Low-light Image Enhancement Method Based On Attention Map Network

Mengfei Wu 1,2, Xucheng Xue 1,*, Taiji Lan 1, Xinwei Xu 1,2

1 Changchun Institute of Optics,Fine Mechanicsand Physics,Chinese Academy of Sciences, No. 3888, Southeast Lake Road, Changchun, Jilin, China
2 University of Chinese Academy of Sciences, Beijing, China

Abstract: Low-light image enhancement is a crucial preprocessing task for some complex vision tasks. Target detection, image segmentation, and image recognition outcomes are all directly impacted by the impact of image enhancement. However, the majority of the currently used image enhancement techniques do not produce satisfactory outcomes, and these enhanced networks have relatively weak robustness. We suggest an improved network called BrightenNet that uses U-Net as its primary structure and incorporates a number of different attention mechanisms as a solution to this issue. In a specific application, we employ the network as the generator and LSGAN as the training framework to achieve better enhancement results. We demonstrate the validity of the proposed network BrightenNet in the experiments that follow in this paper. The results it produced can both preserve image details and conform to human vision standards.

Keyword: Image Enhancement; Neural Network; Attention Mechanism; Attention Map Network; LSGAN

1. Introduction

The imaging capabilities of mobile electronic devices have substantially improved in recent years as a result of the hardware's quick development, and the imaging devices' output quality has also steadily increased. The photographs captured by the image formation system still suffer from considerable issues with inadequate contrast and brightness at night, in scenarios with uneven illumination, and when using excessive camera exposure conditions. The majority of these images also have significant background noise. Such images impact future high-level visual tasks including target detection, recognition, and image segmentation in addition to producing a very bad visual experience. Consequently, it is crucial to utilize the appropriate image algorithms to improve low-light photos. We divide low-light image enhancement approaches into two categories:

- Deep Learning-Based Techniques
- Traditional Image Enhancement and its improvement strategies.

We'll go into more detail regarding the advantages, disadvantages, and current state of the research on similar techniques in this section.

A. Traditional Image Enhancement Methods

For low-light photos, there are numerous traditional methods of image enhancement. The following four types are mostly introduced in this article.

1). Gray-scale Transformation-Based Image Enhancement Methods

The gray value of the original image can be transformed by the gray transformation function in image processing technology to improve the image contrast since the low-illumination image has the characteristics of an uneven pixel distribution. Typical techniques for gray-scale transformation include Gamma Correction, logarithmic transformation, etc. Although simple to use, these techniques typically have little impact. Therefore, some academics suggest improving them by using the partition transformation method [1, 2] on the basis of these methodologies. This technique primarily splits the original image's gray value into many intervals by a threshold and then processes each interval using the appropriate transformation function.

2). Method of Image Enhancement Based on Histogram

One of the techniques that practically most frequently applied is histogram equalization (HE). The specific measure in this method is to enlarge the histogram of the original image so that its gray level covers a wider gray range. This method is mostly based on the histogram distribution information of the image. The drawback of HE is that it can overexpose, cause color alterations, and eat up fine image detail. These issues arise when the method blends minor features with a sparse amount of pixels while also extending the contrast of the original image. [3-5] suggested a better histogram equalization method to address this drawback and produced promising outcomes.

3) Retinex theory-based technique for Image Enhancement

The Retinex algorithm's fundamental premise is that variations in light intensity have no impact on an object's color. The illumination map and reflection map can be separated out of the image that the optical system captured. These algorithms' central idea is to estimate the illumination intensity of the image using the suggested prior information [6-9].

4) Partial Differential Equation-based technique for Image Enhancement

Recent years have seen the emergence of a class of methods for image enhancement based on partial differential equations (PDE). These algorithms' steps are as follows:

- Taking some image enhancement strategy and mathematically model the process, which produces an energy functional.
- A partial differential equation is created by solving the functional extremum using the tools of calculus of variations.
Gradient descent flow is typically used for iterative solutions to partial differential equations that can be solved numerically.

The main idea behind the image enhancement method based on the theory of PDE is that the original image's gradient field must be improved in order to obtain a target gradient field, and that an image must then be found with the best approximation of both the gradient field and the target gradient field in order to achieve image enhancement [10–12]. The particular gradient processing strategy has a significant impact on the actual outcome of the gradient field-based image enhancement technique, and a poor gradient processing strategy may even have the opposite effect. Other common drawbacks of such methods include the extensive computation required during the iterative process and the "staircase effect" produced by the image.

B. Deep Learning-based technique for image enhancement

The main idea of image enhancement methods based on deep learning is to use the highly nonlinear characteristics of neural networks to train a network that can be used for image enhancement. Combining the research findings in this area, we discovered that the majority of deep learning-based image enhancement techniques fall into one of the three training-form categories listed below: supervised learning, weakly supervised learning, or self-supervised learning. These three method categories will be covered in this section.

1) Supervised Learning Methods (SL)

Many researchers train the Supervised Learning-based network using both low-light and normal-light images. We refer to these pictures as paired datasets. [13] was the first achievement to propose the use of neural networks for image enhancement. To process low-light image patches, the author of this paper proposes a seven-layer sparse autoencoder with hidden layers. The author then uses a specific step to combine the processed image patches into a full image. The CRENET network structure, which can be used with traditional methods or Learning-based techniques, was proposed by [14]. On the basis of maintaining the initial enhancement results further improve the image's contrast and brightness.

In [15], a two-stage augmentation network is suggested to better balance the image's smooth and texture components.[16–18] achieved image enhancement by combining Retinex and deep learning. The specific approach is to divide the image into illumination map and reflection map using Retinex theory. The illumination maps are generated by network through paired training. We can simply multiply the reflectance maps and illumination maps which generated by the enhancement network to obtain an enhanced image. [19] proposed an end-to-end image enhancement network that associates the low-light images with the label images rather than learning the mapping from low-light images to normal-light images directly, which also produces better results.

2) Weakly Supervised Learning Methods

For supervised learning-based image methods, It is very challenging to find paired images of the same scene taken at various times of the "low-light and normal-light" datasets. Since most of the paired datasets currently available are synthetic datasets or multi-exposure datasets. Thus reducing the algorithm's requirements for the dataset is especially important. The earliest instance of using unpaired images to improve low-light image quality is [20] and U-Net [21] network serves as the primary network of the GAN generator in this achievement. The authors use some skip connections in U-Net and the attention module in those skip connections enhances the feature map's intricate features significantly.

The methods adopted by [20] and [22] need to perform a matrix dot product operation on the output of the decoding layer and the attention maps which are proposed in the paper during the generator decoding process. That means the decoder needs to generate images under the guidance of the attention map.

3) Self-Supervised Learning Methods

[23] was the first to suggest using a self-supervised Retinex model with maximum information entropy to enhance low-light images. The author chooses to replace the V channel information in the HSV image with the maximum value channel, because they believe that the image can be denoised using its own information of maximum value channel and that the simplified Retinex model and the HSV model are somewhat equivalent. Results from the self-supervised method proposed in [23] can be obtained on a GPU in a matter of minutes.

4) Contribution in this paper:

Our primary contributions in this paper are the following two ideas:

(1) We suggest a new attention module that combines channel attention with multi-head self-attention in order to significantly enhance image detail features.
We propose a network of attention maps that incorporates the overall feature data of images. The network can support the attention module during the enhancement process, and the processed image effect is more according with human vision.

2. Proposed method

In the principle described in this paper, image enhancement is accomplished using the attention map guidance method. Generator network and discriminator network are the two sections of our training network framework.

**Generator:** We use the U-Net as the backbone network of Generator, the network realizes the conversion of the image from “low light” to “normal light” by using "encoding-decoding" for low-light images. We embed an Attention Map Net into the generator (AMNet). The network can direct the training of the backbone network and speed up the network's overall convergence while improving image details. Figure 2 depicts the network's overall structure. The network structure of generator is known as BrightNet.

**Discriminator:** Many convolutional layers are used by the discriminator as decoder layers to extract features layer by layer until the final feature map size is 1x1. The feature map is then stretched and sent to the fully connected layer to produce the judgment result. Backbone of Generator uses the U-Net as backbone network. U-Net can extract higher-dimensional semantic information and compress images more effectively than CNN alone. We achieve feature fusion of various depths using the cross-layer skip connection.

The encoders and decoders of Generator, which both have four sub-layers, use a BatchNorm layer for normalization and a ReLU function as the activation function. Each encoder layer has two convolutional layers and one maximum pooling layer. Since experiments show that upsampling with inverted convolution will result in a "blocky" graininess that is not conducive to the generation of details, we use the upsampling layer in the decoder sub-layers rather than the transposed convolutional layers. A convolutional layer is required after upsampling in order to combine the spatial and channel information of the image. The encoding layer also makes use of the BatchNorm layer and ReLU activation function. The last layer of the backbone network adopts the Tanh activation function in accordance with the input image data normalized by the (-1,1) interval, which can hasten network convergence.

**Attention Module:** As mentioned above, the generator's backbone network is a symmetric U-Net network. To prevent the issue of information loss during layer-by-layer down-sampling, we add some skip connections at symmetrical points in the backbone network.

The multi-head self-attention(MHSA)[24,25] of Transformer and channel attention mechanism are both absorbed in the Attention Module which uses channel stacking to achieve feature fusion. In this module, we set up multiple convolution branches in order to accommodate the various theories for extracting information from various convolution kernels. Figure 3 depicts the precise structure of this module. We use the MHSA to choose features on their own after extracting features using five convolution channel paths of 3x3, 5x5, 7x7, 11x11,13x13. It should be mentioned that various patch sizes correspond to various feature map sizes. Prior to using MHSA, it is also necessary to use a convolutional layer which kernel size is 1x1 to reduce the dimensionality of features maps. After improving multi-branch features with multi-head self-attention, the attention map is expanded using a two-layer convolutional neural network with a channel number of 1. Then we get the attention map by Sigmod activation function in the backbone network. Equations (1) to (2) describe the a MHSA procedure.

\[ Q, K, V = \text{Chunk}([c_{1x1}, c_{3x3}, c_{3x3}, \ldots, c_{3x3}]) \]

\[ \text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d}} \right)V \]

Finally, the feature map processed by multi-head self-attention also needs to be processed by the channel attention mechanism, and then passed through a network sub-layer which contains a single convolutional layer and a LeakyReLU activation function. The channel attention we used is derived from CBAM, and its structure is shown in Figure 4. Formula (3) illustrates the precise channel attention process:

\[ M_c(F) = \sigma(MLP(\text{AvgPool}(F)) + MLP(\text{MaxPool}(F))) \]

\[ F_{\text{avg}} \text{ and } F_{\text{max}} \text{ stand for maximum and average pooling features, respectively. The “shared MLP” shown in Figure 4 generates the final channel attention feature map.} \]

\[ \sigma \text{ represents the Sigmoid activation function that was used to produce the weight score.} \]

**Attention Map Net:** As mentioned in the introduction, the authors of [20] multiply an attention map pixel by pixel in all U-Net decoder layer output results. According to experiments, if we use this operation can greatly speed up the network convergence speed in paired images training, whereas it is difficult for the training process to converge if this operation is missing in unpaired images dataset. Our attention maps are learnable in comparison to [20] and [22], which make network have stronger generalization ability. The network that creates the attention map is known as the Attention Map Net (AMNet), and it is composed of four convolutional layers that are densely connected before each convolutional layer is followed by a GAM [26] attention module for focusing features. The obtained four feature maps are superimposed on the channel after the attention.
Fig 2. The network structure of the Generator

maps of the four corresponding convolutional layers have been obtained. The channel data is then fused through a sublayer to produce the attention map corresponding to the output channel of the U-Net encoder layer. A BN layer, a ReLU activation function, and a convolutional layer are among the sublayers. Finally, the obtained attention maps are converted into different sizes by using the maximum pooling layer, and the corresponding outputs of each decoder layer of U-Net are respectively multiplied pixel by pixel. The most significant distinction between this paper and [20, 22] is that the attention map produced by AMNet is multi-channel.

C. Loss Function

The loss function mainly includes three parts:

L2 Loss: This function is used for the constraint of pixel space, and it can suppress the phenomenon of image color distortion in color space.

\[
L_{\text{color}} = \frac{1}{C_i H_j W_i} \| \phi(i,j) - \rho(i,j) \|_2^2
\]  

(4)

Where \( C_i H_j W_i \) represents the number of image channels and the corresponding position of the pixels, \( \phi(i,j) \) and \( \rho(i,j) \) represent the corresponding position pixel values of the low-light image and the label image. The loss is classified as \( L_{\text{color}} \).

Perceptual Loss: The requirements of machine vision cannot be met by merely satisfying the L2 loss function. Thus a perceptual loss function must be added that can be used to assess the feature content loss of generated images and paired images. In particular experiments, we output the feature maps of normal light images and low light images using the "Conv1-1," "Conv2-1," "Conv3-1," and "Conv4-1" convolutional layers of the VGG19 network which is pre-trained by ImageNet dataset. Then we use the L2 loss function measures the difference between these two feature maps. This loss is classified as \( L_{\text{perception}} \).

Adversarial Loss: The logarithmic form of adversarial loss is used by the traditional GAN, but this adversarial loss has the issue of unstable training, particularly when using the sigmoid activation function in the discriminator's final layer, which may result in the issue of gradient disappearance. Therefore, we use the Least Squares adversarial loss (LSGAN) [27] function to train our model, the particular form of LSGAN loss is shown in (5):

\[
L_D = E_{X \sim P_{\text{real}}}(D_Ra(x_r, x_l) - 1)^2 + E_{X \sim P_{\text{fake}}}(D_Ra(x_r, x_l))^2
\]

\[
L_G = E_{X \sim P_{\text{real}}}(D_Ra(x_r, x_l) - 1)^2
\]

(5)

The overall loss function of the network is:

\[
\text{Loss} = L_G + L_{\text{color}} + L_{\text{perception}}
\]

(6)

Fig 5. Discriminator Structure
Fig. 6. Comparison of the effects of different algorithms.
### 3. Experiment

#### A. Dataset

The synthetic image dataset of LoL [17] is used for the generative network training. About 1000 artificial "low-light and normal-light" paired images are included in this dataset. We resize these 1000 pairs of paired images to 360 × 360 in size and sent all to the network for training.

In the testing phase, we select the DarkFace [28] dataset and the SID [29] dataset as the validation set. DarkFace is a dataset for nighttime face detection; it consists of 6000 nighttime images with a resolution of 1080 × 720 and Face tag data. The dataset "FUJI" and "SONY" of SID has corresponding long exposure images as well as extremely low light images created by combining several short exposures.

#### B. Experiment Detail

The experimental code in this paper is implemented based on Pytorch 1.8.0. The entire algorithm can be easily duplicated since the initialization weights and biases of each network layer are finished by default in accordance with the Pytorch framework and the random seed value is not set.

During the training phase, we use the Adam optimizer, set the batch size to 2 and learning rates to 0.0001 for both the generator and discriminator. The values of for Adam are

---

### Tab1. Indicators Values of different algorithms

| Method      | NIQE | CEIQ | LOE | EN  | SD  |
|-------------|------|------|-----|-----|-----|
| BPDHE       | 2.828| 2.01 | 148.77 | 6.15 | 36.47 |
| GLADNet     | 2.364| 5.38 | 188.31 | 7.35 | 45.15 |
| KinD        | 2.549| 5.23 | 256.05 | 7.28 | 43.49 |
| LIME        | 2.774| 6.54 | 339.88 | 7.69 | 59.79 |
| MBBLEN      | 2.994| 5.13 | 162.27 | 7.32 | 52.02 |
| SCI         | 2.721| 4.18 | 189.59 | 7.19 | 55.92 |
| MF          | 2.578| 4.49 | 436.01 | 6.97 | 36.93 |
| PM-SIRE     | 3.418| 4.18 | 420.98 | 6.23 | 87.42 |
| RetinexNet  | 4.036| 5.38 | 490.51 | 7.21 | 41.60 |
| Zero-DCE    | 2.732| 4.84 | 215.18 | 7.14 | 41.48 |
| Ours        | 2.494| 5.21 | 126.34 | 7.54 | 57.06 |

---

**Fig7. Comparison of the effects of different algorithms**

**Label**

**Input**

**FM**

**GLADNet**

**KIND**

**BPDHE**

**LIME**

**RetinexNet**

**SCI**

**Zero-DCE**

**MBBLEN**

**PM-SIRE**

**Ours**
0.9 and 0.999 respectively. The overall network training process is fairly stable because of the adversarial loss function of LSGAN. Additionally, the GPU we used is Tesla m40. It should be noted that the "Darkface" dataset is chosen to calculate the objective quality evaluation of the image enhancement effect, while the "FUJI" dataset in [29] is used to test the network's robustness.

| Input | Backbone Only | Without Attention | Without Attention | BrightenNet |
|-------|---------------|-------------------|-------------------|-------------|
| ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) | ![Image](image4.png) | ![Image](image5.png) |
| ![Image](image6.png) | ![Image](image7.png) | ![Image](image8.png) | ![Image](image9.png) | ![Image](image10.png) |
| ![Image](image11.png) | ![Image](image12.png) | ![Image](image13.png) | ![Image](image14.png) | ![Image](image15.png) |
| ![Image](image16.png) | ![Image](image17.png) | ![Image](image18.png) | ![Image](image19.png) | ![Image](image20.png) |
| ![Image](image21.png) | ![Image](image22.png) | ![Image](image23.png) | ![Image](image24.png) | ![Image](image25.png) |

**Fig 8.** Ablation Experiment Images

C. Objective assessment

In comparison experiments, this paper introduces six enhanced methods based on learning process (GLADNet, KinD, MBLLEN, RetinexNet, SCI, and Zero-DCE) and four state-of-the-art improved algorithms based on traditional methods (BPDHE, LIME, MF, and PM-SIRE). To demonstrate the enhancement effect of BrightenNet, the visual effect of the image is shown in Figure 6. Figure 6 shows that some image features produced by the MBLLEN, PM-SIRE, and SCI algorithms have excessive exposure. These overexposed regions obscure the image's original fine details, but the technique described in this paper can help to minimize this issue. Figure 6 shows that BrightenNet has some enhancement effects for the SID validation image, improving the image contrast and brightness.

However, proposed network only utilize the data from a single frame. Compared with the multi-frame image enhancement network, the effect of labeling images cannot be achieved temporarily. Since it is not very scientific to assess image quality merely subjectively, this paper uses objective evaluation indicators. NIQE, CEIQ, LOE, EN, and SD are some of the specific indicators. We determined the objective indicator values and plotted them in Table 1 after 30 epochs training.

In addition to the aforementioned indicators, Table 1 also includes the Entropy (EN) and Standard Deviation (SD) indicators that are frequently used in traditional image processing. EN is generally used to evaluate the richness of information contained in an image, and it is often used for quality evaluation of single-channel grayscale images. In the evaluation phase, we first calculated the three-channel EN value of the image and then calculated the arithmetic average. Since the test image data selected in this paper are all uint8 type (unsigned 8-bit binary integer data), the maximum value of the EN is 8; The SD of the image is an indicator describing the details of the image. The SD value of the image reflects the degree of dispersion between the pixel value of the image and its mean value. The quality and clarity of the image improve with increasing SD value.
The results of the algorithm in this paper have the smallest value among the LOE indicators, which suggests that BrightenNet has better image feature retention capabilities.

D. Ablation Study

Figure 6 demonstrates that the image enhancement effect cannot be obtained by using U-Net alone without the addition of any modules. According to the experimental findings, the network’s produced images only have blurred edges and few other details. The images also have significant color misalignment. The backbone network still does not exhibit the initial enhancement effect after an additional 100 training epochs.

The network also has a good enhancement effect, but it is accompanied by the block effect of the image if the attention map network is not included and only the backbone network and the attention module are included. The multi-head self-attention mechanism of the Transformers’ preprocessing process is what causes the block phenomenon. Every image needs to be evenly divided into multiple patches during the specific implementation process when using the multi-head self-attention mechanism. Because every patch must be through the embedding layer and the linear layer, Patch boundaries may be generated. Although this method creates relatively distinct boundaries between different portions of the image, as the training period is extended, these boundaries gradually vanish. According to our tests and experiments, the minimum training epoch required to eliminate the block boundary is 400 epochs.

The result of removing the attention module is the effect shown in column 4 of the ablation experiment. In the early stages of training, an overall improvement effect can only be achieved by using the attention map network. The image will gradually become clearer as the number of training epochs rises, even though the details may initially be hazy.

In the last column of Figure 6 is the final effect of our algorithm, we can see that the final result we get both avoids the blocky effect of the image and has a good enhancement effect, we think this is because our generator network combines the advantages of Attention Module and Atention Map Net. The Attention Map Net is responsible for improving the overall brightness of the image and Attention Module enhances the detail feature of image. Attention Map Net maintains the relationship between the features of each patch of the image during the learning process.

The initial enhancement effect of BrightenNet can be observed in the first epoch and is fairly obvious. It should be noted that each convolutional layer of the Generator in this paper adopts a convolution kernel with a convolution kernel size of $5 \times 5$, and the use of a $3 \times 3$ convolution kernel will greatly reduce the speed of network convergence. However, this phenomenon is not unlimited, and when we tried larger convolution kernel sizes, the network did not bring better results.

4. Conclusion

In this paper, we propose an attention module and attention map network-based image enhancement network. To speed up the convergence of the network, we use U-Net as the backbone network, the attention module to enhance the local details of the image, and the attention map network to enhance the entire image. Experiments demonstrate that BrightenNet performs well across a wide range of indicators, and when compared to other approaches, our suggested network is more reliable in a variety of image data sets. Our future work will focus on creating enhancement networks that can create images of a higher caliber and can be used in conjunction with more challenging visual tasks.

5. Reference

[1] M. Sahnoun, F. Kallel, M. Dammak, C. Mhiri, K.B. Mahfoudh, A.B. Hamida, A comparative study of MRI contrast enhancement techniques based on Traditional Gamma Correction and Adaptive Gamma Correction: Case of multiple sclerosis pathology, 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATISIP), 2018, pp. 1-7.

[2] H. Singh, A. Kumar, L.K. Balyan, G.K. Singh, Dark image enhancement using optimally compressed and equalized profile based parallel gamma correction, 2017 International Conference on Communication and Signal Processing (ICCSP), 2017, pp. 1299-1303.

[3] W. Chao, Y. Zhongfu, Brightness preserving histogram equalization with maximum entropy: a variational perspective, IEEE Transactions on Consumer Electronics, 51 (2005) 1326-1334.

[4] S. Kansal, S. Purwar, R.K. Tripathi, Trade-off between mean brightness and contrast in histogram equalization technique for image enhancement, 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), 2017, pp. 195-198.

[5] Z. Wei, H. Lidong, W. Jun, S. Zebin, Entropy maximisation histogram modification scheme for image enhancement, IET Image Processing, 9 (2015) 226-235.

[6] C. Liu, I. Cheng, Y. Zhang, A. Basu, Enhancement of low visibility aerial images using histogram truncation and an explicit Retinex representation for balancing contrast and color consistency, ISPRS Journal of Photogrammetry and Remote Sensing, 128 (2017) 16-26.

[7] M.K. Ng, W. Wang, A Total Variation Model for Retinex, SIAM Journal on Imaging Sciences, 4 (2011) 345-365.

[8] L. Sun, Y.-M. Huang, A modulus-based multigrid method for image retinex, Applied Numerical Mathematics, 164 (2021) 199-210.

[9] R. Zhang, X. Feng, L. Yang, L. Chang, C. Xu, Global sparse gradient guided variational Retinex model for image enhancement, Signal Processing: Image Communication, 58 (2017) 270-281.

[10] L. Huang, W. Zhao, B.R. Abidi, M.A. Abidi, A Constrained Optimization Approach for Image Gradient Enhancement, IEEE Transactions on Circuits and Systems for Video Technology, 28 (2018) 1707-1718.

[11] X. Sun, H. Liu, S. Wu, Z. Fang, C. Li, J. Yin, Low-Light Image Enhancement Based on Guided Image Filtering in Gradient Domain, International Journal of Digital Multimedia Broadcasting, 2017 (2017) 9029315.

[12] L. Yu, H. Xu, Y. Xu, X. Yang, Robust single image super-resolution based on gradient enhancement, Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, 2012, pp. 1-6.

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

[14] Y. Zhang, X. Di, B. Zhang, R. Ji, C. Wang, Better Than
[1] M. Sahnoun, F. Kallel, M. Dammak, C. Mhiri, K.B. Mahfoudh, A.B. Hamida, A comparative study of MRI contrast enhancement techniques based on Traditional Gamma Correction and Adaptive Gamma Correction: Case of multiple sclerosis pathology, 2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIPI), 2018, pp. 1-7.

[2] H. Singh, A. Kumar, L.K. Balyan, G.K. Singh, Dark image enhancement using optimally compressed and equalized profile based parallel gamma correction, 2017 International Conference on Communication and Signal Processing (ICSSP), 2017, pp. 1299-1303.

[3] W. Chao, Y. Zhongfu, Brightness preserving histogram equalization with maximum entropy: a variational perspective, IEEE Transactions on Consumer Electronics, 51 (2005) 1326-1334.

[4] S. Kansal, S. Purwar, R.K. Tripathi, Trade-off between mean brightness and contrast in histogram equalization technique for image enhancement, 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIIP), 2017, pp. 195-198.

[5] Z. Wei, H. Lidong, W. Jun, S. Zebin, Entropy maximisation histogram modification scheme for image enhancement, IET Image Processing, 9 (2015) 226-235.

[6] C. Liu, I. Cheng, Y. Zhang, A. Basu, Enhancement of low visibility aerial images using histogram truncation and an explicit Retinex representation for balancing contrast and color consistency, ISPRS Journal of Photogrammetry and Remote Sensing, 128 (2017) 16-26.

[7] M.K. Ng, W. Wang, A Total Variation Model for Retinex, SIAM Journal on Imaging Sciences, 4 (2011) 345-365.

[8] L. Sun, Y.-M. Huang, A modulus-based multigrid method for image retinex, Applied Numerical Mathematics, 164 (2021) 199-210.

[9] R. Zhang, X. Feng, L. Yang, L. Chang, C. Xu, Global sparse gradient guided variational Retinex model for image enhancement, Signal Processing: Image Communication, 58 (2017) 270-281.

[10] L. Huang, W. Zhao, B.R. Abidi, M.A. Abidi, A Constrained Optimization Approach for Image Gradient Enhancement, IEEE Transactions on Circuits and Systems for Video Technology, 28 (2018) 1707-1718.

[11] X. Sun, H. Liu, S. Wu, Z. Fang, C. Li, J. Yin, Low-Light Image Enhancement Based on Guided Image Filtering in Gradient Domain, International Journal of Digital Multimedia Broadcasting, 2017 (2017) 9029315.

[12] L. Yu, H. Xu, Y. Xu, X. Yang, Robust single image super-resolution based on gradient enhancement, Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, 2012, pp. 1-6.

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

[14] Y. Zhang, X. Di, B. Zhang, R. Ji, C. Wang, Better Than Reference in Low-Light Image Enhancement: Conditional Re-Enhancement Network, IEEE Transactions on Image Processing, 31 (2022) 759-772.

[15] J. Cai, S. Gu, L. Zhang, Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images, IEEE Transactions on Image Processing, 27 (2018) 2049-2062.

[16] L. Shen, Z. Yue, F. Feng, Q. Chen, S. Liu, J. Ma, MSRnet: Low-Light Image Enhancement Using Deep Convolutional Network, arXiv e-prints, (2017) arXiv:1711.02488.

[17] C. Wei, W. Wang, W. Yang, J. Liu, Deep Retinex Decomposition for Low-Light Enhancement, arXiv e-prints, (2018) arXiv:1808.04560.

[18] Y. Zhang, J. Zhang, X. Guo, Kindling the Darkness: A Practical Low-light Image Enhancer, arXiv e-prints, (2019) arXiv:1905.04161.

[19] R. Wang, Q. Zhang, C.W. Fu, X. Shen, W.S. Zheng, J. Jia, Underexposed Photo Enhancement Using Deep Illumination Estimation, 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 6842-6850.

[20] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, Z. Wang, EnlightenGAN: Deep Light Enhancement Without Paired Supervision, IEEE Transactions on Image Processing, PP (2021) 1-1.

[21] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Springer International Publishing, Cham, 2015, pp. 234-241.

[22] Q. Yang, Y. Wu, D. Cao, M. Luo, T. Wei, A Lowlight Image Enhancement Method Learning from Both Paired and Unpaired Data by Adversarial Training, Neurocomputing, 433 (2020).

[23] Y. Zhang, X. Di, B. Zhang, Q. Li, S. Yan, C. Wang, Self-supervised Low Light Image Enhancement and Denoising, arXiv e-prints, (2021) arXiv:2103.00832.

[24] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, arXiv e-prints, (2020) arXiv:2010.11929.

[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L. Kaiser, I. Polosukhin, Attention Is All You Need, arXiv e-prints, (2017) arXiv:1706.03762.

[26] Y. Liu, Z. Shao, N. Hoffmann, Global Attention Mechanism: Retain Information to Enhance Channel-Spatial Interactions, arXiv e-prints, (2021) arXiv:2112.05561.

[27] X. Mao, Q. Li, H. Xie, R.Y.K. Lau, Z. Wang, S.P. Smolley, Least Squares Generative Adversarial Networks, 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2813-2821.

[28] J. Liang, J. Wang, Y. Quan, T. Chen, J. Liu, H. Ling, Y. Xu, Recurrent Exposure Generators for Low-Light Face Detection, arXiv e-prints, (2020) arXiv:2007.10963.

[29] C. Chen, Q. Chen, J. Xu, V. Koltun, Learning to See in the Dark, arXiv, (2018) arXiv:1805.01934.