ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

Research on Low Illumination Image Enhancement Algorithm Based on Convolutional Neural Network

Xiaoting Niu*, Guanghai Zheng

School of Software, Dalian Jiaotong University, Dalian, China
905922247@qq.com
**Corresponding author*

Abstract: *Low illumination images have insufficient local and global light exposure, loss of structural and detail information, and are prone to generating a large amount of noise. The overall image is grayish or even completely dark, and people often cannot recognize the content of the image with the naked eye. Image enhancement technology aims to enhance image brightness, adjust image contrast, restore hidden details in the Fark, and enhance the utilization value of images through corresponding technical means.The traditional low illumination image enhancement methods mainly focus on Histogram equalization and Retinex methods. Based on the drawbacks of the traditional methods, this paper studies the low illumination image enhancement algorithm based on Convolutional neural network, builds a mathematical model, and lays the foundation for subsequent experimental research and application.*

Keywords: *Convolutional Neural Network, CNN, Low Illumination Image, Enhancement Algorithm, Evaluation Index, Dilation Convolution, Loss Function, Network Optimization*

1. Introduction

As a medium for information transmission, images can most intuitively display the scene information they contain. With the development of science and technology and the improvement of people's living standards, image acquisition devices such as mobile phones, digital cameras, and DSLRs have entered daily life. Through these devices, various images can be collected anytime and anywhere to meet the needs of entertainment and information transmission. However, due to the different equipment used for image capture and the different conditions under which the images are taken, the quality of the images obtained varies. In scenes such as thick fog, cloudy days, dark nights, or areas covered by shadows, and in situations where the image capture equipment has poor light filling ability, some relatively dark images are collected, which are generally referred to as low illumination images. Low illumination indicates weak illumination, and the surface area of objects in the scene is illuminated to a low level [1,2]. Low illumination images have insufficient local and global light exposure, loss of structural and detail information, and are prone to generating a large amount of noise. The overall imaging of the image is grayish or even completely dark, and people often cannot recognize the content of the image clearly with the naked eye, resulting in extremely low utilization value of such images.

Low illumination image enhancement refers to enhancing the image quality of images captured under dim light conditions, making them clearer or closer to the naked eye effect. Low illumination image enhancement technology aims to improve image brightness, adjust image contrast, restore hidden details in the dark, restore image edge information, and maintain color balance for color images through corresponding technical means. Traditional algorithms process all content in the image according to the same rules while enhancing low illumination in the video, which can easily lead to excessive stretching of the bright grayscale in the image, resulting in overexposure. In recent years, as the computing performance of hardware devices has been greatly improved, deep learning has exploded in many fields, bringing a new solution to the field of computer vision. Low illumination image enhancement technology based on Convolutional neural network has been developed rapidly, enhancing the input low illumination images, improving image quality, reducing computing speed, and providing medical image diagnosis Visual tasks such as object detection and image segmentation bring performance improvements and improve people's experience of using images.

ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

2. Traditional Low Illumination Image Enhancement Methods

The traditional low illumination image enhancement methods mainly focus on Histogram equalization and Retinex, and the subsequent methods are mainly aimed at improving these two methods. Image histogram is a fundamental attribute of an image and a statistical feature of the distribution of image pixel data. The Retinex method is based on the Retinex theoretical model, which believes that the observed image can be decomposed into illuminated and reflected images. The enhanced results are obtained by processing and restoring the decomposed images first.

2.1 Method Based on Histogram Equalization

The Histogram equalization method infers the gray cumulative distribution function of the image through the histogram of the image, and uses this as the input of the mapping function to adjust the gray level of the input image to ensure the uniform distribution of gray areas in the image, and improves the contrast and brightness of the image by dynamically adjusting the range.

For an input image *X*, the general distribution function $P(X_k)$ is defined as:

$$P(X_k) = \frac{n_k}{n} \tag{1}$$

In the above equation, *n* is the number of pixels in the entire image, n_k is the number of pixels with a grayscale of X_k , and $P(X_k)$ is the probability of the *k-th* grayscale appearing. Then the transformation function of the traditional Histogram equalization method is the cumulative distribution function:

$$T(X_k) = \sum_{j=0}^{k} P_r(r_j) = \sum_{j=0}^{k} \frac{n_j}{n} \tag{2}$$

In the above equation, $0 \le r_j \le 1$ is the maximum grayscale level of the original image.

Because the gray level of the low illumination image is less, when the traditional Histogram equalization method is used to process the low illumination image, the gray level change range of the processed image is difficult to reach the maximum gray level change range allowed by the image format, and the image information may be lost because the gray level of the image may be merged too much. Therefore, it is more reasonable to use Histogram equalization method with position correction.

The *i*-th level grayscale *X* in the original image histogram is corrected to:

$$m : (l-1-m) = \sum_{k=0}^{i-1} P_r(X_K) : \sum_{k=i+1}^{l-1} P_r(X_K) \tag{3}$$

Organized:

$$m = (l-1)\frac{\sum_{k=0}^{i-1} P_r(X_K)}{\sum_{k=0}^{i-1} P_r(X_K) + \sum_{k=i+1}^{l-1} P_r(X_K)} \tag{4}$$

In the above equation, $\sum_{k=0}^{i-1} P_r(X_K)$ and $\sum_{k=i+1}^{l-1} P_r(X_K)$ are the cumulative distribution functions of the left and right histograms of gray level X_i . Due to $\sum_{k=0}^{l-1} P_r(X_K) = 1$ and $P(X_k) = \frac{n_k}{n}$ P, the above equation can be transformed into:

ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

$$m=(l-1)\frac{\sum_{k=0}^{i-1}P_r(X_K)}{1-P_r(X_i)}=(l-1)\frac{\sum_{k=0}^{i-1}n_k}{n-n_i} \tag{5}$$

The mapping function at this point is:

$$g(X_i)=INT\left\{(l-1)\frac{\sum_{k=0}^{i-1}n_k}{n-n_i}\right\}, i=1,2,\cdots,l-2 \tag{6}$$

The improved Histogram equalization algorithm extends the dynamic range of output image gray, and improves the visual effect. The Histogram equalization method is simple to implement and has strong applicability. It has a good effect in contrast enhancement, but it is easy to lose details due to excessive enhancement.

2.2 Method Based on Retinex Theory

Retinex theory is a color constancy perception theory proposed by Edwin H. Land, which is achieved by simulating the imaging principles of the visual cortex of the human brain. The core idea of this theory is that the reflection ability of most objects determines the apparent characteristics of objects that people observe, which are slightly related to the intensity of incident light on the surface of the object [3]. The algorithm based on Retinex theory can achieve brightness improvement, detail enhancement, and color fidelity, which has been favored by many scholars and widely applied in the field of image enhancement.

The basic description of Retinex theory is [4]: for the grayscale value and dynamic range of an image, which are affected by incident light and correspond to the low-frequency component of the incident light in the frequency domain of the image, they are called illuminated images; The reflectivity of the high-frequency component of light has a significant impact on the object, and the reflection properties of the image itself determine its essential content, which is called a reflection image. Based on this theory, image enhancement is achieved by first decomposing the original image into reflection and illumination components, and then removing the influence of illumination components to retain the inherent attributes of the object, thereby achieving enhancement.

The image seen by the human eye can be expressed as:

$$S(x,y)=R(x,y)\times L(x,y) \tag{7}$$

In the above equation, $L(x,y)$ represents the illumination component of the surrounding light intensity information, $R(x,y)$ represents the reflection component of the object's inherent properties, and $S(x,y)$ represents the image observed by the observer.

The illumination and reflection components in the above equation are both unknown, making it difficult to solve. Considering the mechanism by which the human visual system perceives information, this problem can be transformed into the logarithmic domain. According to the arithmetic of logarithms in mathematics, the above equation can be transformed into:

$$\log(S(x,y))=\log(R(x,y)\times L(x,y))=\log(R(x,y))+\log(L(x,y)) \tag{8}$$

$$s(x,y)=\log(S(x,y)) \tag{9}$$

$$r(x,y)=\log(R(x,y)) \tag{10}$$

$$l(x,y)=\log(L(x,y)) \tag{11}$$

$$s(x,y)=r(x,y)+l(x,y) \tag{12}$$

In order to preserve the reflection component of an object well, the illumination component should be eliminated or reduced to the greatest extent possible. However, the reflection component is generally estimated through a filter, which cannot meet the needs of the human visual system. Using the

ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

reflection component as the output is a serious problem. According to different methods for estimating lighting components, many Retinex algorithms have emerged, and their essence is the same. The basic algorithm flow is shown in Figure 1.

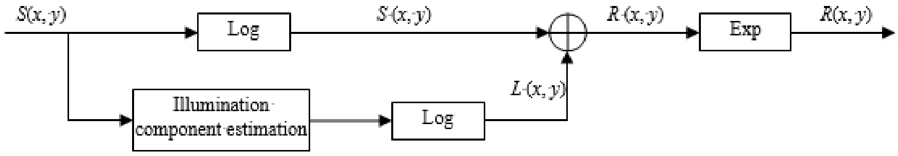

Figure 1: Basic flow of Retinex algorithm

3. Low Illumination Image Enhancement Algorithm Based on Convolutional Neural Network

Convolutional neural network is an artificial neural network model, which has been in the slow laboratory development stage for a long time until Alex Net came out, making Convolutional neural network start to emerge in the field of computer vision, and has been applied to various fields of artificial intelligence at present [5]. Since the development of Convolutional neural network, it has formed a mature and stable structure, as shown in Figure 2.

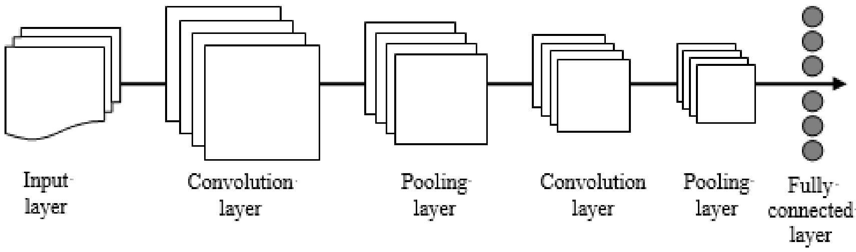

Figure 2: Convolutional neural network architecture

The Convolutional neural network is composed of multiple convolutional layers, each of which contains multiple two-dimensional feature vectors. The corresponding network architecture is obtained by modeling human neurons, as shown in Figure 3.

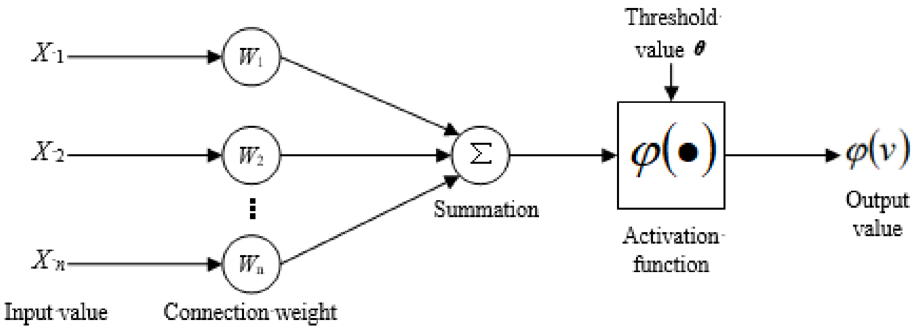

Figure 3: Artificial neuron model

Classical Convolutional neural network models include LeNet, AlexNet, VGG, NiN, ResNet, GoogLeNet, DenseNet, U-Net, CAN, etc. This article selects the CAN network model. The biggest feature of the CAN network is dilation convolution, which does not use maximum pooling operation to avoid information loss caused by maximum pooling.

3.1 Dilation Convolution

In the traditional Convolutional neural network, the pooling layer is used to maintain the feature invariance and avoid over fitting, but it will greatly reduce the spatial resolution and lose the spatial information of the feature map. When deepening the layer of Convolutional neural network, the network needs more parameters, which will lead to more computing resource consumption. Expansion convolution solves this problem well. Inflation convolution is a convolutional operator that uses different dilation factors and the same filter in different ranges [6].

To understand the relationship between the expansion rate *d* and the image output size *o*, it is necessary to consider the impact of *d* on the effective kernel size. The kernel with size *k* is inflated by factor *k* and has an effective size:

$$\hat{k} = k + (k-1)(d-1) \tag{13}$$

The calculation formula for image output size *o* is:

$$o = \left[\frac{i + 2p - \hat{k}}{s} \right] + 1 \tag{14}$$

In the above equation, *i* represents the input image size, *k* represents the filter size, *p* represents padding, and *s* represents step size.

The relationship between expansion rate *d* and image output size *o* is expressed as:

$$o = \left[\frac{i + 2p - k - (k-1)(d-1)}{s} \right] + 1 \tag{15}$$

The expansion convolution increases the Receptive field of the convolution kernel while keeping the number of parameters unchanged, so that each convolution output contains a large range of information, while ensuring that the size of the output feature mapping remains unchanged [7]. The recurrence formula of Receptive field is:

$$RF_{i+1} = RF_i + (\hat{k}_i - 1) \times S_i \tag{16}$$

In the above formula, RF_i represents the Receptive field of layer *i*, and S_i represents the product of steps of layer *i* and all previous layers. The calculation formula is:

$$S_i = \prod_{j=1}^{i} Stride_j = S_{i-1} \times Stride_i \tag{17}$$

3.2 Loss Function

The optimization of Convolutional neural network is related to the selection of Loss function. Based on CAN convolutional network model, L_1 Loss function and *SSIM* Loss function are used. L_1 The Loss function is expressed as [8]:

$$L_1(\Omega) = \frac{1}{N} \sum_{i=1}^{N} \left\| \left(X_{org} + F(X_{org}, \Theta) - Y_{ref} \right) \right\|_1 \tag{18}$$

In the above equation, $F(X_{org}, \Theta)$ is the predicted output image of the network, X_{org} is the low illumination image, Y_{ref} is the reference image, and $\|\bullet\|_1$ is the first norm.

Adopting L_1 loss instead of L_2 loss mainly takes into account that the network mainly focuses on the details of the dark regions. After merging with the dark regions of the input image, it is necessary to restore the details. If loss is used, it will cause the details of the dark regions to be too smooth, leading to blurring in the restored dark regions. In addition to using L_1 loss, *SSIM* Loss function is also used to constrain the overall structure information of the image. The *SSIM* Loss function is expressed as:

$$L_{SSIM}(\Omega) = \frac{1}{N} \sum_{i=1}^{N} [1 - SSIM(z_i, y_i)] \tag{19}$$

In the above formula, z_i represents the *i*-th output image, and y_i represents the *i*-th reference image. After many experiments, the weight of the *SSIM* Loss function is set to 0.1, which can achieve the best effect. Then the Loss function of the whole network can be expressed as:

ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

$$Loss = L_1 + 01. \times L_{SSIM} \tag{20}$$

3.3 Network Optimization

During the training process of deep neural networks, the parameters of each network layer are constantly changing, and the input distribution of each layer of the network is constantly changing. Different input distributions may need to be retrained. In addition, we have to use smaller parameter initialization and smaller Learning rate to train the model, so as to avoid the network output falling into the saturation area, which will cause the gradient to disappear, and the depth model is generally difficult to train [9]. How to make each input distribution of the network layer consistent and avoid output saturation, thereby accelerating training. Based on this, Batch Normalization method is proposed to normalize the input distribution into a standard normal distribution, solving the above problems.

How to effectively implement BN, the first simplification is to standardize each dimension of the input vector independently, and the expectation and variance are calculated throughout the entire training set, which may reduce the nonlinear expression ability of the model. The second simplification is to estimate the mean and variance of each activation within each mini-batch, and BP can be used to update network parameters. The BN algorithm process consists of input and output.

Input: Refers to the quantity data $B = \{x_1, x_2, \cdots, x_m\}$, parameter γ, β.

Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

The average of this batch of data is:

$$\mu_B \leftarrow \frac{1}{m}\sum_{i}^{m} x_i \tag{21}$$

The variance of this batch of data is:

$$\sigma_B^2 \leftarrow \frac{1}{m}\sum_{i}^{m} (x_i - \mu_B)^2 \tag{22}$$

Whitening each data in batch data:

$$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}} \tag{23}$$

Zoom and pan to obtain the final output:

$$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i) \tag{24}$$

Numerous experiments have shown that the Batch Normalization algorithm can learn parameter γ, β, enabling it to recover the original input data, increase generalization ability, and prevent overfitting.

4. Conclusions

High quality images contain richer information and have broader applications in real life. As the most important tool in visual tasks, Convolutional neural network has developed rapidly in recent years. In this paper, the Convolutional neural network is used to study the problems of low brightness, serious loss of details, and visual perception in low illumination images. The research results of this article not only improve brightness and contrast, but also effectively avoid color distortion. The enhancement effect is better than the current mainstream low illumination image enhancement algorithms, and has certain promotion value. However, the method proposed in this article is aimed at low illumination images in ideal conditions, and lacks strong denoising ability for low illumination images with strong noise. Low illumination images in real scenes often have complex noise problems, and subsequent research needs to add noise reduction modules to solve the noise problem in dark areas. In addition, it is necessary to expand the dataset and add the existing dataset to low illumination images containing noise, as well as normal images, to constrain the network's ability to obtain noise reduction.

ISSN 2616-5775 Vol. 6, Issue 9: 115-121, DOI: 10.25236/AJCIS.2023.060917

References

[1] Kumar A M, Sekhar C P. A Review Paper on Low Light Image Enhancement Methods for Un-uniform Illumination[J]. IFAC PapersOnLine, 2022, 55(10): 287-292.
[2] Suresh T, Domnic S. Deep learning model for person identification and verification on low illumination images[P]. National Institute of Technology,Department of Computer Applications, Tiruchirappalli, India, 2022, 31: 062009-062009.
[3] Rasheed Muhammad Tahir, Guiyu G,Daming S, et al. An Empirical Study on Retinex Methods for Low-Light Image Enhancement[J]. Remote Sensing,2022,14(18): 4608-4608.
[4] B. W. Li. Research on Enhancement Algorithms of Low Illumination Images Based on Convolutional Neural Networks[D]. Nanjing University of Posts and Telecommunications, 2021.
[5] Gil Eddie M, Steelman Zachary A, Sedelnikova Anna, Bixler Joel N. Comparing the segmentation of quantitative phase images of neurons using convolutional neural networks trained on simulated and augmented imagery[J]. Neurophotonics, 2023, 10(3): 035004-035004.
[6] Y. Qin. Research on Denoising and Enhancement of Low light Image Based on Convolutional Neural Network[D]. Chang'an University, 2019.
[7] Shapovalova S, Moskalenko Yu, Baranichenko O. Increasing the Receptive Field of Neurons in Convolutional Neural Networks [J]. Cybernetics and Systems Analysis, 2023, 59(2): 339-345.
[8] K. J. Wang. Research on Low-light Image Enhancement based on Convolutional Neural Network [D]. Guilin University of Electronic Technology, 2022.
[9] Houssaine E H, NourEddine J, Mohamed E. A deep learning framework for time series classification using normal cloud representation and convolutional neural network optimization[J]. Computational Intelligence, 2022, 38(6): 2056-2074.