Journal of Information Security and Applications 38 (2018) 71–84

Contents lists available at ScienceDirect

Journal of Information Security and Applications

journal homepage: www.elsevier.com/locate/jisa

Bi-directional extreme learning machine for semi-blind watermarking of compressed images

Anurag Mishra[a], Ankit Rajpal[b,*], Rajni Bala[c]

[a] Department of Electronics, Deen Dayal Upadhyaya College, University of Delhi, Delhi, India
[b] Department of Computer Science, University of Delhi, Delhi, India
[c] Department of Computer Science, Deen Dayal Upadhyaya College, University of Delhi, Delhi, India

ARTICLE INFO

Article history:

Keywords:
Semi-blind watermarking
Bi-directional extreme learning machine
Normalized correlation (NC)
Peak signal to noise ratio (PSNR)
SSIM_Index
Bit error rate (BER)
Root mean square error (RMSE)

ABSTRACT

Bi-directional Extreme Learning Machine (B-ELM) is a newly developed single layer feed-forward network capable of fast training with few hidden neurons. It is also reported to show better generalization capability as compared to its old counterpart ELM. In the past, it has never been applied to image processing data-sets and particularly to any of its applications. In this work, B-ELM is successfully used to carry out watermarking of JPEG compressed images by inserting a binary watermark into it. Two invertible activation functions – Sine and Sigmoid are tested in this work. The RMSE is plotted as a function of number of hidden neurons. As observed in case of other applications, this plot indicates that Sigmoid is better placed in comparison to Sine function. The robustness of embedding scheme is examined by applying seven different attacks over signed images. These results prove that the proposed scheme is robust enough against the selected attacks. The computed processing time for embedding and extraction in milliseconds indicates that this scheme is suitable for developing real time watermarking applications for videos.

1. Introduction

Digital watermarking is an efficient technique to establish ownership of copyright, content authentication, broadcast monitoring and device control. The technique starts with embedding a small amount of copyright content such as a logo or a registered watermark into a host signal either in a visible or an invisible manner. The research on watermarking of host signal may be classified into spatial domain [1] and transform domain [2]. Robustness of the embedding scheme being an important requirement, transform domain embedding is more popular and is also proved to be versatile. The issue of fragility or robustness of watermarking scheme is however, independent of extraction of watermark. On one hand, the watermarking may be fragile yet blind, on the other hand it may be robust, yet informed. Qin et al. [3,4] have proposed a fragile image watermarking scheme based on reference-data interleaving mechanism and pixel-wise recovery based on overlapping embedding strategy.

As said above also, the recovery of watermark is broadly classified as: Public or Informed, Semi-Blind and Completely Blind. In case of public watermarking scheme, the original image is needed to extract the watermark from the signed image. The semi-blind method utilizes the key and the evolved model of the corresponding machine to extract the watermark from the signed image without using original image. A complete blind method does not make use of the original image, key or evolved model of the machine to extract the watermark from the signed image [5].

Presently, the research on image watermarking is focused towards optimization of the twin criteria : (1) The visual quality of the signed and attacked images, and (2) The robustness of the embedding scheme which is examined after carrying out certain image processing attacks over the signed content [6]. Among these two criteria, the visual quality is commonly assessed by computing two full reference metrics – Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM_Index). On the other hand, Normalized Correlation (NC) and the Bit Error Rate (BER) are used to examine the similarity between the embedded and the extracted watermarks. Thus, the issue robustness is a function of computed numerical values of NC and BER. In the present work, the extraction of watermark from signed and attacked images is carried out by utilizing the original B-ELM model constructed during embedding without involving the original image. Hence, it is semi-blind in nature. The robustness of the embedding scheme assumes even more significance because this is used to ascertain the ownership verification and content authentication. It is widely believed that

* Corresponding author.
E-mail address: ankit.rajpal@ieee.org (A. Rajpal).

https://doi.org/10.1016/j.jisa.2017.11.008
2214-2126/

if any watermarking scheme is biased towards robustness then it shall be at the expense of visual quality and vice versa. Therefore, it becomes even more important to minimize the trade-off between these two parameters. The third quantity, i.e., the capacity of the host signal to hold the watermark successfully is ignored as the size of the embedded information (watermark) is usually very small to create any significant perturbation in the host signal coefficients.

Over last many years, these two parameters are being optimized by extensive use of soft computing techniques. This is particularly so because these techniques are conventionally used to train datasets pertaining to different problems and applications. The training usually yields optimal or sub-optimal solutions which are found to be helpful to reach specific conclusions. These techniques include different types of neural networks, Fuzzy Inference Systems (FIS), Support Vector Machines (SVMs) of different configurations, meta-heuristic techniques and swarm intelligence. The neural networks which are used for this purpose include Back-propagation Networks (BPN) [7,8], Radial-Basis Function Networks (RBF) [9] and several other variants of Single-Layer Feed-Forward Neural Networks (SLFNs). Among these configurations, the BPN are found to be slow especially to large datasets. They are usually found to be trapped within the local minima and therefore, the training is very slow or does not complete at all. Therefore, these networks are hardly suitable for the design of information security applications such as image watermarking. On the contrary, the RBF networks are faster and can be trained in an efficient manner as compared to their BPN counterparts. This is generally true for all applications including those based on image processing techniques. Although, both these networks are adaptive and suitable for training the datasets yet the training is slow to carry out processing on a real time scale.

The FIS based schemes are not adaptive in nature. But, the use of fuzzy inference rules by these systems make it more close to natural imitation of the processes involved in embedding and extraction of watermarks from images. This is true particularly to the use of Human Visual System (HVS) based watermarking applied in conjunction with Mamdani or Sugeno type FIS. In this case, more specifically, the visual quality of the signed and attacked images is found to be better as compared to their neural networks counterparts [10]. The issue of robustness using different FIS has also been examined by different groups. As fuzzy based algorithms are not adaptive in nature, therefore, developing a robust image watermarking application based on FIS is comparatively difficult. The FIS based watermarking scheme is also slow in respect of embedding and extraction processes. This shortcoming is avoided by amalgamation of neural networks which are adaptive and FIS which are more close to reality by making it as a neuro-fuzzy architecture [10,11]. This is done with an expectation to give the best of both.

The third category of techniques is based on Support Vector Regression (SVR). Shen et al. [12] have proposed a novel SVR based watermarking scheme which has shown good generalization ability. The watermark can be successfully extracted unless the watermarked image is damaged severely. However, the architecture of basic SVM is also found to be slow for image processing applications such as image watermarking. It usually consumes a time span of the order of few minutes to complete the embedding of a watermark in a regular grayscale or color image. The faster variants to SVR based watermarking are also available. These include Finite Newton Support Vector Regression (FNSVR) [13] and Least Square Support Vector Regression (LSSVR) [14]. Mehta et al. [13] concluded that the processing time of FNSVR based watermarking is less and their algorithm is stable which converges only after few iterations while the LSSVR is not as fast as FNSVR but it is found to be faster than regular SVR [12].

The major shortcoming of all these techniques is that these are only suitable for discrete digital images and can hardly be applied to sequence of frames composed in uncompressed or compressed video. This is primarily because these do not complete training, embedding and extraction process on a real time scale. Faster variants of SLFN are available now-a-days which are capable to finish off training in millisecond time span and hence can be successfully applied to videos as well. A brief evolution to these techniques is given below.

Huang et al. [15,16] proposed Extreme Learning Machine (ELM) which is a fixed-network architecture. The basic idea of ELM is to begin with a fixed number of hidden neurons before building a training model. Input weights are randomly chosen and output weights are analytically computed by Moore Penrose Matrix Inverse [17]. The ELM is batch learning algorithm which computes a unique optimal solution without any necessity of iterations. The easy parameter selection and fast learning effectiveness makes the ELM suitable for various research domains and their datasets. Mishra et al. [18] have proposed an ELM based informed image watermarking scheme in DCT domain. Rajpal et al. [19,20] have proposed a semi-blind watermarking scheme using ELM in DWT domain. Liang et al. [21] have developed an online sequential learning algorithm (OS-ELM) which learns data one-by-one or chunk-by-chunk with fixed or varying chunk size. OS-ELM has also been used for watermarking of grayscale and colored images [22,23].

Miche et al. [24] have proposed an Optimally Pruned ELM (OP-ELM) which is similar to ELM in its initial steps but later applied Multiresponse Sparse Regression (MRSR) algorithm to rank good hidden neurons and further validated the model by Leave-One-Out (LOO) algorithm. The accuracy of OP-ELM is better than other well-known methods such as Support Vector Machine (SVM), Multi Layer Perception (MLP), or Gaussian Process (GP). Although, the original ELM is much faster than OP-ELM but in the case of few datasets the ELM fails to show good generalization capability. On the other hand, the OP-ELM remains robust to all tested datasets.

Huang et al. [25] have also proposed an increased network architecture known as Incremental ELM (I-ELM) which reduces network structure and training time. Hidden nodes are randomly generated and output weights are analytically determined in this SLFN by using I-ELM. However, there is no need to recalculate the output weights of the existing nodes whenever a new hidden node is added.

The I-ELM is further improved upon by Huang et al. [26] in its enhanced version called EI-ELM. The basic distinction between I-ELM and EI-ELM is that EI-ELM picks the optimal hidden nodes among several randomly generated hidden nodes which leads to minimization of residual error. Compared with the original I-ELM, the EI-ELM can achieve faster convergence rate and much more compact network architectures.

Feng et al. [27] have proposed a simple and efficient method called Error Minimized ELM (EM-ELM) which automatically determines the number of hidden nodes in generalized SLFNs. This new method concluded that the random hidden nodes can be added one by one or group by group (with varying group size) and the output weights can be incrementally updated efficiently during the growth of the networks. Furthermore, the performance of EM-ELM has been compared with other methods - Resource Allocation Network (RAN), Minimum Resource Allocation Network (MRAN), I-ELM as well as the original ELM. EM-ELM is found to be faster than the traditional ELM with the similar generalization performance [27].

Yang et al. [28] have proposed a new learning algorithm called Bi-directional Extreme Learning Machine (B-ELM). The B-ELM is an incremental learning algorithm whose architecture and the pseudo-code is given in Section 2. According to Yang et al. [28], unlike other incremental variants of ELM, the B-ELM is better in