

Convolutional neural network-based low light image enhancement method

J. Guo¹

¹Department of Information Engineering,
Xiamen Ocean Vocational College, Xiamen, 361012, China

Abstracts

Low-light image augmentation has become increasingly important with the advancement of computer vision technologies in a variety of application settings. However, noise and contrast reduction frequently have an impact on image quality in low-light situations. In this paper, a convolutional neural network-based technique for low-light picture augmentation is put forth. The stability of local binary features under variations in illumination is the study's initial method of providing directional advice for the enhancement algorithm. Second, the addition of a channel attentiveness mechanism improves the network's capacity to acquire low-light image features. The proposed model of the study performed better on average in the two dataset tests when compared to the contrast-constrained adaptive histogram equalization algorithm and the bilateral filtering algorithm. Additionally, the recall and DICE coefficient performed better in the tests as well, improving by 16.24% and 4.98%, respectively. The proposed method outperformed all others in the picture enhancement studies, according to the experimental findings, proving the validity of this study. The purpose of the study is to offer a reference framework for low-light image enhancing techniques.

Keywords: computer vision, image enhancement, image quality, convolutional neural networks.

Citation: Guo J. Convolutional neural network-based low light image enhancement method. *Computer Optics* 2024; 48(5): 745-752. DOI: 10.18287/2412-6179-CO-1415.

Introduction

Image processing has grown in importance as a research area in the Computer Vision (CV) and Image Analysis (IA) sectors in the current digital era [1]. Particularly, the subsequent IA and recognition accuracy are directly impacted by the quality of photos taken in low light circumstances [2]. As a result, Low Light Image Enhancement (LLIE) approaches are essential for enhancing visibility and Image Quality (IQ) [3]. Low light, however, frequently results in noise, decreased contrast, and loss of detail in photographs, which makes image improvement extremely difficult [4]. In recent years, Convolutional Neural Networks (CNNs) have displayed outstanding performance in the field of image processing [5]. Through the use of resilient under changes in illumination Local Binary Pattern (LBP) based features, CNNs have been optimised in this study. The work added channel Attention Mechanisms (AM) to the CNN's architecture in order to aid in the capture of low-light picture characteristics. Additionally, the network in the study includes a Feature Aggregation Module (FAM). By examining feature correlations, FAM improves feature representation, enabling more precise feature extraction and application from images. There are five key sections to the research, the first of which is an overview; the second part introduces related work at home and abroad; the third part is divided into two subsections, the first subsection introduces CNN optimization based on LBP feature extraction, and the second subsection introduces the LLIE model construction method based on Multi Modular Network (MM-Net); the fourth part of the research experiments are conducted on the optimized

LLIE model to verify its performance; the fifth section concludes the study and the outlook for future research. The objective of this research is to develop an LLIE technique that enhances image contrast, reduces noise, and improves image detail. This technique is especially beneficial for low-light CV tasks, such as target recognition and IA.

1. Related work

One of the top areas for research in the realm of digital image has long been LLIE approaches [6]. To address the issue of poor visibility and noise in low light photos, Yang et al. suggested a Retinex decomposition-based low light image improvement technique. The strategy surpassed the most recent state-of-the-art picture enhancing techniques, according to the experimental findings when the researchers combined several techniques for image segmentation [7]. A feature-level attention model was created by Yang et al. to enhance the visual appeal of underwater photos. The model was a multi-scale grid CNN that can take into account various kinds of information when learning. The model's effectiveness was demonstrated by the researchers through extensive tests utilising benchmark and real-world underwater photos [8]. He et al. concluded that some images taken in outdoor environments suffered from colour distortion and missing details due to the effects of variable environments. The outcomes of the experiment demonstrated that the method worked better than the other strategies listed therein [9]. Azam and colleagues proposed an image enhancement approach that combined discrete wavelet transform and principal component averaging techniques for image fusion with multiresolution stiff alignment techniques for

multimodal image alignment. The researchers hope that this method will further help doctors to make accurate diagnosis through lesion images. Experimental results indicated that the method provided better IQ and could better assist doctors in medical diagnosis [10]. Lore and colleagues introduce a technique employing a depth-based autoencoder to identify signal attributes in low-light images and intelligently amplify the image without over-saturating brighter areas of a high dynamic range image. The technique is trained using a variation of the stacked-sparse denoising autoencoder, enabling it to learn from artificially produced low-light and noisy image examples. The method exhibits a high degree of reliability in improving natural low-light environments and hardware degraded images, as evidenced by the results [11]. Wang and colleagues present a new neural network to enhance underexposed photographs. The approach differs from earlier studies by incorporating intermediate illumination that connects the inputs to the desired enhancement outcome, leading to better capacity in the network to acquire complex photographic adjustments. This improvement allows the network to restore defined details, enhanced contrast, and realistic colors in the final output. Numerous experiments have shown that the network can effectively process images that were previously challenging [12].

CNN, one of the representatives of deep learning, has been one of the more popular optimization methods in CV. Li et al. found that multiplicative noise always accompanies synthetic aperture radar and laser imaging processes. Therefore, the researchers used an alternating direction multiplicative algorithm based on a deep convolutional network denoising prior to solve this problem. After experiments, the final experimental results showed that the images obtained by this method were visually better [13]. Wu et al. found that the random placement of parts during fabrication made it difficult for the robot to identify and manipulate them. The grasping process used CNN to extract the key points of the dispersed parts, which improves the success rate. Experimental results demonstrated that the method was effective in helping the robot to identify and perform operations [14]. Shalash proposed a system for estimating driver fatigue using only one EEG signal channel. The system converted the received black and white EEG into color images and then used a CNN to identify whether the driver is fatigued or not. The researchers found the three most accurate EEG signal channels. The system then experimentally estimated the three channels with an accuracy of 94.33%, 92.57% and 93% respectively. According to the testing findings, the system can efficiently and accurately assess the driver's level of weariness [15].

In conclusion, LLIE research has so far produced a sizable number of outcomes as one of the primary study objects in the CV field. However, CNNs have demonstrated strong performance in image processing. The use and optimization of CNNs in LLIE, however, clearly falls short. For this reason, the study proposes a CNN-based approach to LLIE and optimizes the CNN

with the help of LBO feature extraction. The study aims to provide a more efficient method for LLIE.

2. CNN-based LLIE approach

Images acquired in poorly lit environments can be a significant obstacle to subsequent CV tasks due to their low brightness and contrast [16]. The LLIE problem needs to improve IQ as much as possible while maintaining a balance of various factors such as brightness, contrast, artefacts and noise. This study uses the stability of LBP features under illumination changes to provide directional guidance on the various stages of LLIE. The AM introduced in the network structure effectively aids the capture of low-light image features. Furthermore, by incorporating FAM into the network, its ability to enhance feature representation based on correlation between features is enhanced.

2.1. CNN optimization based on LBP feature extraction

In the current image processing field, CNN is a highly favoured network model [17]. It has received a lot of application in fields including image fusion, image segmentation, and target recognition since it demonstrated good performance on CV problems. Typically, a single CNN is often unable to effectively handle LLIE tasks. Therefore, this study chooses to employ a multi-stage and multi-level CNN structure to perform LLIE. The study draws on the robustness of LBP-based features under illumination variations to optimise the CNN. The study first utilised inter-stage recursive computation to train network F to learn a direct mapping from the low-light image to the enhanced image by performing T sub-expansion of this mapping, with stage t derived as shown in Equation (1).

$$x^t = f(x^{t-1}, y). \tag{1}$$

In Equation (1), x denotes the input to the network, i.e. the image block; y denotes the output predicted by the neural network; and $f()$ is a function in the neural network for mapping from input x to output y . The experiment feeds low-light images through the AM for comprehensive feature extraction. The different features are then integrated into the FAM. The network then passes through a recurrent and repetitive layer, and uses the recursive residual layer to extract local features. The recursive layer shares deep features at each stage and the recursive residual layer uses a recursive approach. Last, the increased image is output from the network. There are therefore five layers in the P-FANet network: the output layer, the feature aggregation, the residual layer and the output layer, the recursive layer. The derivation of the P-FANet network at stage t is shown in Equation (2).

$$\begin{cases} x^t = f_{OUT}(f_{RB}(r^t)) \\ r^t = f_{RC}(r^{t-1}, s^t) \\ s^t = f_{IN}(x^{t-1}, y) + f_{FA} \end{cases} . \tag{2}$$

In Equation (2), f_{IN} denotes the input layer; f_{FA} denotes the feature aggregation layer; f_{RC} denotes the cyclic layer; f_{RB} denotes the residual layer; and f_{OUT} . B denotes the output layer. From Equation (2), it can be seen that f_{IN} , f_{RB} and f_{OUT} are fixed in each recurrent stage, thus increasing the utilisation of parameters and thus improving the performance of the network. In the residual layer f_{RB} , the study designed the residual blocks in recursive form to perform the extraction of depth features while ensuring the speed of convergence. The jump connection between residual blocks is shown in Equation (3).

$$f_{t+1}(x) = ReLU(f_t(x) + f_{t-1}(x)). \quad (3)$$

In the output layer, it was found that using a single Structural Similarity Index measure (SSIM) as the loss function had a significant enhancement effect on the network performance. For this purpose, the study calculated the loss of the output image after each stage of enhancement with the image under normal illumination, as shown in Equation (4).

$$L = -SSIM(x^t, x^{gt}). \quad (4)$$

In Equation (4), x^t denotes the output of stage t ; x^{gt} denotes the image under normal illumination; and $-SSIM \in [-1, 0]$. In recent years, LBP has become popular in CV and image processing. LBP is effective in capturing the local texture of an image and is robust to light variations, exhibiting grey scale invariance [18]. This means that lighting changes do not affect the image description. The study therefore exploits the light insensitivity of LBP to enhance low-light images. The LBP feature calculation process for the images is shown in Fig. 1.

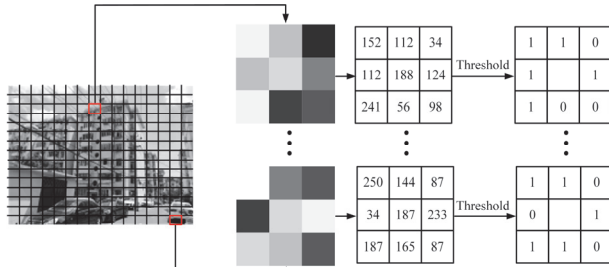


Fig. 1. LBP feature calculation process of images

As shown in Fig. 1, the process of calculating the LBP features of an image starts by dividing the window into smaller blocks, and then within each block, the LBP value for that block is calculated by comparing it with neighbouring pixels in clockwise order to obtain an 8-bit binary number. Next, the histogram is calculated for each block. The LBP algorithm provides an LBP code for each pixel, so that the LBP feature obtained after LBP processing of the image is itself a map. With a deeper understanding of how the LBP operation works, it can be seen that LBP focuses mainly on the peripheral pixels of the image, which can lead to a loss of local features. The image can be divided into blocks to help with this difficulty, and the average greyscale of each block can

then be used to represent the greyscale of that block. However, this method still ignores the local features within the block and may reduce the effectiveness of image enhancement. To ensure a balanced approach between global and local information, our study suggests incorporating global information during the extraction of local information and fusing the two. Additionally, the LBP features are re-evaluated for each region of the image. Fig. 2 depicts the enhanced LBP feature computation procedure.

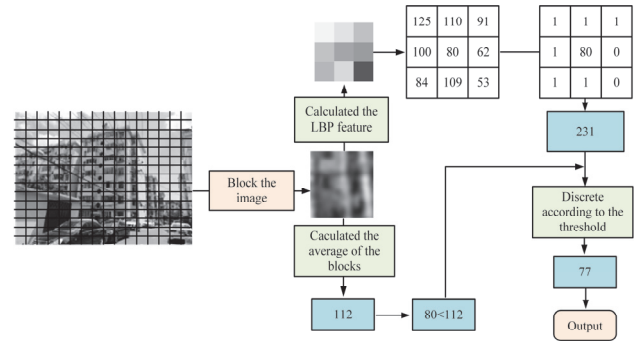


Fig. 2. Improved LBP feature calculation process

As shown in Fig. 2, the threshold value introduced when recoding the LBP, was also determined, and the expression for the threshold value is shown in Equation (5).

$$M = \frac{1}{p} \sum_{i=0}^{p-1} |-g_c + g_i|. \quad (5)$$

In Equation (5), M denotes the threshold value; g_c and g_i denotes the grey scale value of the central pixel point and the boundary pixel point, respectively; and p denotes the number of neighbouring pixels of the middle pixel. The expression is shown in Equation (6).

$$S(g_p - g_c) = \begin{cases} 0, & |-g_c + g_p| \leq M \\ 1, & |-g_c + g_p| \geq M \end{cases}. \quad (6)$$

In Equation (6), g_p is the greyscale value of the field pixel point. The binary code obtained from Equation (6) is converted into a decimal number as shown in Equation (7).

$$LBP_{p,R}^* = \sum_{i=1}^{p-1} S(g_p - g_c) 2^p. \quad (7)$$

2.2. MM-Net based LLIE model

To achieve a well-balanced approach to incorporating global and local information, our study recommends integrating global information into the extraction of local information and fusing the two. Furthermore, it is suggested to re-evaluate the LBP features for every region of the image. However, it does not perform well in extremely dark conditions [19]. In the feature extraction phase, improved LBP features are used to recover texture details. In the feature aggregation stage, deep contrast and colour information needs to be integrated by FAM. Global averaging is used in the study, followed by a fully

connected layer to reduce the feature size to 1/16th of its original size, a ReLU activation function, another fully connected layer, and a Sigmoid function to normalize the weights. In this way, the features of each channel can be compressed to a single value, a process that can be expressed in Equation (8).

$$F_{sq}(U_c) = \frac{1}{W \times H} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j). \tag{8}$$

In Equation (8), U_c denotes the feature mapping of channel C ; $u_c(i, j)$ denotes the attribute value at position (i, j) . The vector of weights that can be obtained after two fully connected layers is $1 \times 1 \times C$, which represents a feature map with one value per channel, i.e., each feature in a "channel" is compressed into a single value. This is usually obtained by global average pooling, i.e., each channel of a feature map of size $W \times H$ is averaged to obtain a value of 1×1 . If there are C such channels, then the size of the final feature map obtained is $1 \times 1 \times C$. The weights are then expanded in the width and height planes to give weights of $W \times H \times 3C$. In this step, the weights of 1×1 are first copied and expanded over the entire space of $W \times H$, which gives the weights of $W \times H \times C$. Next, this operation was performed three times, resulting in three times the number of channels, i.e., $W \times H \times 3C$. The relationship between the expanded weights and the feature mapping is shown in Equation (9).

$$F_{scale} = (U_c, \omega_c) = \omega_c \cdot U_c. \tag{9}$$

In Equation (9), ω_c denotes the estimated value of channel U_c . Finally, to further enhance the feature representation capability, the study reduced the number of

channels to $W \times H \times C$ using a 1×1 convolution. The FAM is shown in Fig. 3.

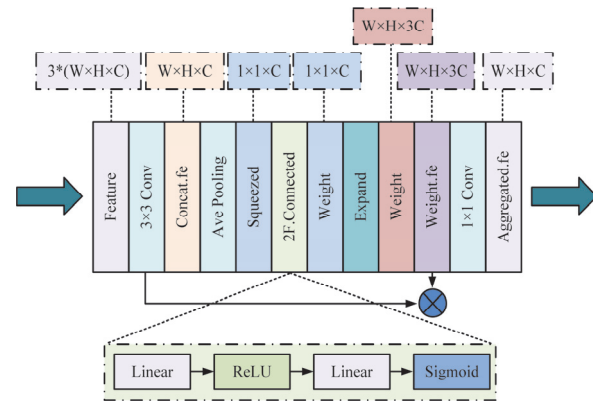


Fig. 3. Feature aggregation module

Fig. 3 depicts the flow of the feature aggregation stage. The shallow and deep characteristics of the image must be successfully combined during feature aggregation. Four feature maps acquired during the feature extraction process function as the input for feature aggregation. Downsampling these feature maps leads to a 512-channel feature map creation while containing the major portion of the feature data obtained at each stage. The feature aggregation stage's flow is shown in Fig. 3. The feature aggregation process must correctly blend the shallow and deep aspects of the image. The feature aggregation method takes as its input a group of four feature maps obtained during the feature extraction step. By downsampling these feature maps, a 512-channel feature map is produced. The majority of the feature data that was obtained at each stage is ensured to be present in the 512-channel feature map using this technique.

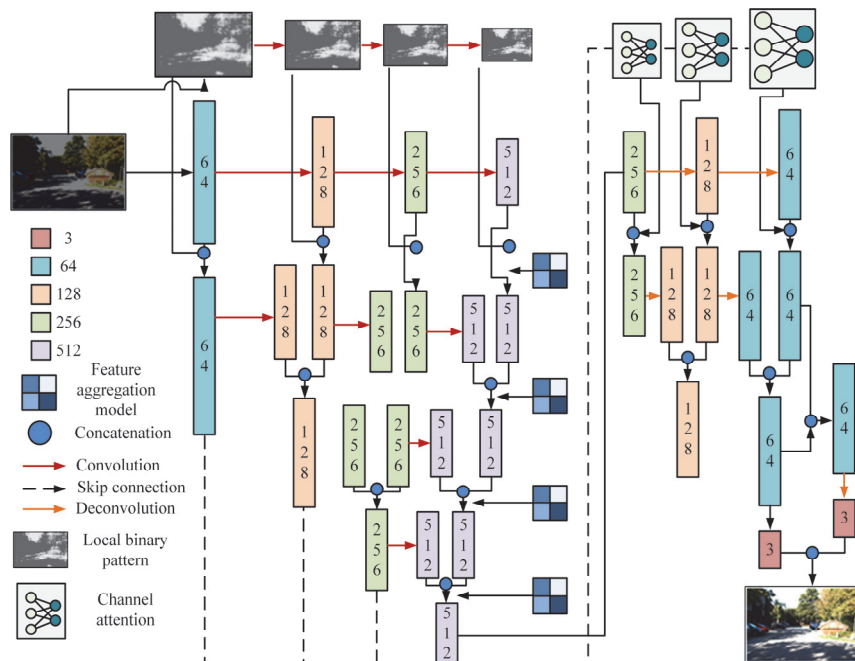


Fig. 4. MM – Net network model

Local and shallow features, such as colour and texture, are collected from the bottom up whereas global and deep features, as seen in Fig. 4, are obtained from the top down within the MM-Net network framework. The study also makes use of the stability of LBP features during variations in illumination. The study effectively integrates multiple features during the feature aggregation stage, simplifying the process of combining various features through the nesting of FAM throughout the downsampling phase. Furthermore, the study introduces channel AM, a mechanism that permits the

network to obtain image channel information, compensating for details lost during downsampling. This mechanism allows the network to learn more precisely. Depending on the importance of each channel in the task, AM reassigns differentiated weights to each channel, which helps guide the network to focus on valuable information while downplaying irrelevant content. This research study incorporates a channel attention module in the image enhancement phase for each of the 256, 128 and 64 channel feature mappings. The channel AM introduced by the study is shown in Fig. 5.

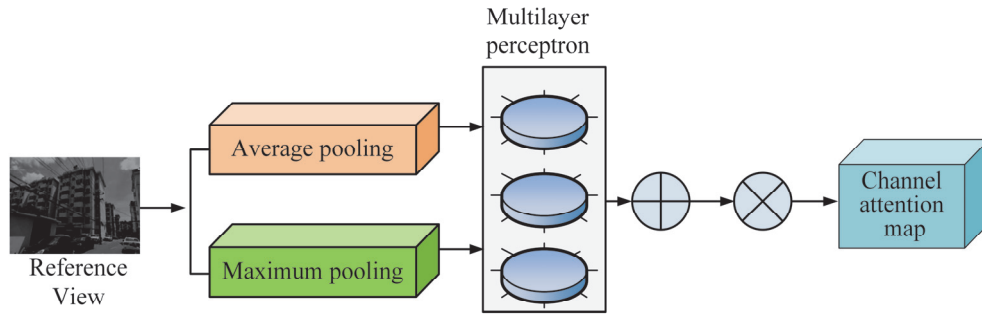


Fig. 5. Channel attention mechanism

The local information of each channel in the feature map is initially collected by global pooling during the channel AM process, as shown in Fig. 5. Then, the feature data of each channel in the feature map is formed at the fully connected layer and the weights of each channel are readjusted. Finally, after linear weighting, it proceeds to the next stage. The squeeze process, excitation process, and attention process are the three main divisions of the channel AM. where Equation (10)'s representation of the squeezing process is used.

$$z_c = F_{sq}(u_c) = \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \frac{1}{H \times W}. \quad (10)$$

In Equation (10), $U \in R^{H \times W \times C}$, the height and width of the image is denoted by H and W respectively; the feature information is denoted by $u_c(i, j)$. The calculation of the excitation process is shown in Equation (11).

$$S = \sigma(W_2 \delta(W_1 z_c)). \quad (11)$$

In Equation (11), δ denotes the ReLU activation function; σ denotes the Sigmoid function; W_1 and W_2 denote the weights of different dimensions; and S denotes the obtained excitation weights, using S to activate the individual image feature channels. Equation (12) depicts the process.

$$\tilde{X}_c = u_c \cdot S = F_{scale}(u_c, S_c). \quad (12)$$

In Equation (12), u_c denotes the channel value; S_c denotes the weighting value. In addition, subtle noise during image enhancement can significantly affect the quality of the enhancement, and numerous enhancement algorithms may amplify noise within the image. TVLoss is a widely used regularization term that acts as a noise

suppressor by combining it with other loss functions. It is calculated as shown in Equation (13).

$$J_{T0}(u) = \int_{\Omega_u} |\nabla u| dx dy = \int_{D_u} \sqrt{u_x^2 + u_y^2} dx dy. \quad (13)$$

In Equation (13), $u(x, y)$ denotes the pixel value, D_u denotes the support domain; $x, y \in \Omega$; and ∇ denotes the gradient operator.

3. Performance verification of improved LLIE models

To ensure that the experimental environment does not produce errors in the results, the same computer equipment was used for the simulation tests. Tab. 1 displays information on the experimental setting.

Tab. 1. Experimental environment information

Name	Configuration
Video card	GTX 1080ti
CPU	Inter(R)Core(TM)i5-7200U
Gpu-accelerated library	CUDA 10.0
Memory	64 GB
Operating system	Windows 10
Platform	MATLAB R2014a

Low-light images from publically accessible datasets, including the NPE dataset and the ExDark dataset, were used for the trials to verify the performance of the MM-NetLLIE model put out in the article. The datasets were split into 3 sets: 1300 images for training, 100 for validation, and 100 for testing. The model's initial learning rate was 0.0001, and the study was conducted with 3 different Epoch lengths: 60, 80, and 100, while keeping the images at $512 \times 512 \times 3$. Empirical evidence indicated that image enhancement was most effective in terms of color, brightness, and contrast when the Epoch was set to 80. The

comparative methods encompass approaches derived from Literature 11, Literature 12, and the MM-Nett model proposed in this study for low-light image enhancement. Fig. 6 displays the recall comparison results.

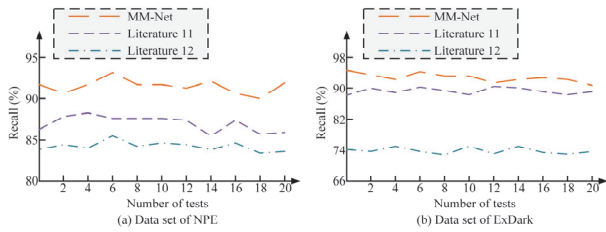


Fig. 6. Comparison of algorithm recall

In the LLIE study, recall indicates the proportion of correctly enhanced pixels of an image to the total pixels. In Fig. 6(a), the MM-Net algorithm has the best recall performance in all 20 experiments, which is significantly higher than the other two types of image enhancement algorithms; the Literature 11 algorithm has a slightly higher recall than the Literature 12 algorithm. In Fig. 6(b), the MM-Net algorithm has the best recall performance in all 20 experiments, slightly higher than the Literature 11 algorithm. In the two datasets, the mean recall rates of MM-Net, Literature 11 and Literature 12 were 94.51%, 88.12% and 78.94% respectively. The experimental findings demonstrate that the MM-Net method enhances images more accurately than both the Literature 11 and the Literature 12 algorithms. Comparing the accuracy of image enhancement of the three algorithms, the results are shown in Fig. 7.

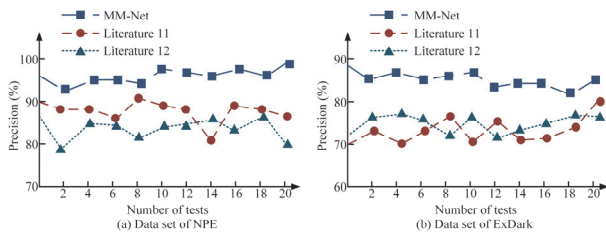


Fig. 7. Comparison of algorithm precision

In terms of image enhancement, accuracy is the intersection of the quantity of correctly enhanced pixels and the actual value. The accuracy rates of various image enhancement techniques in the NPE dataset are contrasted in Fig. 7a. In this figure, the average accuracy of Literature 12 is 86.98%, the average accuracy of Literature 11 is 83.13%, and the accuracy of the MM-Net algorithm is above 91% in all 20 experiments. In the

ExDark dataset, Fig. 7b compares the accuracy of various picture enhancement techniques. The MM-Net algorithm, which has an accuracy rate above 82%, outperformed the other two types of algorithms in all 20 experiments, as shown this figure; the accuracy rates of Literature 12 and Literature 11 are comparable. According to the experimental findings, the MM-Net method performs more accurately than the Literature 12 algorithm and the Literature 11 algorithm when it comes to picture improvement. Fig. 8 displays the results of comparing the DICE coefficients of the three picture enhancement techniques.

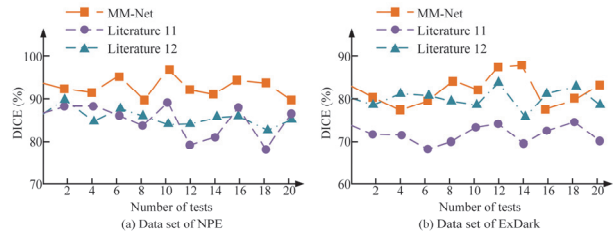


Fig. 8. Comparison of algorithm DICE

The DICE coefficient, which calculates the intersection of the enhancement result and the true value twice, is used to assess the success of image enhancement. Fig. 8(a) shows a comparison of the DICE coefficients of different image enhancement algorithms on the NPE dataset. As can be seen from the Fig., the DICE coefficients of the MM-Net algorithm are both higher than the other two types of algorithms, with DICE coefficients above 90%; the DICE coefficients of the Literature 12 and Literature 11 algorithms are close to each other with no significant difference. In the photos from the ExDark dataset, Fig. 8(b) compares the DICE coefficients of various image improvement techniques. In this figure, the DICE coefficients of MM-Net and Literature 11 algorithms are significantly higher than those of Literature 12 algorithm. The experimental results show that in the NPE dataset, the MM-Net algorithm enhances the image DICE coefficient significantly better than the Literature 12 algorithm and the Literature 11 algorithm; in the ExDark dataset, the DICE coefficient is better than the Literature 12 algorithm, but the difference with the Literature 11 algorithm is not obvious. The average values of accuracy, recall and DICE coefficients in the experiments were calculated separately, and the specific data results were obtained as shown in Tab. 2.

Tab. 2 The average of accuracy rate, recall rate and DICE coefficient

Data set	Precision (%)			Recall (%)			DICE (%)		
	Literature 12	Literature 11	MM-Net	Literature 12	Literature 11	MM-Net	Literature 12	Literature 11	MM-Net
NPE	86.88	80.26	94.16	87.31	87.22	91.72	83.45	84.61	92.39
ExDark	70.08	72.65	82.67	72.45	90.62	94.31	82.12	72.23	83.72

As shown in Tab. 2, MM-Net has the best accuracy, recall and DICE coefficient among the compared algorithms in the NPE dataset. In both datasets, compared to the Literature 12 and Literature 11 algorithms, MM-Net

improved the accuracy mean by 10.52% and 13.42%; the recall by 16.24% and 4.98%; and the DICE coefficient by 6.32% and 12.87%, respectively. The MM-Net algorithm outperformed all other algorithms in the picture

enhancement studies, as shown by the data in the table, proving the correctness of this research.

The Peak Signal-to-Noise Ratio (PSNR) is a very important evaluation metric in LLIE research, as the PSNR indicates better enhancement by evaluating the deviation from the pixel to pixel. The study uses two datasets to train the image enhancement model, and the PSNR curves are shown in Fig. 9.

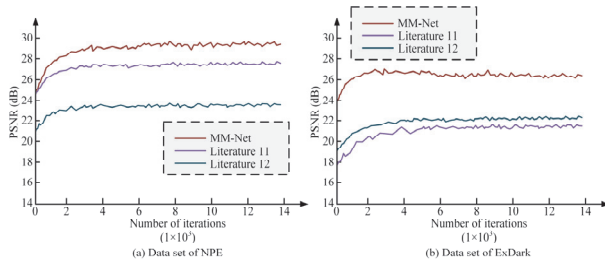


Fig. 9. Comparison of PSNR training results

The initial PSNR of the MM-Net and Literature 11 are comparable to one another and greater than that of the Literature 12 algorithm, as shown in Fig. 9a, which compares the training outcomes of PSNR in the NPE dataset. The initial PSNR of MM-Net is 24.85dB; the initial PSNR of Literature 11 is 24.79dB; and the initial PSNR of Literature 12 is 21.12dB. The training comparison in the ExDark dataset is shown in Fig. 9(b). The graph shows that while BM3D and NLM reach convergence only after 4500 iterations, the MM-Net algorithm reaches it after only 2500. The MM-Net strategy shows quicker convergence than the other two algorithms, and the experimental results validate its effectiveness.

To compare the structural similarity (SSIM) performance of the algorithms, the study added additive Gaussian white noise with standard deviations of 15, 25 and 50 to the original test images to generate the image enhancement effects used to test the algorithms in a noisy environment. The average SSIM box plots of the different algorithms on the test set are plotted as shown in Fig. 10.

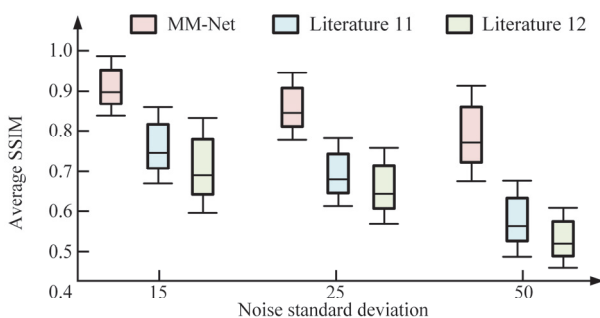


Fig. 10. Average SSIM of different algorithms on the test set

The average SSIM performance of the algorithm for various noise standard deviations is shown in Fig. 10. The improved image has the highest resemblance to the original image, according to the Fig., where MM-Net has the highest SSIM values at noise standard deviations of 15, 25, and 50. At noise standard deviations of 15, 25, and 50, MM-Net had an average SSIM improvement over

Literature 11 of 17.12%, 23.83%, and 54.29%, respectively. At noise standard deviations of 15, 25, and 50, respectively, MM-Net increased the average SSIM by 31.11%, 30.77%, and 53.18% when compared to Literature 12. Overall, MM-Net performs the best in terms of enhancement, followed by Literature 11, and Literature 12 performs rather poorly in terms of denoising.

Conclusion

Images taken in low-light situations frequently experience quality degradation, including increased noise and decreased contrast, which presents difficult IA and processing problems. This study suggests a CNN-based LLIE strategy to address these issues. The improvement approach is based on stability analysis of LBPs, and channel AM effectively improves the network's capacity to acquire low-light picture data. The expression of characteristics was enhanced further by integrating FAM. In the experimental results, the mean recall rates of MM-Net, Literature 11, and Literature 12 were 94.51%, 88.12%, and 78.94%, respectively, outperforming the Literature 11 approach and the Literature 12 algorithm. In particular, M-Net boosted the average SSIM by 17.12%, 23.83%, and 54.29% with noise standard deviations of 15, 25, and 50, respectively. When compared to Literature 12, MM-Net enhanced the average SSIM by 31.11%, 30.77%, and 53.18% with noise standard deviations of 15, 25, and 50, respectively. The results showed that the study, especially in loud conditions, showed higher performance when mentioning LLIEs. To improve the efficiency of LLIE even further, future research can investigate additional feature extraction and fusion techniques.

References

- [1] Gao K, Akbarpour HA, Fraser J, Nouduri K, Bunyak F, Massaro R, Seetharaman G, Palaniappan K. Local feature performance evaluation for structure-from-motion and multi-view stereo using simulated city-scale aerial imagery. *IEEE Sens J* 2020; 21(10): 11615-11627. DOI: 10.1109/JSEN.2020.3042810.
- [2] Fan X, Lei J, Liang J, Fang Y, Cao X, Ling N. Unsupervised stereoscopic image retargeting via view synthesis and stereo cycle consistency losses. *Neurocomputing* 2021; 447(11): 161-171. DOI: 10.1016/j.neucom.2021.02.079.
- [3] Zhang S, Li H, Kong W. Object counting method based on dual attention network. *IET Image Process* 2020; 14(8): 1621-1627. DOI: 10.1049/iet-ipr.2019.0465.
- [4] Lévy B, Mohayaee R, Hausegger SV. A fast semi-discrete optimal transport algorithm for a unique reconstruction of the early Universe. *Mon Not R Astron Soc* 2021; 501(1): 1165-1185. DOI: 10.1093/mnras/stab1676.
- [5] Belizario IV, Linares OC, Neto J. Automatic image segmentation based on label propagation. *IET Image Process* 2021; 15(15): 2532-2547. DOI: 10.1049/ipr2.12242.

- [6] Sandoub G, Atta R, Ali HA, Abdel-Kader RF. A low-light image enhancement method based on bright channel prior and maximum colour channel. *IET Image Process* 2021; 15(8): 1759-1772. DOI: 10.1049/ipr2.12148.
- [7] Yang J, Xu Y, Yue H, Jiang Z, Li K. LLIE based on Retinex decomposition and adaptive gamma correction. *IET Image Process* 2021; 15(5): 1189-1202. DOI: 10.1049/ipr2.12097.
- [8] Yang H-H, Huang K-C, Chen W-T. LAFFNet: A lightweight adaptive feature fusion network for underwater image enhancement. *IEEE International Conference on Robotics and Automation* 2021; Corpus ID: 233715102. DOI:10.1109/ICRA48506.2021.9561263
- [9] He K, Tao D, Xu D. Adaptive colour restoration and detail retention for image enhancement. *IET Image Process* 2021; 15(14): 3685-3697. DOI: 10.1049/ipr2.12223
- [10] Azam MA, Khan KB, Ahmad M, Mazzara M, Khattak D, Multimodal K. Medical image registration and fusion for quality enhancement. *Comput Mater Contin* 2021; 68(1): 821-840. DOI: 10.32604/cmc.2021.016131.
- [11] Lore KG, Akintayo A, Sarkar S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recogn* 2017; 61: 650-662. DOI: 10.1016/j.patcog.2016.06.008.
- [12] Wang R, Zhang Q, Fu C W, Shen X, Zheng W S, Jia J. Underexposed photo enhancement using deep illumination estimation. *Proc IEEE/CVF Conf on Computer Vision and Pattern Recognition* 2019: 6849-6857. DOI: 10.1109/CVPR.2019.00701.
- [13] Li Y, Hu J, Ni G, Zeng T. Deep CNN denoiser prior for blurred images restoration with multiplicative noise. *Inverse Probl Imag* 2023; 17(3): 726-745. DOI: 10.3934/ipi.2022075.
- [14] Wu X, Li P, Zhou J, Liu Y. A cascaded CNN-based method for monocular vision robotic grasping. *Industrial Robot* 2022; 49(4): 645-657.
- [15] Shalash WM. A deep learning CNN model for driver fatigue detection using single EEG channel. *J Theor Appl Inf Technol* 2021; 99(2): 462-477.
- [16] Zeng Z, Sun S, Sun J, Yin J, Shen Y. Constructing a mobile visual search framework for Dunhuang murals based on fine-tuned CNN and ontology semantic distance. *Electron Libr* 2022; 40(3): 121-139. DOI: 10.1108/EL-09-2021-0173.
- [17] Wu J, Zhang Y, Luo C, Yan L, Shen X. A modification-free steganography algorithm based on image classification and CNN. *Int J Digi Crime Forens* 2021; 13(3): 47-58. DOI: 10.4018/IJDCF.20210501.0a4.
- [18] Yang Y, Song X. Research on face intelligent perception technology integrating deep learning under different illumination intensities. *Journal of Computational and Cognitive Engineering* 2022, 1(1): 32-36. DOI: 10.47852/bonviewJCCE19919.
- [19] Nsugbe E. Toward a self-supervised architecture for semen quality prediction using environmental and lifestyle factors. *Artif Intell Appl* 2023; 1(1): 35-42. DOI: 10.47852/bonviewAIA2202303.

Author's information

Jian Guo (b. 1973) graduated from Steel and Metallurgy from Guizhou University of Technology in 1994 and a master's degree in Computer Application from Guizhou University in 2006, majoring in Digital Image Technology Research. Work experience: from 1994 to 2003, worked as an assistant engineer at Shuicheng Iron and Steel (Group) Company. From 2006 to present, worked as an associate professor at the School of Information Engineering, Modern Education Technology Center, Xiamen Ocean Vocational and Technical College. He has published 10 academic papers, 4 academic works and textbooks, 8 research projects, 1 patent, and 2 academic awards. E-mail: mycyber@126.com

Received August 17, 2023. The final version – November 03, 2023.