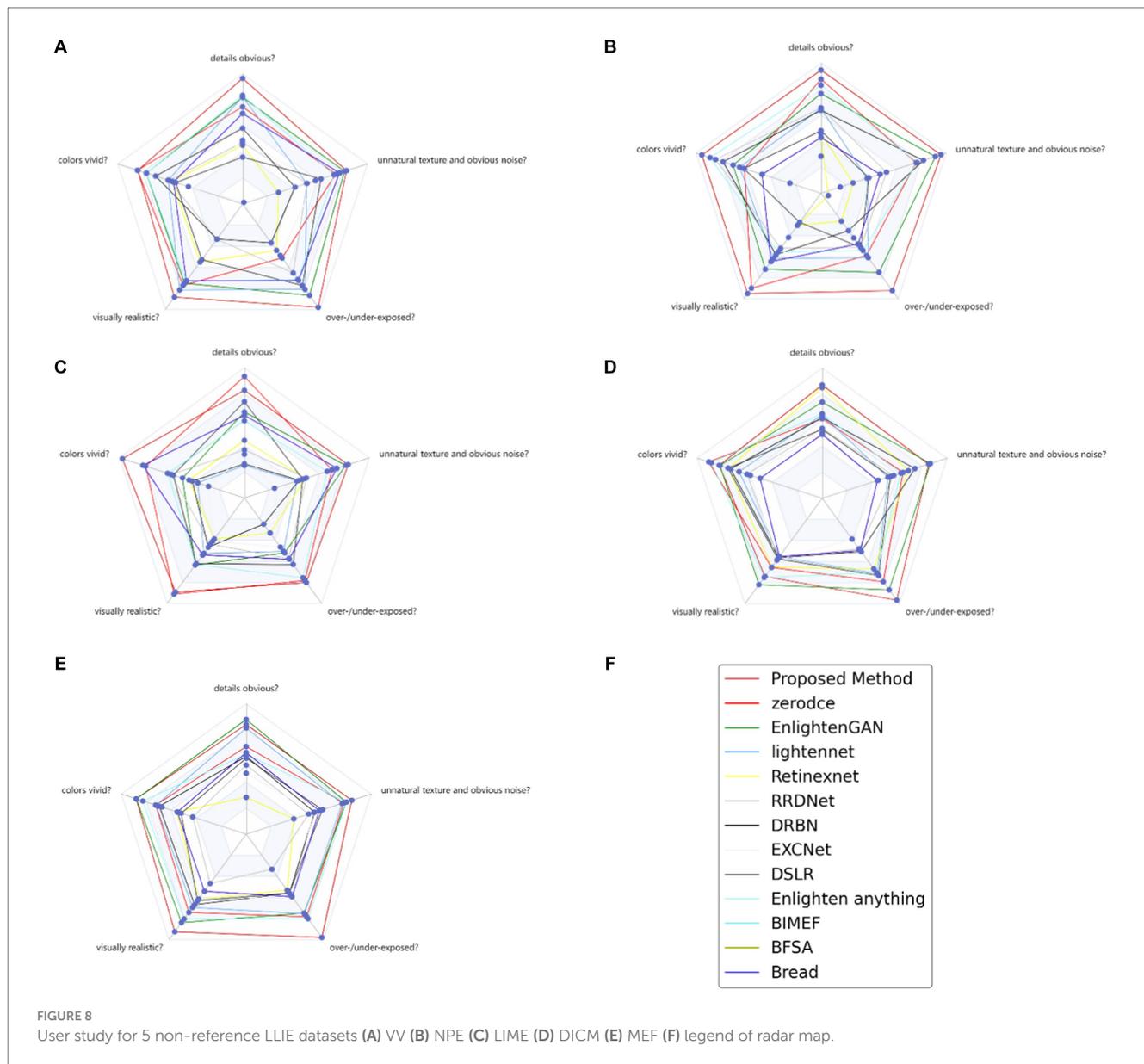


TABLE 6 Memory occupation of the model and testing process (red bold for the best, black bold for the second best).

Detection methods	Testing memory	Model memory
RRDNet	1,040,384 (1.04 MB)	511 KB
zero-DCE	29,149,184 (29.1 MB)	315 KB
DRBN	53,602,358 (53.6 MB)	2.2 MB
EXCNet	22.9 MB (in Tensorflow)	157.2 MB
LightenNet	366.6 MB (in MATLAB)	108 KB
DSLRL	40,600,532 (40.5 MB)	3.2 MB
BIMEF	125.8 MB (in MATLAB)	/
LLFLOW	239.6 MB	20.9 MB (smallest version)
Enlighten anything	137,830,912 (131.6 MB)	144 MB
Proposed (Spiking+convLSTM)	22,233,600 (22.23 MB)	151 KB
EnlightenGAN	590,612,992 (590.61 MB)	33,774 KB
Bread	2451.456 MB	6.6 MB
BFSA	3999.744 MB	230.1 MB

methods using a paired *t*-test (Guo et al., 2020). The results revealed that the effect of EnlightenGAN was the least different from that of the proposed method, except for zero-DCE.

4 Discussion

The dark light image enhancement method proposed in this paper has been tested by ablation experiments of different image evaluation indices in different datasets and performance comparison experiments, which have verified its performance superiority. In terms of space proportion, the model in this study is a single model, which does not need to involve a pre-training model or other model framework fusion methods. Enlighten Anything involves the pre-training weight of the SAM model. Compared with EnlightenGAN and related reinforcement learning methods, the new method has relatively low training configuration requirements and difficulty. The limitation of this method is that it is time-consuming at an average of 0.007 s, as determined by the LOL test dataset, which is marginally less than EnlightenGAN. After testing, it was found that the convLSTM structure occupied 0.006 s during testing. However, it still enhances images at 140 fps, which exceeds the real-time demand of 30 fps.

5 Conclusion

Originating from the further introduction of spiking coding mechanisms into DL, a novel network exhibits better performance based on DCENet by spiking encoding and convLSTM. Intensity-to-latency conversion, which is a spiking-coding methodology, can be used to gradually acquire the structural characteristics of an image. The multiple subgraphs generated by this method relate to the time step defined by spiking coding, and convLSTM is suitable for solving the image sequence problem and introducing the relationship information between multiple images into the network structure. Furthermore, the simplified DCENet structure without supervision achieved a superior result in terms of improvement. The performance comparison of this method with nine conventional methods in terms of five metrics was validated. The ablation study proved the necessity of the various parts of the structure, such as network and training losses. The proposed method yielded the best values with PSNR, SSIM, MSE, UQI, and VIFP. The proposed model occupies only 151 KB, which will better meet the algorithm integration and practical application requirements on a small chip.

6 Scope

The dark light enhancement method used in the study is closely related to the bionic neural networks and learning systems section of the special issue. The relationship between dark light enhancement and neural networks is that neural networks can be applied to tasks with dark light enhancement. Dark light enhancement is an image processing technique designed to improve the visibility of images taken in low-light conditions. By learning a large amount of training data, a neural network can automatically learn and extract the features in the image and perform enhancement processing on the image to improve the quality and visibility of the image.

By using a neural network, a dark light-enhanced model can be built, which is capable of receiving an input image and producing the enhanced image as output. The neural network can automatically learn and fit the mapping relationship between the input image and the output image through the connection and weight adjustment between the multiple layers of neurons. By training and optimizing the neural network, it can enhance the dark light image and have better generalization ability for different input images.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XW: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. QW: Project administration, Resources, Writing – review & editing. LZ: Conceptualization, Writing – original draft, Writing – review & editing. YQ: Software, Writing – review & editing. FY: Conceptualization, Visualization, Writing – review & editing. JY: Supervision, Validation, Writing – review & editing. QL: Formal analysis, Methodology, Writing – review & editing. RX: Investigation, Validation, Writing – review & editing. ZX: Data curation, Formal analysis, Writing – review & editing. ST: Conceptualization, Formal analysis, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bao, C., Guo, Z., Yao, W., Ding, X., and Zhang, D. (2022). A novel low-light enhancement via fractional-order and low-rank regularized retinex model. *Comput. Appl. Math.* 42:7. doi: 10.1007/s40314-022-02140-6
- Cai, J., Gu, S., and Zhang, L. (2018). Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Trans. Image Process.* 27, 2049–2062. doi: 10.1109/tip.2018.2794218
- Chen, X., Li, J., and Hua, Z. (2022). Retinex low-light image enhancement network based on attention mechanism. *Multim. Tools Appl.* 82, 4235–4255. doi: 10.1007/s11042-022-13411-z
- Chen, W., Wang, W., Yang, W., and Liu, J. (2018). Deep Retinex decomposition for low-light enhancement. *arXiv*. doi: 10.48550/arXiv.1808.04560
- Dar, K. A., and Mittal, S. (2020). A dynamic fuzzy histogram equalization for high dynamic range images by using multi-scale Retinex algorithm. *SSRN electron. J.* doi: 10.2139/ssrn.3565891
- Fan, C.-M., Liu, T.-J., and Liu, K.-H. (2022). Half wavelet attention on m-net+ for low-light image enhancement. *Neural Inf. Process. Syst.*, 3878–3882. doi: 10.1109/ICIP46576.2022.9897503
- Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., et al. (2020a). Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. [online] *arXiv.org*. doi: 10.48550/arXiv.2001.06826
- Guo, J., Ma, J., Garcia-Fernandez, A. F., Zhang, Y., and Liang, H. (2023). A survey on image enhancement for low-light images. *Heliyon* 9:e14558. doi: 10.1016/j.heliyon.2023.e14558
- Guo, X., Wang, Y., Zhou, N., and Zhu, X. (2020b). Optimal weighted two-sample t-test with partially paired data in a unified framework. *Journal of Applied Statistics* 48, 961–976. doi: 10.1080/02664763.2020.1753027
- Hu, Q., and Guo, X. (2022). Low-light image enhancement via breaking down the darkness. *Int. J. Comput. Vis.* 131, 48–66. doi: 10.1007/s11263-022-01667-9
- Ignatov, A., Kobyshev, N., Timofte, R., Vanhoey, K., and Van Gool, L. (2017). DSLR-quality photos on Mobile devices with deep convolutional networks. *Int. Conf. Comput. Vis.* doi: 10.1109/iccv.2017.355
- Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., et al. (2021). Enlightengan: deep light enhancement without paired supervision. *IEEE Trans. Image Process.* 30, 2340–2349. doi: 10.1109/tip.2021.3051462
- Jiang, Z., Li, H., Liu, L., Men, A., and Wang, H. (2021). A switched view of retinex: deep self-regularized low-light image enhancement. *Neurocomputing* 454, 361–372. doi: 10.1016/j.neucom.2021.05.025
- Jiang, N., Lin, J., Zhang, T., Zheng, H., and Zhao, T. (2023). Low-light image enhancement via stage-transformer-guided network. *IEEE Trans. Circuits Syst. Video Technol.* 33, 3701–3712. doi: 10.1109/TCSVT.2023.3239511
- Ke, X., Xin, Y., Baocai, Y., and Rynson, W.H.L. (2020). *Learning to restore low-light images via decomposition-and-enhancement*. IEEE WA, USA
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., et al. (2023). Segment Anything. In: *ICCV 2023*. Available at: <https://arxiv.org/pdf/2304.02643v1.pdf>
- Li, J., Feng, X., and Hua, Z. (2021). Low-light image enhancement via progressive-recursive network. *IEEE Trans. Circuits Syst. Video Technol.* 31, 4227–4240. doi: 10.1109/tcsvt.2021.3049940
- Li, C., Guo, C., and Chen, C. L. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 1, 4225–4238. doi: 10.1109/tpami.2021.3063604
- Li, C., Guo, C., Han, L.-H., Jiang, J., Cheng, M.-M., Gu, J., et al. (2022). Low-light image and video enhancement using deep learning: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 9396–9416. doi: 10.1109/tpami.2021.3126387
- Li, C., Guo, F., Porikli, J., and Pang, Y. (2018). Lightnet: a convolutional neural network for weakly illuminated image enhancement. *Pattern Recogn. Lett.* 104, 15–22. doi: 10.1016/j.patrec.2018.01.010
- Liu, X., Xie, Q., Zhao, Q., Wang, H., and Meng, D. (2023). Low-light image enhancement by retinex-based algorithm unrolling and adjustment. *IEEE Trans. Neural Netw. Learn. Syst.* doi: 10.1109/TNNLS.2023.3289626
- Liu, J., Xu, D., Yang, W., Fan, M., and Huang, H. (2021). Benchmarking low-light image enhancement and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.* 129, 1153–1184. doi: 10.1007/s11263-020-01418-8
- Long, L., Dian, J., Nan, A., Jinyuan, L., Xin, F., Zhongxuan, L., et al. (2023). Bilevel Fast Scene Adaptation for Low-Light Image Enhancement. *International Journal of Computer Vision*. doi: 10.1007/s11263-023-01900-z
- Lore, K. G., Akintayo, A., and Sarkar, S. (2017). LLNet: a deep autoencoder approach to natural low-light image enhancement. *Pattern Recogn.* 61, 650–662. doi: 10.1016/j.patcog.2016.06.008
- Lv, F., Lu, F., Wu, J., and Lim, C. (2018). MBLEN: Low-light image/video enhancement using CNNs. In *British machine vision conference*. Newcastle, UK
- Martinez-Seras, A., Del Ser, J., Lobo, J. L., Garcia-Bringas, P., and Kasabov, N. (2023). A novel out-of-distribution detection approach for spiking neural networks: design, fusion, performance evaluation and explainability. *Inf. Fusion* 100:101943. doi: 10.1016/j.inffus.2023.101943
- Mozafari, M. N. D. A., and Ganjtabesh, M. T. M. (2019). Spyketorch: efficient simulation of convolutional spiking neural networks with at most one spike per neuron. *Front. Neurosci.* 13:625. doi: 10.3389/fnins.2019.00625
- Rao, Y., Lin, W., and Chen, L. (2010). Image-based fusion for video enhancement of night-time surveillance. *Opt. Eng.* 49:120501. doi: 10.1117/1.3520553
- Rong, Y., Liu, W., Zhang, Y., Zhi, Q., Zhao, D., and Zhang, B. (2018). Deepexposure: learning to expose photos with asynchronously reinforced adversarial learning. *Neural Inf. Process. Syst.* 31, 2149–2159.
- Sheikh, H. R., and Bovik, A. C. (2006). Image information and visual quality. *IEEE Trans. Image Process.* 15, 430–444. doi: 10.1109/tip.2005.859378
- Shen, Z., Xu, H., Jiang, G., Yu, M., Du, B., Luo, T., et al. (2023). Pseudo-retinex decomposition-based unsupervised underwater image enhancement and beyond. *Digit. Signal Process.* 137:103993. doi: 10.1016/j.dsp.2023.10399
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W. K., and Woo, W. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Advances in neural information processing systems*. doi: 10.48550/arxiv.1506.04214
- Singh, K., Pandey, A., Agarwal, A., Agarwal, M. K., Shankar, A., and Parihar, A. S. (2024). Frn: fusion and recalibration network for low-light image enhancement. *Multimed. Tools Appl* 83, 12235–12252. doi: 10.1007/s11042-023-15908-7
- Tang, L., Ma, J., Zhang, H., and Guo, X. (2022). DRLIE: Flexible Low-Light Image Enhancement via Disentangled Representations. *IEEE transactions on neural networks and learning systems* 1–14. doi: 10.1109/tnnls.2022.3190807
- Tang, H., Zhu, H., Tao, H., and Xie, C. (2023). An improved algorithm for low-light image enhancement based on retinexnet. *Appl. Sci.* 12:7268. doi: 10.3390/app12147268
- Wang, Z., and Bovik, A. (2002). A universal image quality index. *IEEE Signal Process. Lett.* 9, 81–84. doi: 10.1109/97.995823
- Wang, Q., Fu, X., Zhang, X.-P., and Ding, X. (2016). A fusion-based method for single backlit image enhancement. In 2016 IEEE international conference on image processing (ICIP). 4077–4081. Phoenix, AZ, USA
- Wang, Y., Wan, R., Yang, W., Li, H., Chau, L.-P., and Kot, A. (2022). Low-light image enhancement by normalizing flow. *Proc. Int. AAAI Conf. Artif. Intell.* 36, 2604–2612. doi: 10.1609/aaai.v36i3.20162
- Wu, T., Wu, W., Yang, Y., Fan, F.-L., and Zeng, T. (2023). Retinex image enhancement based on sequential decomposition with a plug-and-play framework. *IEEE Trans. Neural Netw. Learn. Syst.*, 1–14. doi: 10.1109/tnnls.2023.3280037
- Xu, S., Zhang, J., Qin, X., Xiao, Y., Qian, J., Bo, L., et al. (2022). Deep retinex decomposition network for underwater image enhancement. *Comput. Electr. Eng.* 100:107822. doi: 10.1016/j.compeleceng.2022.107822
- Yang, W., Wang, S., Fang, Y., Wang, Y., and Liu, J. (2020). From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 3063–3072. doi: 10.1109/cvpr42600.2020.00313
- Ying, Z., Li, G., and Gao, W. (2017). A bio-inspired multi-exposure fusion framework for low-light image enhancement. *arXiv*. doi: 10.48550/arxiv.1711.00591
- Yusob, B., Mustafa, Z., and Sulaiman, J. (2018). Anomaly detection in time series data using spiking neural network. *Adv. Sci. Lett.* 24, 7572–7576. doi: 10.1166/asl.2018.12980
- Zhang, H., Fan, X., and Zhang, Y. (2023a). Energy-efficient spiking segmenter for frame and event-based images. *Biomimetics* 8:356. doi: 10.3390/biomimetics8040356
- Zhang, H., Li, Y., He, B., Fan, X., Wang, Y., and Zhang, Y. (2023b). Direct training high-performance spiking neural networks for object recognition and detection. *Front. Neurosci.* 17:1229951. doi: 10.3389/fnins.2023.1229951
- Zhang, L., Liu, X., Shen, Y., Zhang, S., and Zhao, S. (2019). Zero-Shot Restoration of Back-lit Images Using Deep Internal Learning. Proceedings of the 27th ACM International Conference on Multimedia, Nice, France.
- Zhang, H., and Ma, J. (2023). Iid-mef: a multi-exposure fusion network based on intrinsic image decomposition. *Inf. Fusion* 95, 326–340. doi: 10.1016/j.inffus.2023.02.031
- Zhang, Y., Zhang, J., and Guo, X. (2019). Kindling the darkness: a practical low-light image enhancer. *MM '19: Proceedings of the 27th ACM international conference on multimedia*, 1632–1640. doi: 10.1145/3343031.3351069
- Zhou, D., Yang, Z., and Yang, Y. (2023). Pyramid diffusion models for low-light image enhancement. *Neural Inf. Process. Syst.* doi: 10.48550/arXiv.2305.10028
- Zhu, A., Zhang, L., Shen, Y., Ma, Y., Zhao, S., and Zhou, Y. (2020). “Zero-shot restoration of underexposed images via robust retinex decomposition” in *IEEE Inter. Conf. Multimedia Expo(ICME)*, 1–6. doi: 10.1109/ICME46284.2020.9102962