Using a Full Counterpropagation Neural Network for Image Watermarking

Chuan-Yu Chang, Sheng-Jyun Su, Hung-Jen Wang

Graduate School of Computer Science and Information Engineering
National Yunlin University of Science & Technology
chuanyu@yuntech.edu.tw

Abstract

Digital watermarks are an important technique for protection and identification that allows authentic watermarks to be hidden in multimedia such as image, audio, and video. Watermarking has been developed to protect digital media from being illegally reproduced and modified. Embedding and extracting watermark used to require complex procedures. These include randomizing the watermark, choosing positions to embed and extract it, embedding the randomized watermark into the specific positions, and extracting it from the specific positions. In this paper, we propose a novel method called Full Counter-propagation Neural Network (FCNN) for digital image watermarking, in which the watermark is embedded and extracted through specific FCNN. Different from the traditional methods, the watermark is embedded in the synapses of FCNN instead of the cover image. The experimental results show that the proposed method is able to achieve robustness, imperceptibility and authenticity in watermarking.

Keywords: digital watermark, full counterpropagation neural network, information hiding

1. Introduction

The rapid development of computer network and multimedia technology makes it easier to assess digital media. Since the problem of illegal reproduction and modifications has become more serious than before, it is important to protect the intellectual property of digital media. To tackle the problem, digital watermark has been proposed as a means to identify the owner of digital media such as text, image, video and audio. The watermarking technique embeds authors' information into digital media and provides the corresponding authentication mechanism. Satisfactory digital watermarking must meet the following requirements: (1) Robustness: the watermark must be difficult to delete and should be resistant to standard image processing operations such as filtering, cropping, loose compression, etc. (2) Imperceptibility: the watermark embedded in the image should not be caused obvious visual degradation of images. (3) Security: the attacker should not be able to detect the embedded watermark using common statistical analysis or correlative attacks. In general, the proposed image watermarking techniques are divided into two groups, namely embedding watermark in spatial domain and embedding watermarking in frequency domain, depending on the processing domain of cover image that the watermark is embedded in. In the spatial domain, Schyndel *et al*. [2] proposed two digital watermark techniques. The first technique is based on manipulate the bit plane of the LSB, while the second utilizes linear addition of the watermark to cover image. However, these methods are not robust enough. Hwang *et al*. [3] presented a watermark technique based on color space transformation. The color space of cover image is transformed from RGB to HSI space and then embedded watermark into saturation channel. This method is able to resist some attacks. In the frequency domain, Ahmidi *et al*. [4] proposed a color image watermark technique based on DCT. This method embedded the watermark into middle frequency coefficient of transformed blocks. Deng *et al*. [5] proposed a method of embedding the watermark in the DC component of transformed blocks.

Neural networks have been suggested as alternative approaches owning to high fault tolerance and potential for adaptive training [6-8]. Davis *et al*. [6] proposed a method based on neural networks to find maximum-strength watermarks for DWT coefficients. Mei *et al*.[7] put forward a three-layer neural networks to determine the strength of watermarking for image DCT coefficients. Zhang [8] *et al*. proposed a RBF neural network to achieve maximum-strength watermark according to the frequency component of the cover image.

The traditional methods mentioned above require complex embedding and corresponding extraction

procedures. In this paper we proposed a specific full counterpropagation neural network (FCNN) for watermarking. Different from the traditional methods, the watermark is embedded in the synapses of FCNN rather than the cover image. Hence the quality of the watermarked image is almost same as the original cover image. In addition, the quality of the extracted watermark image does not degrade after most attacks, because the watermark is stored in the synapses. [1] In our method, FCNN simplifies the complex embedding and corresponding extraction procedures. The experimental results show that there was no need for these procedures in FCNN-based watermarking. Furthermore, the proposed method is able to achieve robustness, imperceptibility and authenticity in watermarking.

The following sections discuss the application of FCNN in watermarking in the proposed method. Section 2 introduces the FCNN, the algorithm of embedding, and the method of extraction. Section 3 summarizes the experimental results, and conclusions are given Section 4.

2. Full Counterpropagation neural network for watermarking

In this paper, a novel full counterpropagation neural network (FCNN) is proposed for image watermarking. The full counterpropatation neural network is a supervised-learning network with capacity of bidirectional mapping. Figure 1 shows the conceptual diagram of the FCNN.

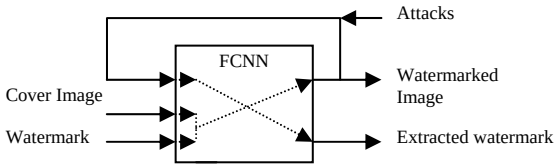


Figure 1. The conceptual diagram of the FCNN

The traditional watermarking methods require complex embedding and corresponding extraction procedures. However, the proposed watermarking method integrates the embedding and extraction procedure into a full counterpropagation based neural network. In order to ensure that the proposed watermarking method has capability of embedded and extracted watermarks, the cover image and watermark image are input to the FCNN simultaneously during the embedding process. After the network's evolution, a watermarked image was generated. On other hand, the same FCNN was used for extracting the corresponding watermark from the watermarked images with or without being attacked.

The purpose of this paper is to present a full counterpropagation neural network for watermarking. The FCNN is designed to learn bidirectional mappings. Through the process of supervised training, the FCNN adaptively constructs a lookup table approximating the mapping between the presented input/output training pair: cover image X and watermark image Y. After being trained, the FCNN can be used to extract the corresponding watermark Y* if the embedded image X* is known. Figure 2 shows the architecture of the proposed FCNN.

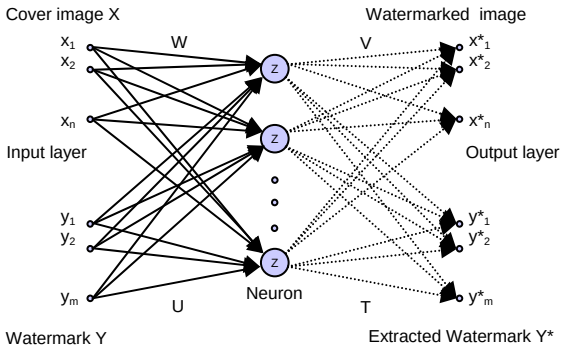


Figure 2. The architecture of the full counterpropagation neural network for watermarking

The two-dimensional cover image X and watermark image Y can be written in vector form as:

$$X = \{x_1, x_2, \cdots, x_n\} \quad (1)$$

$$Y = \{y_1, y_2, \cdots, y_m\} \quad (2)$$

where n is the number of pixels of the cover image X, and m is the number of pixels of the watermark image Y. The input vectors X and Y are connected to neuron Z_i with weights W and U respectively.

$$W = \{w_{1,1}, w_{1,2}, w_{1,3} \cdots, w_{n,i}\} \quad (3)$$

$$U = \{u_{1,1}, u_{1,2}, u_{1,3} \cdots, u_{m,i}\} \quad (4)$$

where $w_{n,i}$ denotes the weight between i-th neuron and input x_n. Similarly, $u_{m,i}$ denotes the weight between i-th neuron and input y_m.

Accordingly, the total input of the i-th neuron is defined as

$$Z_i = \sum_{k=1}^{n} (x_k - w_{i,k})^2 + \sum_{k=1}^{m} (y_k - u_{i,k})^2 , \quad (5)$$

which represents the distance between the input x, y pair and the i-th neuron. The activation function for each neuron is given by

$$\Gamma_i = \begin{cases} 1 & \text{if } Z_i \text{ is smallest for all i} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

therefore, the j-th output of X^* and Y^* can be obtained as

$$x_j^* = \sum_{i=1}^{n} \Gamma_i v_{j,i} \quad (7)$$

and

$$y_j^* = \sum_{i=1}^{n} \Gamma_i t_{j,i} \quad (8)$$

where $v_{j,i}$ denotes the weight between i-th neuron and output x_j^*. Similarly, $t_{j,i}$ denotes the weight between i-th neuron and output y_j^*. The synaptic weights can be written in a vector form as

$$V = \{v_{1,1}, v_{1,2}, v_{1,3}, \cdots, v_{n,i}\} \quad (9)$$

and

$$T = \{t_{1,1}, t_{1,2}, t_{1,3} \cdots, t_{m,i}\} \quad (10)$$

where *n* and *m* are the number of pixels of the cover image X and watermark image Y respectively. The output errors of FCNN are calculated by

$$E_c = \sum_{i=1}^{n} \left| x_i^* - d_i \right| \quad (11)$$

$$E_m = \sum_{j=1}^{m} \left| y_j^* - d_j \right| \quad (12)$$

where d_i denotes the i-th pixel value of desired watermarked image (input cover image) and d_j denotes the j-th pixel value of desired watermark image (input watermark image). If the output error is less than a predefined threshold, the network converges. Otherwise, the input weight vectors W and U associated with the winning neuron ($\Gamma_i = 1$) are updated by

$$w_{i,j}(k+1) = [1-\alpha(k)]w_{i,j}(k) + \alpha(k)x_i, \quad i = \arg(\Gamma_i = 1) \quad (13)$$

$$u_{i,j}(k+1) = [1-\alpha(k)]u_{i,j}(k) + \alpha(k)y_i, \quad i = \arg(\Gamma_i = 1) \quad (14)$$

where $\alpha(k)$ is the learning rate of input layer. In addition, the learning rate $\alpha(k)$ is suitably decreasing functions of learning time k. The learning function $\alpha(k)$ can be specified as follows:

$$\alpha(k) = \alpha(0)\exp\left(-\frac{k}{k_0}\right) \quad (15)$$

where $\alpha(0)$ is initial learning rates, k_0 is a positive constant. Similarity, the output weight vectors V and T associated with the winning neuron are updated by

$$v_{i,j}(k+1) = v_{i,j}(k) + \beta(k)[x_j - v_{i,j}(k)]\Gamma_i, \quad j = 1,2,...,n \quad (16)$$

$$t_{i,j}(k+1) = t_{i,j}(k) + \beta(k)[t_j - t_{i,j}(k)]\Gamma_i, \quad j = 1,2,...,m \quad (17)$$

where $\beta(k)$ is the learning rate of output layer.

After the FCNN converged, the watermarked image is obtained as

$$X^* = \{x_1^*, x_2^*, \cdots, x_n^*\} \quad (18)$$

After the watermarked image is obtained, the same FCNN is used to extract the corresponding watermark from the watermarked images with or without being attacked. The extracted watermark image is obtained as

$$Y^* = \{y_1^*, y_2^*, \cdots, y_m^*\} \quad (19)$$

2.1 Embedding algorithm

The watermark embedding approach is summarized as follows:

Input: The cover image X and watermark Y

Output: watermarked image X*

Step 1. Arbitrarily assigns the initial weights to U, V, W, and T.

Step 2. Use Eq.(5) and Eq.(6) to calculate the output of each neuron.

Step 3. Use Eq.(7) and Eq.(8) to calculate the network outputs

Step 4. Use Eq.(11) and Eq.(12) to calculate the output errors of FCNN, if E_c and E_m is less than a predefined threshold value, the network converged and use Eq.(7) and Eq.(18) to obtain the watermarked image X*, else go next step.

Step 5. Use Eq.(13) and Eq.(14) to update the weight vectors W and U, and the learning rate $\alpha(k)$ by Eq.(15).

Step 6. Use Eq.(16) and Eq.(17) to update the weight vectors V and T. Then go to step2.

2.2 Extracting algorithm

To extract watermark from the trained FCNN described in Section 2.1, the watermark extracting approach is summarized as follows:

Input: Watermarked image X*

Output: Extracted watermark Y*

Step 1. Use Eq.(5) and Eq.(6) to calculate the neuron output.

Step 2. Use Eq.(8) and Eq.(19) to obtain the extracted watermark image Y*.

3. Experimental results

In order to show that the proposed FCNN has good performance for watermarking, four experiments are proposed to demonstrate robustness, imperceptibility and authenticity achieved by the application. In our experiments, the cover images including Jet and Couple, and the image size is 256×256. In addition, the watermark image is the school badge of NYUST with image size 32×32. Figures 3(a-c) show the cover images and watermark image respectively. To evaluate the robustness of conventional LSB method[2] and proposed FCNN method, seven types of attacks were used to attack watermarked image. These include 3×3 averaging filter, crop left-top 1/4 part of watermarked image, 3×3 Laplacian mask, jpeg compression, rotating 90°, 2×2 mosaic and Gaussian.

(a) (b)

(c)

Figure 3. Cover image and watermark image, (a) Cover Jet Image. (b) Cover Couple Image. (c) Watermark.

The Peak Signal to Noise Ratio (PSNR) was used to evaluate the quality of watermarked image and extracted watermark, which can be represented as:

$$PSNR(dB) = 10\log_{10}\frac{X^{2}{}_{peak}}{\sigma_e^2} \tag{20}$$

where σ_e^2 is defined as

$$\sigma_e^2 = \left(\frac{1}{MN}\right)\sum_{i=1}^{M}\sum_{j=1}^{N}\left(X_{ij} - Z_{ij}\right)^2 \tag{21}$$

where $M \times N$ is the size of cover image, X_{ij} is gray level of (i,j)th pixel of cover image. Similarly, Z_{ij} denotes the gray level of (i,j)th pixel of watermarked image. X^2_{peak} denotes the squared peak value of cover image. The higher PSNR is, the more similar embedded image and the cover image are.

(a) (b)

Figure 4. Watermarked image, (a) watermarked Jet Image. (b) watermarked Couple image.

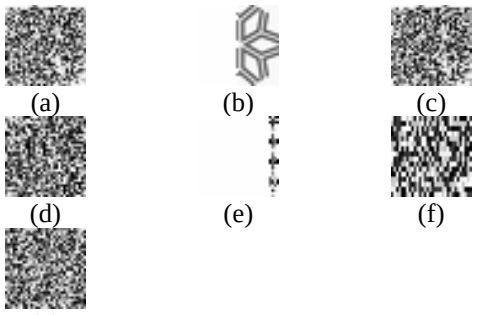
3.1 Experimental 1: Robust testing for LSB Method

Since the watermark image is a 32*32 gray-level image. LSB method [2] requires 8192 pixels of cover image to embed the watermark. Figure 4(a) and 4(b) show the watermarked images of jet and couple image respectively. To evaluate the robustness of LSB method, seven kinds of noise were added to the watermarked images. The PSNR values of the marked images by LSB method are shown in Table 1.

Table 1. PSNR of Watermarked Image by LSB method

Image	PSNR
Jet	65.43 dB
Couple	65.43 dB

Figures 5(a-g) show the extracted watermarks of marked jet image attacked by (a) 3x3 averaging filter, (b) crop left-top 1/4, (c) 3x3 Laplacian mask , (d) jpeg compression , (e) rotating 90°, (f) 2x2 mosaic, and (g) Gaussian noise. Figures 6(a-g) show watermarks extracted from the marked Couple image under the same attacks as previous experiment. Obviously, the LSB method cannot extract complete watermark result in messy pattern. In other words, the hidden watermark has been destroyed by the attacks. Table 2 shows the PSNR values of the extracted watermarks of marked images under the seven types of attacks. From Table 2, the PSNR values are between 5.41 dB to 12.9dB. The low PSNR values indicate that the LSB method is unable to resist attacks by 3x3 averaging filter, crop left-top 1/4, 3x3 Laplacian mask, jpeg compression, rotating 90°, 2x2 mosaic and Gaussian noise.

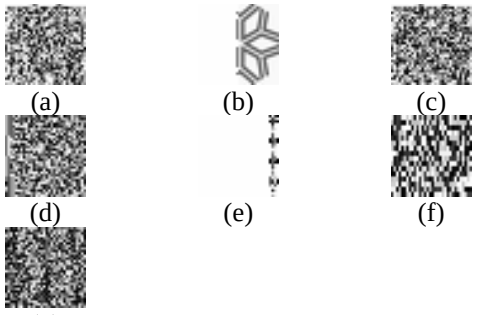


(a) (b) (c)

(d) (e) (f)

(g)

Figure 5 The extracted watermarks of watermarked Jet Image.



(a) (b) (c)

(d) (e) (f)

(g)

Figure 6 The extracted watermarks of watermarked Couple Image.

Table 2 PSNR of extracted watermark from watermarked Jet and Couple image by LSB method

Attack	PSNR of extracted watermarked from embedded Jet Image	PSNR of extracted watermark from embedded Couple Image
Blurred	7.22 dB	7.05 dB
Cropped	12.9 dB	12.9 dB
Sharpened	8.76 dB	8.12 dB
JPEG Compressed	6.48 dB	6.92 dB
Rotated	8.74 dB	8.74 dB
Mosaic	5.41 dB	6 dB
Gaussian Noise	6.83 dB	6.16 dB

3.2 Experimental 2: Robust testing for proposed Method

In this evaluation, the initial learning rates α and β was set to 0.95 and 0.7 respectively. The number of neurons is 32 and constant k_0 was set to 10. The threshold value of 1 was established to terminate training. Figures 7(a) and 7(b) show the watermarked Jet and Couple images respectively. The PSNR values of watermarked images by FCNN method are shown in Table 3. The high PSNR values indicate that the watermarked image is the most similar to the cover image.

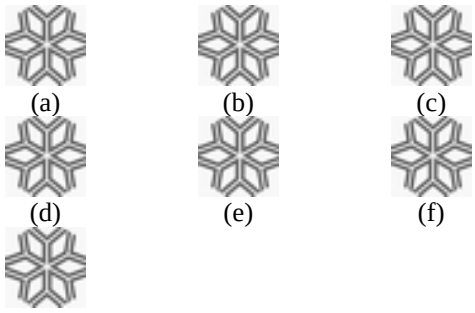
(a) (b)

Figure 7. Watermarked image, (a) Watermarked Jet image. (b) Watermarked Couple image.

Table 3. PSNR of Watermarked Image by proposed FCNN

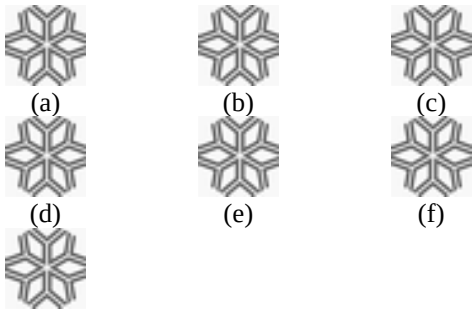
Image	PSNR
Jet	61.28 dB
Couple	59.14 dB

Figures 8(a-g) demonstrate the extracted watermarks from the watermarked Jet image attacked by (a) 3x3 averaging filter, (b) crop left-top 1/4, (c) 3x3 Laplacian mask, (d) jpeg compression, (e) rotating 90°, (f) 2x2 mosaic and (g) Gaussian noise. Figures 9(a-g) show the extracted watermarks of the marked 'couple' image attacked in the same ways as in the previous experiment. All the watermarks were completely extracted by means of FCNN, which evidences robustness facilitated by our method.



(a) (b) (c) (d) (e) (f) (g)

Figure 8 The extracted watermarks of watermarked Jet image.



(a) (b) (c) (d) (e) (f) (g)

Figure 9. The extracted watermarks of watermarked Couple image.

Table 4 shows the PSNR values of extracted watermarks after various attacks. It indicates the proposed method is robust enough to resist attacks such as 3x3 averaging filter, crop left-top 1/4, 3x3 Laplacian mask, jpeg compression, rotating 90°, 2x2 mosaic and Gaussian noise

Table 4. PSNR of Extracted Watermark from watermarked Jet and Couple Image by Proposed FCNN

Attack	PSNR of watermarks extracted from the embedded Jet Image	PSNR of watermarks extracted from the embedded Couple Image
Blurred	50.13 dB	50.06 dB
Cropped	50.13 dB	50.06 dB
Sharpened	50.13 dB	50.06 dB
JPEG Compressed	50.13 dB	50.06 dB
Rotated	50.13 dB	50.06 dB
Mosaic	50.13 dB	50.06 dB
Gaussian Noise	50.13 dB	50.06 dB

3.3 Experimental 3: Imperceptibility for proposed method

The setting of the threshold value is the major parameter related to the quality of watermarked

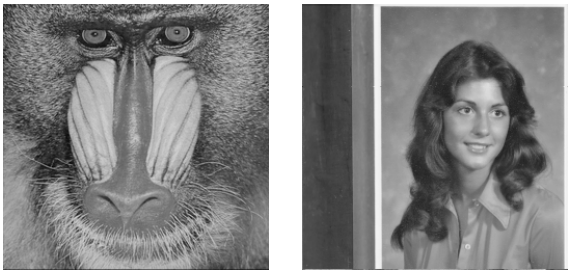
image. Table 5 shows the PSNR values and number of epoch of FCNN for different threshold values. We can obtain the higher PSNR from the table when the threshold is set to a small value. In other words, a small threshold value obtains a watermarked image that is the most similar to the cover image. However, the smaller threshold value implies more time for network training. The threshold value was selected to offer the best trade-off between imperceptibility and training time.

Table 5. PSNR of Watermarked Image

Threshold value	Embed watermark to Jet Image		Embed watermark to Couple Image	
	Epoch	PSNR	Epoch	PSNR
5	4	50.8	4	48.68
1	5	61.26	5	59.15
0.1	7	82.19	7	80.05
0.01	9	103.09	9	100.95

3.4 Experimental 4: Authenticity test

Raw cover image were used to test the authenticity of the proposed FCNN. Figure 10(a) and 10(b) are non-watermarked Baboon and Girl images. Figure 11(a) and 11(b) show the extracted watermarks, which are not watermarked images. As the results show, we cannot extract watermarks from the raw cover images. Hence the proposed method is able to extract corresponding watermark from marked images, but not from unmarked images.



(a) (b)

Figure 10. Non-training cover images: (a) the baboon, (b) the girl.

(a) (b)

Figure 11. The extracted watermarks of Fig. 10(a) and Fig. 10(b)

4. Conclusions

In this paper, a specific designed full counterpropagation neural network has been presented for digital image watermarking. Different from the traditional methods, the watermark was embedded in the synapses of FCNN instead of the cover image. The quality of the watermarked image was almost the same as the original cover image. In addition, because of the watermark was stored in the synapses, most of the attacks could not degrade the quality of the extracted watermark image. This shows that the proposed FCNN could resist various attacks. In addition, the watermark embedding procedure and extracting procedure is integrated into the proposed FCNN. By doing so, the proposed approach simplifies traditional procedures. The experimental results show that our application achieved robustness, imperceptibility and authenticity in digital watermarking.

Acknowledgment

This work was supported by the National Science Council, Taiwan, R.O.C., under grant nr. NSC 92-2213-E-224-041.

References

[1] Fredric M. Ham and Ivica Kostanic, *Principles of Neurocomputing for Science & Engineering*, McGraw-Hill, Singapore, 2001.

[2] R.G.van Schyndel, A.Z. Tirkel and C.F. Osborne, "A digital watermark," *Proceeding of IEEE International Conference on Image Processing*, vol: 2, pp: 86-92, Nov, 1994.

[3] Ren-Junn Hwang, Chuan-Ho Kao and Rong-Chi Chang, "Watermark in color image," *Proceeding of First International Symposium on Cyber Worlds*, pp: 225-229, Nov, 2002.

[4] Ahmidi N., Safabakhsh R., "A Novel DCT-based Approach for Secure Color Image Watermarking," *Proceedings ITCC 2004 International Conference on Information Technology: Coding and Computing*, vol. 2, pp. 709 – 713, Apr, 2004

[5] Fengsen Deng and Bingxi Wang, "A Novel Technique for Robust Image Watermarking in the DCT Domain," *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, vol. 2, pp. 1525 – 1528, Dec. 2003

[6] K.J Davis and K. Najarian, "Maximizing Strength of Digital Watermarks Using Neural Networks," *Proceeding of International Joint Conference on Neural Networks*, vol. 4, pp.2893 – 2898, July, 2001.

[7] Shi-chun Mei, Ren-hong Li, H. Dang and Yun-kuan Wang, "Decision of image watermarking strength based on artificial neural-networks," *Proceedings of the 9th International Conference on Neural Information Processing*, vol. 5, pp. 2430 – 2434, Nov, 2002.

[8] Zhang Zhi-Ming, Li Rong-Yan, Wang Lei, "Adaptive Watermark Scheme with RBF Neural Networks," *Proceedings of the 2003 International Conference on Neural Networks and Signal Processing*, vol. 2, pp. 1517 – 1520, Dec, 2003.