Research on image enhancement algorithm based on deep neural network algorithm

Qiang wen Lin¹, Liu mei Li¹*¹
¹Academic affairs office, Guangzhou College of Commerce, Guangzhou 511363, Guangdong, China

Corresponding author E-mail: 20102015@gcc.edu.cm

Abstract. Image denoising and enhancement technology plays an important role in remote sensor image processing. Due to the low brightness, signal-to-noise ratio and contrast of image processing, the traditional remote sensor is limited in obtaining object information. Based on the principle of biological segmentation algorithm, neural network algorithm technology can carry the advantages of image evaluation algorithms such as PSNR, IFC and SSIM, through the end-to-end training of nonlinear infinite approximation, and carry out the research of image fitting low light level image enhancement and denoising. On this basis, the effectiveness of the research results is evaluated. The results show that compared with the traditional remote sensor imaging technology, the PSNR value of the image constructed by the depth neural network algorithm is significantly improved by about 15.2401, the IFC value is improved by about 0.4523, and the SSIM value is improved by about 0.4500, which improves the low light level imaging quality to a higher extent, and provides excellent technical guidance for the remote sensor to obtain the object information at night.

Keywords: Deep Neural Network, Peak Signal To Noise Ratio, Information Fidelity, Structural Similarity

1. Introduction
The deep neural network algorithm is a nonlinear system; it can carry more processing algorithms from the input system to the output system; through the intermediate processing algorithm, it can realize the adaptive requirements of any object; so as to achieve the goal of any signal model [1]. The traditional enhancement algorithm of remote sensor can not take into account the brightness, contrast and signal-to-noise ratio of the object at the same time. Therefore, training a deep neural network algorithm can effectively carry any algorithm of image processing, and solve the problem of enhancement and denoising of low light level remote sensor by nonlinear infinite approximation fitting.
2. Network image enhancement algorithm model construction

2.1. Establishment of network image algorithm structure model

Deep neural network is a network structure composed of encoder and decoder. Two convolution layers and one pooling layer form a set of operation processes [2]. The encoder is a typical convolution neural network structure, and the decoder is a deconvolution neural network structure. The structure details are shown in Figure 1.

As shown in Figure 1, the specific coding process of the neural network system is shown. In order to ensure that the size of the image remains unchanged in each convolution process, it is necessary to set the size of the convolution kernel in advance and add the Lrelu activation function after the convolution. Then, we enter the maxpool layer of 2×2 with the step size of 2. Each lower layer samples 1/4 of the original image and increases the number of channels of the sampled image [3]. The deconvolution operation on the right is the same as that on the left. In order to ensure the integrity of the image information after deconvolution operation, in the process of image decoding and shrinking, the high-resolution local information generated by the deconvolution operation is spliced and fused with the feature image, and the fused feature image is convoluted twice to form a more abstract feature image. Finally, the final feature map is obtained by reducing the dimension of the feature map with 1×1 convolution kernel. In order to expand the number of training samples, increase the diversity of training samples, avoid over fitting, improve the performance of the model, and make the network more robust [4]. The data enhancement methods of random clipping, vertical flipping, horizontal flipping, rotation and scale transformation are studied. With the deepening of network learning, the network fitting ability will be stronger and stronger, the image features will be more and more, and the learning ability will be stronger and stronger. Therefore, the loss function and $L_2$ regularization are introduced to converge the error in this process, and the formula can be expressed as (1).

$$L(w, b) = \frac{1}{2N} \sum_{i=1}^{N} |\hat{y}_i(w, b) - y_i| + \frac{\lambda}{2} ||w||^2$$  \hspace{1cm} (1)

In formula (1), $n$ is the total number of training samples; $b$ is the bias parameter; $w$ is the calculation weight parameter; $y$ is the $i$-th reference image, $r$ is the weight penalty factor; $\hat{y}(w, b)$ is the $i$-th sample to predict the reconstructed image. Firstly, the input image of one iteration
of the network is the original short exposure image; secondly, the error is calculated, and the trained image is combined with the ground truth image by L1; finally, the average error is optimized by using adama, and the network parameters are updated to complete the iteration [5]. In order to reduce the error rate of the image, $512 \times 512$ image regions are randomly clipped for enhancement training. Set the initial learning rate as 0.0001, take 2000 learning frequency as the bound, 4000 times as the total number of training, and carry out iterative training simulation. The training error value can be obtained as shown in Figure 2.

![Figure 2. Loss function decline curve](image)

As shown in Figure 2, after the iterative training setting, the image loss rate decreases significantly after 2000 times, and the training ends after 4000 times. After the neural network training structure setting, the image generation loss rate can be reasonably controlled at about 0.05, meeting the image generation loss conditions.

### 2.2. Network low light level image quality evaluation method

Image quality evaluation can be divided into subjective evaluation and objective evaluation. On the basis of fully considering the operation cost, easy operation and embedded implementation factors, this paper studies the image quality evaluation methods from Peak signal to noise ratio (PSNR), Information fidelity criteria (IFC) and Structural similarity index measurement system (SSIM) to evaluate the image quality [6]. PSNR evaluates the image by comparing the error values between the original image and the generated image pixels. The evaluation formula is shown in (2).

$$PSNR = 10 \log_{10} \left( \frac{2^k - 1}{MSE} \right)$$

In formula (2), $m$ and $n$ respectively represent the width and height of the image, $m(i, j)$ and $n(i, j)$ respectively represent the gray value of the original image and the distorted image on the coordinate $(i, j)$. $k$ is the gray level of the image. The greater the distortion of the image, the smaller the PSNR. On the contrary, the higher the PSNR, the lower the distortion of the image [7].

IFC method can evaluate the quality of distorted image by comparing the distorted image with "good image". GSM is represented by two independent random fields, that is, for GSM, $C$ is represented by formula (3).

$$C = S \cdot \mu = \{ S_i \cdot U_i : i \in I \}$$

In formula (3), $i$ is the spatial subscript of random field; $\mu$ is Gaussian random field; 0 and $C_u$ are mean and variance respectively; $C_1$ and $U_1$ are vector values of $M$ dimension. GSM can also describe the process of image distortion and degradation in the process of image basic evaluation. The calculation model is formula (4).

$$D = gC + V = \{ g \cdot C_i + V_i : i \in I \}$$
In formula (4) distortion degradation calculation model, \( D = \{ D_i : i \in I \} \) is a corresponding subband in an iterative process of the distorted image; \( C \) is the subband of the input image; \( V = \{ V_i : i \in I \} \) is the additive noise, whose average value is about 0, through which the interference of noise on the image can be detected; variance is \( C_v = \delta_2 I \); \( g = \{ g_i : i \in I \} \) is the attenuation factor, which indicates the signal energy state of the distorted image lost. Because for each \( i \) and \( j \), \( D_i \) is independent of \( S_i \) and both of them are independent of \( S_j \), and \( C_i \) and \( V_i \) obey Gaussian distribution, then each subband can be expressed as formula (5).

\[
I(C^N; D^N / s^N) = \sum_{i=1}^{N} (h(g_i C_i + V_i / s_i) - h(V_i)) = \frac{1}{2} \sum_{i=1}^{N} \log_2 (1 + (\frac{g_i s_i \sigma_{s_i}^2}{\sigma_v^2}))
\]

In formula (5), \( h(x) \) represents the differential entropy of any random variable \( x \), where \( x \) obeying Gaussian distribution can be expressed as \( h(X) = \frac{1}{2} \log_2 (2\pi e \sigma^2) \). At this time, IFC is the cumulative result of all subbands, which is expressed as formula (6).

\[
IFC = \sum_{k=xsubbands} I(C_i^{N_i,k}; D_i^{N_i,k} / s_i^{N_i,k})
\]

Formula (6) is the image evaluation algorithm model considering human visual system. Compared with the peak signal-to-noise method, IFC evaluation method can keep the same with human visual sense, and can combine the advantages of objective evaluation and subjective evaluation at the same time, and it can be extended to the image evaluation system without parameters.

SSIM evaluation system mainly includes the adjustment of image structural features, contrast and brightness, which has an effect on image quality, and then affects the extraction of interested parts of the image by human brain [8]. If \( I \) is the brightness of the image, \( C \) is the standard deviation of the contrast of the image, and \( s \) represents the covariance of structural similarity of an image, the relationship between them can be expressed as formula (7).

\[
SSIM(X,Y) = l(X,Y) \cdot c(X,Y) \cdot s(X,Y)
\]

\[
l(X,Y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}
\]

\[
c(X,Y) = \frac{2\delta_{xy} + C_2}{\delta_x^2 + \delta_y^2 + C_2}
\]

\[
s(X,Y) = \frac{1}{H \times W - 1} \sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - \mu_x)(Y(i,j) - \mu_y)
\]

\( X \) and \( Y \) in formula (7) represent the processed image to be generated and the input original image respectively. \( \mu_x \) represents the average result of the image generated after the establishment, \( \mu_x = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X(i,j) \). \( \mu_y \) is the average result of the original image, \( \mu_y = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} Y(i,j) \). The structure similarity comparison calculation process \( sxy \) represents the covariance of the two images to be compared, \( \delta_{xy} = \frac{1}{H \times W - 1} \sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - \mu_x)(Y(i,j) - \mu_y) \).

\( H, W \) represents the brightness and width of the image to be generated. The constant \( C_2 \) is added in the calculation of image brightness and structure similarity to prevent the denominator from appearing 0 in the calculation process.
3. Evaluation of the applicability of network image enhancement algorithm model

3.1. SNR and fidelity evaluation of low light level image

![PSNR and Fidelity Comparison](image)

**Figure 3.** Comparison results of SNR and fidelity between traditional algorithm and neural network algorithm

Figure 3 (a) and Figure 3 (b) are the image after original low light level image equalization (HE), the image after gamma transform (GC), the image after CLAHE processing and the image after convolution neural network training respectively. The SNR and fidelity of several algorithms are compared after enhancement algorithm, Median Filter (MF) and Non Local Mean (NLM). It can be seen from Figure 3 that the SNR and fidelity of the three algorithms after non local averaging are improved. However, compared with the neural network algorithm, the SNR and fidelity of the three algorithms are lower. The SNR and fidelity of the neural network algorithm image are about 27.2300 and 0.7408 respectively. Compared with the traditional algorithm, the average improvement is about 15.2401 and 0.4523 respectively. It shows that the neural network algorithm can guarantee good effect in the aspect of SNR and fidelity in the process of image generation.
3.2. Evaluation of low light level image structure similarity and average structure similarity

Figure 4. Comparison results of structural similarity and average structural similarity between traditional algorithm and neural network algorithm.

Figure 4 (a) and Figure 4 (b) are the comparison results of traditional algorithm and neural network algorithm on image structure similarity and average structure similarity respectively. It can be seen from Figure 4 (a) that the structural similarity of the traditional algorithm decreases after median filtering on the basis of enhancement; after non local mean calculation, the structural similarity increases significantly, but it is significantly lower than that of the neural network algorithm with a structural similarity of about 0.9680, and the reduction range is about 0.4500. As can be seen from Figure 4(b), compared with the structural similarity, the average structural similarity of the traditional algorithm decreases by a low margin, while the average structural similarity of the neural network algorithm always keeps at a high level, which is about 0.9626, indicating that the neural network can maintain a relatively stable structural similarity and ensure the stability of image generation quality.

4. Conclusion

The traditional remote sensor image enhancement technology cannot take into account the factors of contrast, brightness and signal-to-noise ratio in the process of image acquisition, which seriously affects the quality of remote sensor's acquisition of object information. Through its high-intensity approximation fitting function and bearable algorithm technology, neural network algorithm has been comprehensively improved in brightness, contrast, SNR and structural stability. Through the effectiveness evaluation, the SNR, fidelity, structural similarity and average structural similarity of the neural network algorithm are 27.2300, 0.7408, 0.9680 and 0.9626 respectively; Compared with the traditional algorithm, the average values are increased by 15.2401, 0.4523, 0.4500 and 0.4800 respectively. Compared with the traditional image enhancement algorithm, it achieves a higher degree of performance improvement. However, due to the lack of training set collection in the research, the image generation appears hazy. Therefore, it is still necessary to increase the data collection in the future to improve the clarity of image collection.

References

[1] Zhang Hongying, Zhao Jindong. Retinexnet low illumination image enhancement algorithm in HSV space [J]. Progress in laser and optoelectronics, 2020, 57 (20): 294-301.

[2] Li Jianghua, Wang Kun. A low illumination image enhancement method based on convolutional neural network [J]. Journal of Jiangxi University of science and technology, 2020, 41 (05): 73-79.
[3] Hu Jun, Lu Hui Ling, Zhang Jie. Laser image enhancement algorithm based on convolutional neural network [J]. Laser journal, 2020, 41 (09): 147-150.

[4] Cai liming, Li Wei, Gao Yongfa, Zhang Wenliang, Zhang Yuqiang. Classification algorithm of aerospace electronic components based on multi feature image enhancement depth convolution neural network [J]. Navigation and control, 2020, 19 (02): 112-119.

[5] Bao Zhengfeng. Indoor low illumination image enhancement based on convolutional neural network [J]. Information communication, 2019 (09): 20-21.

[6] Wu ruoyou, Wang Dexing, Yuan Hongchun, Gong Peng, Chen Guanqi, Wang Dan. Low illumination image enhancement based on multi branch convolution neural network [J]. Progress in laser and optoelectronics, 2020, 57 (14): 197-207.

[7] Cheng Yu, Deng Dexiang, Yan Jia, fan Cîen. Weak illumination image enhancement algorithm based on convolutional neural network [J]. Computer applications, 2019, 39 (04): 1162-1169.