Multidisciplinary Review Journal

Journal of Critical Reviews

ISSN- 2394-5125 Vol 7, Issue 2, 2020

Review Article

MACHINE LEARNING BASED DIGITAL IMAGE WATERMARKING

K. Bramara Neelima[1], Dr.S. Arulselvi[2]

[1]Research Scholar, Bharath Institute of Higher Education and Research.
[2]Research Supervisor, Bharath Institute of Higher Education and Research.

Received: 13.*11.2019* ***Revised:* 15.*12.2019*** ***Accepted:* 17.*01.2020***

ABSTRACT

Watermarking is the method of adding information in digital media for content protection and authentication. Digital image watermarking is one of the solutions to offer value added security on the top of authentication and data encryption for content protection in digital images. The watermarking methods can be frequential and spatial domain methods. In this work, the frequency domain based watermarking method is implemented with the use of discrete wavelet transform (DWT). Along with the frequency domain technique, here we are utilizing the combination of topical developments of the mathematical techniques and most advanced algorithms of machine learning for digital image watermarking process. The mathematical technique considered here is the principal component analysis (PCA) because of its property of dimensionality reduction which further enhances the robustness to watermarking process. The machine learning algorithm taken into account in this work is the support vector machine (SVM) algorithm to increase the accuracy of watermarking process.

Keywords: Digital Image Watermarking, Principal Component Analysis, Support Vector Machine.

DOI: http://dx.doi.org/10.31838/jcr.07.02.24

INTRODUCTION

The rapid growth of communication and information exchange through digital media in the past three decades, in the form of sound, image and video, consequences the need of copyright protection techniques. Watermarking is the procedure of adding data in digital media for content protection and authentication. The data that embedded is termed as the watermark, and the digital form that contains it, is denoted as watermarked data. In case of transmission of digital data in the form of image, then the process is referred as digital image watermarking. There exist two types of image watermarking techniques based on the watermark visibility criteria in the watermarked image, termed as perceptive and imperceptive watermarking. The watermark is visible on watermarked image in perceptive watermarking and is invisible in imperceptive watermarking. Based on the watermark embedding method, there exist two types of digital image watermarking process, frequency domain and spatial domain watermarking. In spatial domain where the intensity value of the pixels is directly changed by inserting watermark and in frequency domain the discrete coefficients are changed to insert the watermark. There is always a tradeoff between robustness, capacity, complexity and imperceptibility in image watermarking. Spatial domain techniques are simple, more capacitive but high perceptive and less robust. Frequency domains are highly robust, imperceptive and more complex to implement than spatial domain techniques. The watermarking techniques have to provide copyright protection, authentication, content protection, and high resistance to attacks and unauthorized extraction of data.

The attacks on the watermarked image can be simple or difficult to handle since the main purpose of attacks is to extract, remove or modify the watermark. There exist four types of attacks based on the attacking methods functioned on watermark - active attacks, passive attacks, forgery attacks, and collusion attacks. More attacks on copyright protected content are active attacks, which are the result of removing the watermarks from the watermarked image. In passive attacks, the watermark is detected by unauthorized users and protected data is distributed without proper authentication. In forgery attacks, a new watermark is embedded in watermarked image, manipulates the data and distributed as original content. In collusion attacks, the protected content is reconstructed without watermarks based on the instances of data. And also there exist transmission attacks, due to noise in signal processing medium and data loss in compression methods. The watermarking methods should be robust against all types of attacks and protect from insertion of another watermark in data.

The algorithms of Machine learning are intended to predict the outcome from past behavior or observations. Machine learning tasks are categorized into three main groups – unsupervised, supervised, and semi-supervised learning. In this study, we are reviewing the numerous machine learning algorithms that applied for digital image watermarking. There are various classification and pattern recognition methods available for detection and extraction of watermark even though being attacked. The practice of machine learning algorithms on watermarking images improves the accuracy, performance and efficiency of watermark detection and extraction procedures. Some of the algorithms are applied with spatial domain watermarking and few applied along with frequency domain watermarking techniques. The general process of watermarking is showed in figure1.

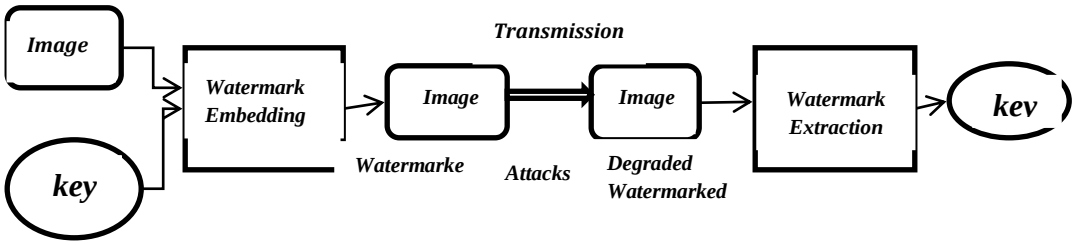

Figure 1: General Watermarking Embedding and Extraction Process

RELATED WORK

Numerous digital image watermarking methods had been suggested in the literature, spatial domain techniques [1-3], frequency domain techniques [4-10], combination with mathematical models [8-13] and recent works combined with machine learning [14-19] models. Visible watermarking [1] is the first spatial domain watermarking method that work on the pixel value of the image. The modifications of pixel intensity values by adding a watermark result in visible or invisible watermarking process. Least significant bit (LSB) modification [2] is the straight forward technique that embeds watermark in least significant bits of image. A digital watermarking method [3] is proposed based on the difference expansion along with LSB. In frequency domain methods, the frequency components are modified by adding the watermark. The components of 2D-DFT [4] are modified by inserting the watermark for extraction and embedding methods. Pranabkumar [5] suggested a watermarking process based on Fractional Fourier transform and 2D-FFT. Ferda Ernawan [6] presents a reliable watermarking method using optimal DCT for copyright protection. Yao yu [7] proposed the DWT based image watermarking algorithm for color images.

The mathematical techniques, such as SVD and PCA, are utilized in combined with frequency transform methods. Jerill George [8] proposed a color image watermarking technique that uses DWT-SVD. Anandkumar [9] proposed a semi visible watermarking scheme based on DWT-PCA. Himanshu [10] suggested a reliable watermark extraction system based on DCT-SVD. Abolfazi [11] proposed a new method for color image watermarking based on the DCT-PCA. Veni [12] proposed a new image watermarking algorithm based on combination of DWT-DCT to improve the imperceptibility and robustness. Maniekansai [13] proposed the image watermarking algorithm based on DWT-DCT-SVD for robust watermarking.

Yen [14] proposed a spatial domain digital image watermarking technique on support vector machine (SVM) model. Jianzhen [15] proposed a rotation scaling translation (RST) invariant image watermarking technique using SVM. Vasta [16] proposed combining DWT and SVD based digital image watermarking technique. Lei li [17] proposed spatial domain based watermarking technique on fuzzy support vector machine (FSVM). Vafaei [18] proposed a robust watermarking method uses artificial neural networks (ANN) in DWT domain. Ramamurthy [19] proposed a method to embed watermark using quantization in DWT based on back propagation neural network (BPNN) and dynamic fuzzy interference system (DFIS). Yahya [20] proposed embedding process through shifted least significant bit in discrete transform domain using support vector machine classification (LSB-DCT-SVM). Sakazawa [21] proposed visual decoding of watermark to a trained deep neural network (DNN) models to protect its copyright.

PROPOSED WORK

Digital image watermarking is one of the solutions to offer value added protection on the top of data encryption and clambering for content protection in digital images. The watermarking detection verifies the ownership and watermark extraction process proves the ownership of image content. The watermarking methods can be spatial and frequential domain methods. In this work, the frequency domain based watermarking method is implemented with the use of DWT. various studies on image watermarking proved that the discrete wavelet transform is more robust against various types of attacks. Hence, DWT coefficients of images are used to add the watermark in images and inverse discrete wavelet transform is used in watermark detection and extraction process. Along with the frequency domain technique, here we are utilizing the combination of topical developments of the mathematical techniques and most advanced algorithms of machine learning for digital image watermarking process. The mathematical technique considered here is the PCA because of its property of dimensionality reduction which further enhances the robustness to watermarking process. The machine learning algorithm taken into account in this work is the SVM algorithm to increase the accuracy of watermarking process.

Discrete Wavelet Transform (DWT)

Wavelet transform decomposes a signal into a set of basis functions, called wavelets. Discrete wavelet transforms decomposes an image into a set of sub-images, which is a series of wavelets that can be stored efficiently than pixel blocks. The wavelet transform converts the spatial domain pixels to frequency domain sub-bands. Haar wavelet transform is used in this work, which decomposes the discrete signals into two sub-signals of half in length. The two dimensional wavelet transform and its inverse transform equations are given below.

2D-DWT is

$$W_{\emptyset}(j_0,m,n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\emptyset_{j0,m,n}(x,y)$$

$$W_{\emptyset}(j,m,n) = \frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\emptyset^{i}_{j,m,n}(x,y), \quad i=\{H,V,D\}$$

Inverse DWT is

$$f(x,y) = \frac{1}{\sqrt{MN}} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W_{\emptyset}(j_0,m,n) \emptyset_{j0,m,n}(x,y) + \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j0}^{\infty} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W_{\emptyset}^{i}(j,m,n) \emptyset_{j,m,n}(x,y)$$

DWT divide an image into four coefficient images in the single level. Each coefficient image contains one of the low frequency bands. With an MXN image 2D-DWT produces four M/2XN/2 constants namely, LL denotes low frequency band, LH represents a horizontal high frequency band, HL is a vertical high frequency band and HH is a diagonal high frequency band. The most prominent data in the image appears in high amplitudes and less prominent information appears in very low amplitudes.

Principal Component Analysis (PCA):

It is a statistical technique used for dimensionality reduction or decorrelation. In Image processing, principal component analysis used as a linear transform that can be mainly used in finding pattern, image compression, finding similarities and differences, object recognition, and motion detection. PCA is a statistical method that utilizes orthogonal transformations to transform a potentially correlated set of data into a linearly uncorrelated set of data which contains principal components. The algorithmic steps for PCA are shown below.

PCA algorithm(X,k): Top k eigen values/ eigen vector

- $X = M \times N$ data matrix
- Each data point x_i= column vector i= 1,2,..m
- $\bar{X} = \frac{1}{m} \sum_{i=1}^{m} x_i$
- X ← subtract mean $\bar{X}$ from each column vector x_i in X
- $X = X - \bar{X}$
- $\sum \leftarrow XX^T$ covariance matrix of X
- $\{\lambda_i, u_i\}_{i=1,2,..N}$ = Eigen vector / Eigen values of $\sum$
- $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_N$
- $Return \{\{\lambda_i, u_i\}_{i=1,2,..k}\}$
- $Top\ k\ principal\ components$

Support Vector Machine (SVM):

Machine learning is the scientific study of statistical models and algorithms that computer system used to perform a specific task without using explicit instructions. Machine learning algorithms are mainly divided into two types; supervised learning and unsupervised learning. SVM is the one of the mostly used supervised machine learning algorithm that analyses the data for classification and regression. SVM works on training examples and can generate a separation hyper plane that separates positive examples and negative examples. SVM is naturally a binary classification model that separates data into two classes. It is suitable for watermarking to detect whether there is a watermark or not. There are three types of kernels used in SVM; radial basis function (RBF), polynomial, and two layer perceptron. RBF kernel is more powerful and efficient than other two kernels hence it is preferred to train the classifier. The SVM kernels are defined below.

Kernel	Equation
Linear	$K(x,y) = x \cdot y$
Sigmoid	$K(x,y) = \tanh(ax.y + b)$
Polynomial	$K(x,y) = (1 + x \cdot y)^d$
KMOD	$K(x,y) = a\left[\exp\left(\frac{\gamma}{\|x-y\|^2+\sigma^2}\right) - 1\right]$
RBF	$K(x,y) = \exp(-a\|x-y\|^2)$
Exponential RBF	$K(x,y) = \exp(-a\|x-y\|)$

Figure 2: From Wikipedia

Watermarking process:

The procedure of embedding watermark in an image is broadly classified into two sorts; invisible watermarking and visible watermarking; based on the visibility of watermark in the embedded image. In visible watermarking of images, the watermark is added on the primary or cover image such that the watermark is calculatedly perceptible to a human observer whereas in case of invisible watermarking, the embedded image contains of watermark that is not observable, but removed by a computer program. The authentication key is the watermark, which is processed by the principal component analysis and is added in the cover image. The invisible watermarking process is applied using the two dimensional DWT on the cover image by placing the watermark at mid regions. The watermarked image is directed through transmission where various attacks are performed on watermarked image. The obtained degraded watermarked image at receiver end is processed for watermark extraction process which is the step by step inverse of embedding process.

The watermarking algorithms are evaluated with respect to two metrics: imperceptibility and robustness. The supposed quality of the original image should not be partial by the occurrence of the watermark. PSNR is used to measure the quality which results the imperceptibility. Generally, PSNR is defined in terms of MSE and is defined as follows.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2$$

$$PSNR = 10\ \log_{10}(255^2/MSE)$$

Where n is the number of pixels in the cover image. y_i and $\tilde{y}_i$ are the original and watermarked images respectively. Robustness is the amount of the immunity of watermark against challenges to degrade, modify, or remove it. The similarity among the original watermark and the watermark extracted at receiver end is defined by correlation factor (CF) and is defined as follows.

$$\rho = \sum_{i=1}^{m} w_i w_i' \Big/ \sqrt{\sum_{i=1}^{m} w_i} \sqrt{\sum_{i=1}^{m} w_i'}$$

Where m is the number of pixels in the watermark, w_i and w_i' are the original and extracted watermarks correspondingly. The correlation factor of about 0.75 or above is considered acceptable. The performance metrics of proposed method are related with the similar existing approaches are shown in table1 below.

Table 1: Performance metrics of watermarking algorithms

Algorithm	*DWT*	*PCA*	*DWT-SVM*	*DCT-SVM*	*DWT-PCA-SVM*
MSE	*0.0563*	*0.0303*	*0.0392*	*0.0374*	*0.0325*
PSNR	*73.1207*	*78.502*	*75.126*	*78.367*	*78.734*
CF	*0.7932*	*0.7663*	*0.8954*	*0.915*	*0.962*

CONCLUSION

The watermarking detection verifies the ownership and watermark extraction process proves the ownership of image content. The watermarking techniques have to provide copyright protection, authentication, content protection, and high resistance to attacks and unauthorized extraction of data. In this work, the frequency domain based watermarking method is implemented with the use of DWT, PCA and SVM algorithm. Proposed watermarking algorithm is estimated with detail to imperceptibility and robustness, in terms of MSE, PSNR and correlation coefficient. From the observation the suggested process yields better performance related to other algorithms.

REFERENCES

1. S. Yip, O.C. Au, C. Ho and H. Wong, "Lossless Visible Watermarking," *2006 IEEE International Conference on Multimedia and Expo*, Toronto, Ont., 2006, pp. 853-856.
2. N. Bansal, V.K. Deolia, A. Bansal and P. Pathak, "Digital Image Watermarking Using Least Significant Bit Technique in Different Bit Positions," *2014 International Conference on Computational Intelligence and Communication Networks*, Bhopal, 2014, pp. 813-818.
3. A. Bamatraf, R. Ibrahim and M.N.B.M. Salleh, "Digital watermarking algorithm using LSB," *2010 International Conference on Computer Applications and Industrial Electronics*, Kuala Lumpur, 2010, pp. 155-159.
4. C. Pun, "A Novel DFT-based Digital Watermarking System for Images," *2006 8th international Conference on Signal Processing*, Beijing, 2006.
5. P.K. Dhar and I. Echizen, "Robust FFT Based Watermarking Scheme for Copyright Protection of Digital Audio Data," *2011 Seventh International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, Dalian, 2011, pp. 181-184
6. F. Ernawan and M.N. Kabir, "A Robust Image Watermarking Technique With an Optimal DCT-Psychovisual Threshold," in *IEEE Access*, vol. 6, pp. 20464-20480, 2018.
7. G. Sun and Y. Yu, "DWT Based Watermarking Algorithm of Color Images," *2007 2nd IEEE Conference on Industrial Electronics and Applications*, Harbin, 2007, pp. 1823-1826.
8. J. George, S. Varma and M. Chatterjee, "Color image watermarking using DWT-SVD and Arnold transform," *2014 Annual IEEE India Conference (INDICON)*, Pune, 2014, pp. 1-6.
9. A. Kumar and M. Gupta, "Semi visible watermarking scheme based on DWT and PCA," *2015 International Conference on Green Computing and Internet of Things (ICGCIoT)*, Noida, 2015, pp. 986-990.
10. Himanshu, S. Rawat, B. Raman and G. Bhatnagar, "DCT and SVD Based New Watermarking Scheme," *2010 3rd International Conference on Emerging Trends in Engineering and Technology*, Goa, 2010, pp. 146-151.
11. S. Abolfazl Hosseini and A. Saboori, "A new method for color image watermarking based on combination of DCT and PCA," *2015 International Conference on Communications, Signal Processing, and their Applications (ICCSPA'15)*, Sharjah, 2015, pp. 1-5.
12. M. Veni and T. Meyyappan, "DWT DCT based new image watermarking algorithm to improve the imperceptibility and robustness," *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*, Karur, 2017, pp. 1-6.
13. M. Kansal, G. Singh and B.V. Kranthi, "DWT, DCT and SVD Based Digital Image Watermarking," *2012 International Conference on Computing Sciences*, Phagwara, 2012, pp. 77-81.
14. Shwu-Huey Yen and Chia-Jen Wang, "SVM Based watermarking technique", *Tamkang Journal of Science and Engineering*, vol. 9, no.2, 2006.
15. Wu Jianzhen, "A RST Invariant Watermarking Scheme Utilizing Support Vector Machine and Image Moments for Synchronization", *IEEE International Conference on Information Assurance and Security*, 2009.
16. Mayank Vatsa, Richa Singh and Afzal Noore, "Improving biometric recognition accuracy and robustness using DWT and SVM watermarking", *IEICE Electronics Express, vol.2, no.12*, pp.362-367, 2005.
17. Lei Li, Wen-Yan Ding and Jin-Yan Li, "A Novel Robustness Image Watermarking Scheme Based on Fuzzy Support Vector Machine", *IEEE Pattern Recognition and Intelligence System*, 2010.
18. M. Vafaei, H. MahdaviNasab and H. Pourghassem, "A new blind Robust Watermarking method based on Neural

Networks in Wavelet Transform Domain", *World Applied Science Journal*, vol.22, no.11, 2013.
19. Nallagarla Ramamurthy and Dr.S. Varadarajan, "Robust Digital image watermarking scheme with Neural Network and fuzzy logic approach", *International Journal of Emerging Technology and Advanced Engineering*, vol.2, no.9, 2012.
20. SaadiahYahya, Hanizan Shaker Hussain and Fakariah Hani M. Ali, "DCT Domain Stega SVM- shifted LSB Model for highly Imperceptible and robust cover image", *International Conference on Computing and Informatics, vol. 43*, 2015.
21. S. Sakazawa, E. Myodo, K. Tasaka and H. Yanagihara, "Visual Decoding of Hidden Watermark in Trained Deep Neural Network," *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, San Jose, CA, USA, 2019, pp. 371-374.
22. Santosh Nemichand Kale,Sharada Laxman Deore. "Emulsion Micro Emulsion and Nano Emulsion: A Review." *Systematic Reviews in Pharmacy* 8.1 (2017), 39-47. Print. doi:10.5530/srp.2017.1.8