Computers in Biology and Medicine 68 (2016) 37–48

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Analysis of the impact of digital watermarking on computer-aided diagnosis in medical imaging

Jose Juan Garcia-Hernandez*, Wilfrido Gomez-Flores, Javier Rubio-Loyola

Technology Information Laboratory, Center for Research and Advanced Studies of the National Polytechnic Institute, Ciudad Victoria, Tamaulipas, Mexico

ARTICLE INFO

Article history:
Received 4 May 2015
Accepted 27 October 2015

Keywords:
Breast ultrasound
Computer-aided diagnosis
Data security
Segmentation
Watermarking

ABSTRACT

Medical images (MI) are relevant sources of information for detecting and diagnosing a large number of illnesses and abnormalities. Due to their importance, this study is focused on breast ultrasound (BUS), which is the main adjunct for mammography to detect common breast lesions among women worldwide. On the other hand, aiming to enhance data security, image fidelity, authenticity, and content verification in e-health environments, MI watermarking has been widely used, whose main goal is to embed patient meta-data into MI so that the resulting image keeps its original quality. In this sense, this paper deals with the comparison of two watermarking approaches, namely spread spectrum based on the discrete cosine transform (SS-DCT) and the high-capacity data-hiding (HCDH) algorithm, so that the watermarked BUS images are guaranteed to be adequate for a computer-aided diagnosis (CADx) system, whose two principal outcomes are lesion segmentation and classification. Experimental results show that HCDH algorithm is highly recommended for watermarking medical images, maintaining the image quality and without introducing distortion into the output of CADx.

1. Introduction

Clinical data are one of the main sources of information for the diagnosis and treatment of a large number of illnesses and abnormalities. This includes physiological variables such as electrocardiograms (ECG or EKG), electroencephalograms (EEG), electromyograms (EMG), and, also, different types of medical imaging such as computed tomography (CT), ultrasound (US), magnetic resonance imaging (MRI), and X-rays. Medical images (MI) are usually collected in hospitals or research centers due to the specialized nature of the equipment. There are standards, such as DICOM (Digital Imaging and Communications in Medicine), that have enabled the management, storage and printing of MI in standard formats. DICOM-based files facilitate sending MI to specialists or researchers in remote places by electronic means. However, during an information exchange, the potential risk of altering the MI contents increases within the processing tasks that they may be exposed to [1]. In addition, specialists usually capture sensitive, confidential diagnostic information about MI as flat text, which is normally electronically, compromising the confidentiality of the patient with respect to potential illnesses or injuries that the patient may suffer. In most countries, strict regulations apply to protect a patient's confidentiality, such as electronic patient records (EPR), identification, and participation in any trials or programs [2,3]. From the above, the constant need for security procedures to keep MI information unaltered is widely accepted.

MI watermarking has been widely recognized as a technique for enhancing data security, image fidelity, authenticity, and content verification in e-health environments, where MI are stored, retrieved, and transmitted electronically [4]. In addition, it helps control the integrity of MI for protection purposes, as a key objective of MI watermarking is to embed data into the MI so that the watermarked MI contains useful information with a perfect linking with the patient. The embedded information could include the hospital, patient records, diagnostic information, etc., in this way concealing such information. In this context, a number of MI watermarking techniques have been proposed to find appropriate trade-offs between the payload (i.e., the hidden message or watermark) and its imperceptibility (i.e., the quality of the image). Moreover, image degradation due to aggressive or greedy watermarking impacts directly on the image fidelity and, consequently, on key medical activities like specialists' diagnoses and computer-aided diagnosis (CADx) [5]. Hence, finding the appropriate techniques that guarantee an adequate trade-off between payload and imperceptibility are necessary, so that clinical diagnosis can be performed appropriately on watermarked MI. To the best of our knowledge, a deep and detailed analysis of the implications of MI watermarking techniques over CADx has received little attention.

* Corresponding author. Tel.: +52 834 107 0220; fax: +52 834 107 0224.
E-mail addresses: jjuan@tamps.cinvestav.mx (J.J. Garcia-Hernandez), wgomez@tamps.cinvestav.mx (W. Gomez-Flores), jrubio@tamps.cinvestav.mx (J. Rubio-Loyola).

http://dx.doi.org/10.1016/j.compbiomed.2015.10.014
0010-4825/

This study deals with the critical nature of identifying an adequate watermarking approach such that the resulting watermarked MI guarantees the suitability of the CADx systems. The main contribution of the present paper is the identification of an appropriate embedding algorithm, payload parameters, and limits, that guarantee an acceptable trade-off between image quality and computerized segmentation. Due to their importance, our analysis is focused on breast ultrasound (BUS) images, which are crucial for detecting the most common cancer among women worldwide.

The present paper is organized as follows. Section 2 presents related research. Section 3 describes background information on computer-aided diagnosis. Section 4 describes the implemented watermarking techniques for data hiding, which are applied to the BUS images. Section 5 presents the experimental results obtained through the former in terms of payload (i.e., the hidden message or watermark) and medical image quality. Section 6 provides some important discussions of our findings and some conclusions. Finally, Section 7 concludes this paper with a summary.

2. Watermaking approaches

The application of watermarking techniques in MI has attracted attention mainly for authentication and protection purposes, as discussed by Coatrieux et al. [6]. The relevance of watermarking in MI for security and privacy has been pivotal for qualitative enhancement in imaging technology in medical and health information systems, as suggested by Rao and Kumari [4]. In this sense, the zero-watermarking method has been preferred, since the watermark is not embedded physically in the image, but both the master and the secret shares are created out of the host image as well as the watermark image at the sender's end, so that the watermark is reconstructed by the combination of these two at the receiver end [7]. Dong et al. [8] proposed a zero-watermarking algorithm based on the discrete cosine transform (DCT), which combines visual feature vectors, encryption, and third party authentication to address issues of security, confidentiality, and integrity. Also, Seenivasagam and Velumani [9] proposed a general framework for patient authentication and controlled access to Electronic Health Records (EHR) in tele-radiology environments based on the zero-watermarking scheme.

On the other hand, watermarking schemes applied to breast screenings are rather scarce. Li et al. [10] proposed a DCT-based scheme for protecting mammograms on Picture Archiving and Communication Systems (PACSs). The goal was to protect the textual information on mammograms as well as verifying their authenticity and integrity when the data extraction process does not require the availability of the originals. A similar approach is taken by Manaf et al. [11], aiming to identify the best location on the mammogram to embed the patient's information without affecting the quality of the image as well as to develop an authentication technique of watermarking mammogram using the least significant bits (LSB) technique. Likewise, Hajjaji et al. [12] developed a watermarking algorithm based on DCT for hiding information in MRI, radiographic, and echographic data.

Watermarked images could potentially be used for CADx systems to detect and diagnose suspected regions such as breast tumors, for instance. A critical stage within a CADx system is the image segmentation, whereby a region of interest (ROI) is isolated from the adjacent structures for further analysis. In this sense, a watermarked image should preserve the relevant information about the tissue morphology and texture patterns, which are commonly used to detect accurately suspected structures like tumors or injuries. The term "accurately" means that the segmented region corresponds to a specialist's perception. Usually, the segmentation stage is implemented within a computer-aided detection (CADe) module, where a wide variety of techniques, such as edge-based methods, region-based approaches, active contour models, split-and-merge, and morphological watersheds, have been explored [13].

Some segmentation-oriented watermarking schemes have been addressed by Lim and Feng [14], Rathi [15], and Gunjal and Mali [16]. These approaches use segmentation information of medical image contents to embed watermarks in medical images; however, they are mainly focused on secure distribution [14], protection and authentication [15], and security and copyright protection [16] rather than payload maximization (i.e., maximization of the length of the hidden message or watermark). A recent study of embedding capacity was presented by Al-Qershi and Khoo [17]; however, the impact of the amount of information hidden by distinct watermarking algorithms on the accuracy of detected abnormalities has not yet been studied. Therefore, we will here analyze the effect of watermarking on the outcomes of a CADe system developed for BUS images.

3. Computer-aided diagnosis

Computer-aided diagnosis (CADx) systems have been developed to assist specialists in locating and diagnosing potential abnormalities by using the outcomes from a computerized analysis of the medical images. Commonly, CADx systems involve four main stages [18]: image preprocessing, lesion segmentation, feature extraction, and lesion classification. Many types of CADx systems have been proposed to diagnose different lesions in medical imaging, such as ultrasound (US), X-ray, computed tomography (CT), and magnetic resonance imaging (MRI). In addition, CADx has been used to analyze various human organs, such as the breast, lung, colon, brain, liver, and vascular system [19].

Regarding breast ultrasound (BUS), CADx systems have been used to segment masses that are then depicted so as to quantify their malignancy. Hence, lesion segmentation is a critical stage within CADx systems, whereby the lesion should be separated accurately from the background and other structures [18].

As discussed in Section 2, an important aspect of digital watermarking consists in hiding meta-data in the image, where the embedded information remains invisible to the user and the quality of the watermarked image is preserved [20]. In this sense, when a watermarked BUS image is analyzed by a CADx system, the classification performance should be similar to that of the classification without watermarking. For this reason, a watermarking approach should avoid critical image data distortion.

Hence, in this study, the goal is to quantify the impact of watermarking on the performance of a CADx for BUS that was developed by our research group. It follows the basic aforementioned stages [21–23]:

1. *Image preprocessing*: Contrast enhancement is performed by means of the contrast-limited adaptive histogram equalization (CLAHE) technique and speckle is reduced by using an anisotropic diffusion filter guided by texture descriptors. Next, using a constraint Gaussian function, the lesion region is enhanced to attenuate distant pixels (far from the lesion) whose intensity values are similar to those of the tumor region.
2. *Segmentation*: An iterative threshold procedure creates binary masks for the marker-controlled watershed transformed to create potential lesion-like margins. The goal is to preserve the fine details of the tumor boundary. Then, the margin that maximizes the average radial derivative (ARD) function approximates the final lesion contour.
3. *Feature extraction*: Five morphological descriptors are computed from binary blobs to describe the following features: elliptic-