Image Denoising Based on Dual-tree Complex Wavelet Transform and Convolutional Neural Network

Yilu Zhou* and Xiaojin Fu
Shanghai Dianji University, Shanghai, China

*Corresponding author email: 997963959@qq.com

Abstract. In view of the problem that the noise in the image will have a negative impact on the subsequent image processing, this paper proposes a noise reduction algorithm based on the dual-tree complex wavelet transform and convolutional neural network by using the characteristics of the DTCWT and CNN. Firstly, the two-dimensional image is decomposed into two parallel wavelet trees by double tree complex wavelet transform, which forms six high-frequency detail parts and two same low-frequency parts in different directions. Then, it is put into the designed deep convolution neural network for training. Finally, the six high-frequency components and one of the low-frequency components are recombined into the image after noise reduction by using the inverse transform of double tree complex wavelet. The simulation results show that PSNR and MS-SSIM of the image are better than other noise reduction algorithms, and the details of the original image can be restored well while noise reduction. It has practical application value in the case that the collected image has a large noise due to various reasons.

1. Introduction
Nowadays, more and more scholars have invested in computer research, which directly promotes the development of image processing. Because of some factors in the acquisition of images and transmission of images, noise reduction has become a necessary part in the process of image processing. The main purpose of noise reduction is to decompose the blurred image into different signals, separate and remove the noise signal from the original image, and retain the real and effective image information as much as possible. In the past few decades, many scholars have proposed a large number of image denoising methods, but the traditional denoising methods have some shortcomings, such as the increase of entropy after signal transformation, the inability to describe the edge and breakpoint of signal and the inability to get the correlation of signal. Wavelet transform has always occupied a place in the field of image denoising because of its good time-frequency characteristics[1]. Image denoising based on deep learning has also been widely used and developed. Different from traditional denoising, it has gradually become a popular algorithm from shallow model to depth model, from noisy image mapping denoised image to noisy image mapping noisy image[2-4]. Therefore, the integration of these two methods can magnify their respective advantages. This idea can further promote the development of image denoising.

2. Basic Theory
2.1. Dual-tree Complex Wavelet Transform
Since complex wavelet transform is input in complex form, it is difficult for general complex wavelet transform to match the corresponding reconstruction filter. Therefore, Kingsbury et al. proposed dual-
tree complex wavelet transform (DTCWT) in 1998 and improved it on DTCWT in 2003[5], and proposed Q-shift DTCWT. It can accomplish the half sampling period delay of the filter between two trees in a more ingenious way. It can solve the problem of complete reconstruction on the basis of the advantages of wavelet transform, and can also guarantee the amplitude frequency characteristics better. Similar to the one-dimensional DTCWT, for the two-dimensional case, because there is \( \psi(x,y) = \psi(x)\psi(y) \), then \( \psi(x) = \psi_h(x) + j\psi_g(x) \), after substituting \( \psi(x) \) and \( \psi(y) \), the complex wavelet can be expressed as:

\[
\psi(x,y) = [\psi_h(x) + j\psi_g(x)][\psi_h(y) + j\psi_g(y)]
\]

\[
= \psi_h(x)\psi_h(y) - \psi_g(x)\psi_g(y) + j[\psi_g(x)\psi_h(y) + \psi_h(x)\psi_g(y)]
\]

(1)

\( \psi_h(x) \) and \( \psi_g(x) \) denotes the real part and imaginary part of complex wavelet respectively, so the two decomposition trees of DTCWT become various independent complex wavelet transforms. The real part and the virtual part obtained by the DTCWT can be represented by tree A and tree B respectively. The low pass and high pass filters of tree A can be expressed as \( h_0(n) \) and \( h_1(n) \). The low pass and high pass filters of tree B can be expressed as \( g_0(n) \) and \( g_1(n) \), \( \downarrow_2 \) Represents sampling at intervals. Figure 1 shows the decomposition diagram of two-dimensional DTCWT.

![Figure 1. Decomposition diagram of two-dimensional dual-tree complex wavelet transform.](image)

From Figure 1, we can find that the first level dual tree complex wavelet decomposition can get two low-frequency signals and six high-frequency signals, and each low-frequency signal of the first level can be decomposed into two low-frequency signals and six high-frequency signals of the second level, so no matter how deep the decomposition is, its overall data redundancy is 4:1.

It has good two-dimensional directional selectivity, but it is also applicable in higher dimensions. The DTCWT has selectivity in six directions \( \pm 15^\circ, \pm 30^\circ \) and \( \pm 75^\circ \). The image after DTCWT is shown in Figure 2.
2.2. Convolutional Neural Network

In 1962, biologists Hubel and Wiesel discovered that there are many cells with complex structures in the visual cortex of cat brain. Because of the characteristics of these cells, he can use the sensitivity of space to mine the correlation of local space in the image. Then, according to this breakthrough point, a new method is proposed, which is composed of input layer, convolution layer, pooling layer, filtering layer The CNN structure consists of full connection layer and output layer. In recent years, CNN has been widely used in the field of image processing, residual learning was originally proposed to solve the problem of gradient explosion, and then because it can greatly improve the efficiency of the algorithm, it is widely used in the model. In 2016, Kai Zhang et al. Proposed a new feedforward denoising convolutional neural network\[6\], which uses residual learning and batch normalization, combined with small batch SGD. When a batch of images enter the convolution layer, first calculate their mean and variance, then normalize them, and then input the normalized data into the convolution layer; For the output data of convolution layer, the output is restored to its original distribution by learning the mean and variance of the output data, and then input the activation function. It can improve the effect of Gaussian denoising with specific noise level.

3. Denoising Algorithm Based on Dual Tree Complex Wavelet Transform and Convolutional Neural Network

3.1. Algorithm Structure

This algorithm is mainly composed of seven identical network models. We set the size of convolution filter to 3×3, but all pooling layers are removed. It consists of three types of layers, each of which is given a DNCNN with depth D, as shown in the figure 3. (i) Conv+ReLU: for the first layer, 64 filter of 3×3×C size are used to generate 64 feature maps. Then the rectified linear unit (ReLU, max(0, ·)) is used for nonlinear analysis. C represents the number of channels of the image. When C = 1, it is a gray image, and when C = 3, it is a color image. (II) Conv + BN + ReLU: corresponding to 2 ~ (D-1) layer, use 64 filter of 3×3×64 size to batch normalization is added between convolution and ReLU. (III) Conv: corresponding to the last layer, e filter of 3×3×C size is used to reconstruct the output. Because the residual learning formula is used to learn R (y) and batch normalization is used to accelerate the training, the performance of noise reduction can be effectively improved.
3.2. Algorithm Flow

The flow of the algorithm is as follows:

- The input image is cut into standard 512 * 512 format. The text should be set to single line spacing.
- DTCWT is applied to the image block, and the components of each angle are put into the designed network for training, and the loss function is used to adjust to achieve the purpose of convergence.
- After 7 different training models are obtained, they can be tested.
- The image to be tested is transformed by dual tree complex wavelet transform and put into their respective models for denoising, and seven denoised components are obtained.
- Inverse DTCWT is used to get the final denoised image.

4. Experimental Results and Analysis

The experimental training data set is bsds68 image set. Bsds68 data set only contains 68 grayscale images, which is far less than the number of images needed for training. Therefore, the 68 images are used to expand the data set by clipping, brightness contrast adjustment and flipping, and finally 7752 images are obtained as the training set. The experimental data set is set12 image set. There are 12 grayscale images, we only select two of them to test. The hardware of the computer is configured as AMD R5-2600X and NVIDIA GeForce GTX1660ti. The operating system is Win10. It is trained by using the Keras deep learning framework. It is based on the TensorFlow, Theano and CNTK backend, which is equivalent to their upper interface. Since TensorFlow is more portable, efficient and scalable,
it is selected as the backend of Keras. The framework supports CUDA (computing platform developed by NVIDIA) and Cudnn (deep convolution neural network accelerator) to use GPU for computing, which can effectively reduce the training time. The whole experimental program is written by python. Because some noise reduction effects are laborious to be distinguished by human eye, two algorithms are used to test the noise reduction effect of the image.

PSNR (Peak Signal to Noise Ratio) is a standard to estimate the similarity between the original image and the noise reduced image. However, because it can not evaluate the image quality according to the inability of each pixel point, the human naked eye may get different senses from the results. PSNR is defined by mean square deviation (MSE), and the specific formula of PSNR can be expressed as follows:

\[
PSNR = 10 \cdot \log_{10} \left( \frac{MAX_f^2}{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \| I(i,j) - K(i,j) \|^2} \right)
\]

I and K represent two m × n monochromatic images, \( MAX_f \) represents the maximum value of the image color, and usually uses 8-bit sampling points, its value is 255.

MS-SSIM is an index to measure the similarity between two pictures. It is changed according to SSIM. The width and height are reduced by \( 2^{M-1} \). Its formula can be expressed as follows:

\[
SSIM(x, y) = [l_M(x, y)]^{a_M} \cdot \prod_{j=1}^{M} [c_j(x, y)]^{b_j} \cdot [s_j(x, y)]^{y_j}
\]

In this formula \( a, \beta, \gamma \) is used to adjust the weight of each component and usually to simplify the formula calculation, let

\[
a_j = \beta_j = \gamma_j, \sum_{j=1}^{M} \gamma_j = 1
\]

The contrast algorithms selected in this paper are median filter (MF), Gaussian low-pass filter (GLP), wavelet de-noising (WDD)[7], DNCNN[8] and wavelet transform and neural network based image de-noising algorithm (WT-CNN)[9]. The images used for testing are randomly selected from set12[10]. We select two images to compare the noise reduction effect, in which the noise level is \( \sigma = 25 \). As shown in Figure 4, it can be found that MF, GLP and WDD have poor denoising effect. DNCNN noise reduction process is too large, resulting in partial detail distortion. The effect of WT-CNN is similar to that of this algorithm, but because this algorithm obtains six high-frequency components in different directions through dual tree complex wavelet transform, while WT-CNN only uses wavelet transform to transform the image into row and then to transform the image into column, only four components are obtained, Some details are not as good as the visual effect of this algorithm.
Figure 4. Comparison of denoising performance with different algorithms

Table 1 shows the PSNR values of two Gaussian white noise test images added with $\sigma = 25$ after six different algorithms. It can be seen that the PSNR values of this algorithm are higher than those of other algorithms. Table 2 shows the MS-SSIM values of each algorithm after noise reduction, and the algorithm in this paper is also higher than other algorithms.

Table 1. PSNR of noise reduction results with different algorithms

| Methods | MF | GLP | WDD | DNCNN | WT-CNN | Ours |
|---------|----|-----|-----|-------|--------|------|
| lena    | 24.06 | 27.94 | 31.41 | 31.58 | 33.47 | 34.38 |
| woman   | 21.42 | 24.92 | 28.06 | 29.62 | 31.79 | 32.20 |

Table 2. MS-SSIM of noise reduction results with different algorithms

| Methods | MF | GLP | WDD | DNCNN | WT-CNN | Ours |
|---------|----|-----|-----|-------|--------|------|
| lena    | 0.571 | 0.637 | 0.821 | 0.849 | 0.877 | 0.880 |
| woman   | 0.512 | 0.594 | 0.786 | 0.832 | 0.871 | 0.873 |

5. Conclusion

This paper presents a new fusion algorithm based on the characteristics of dual-tree complex wavelet transform and convolutional neural network. The multi components are obtained by using DTCWT, and each component is put into the corresponding training model to reduce noise. Finally, the final image is obtained by wavelet inversion. The simulation results show that the algorithm has good noise reduction effect and higher detail retention ability. It is of practical value to pre-process the image which needs noise reduction in image processing. At the same time, it makes a certain contribution to the development of noise reduction technology. Because this algorithm only uses one level of DTWCT for image, it can be tried in more levels in the future.

References

[1] WU F, ZHU X, XUANG Ri, et al. Design of cloud and mist removal system from remote sensing images based on dual-tree complex wavelet transform. Journal of Applied Optics. 39(1), 64-70, (2018)
[2] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition. USA: IEEE, 770-778, (2016)
[3] Pan K, Hou L. Remote sensing image denoising based on convolutional neural network. Modern information technology. 4 (12), 60-65, (2020)
[4] Lyu Y, Zhao J, Cao F. Image denoising algorithm based on composite convolution neural network. *Pattern Recognition and Artificial Intelligence*. **30**(2), 97-105, (2017)

[5] Kingsbury N.G. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters. *IEEE Digital Signal Processing Workshop*. **98**(1), 2-5, (1998)

[6] Zhang K, Zuo W, Zhang L. FFDNet: Toward a fast and flexible solution for CNN based image denoising. *IEEE Transactions on Image Processing*. **27**(9), 4608-4622, (2017)

[7] Wang M, Zhou L, Zhou S, et al. Image SVD denoising based on PSNR and wavelet directional feature. *Journal of Applied Optics*. **34**(1), 85-89, (2013)

[8] Chen Q, Shi X, Chai Y. Image denoising algorithm based on wavelet transform and convolutional neural network. *Applied optics*. **41**(02), 288-295, (2020)

[9] Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*. **26**(7), (2017)

[10] Bijalwan A, Ggyal A, Sethi N. Wavelet transform based image denoise using threshold approaches. *International Journal of Engineering & Advanced Technology*. **05**, 218-221, (2012)