Journal of Digital Imaging

Security Protection of DICOM Medical Images Using Dual-Layer Reversible Watermarking with Tamper Detection Capability

Chun Kiat Tan,[1] Jason Changwei Ng,[1] Xiaotian Xu,[1] Chueh Loo Poh,[2] Yong Liang Guan,[1] and Kenneth Sheah[3]

Teleradiology applications and universal availability of patient records using web-based technology are rapidly gaining importance. Consequently, digital medical image security has become an important issue when images and their pertinent patient information are transmitted across public networks, such as the Internet. Health mandates such as the Health Insurance Portability and Accountability Act require healthcare providers to adhere to security measures in order to protect sensitive patient information. This paper presents a fully reversible, dual-layer watermarking scheme with tamper detection capability for medical images. The scheme utilizes concepts of public-key cryptography and reversible data-hiding technique. The scheme was tested using medical images in DICOM format. The results show that the scheme is able to ensure image authenticity and integrity, and to locate tampered regions in the images.

KEY WORDS: Digital watermark, security, image authentication, teleradiology, public-key cryptography

INTRODUCTION

With the advent of teleradiology, there is an increasing need for doctors to transmit images to healthcare professionals all over the globe to seek high-quality diagnosis or second opinions. As a result, medical image security has become an important issue when medical images are being transmitted over open network, where sensitive patient information is exposed to hackers or individuals with malicious intents. Possible security breaches may include tampering of images to include false data which may lead to wrong diagnosis and treatment. There are several mandates and guidelines in place to protect sensitive patient information. The Health Insurance Portability and Accountability Act requires healthcare providers to take measures to ensure the security of medical images so as to protect patient's privacy.[1] The Digital Imaging and Communication in Medicine (DICOM) standard aims to define a technical framework for application entities involved in the exchange of medical data to adhere to a set of security profiles.[2] DICOM standard has incorporated digital signatures into DICOM object which can be used to check the integrity of medical images. However, digital signatures cannot locate where the images have been tampered. Current security measures have their limitations.[3] Cryptography is able to ensure security in terms of storage and transmission; but, once decrypted, the information is no longer protected. Firewalls and access–control methods only protect the images up to the point of the internal networks. Authenticity problems are often a result of human actions such as illegal distribution or human error in transmitting to unauthorized individual. To ensure the authenticity of the images, the two common tools used are digital signature and watermark. A digital signature is the

[1]*From the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, Singapore.*

[2]*From the School of Chemical and Biomedical Engineering, Nanyang Technological University, 70 Nanyang Drive, N1.3 B2-09, 637457, Singapore, Singapore.*

[3]*From the Department of Diagnostic Radiology, Changi General Hospital, Singapore, Singapore.*

Correspondence to: Chueh Loo Poh, School of Chemical and Biomedical Engineering, Nanyang Technological University, 70 Nanyang Drive, N1.3 B2-09, 637457, Singapore, Singapore; tel: +65-6514-1088; fax: +65-6791-1761; e-mail: clpoh@ntu.edu.sg

Online publication 23 April 2010
doi: 10.1007/s10278-010-9295-4

non-repudiation, encrypted version of the message digest extracted from the data to prove integrity and originality.[4] The security of digital signature largely depends on the strength of the hash functions used to validate the signatures. It has been demonstrated that it is possible to generate two datasets with different content but having the same Message-Digest algorithm 5 (MD5) hash.[5] As a result, it is then possible to append arbitrary data to the dataset, and their hash value may still be the same. In mathematical terms, if $\mathrm{MD5}(x) = \mathrm{MD5}(y)$, then $\mathrm{MD5}(x + q) = \mathrm{MD5}(y + q)$[6] (where x and y could represent two different 128-byte datasets and q is an arbitrary dataset of any length). We can then apply these concepts to medical images, for example, by modifying the first 1,024 bits of the pixel values of an image. Consequently, two images can be nearly identical except for the six pixels, and the two images can produce the same MD5 hash. This shows that it could be possible for a hacker to tamper an image to include artifacts that may lead to wrong medical diagnosis while keeping the MD5 of the image unchanged. This type of tampering may also give rise to serious security issues if the image was used in a legal or police investigation.

Watermarking is the practice of imperceptibly adding hidden data to the cover-signal (e.g., image, audio, video, or other work of media) in order to convey the hidden data. In the context of medical images, the hidden data can be used to verify the authenticity of the images. This provides an alternative technique to protect medical images. It allows messages to be indiscernibly embedded into an image by modifying the pixel values.[7] The key characteristics of digital watermarking are imperceptibility, integrity control, and hiding capacity.[8]

Three different types of watermarking methods can be considered for use on digital medical images.[3,7–14] The first method uses non-reversible watermarking techniques that will introduce permanent alterations to the images. Examples of such watermarking techniques include replacing the least significant bit of the image to embed information[9] and a trusted header scheme[10] to embed the hash of the metadata into the image itself. Using non-reversible watermarking and public-key cryptography scheme, Zhou et al. has demonstrated that authenticity and integrity can be verified for digital mammography images.[9] However, because the watermarks are permanently embedded in the image, the proof that the watermarks will not introduce any judicious marks on the image that will cause any image misinterpretation has to be assessed.[3]

The second method is reversible watermarking which allows the watermarks to be fully removed once the watermarks have been detected and verified.[7,11–15] This method allows the image to be restored to its original pixel values. Hence, original images can be used in medical diagnosis. Guo and Zhuang developed a region-based lossless watermarking based on difference expansion embedding technique,[12] similar to the method implemented by Coatrieux et al.[7] This technique enables data to be embedded and fully removed. Coatrieux et al. proposed the use of an estimator signal, which is invariant to the watermarking process, to determine if a pixel block can be modified to embed a bit of information.[7] This is to minimize the distortion to the image. If the pixel block can be modified, the watermarking process will embed a bit of the message by adding or subtracting at least one gray level. Using the same estimator signal, the watermarking process can be reversed to recover the embedded message and to restore the image to its original form. This scheme is secure only to the extent to which the estimator signal is kept secret from unauthorized individuals. Wu et al. implemented a reversible method, but loss is present due to a preprocessing step to reduce distortion caused by the flipped pixels.[13] Reversible watermarking can also be robust even in lossy systems, by using a circular interpretation of histogram bijective transformation.[14]

The third method involves embedding of the watermarks into areas of the medical image known as the regions of non-interest (RONI).[8,11–13] This method can be implemented using non-reversible or reversible watermarking technique. The motivation for this technique is that doctors and radiologists are generally only concerned with the regions of the images that are of diagnostic significance. Hence, regions of non-diagnostic importance can be labeled as the RONI. Thus, watermarks that are being embedded into such regions do not interfere with medical diagnosis.[11] Because there is no interference with image content, invisibility is less strict, and methods that feature higher embedding capacity and robustness could be used.[8] However, RONI using non-reversible watermarking method is still subjected to acceptance by radiologists because original images are still generally preferred for diagnostic purposes. This was based on literature[7,11,12] and from feedbacks provided by our radiologist colleagues. This is the reason why reversible/lossless watermarking was introduced for medical images.

The three watermarking schemes discussed above often involve a tradeoff between robustness, invisibility, and capacity. For example, the RONI approach will leave diagnostic information intact but can only be applied if a RONI exists.[8,11–13] Moreover, embedding capacity depends on the size of the RONI. Because RONI watermarking depends on user selection of the area,[13] there exists the risk of watermark superimposition if the user selects the same area multiple times. Non-reversible watermarking schemes are more robust, but the distortions caused by the watermarks are permanent and more noticeable. These distortions may not be acceptable, especially for medical images, because they may lead to incorrect diagnosis and treatment with life-threatening consequences. Hence, it remains subjective as to whether they can be accepted by doctors for medical diagnosis. As a result, we turn to fully reversible watermarking techniques such as the one proposed by Coatrieux et al.[7] The main advantage of a reversible watermark is that it will ensure that alterations introduced during the embedding process can be removed from the image. Thus, the image original pixel values can be restored, and diagnosis results will not to be interfered with.

Because medical images can be easily modified,[9] it is also important to identify whether tampering has been performed on the images and to locate the regions that have been tampered. Tamper detection can be implemented in reversible watermarking schemes for medical images.[11,13] Tamper detection allows regions of the image which have been modified to be identified automatically. This gives an added protection to the doctors using such images. Guo and Zhuang demonstrated that tamper detection can be performed by verifying block signature after original image has been restored.[11] It is also possible to partially restore tampered regions with highly compressed version of the whole original image or a less compressed version of some regions of more diagnostic importance.[13]

This paper describes a secure and fully reversible watermarking scheme which is capable of verifying authenticity and integrity of DICOM images. The scheme presented is based on the concepts of public-key cryptography and reversible data-hiding developed for medical images. We have adapted the fully reversible watermarking technique by Coatrieux et al.[7] and modified the watermarking scheme to utilize a secret random location signal which is encrypted using public-key in order to make the scheme more secure. Because it is important to detect whether images have been tampered, we have extended the watermarking scheme by incorporating a tampering detection and localization feature using dual-layer watermarking technique. The system was tested using sample medical images [i.e., computed tomography (CT), magnetic resonance images (MRI), ultrasound (US), and X-ray angiography (XA)] in DICOM format.

METHODOLOGY

The method proposed in this paper involves dual-layer watermarking which embeds the patient's metadata (patient's information from DICOM header) and other source information (e.g., a user-defined message) into the image using a reversible watermarking scheme. This scheme is developed to ensure the authenticity and integrity of the medical images. The reversible watermarking algorithm used in this scheme is adapted from the method proposed by Coatrieux et al.[7] but with modifications which include incorporating the concept of public-key cryptography to secure a random location signal and integrating a tamper detection and localization feature. The scheme comprises three main modules—image preprocessing, data embedding and data extraction, and tampering localization.

Image Preprocessing

In the image preprocessing module, underflow and overflow conditions are being taken care of to ensure that the selected image is suitable for the watermarking procedure. Before a digital medical image is watermarked, the image depth (2^p-1 possible gray levels for an image of p bits depth) has to be taken into account.[7] The occurrence of underflow or overflow condition implies that the image pixel range has been exceeded. An underflow will occur if the intended pixel to be watermarked has a pixel of gray value equal to 0. Consequently, subtracting one gray level from this pixel will result in a negative value. On the other hand, an overflow will take place if the intended pixel to be watermarked has a pixel of gray value equals to the maximum allowable pixel value 2^p-1 (e.g., 255 for an 8-bit grayscale image). Hence, adding one gray value to the pixel will exceed the maximum value for a p bits image. As a result, pixels that have pixel gray values 0 or 2^p-1 are not modifiable. DICOM images are generally stored using 16 bits per pixel[2] and imaging modalities usually do not produce images that utilize the full range of pixel values. Thus, we propose to shift all image pixels up by four pixel values. (This will be described in more detail in the next section.) At the receiver's end, the gray levels of the image are restored to their original values by subtracting all the pixels by four after the dewatermarking process has been performed.

Data Embedding and Data Extraction

The embedding process seeks to protect the source information (e.g., metadata) by watermarking it into the image using a random location signal.

For data embedding, the algorithm first divides the image into 2×2 non-overlapping blocks, as shown in Figure 1a. Considering each block of 2× 2 pixels, a binary message (*msg*) is inserted according to the following steps:

1. A random location signal that denotes one of the four (2×2) possible pixel positions where the "estimator" pixel resides on is generated. This random signal is hereby called the random location signal. The use of a random location signal is to ensure that it will be more difficult to decipher which estimator is used. Only one pixel is designated as the estimator in each block.
2. Consider that one pixel a of the block is selected and compared with the estimator.
3. If it satisfies $|\text{estimator}-a| <2$, then the pixel is able to carry one bit and is modified in the following manner:

   (a) If *msg(i)*=1, then change a to $a_w = a+2$
   (b) If *msg(i)=0*, then change a to $a_\mathrm{w} = a-2$

   (A difference of 2 is used in the scheme instead of 1 as proposed by Coatrieux et al.[7] This is to increase the embedding capacity.)
4. If $|\text{estimator} - a| < 2$ is NOT satisfied, then the distance between estimator and a is increased by 2 by changing a.
5. The process repeats from step 4 to 6 with pixels "b" and "c"
6. The process repeats until all the message bits have been embedded or all the blocks had been processed.

For data extraction, the image is once again divided into 2×2 non-overlapping blocks. The location of the estimator is known using the location signal. Referring to Figure 1b, a_w, b_w, and c_w represent pixels in the 2×2 block in the watermarked image. Using the estimator, the pixels can be restored and hidden message extracted according to the following steps:

1. If $a_\mathrm{w} >$ estimator then change a_w to $a_\mathrm{r} = a_\mathrm{w} - 2$
2. If $a_\mathrm{w} <$ estimator, then change a_w to $a_\mathrm{r} = a_\mathrm{w} + 2$
3. The decoder will consider that a bit **"1"** was embedded if $|a_\mathrm{r} - \text{estimator}|2$, and vice versa.

All the pixels will be increased by 4 pixel values to avoid underflow because pixels which are allowed to be modified will be changed by ±2, and this value is increased by a factor of 2 with dual-layer watermarking (to be explained). Hence,

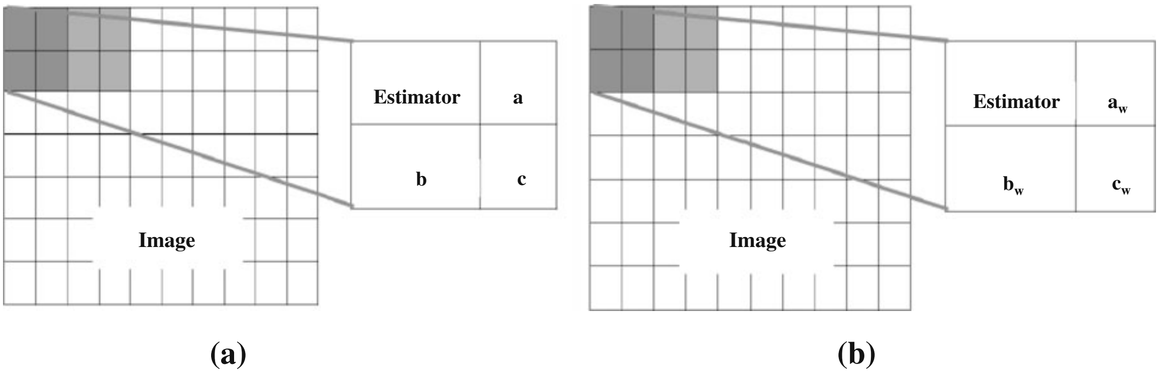


Fig. 1. Illustration of embedding and extraction of message. a Embedding b Extracting.

to avoid overflow, the maximum pixel value allowable for an image to be watermarked in our scheme is calculated by:

$$\text{Maximum pixel limit} = 2^p - 1 - q - r \qquad (1)$$

Where p is the bits depth of the image, q is the increase in all pixel values (i.e., 4), and r is pixel values allowed for modification (i.e., 2×2). Hence, our technique could support a 16-bit image with maximum pixel values of 65,527.

The security of this scheme depends on the ability to keep the estimator location secret. Hence, in order to keep the random location signal secure, a cryptography system known as public-key cryptography[16,17] (asymmetric cryptography) is used to encrypt the signal. The public-key system makes use of a pair of codes (also known as the public and private key) to encrypt a message. The signal which is encrypted using the public key can only be decrypted using the corresponding private key. The main advantage of using the public-key cryptography is that the two keys are mathematically related but it is computationally infeasible to deduce one key from the other. In our scheme, the random location signal is encrypted using the RSA (Rivest, Shamir and Adleman) cryptosystem, developed by Rivest et al.[17] Its security is based on the difficulty of factoring large integers. The RSA program used in our watermarking scheme was developed by Rajatasereekul and Kiettrisalpipop which is available online.[18] This encrypted pattern was watermarked into fixed locations in the image. During the dewatermarking process, this encrypted signal can be retrieved from the fixed locations and subsequently decrypted using a private key.

Thus, in practice, in order for a radiologist (e.g., the sender) to send an image to a doctor (e.g., the recipient) in another hospital, he would encrypt the random location signal with the doctor's public key (which is widely distributed). Upon receiving the image, the doctor can only retrieve the metadata by decrypting the random location signal using his private key, which is kept secret.

Tamper Detection and Localization

Integrity of our proposed system is controlled by incorporating a feature known as tamper detection and localization function. If the areas tampered are not within the region of interest (ROI), the image may still be accepted for diagnosis by the radiologist. Tamper localization is useful because integrity control based on the exact preservation of all parts of the image maybe unnecessarily strict as distortions on the image may also be due to noise originating from the transmission process. Tamper localization will avoid unnecessary requests for retransmission which may increase delay time and slow down the hospital's network. In the event that ROI alterations are indeed performed by a hacker to achieve malicious intents, the tamper localization property would be able to detect such alterations. Thus, the radiologist can be alerted that an attack had been carried out on the hospital information system.

Figure 2 illustrates the tampering localization approach implemented in our proposed system. The image is first divided into 16×16 non-overlapping pixel blocks and Cyclic Redundancy Code, CRC-16, which is an error-checking code, is computed for each of the blocks. These CRC bits form the tamper detection info which will be embedded as a second layer of watermark into the medical image. CRC is chosen because it is computationally less intensive as compared with hash functions which are used by Guo and Zhuang.[11] This consideration is important because a large number of CRC codes may have to be generated, depending on the size of the image. Computational time becomes a crucial issue when watermarking medical images of volumes containing multiple slices of DICOM images. In the tamper detection function, the standard CRC-16-CCITT polynomial is used together with a block size of 16×16 pixels. These parameters are selected based on the tradeoff between the area of detection, strength of detection, and the capacity to embed the tamper localization information.

Using the same watermarking embedding algorithm, each CRC code is embedded into its own block. In the event that the 16 bits of the CRC code of block 1 cannot be embedded into its own block, the remaining bits will be carried over to block 2 to be embedded prior to the embedding of the CRC of the second block itself. If both the remaining bits and the CRC of the second block can be embedded into block 2, block 3's CRC will be embedded into block 3 itself. This method is preferred to simply concatenating the CRC as a string spanning all the blocks because the latter will result in a failure to retrieve the CRC of each block when any of the embedded CRC bits is

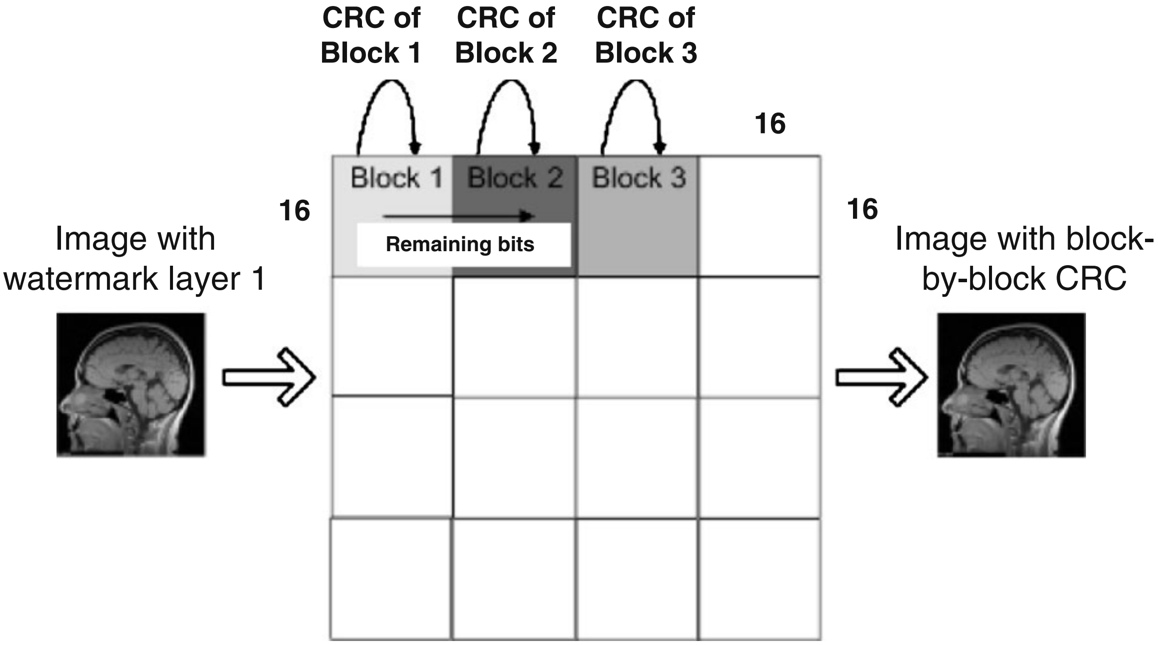


Fig. 2. Illustration of block-by-block CRC embedding.

altered. In the watermark extraction phase, the to-be-authenticated image is divided again into 16× 16 blocks, and the CRC of each block retrieved from the watermarked image will be verified with the CRC of each block of the restored image. Hence, if both CRCs do not match, the block will be identified as being tampered, hence achieving tamper localization.

Dual-Layer Watermarking

Figure 3 illustrates the layer concept used in the proposed reversible watermarking system. Two layers of watermark are embedded in one image for the watermarking process. In layer 1, source information (e.g., Metadata of the image) is first embedded followed by a digital envelope (DE). The DE is created by concatenating the bit stream of the random location signal encrypted using recipient's public key, CRC of the random location signal, and Secure Hash Algorithm (SHA)-256 hash of image. The CRC code of the random location signal is calculated to serve as a check to ensure that decrypted signal is correct. The standard CRC-32 polynomial used in the IEEE 802.3 (Ethernet) is being used in this case. The SHA-256

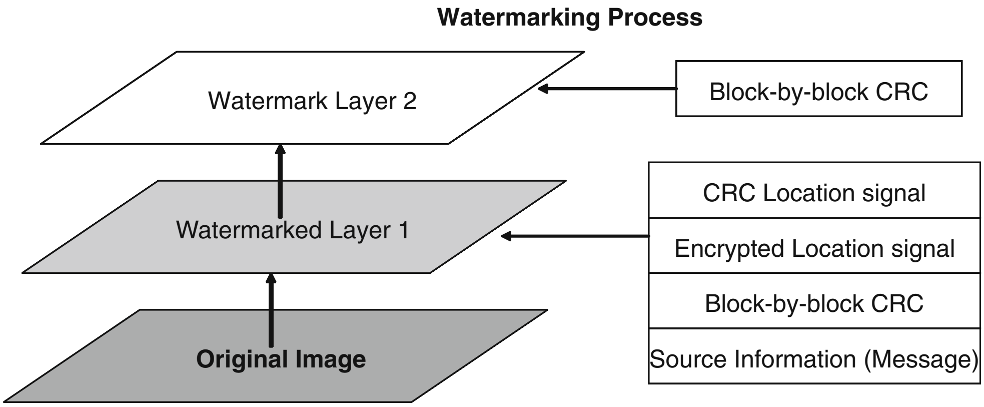


Fig. 3. Dual-layer watermarking scheme.

hash code of the original image data is calculated so that it can be used to verify the success of dewatermarking at the recipient side. The DE is embedded into the image after the last bit of metadata that has been embedded. The tamper detection information is embedded as layer 2 of the watermark as described in the earlier section.

RESULTS

Overview of Performance Measures

Sample medical images (i.e., CT, MRI, US, and XA) in DICOM format were used to test the system. Four important performance metrics were studied. These include:

1. Embedding capacity: A measure of embedding capacity is necessary to ensure that sufficient authentication information can be embedded into the image.
2. Imperceptibly**:** This is to test the quality of the medical images in terms of the invisibility of the watermark.
3. Run time: The time taken for the watermarking and dewatermarking process of an image should be accessed to ensure that it does not slow down the hospital's information system.
4. Robustness to tampering: This measure addresses the effectiveness of the tamper detection and localization function to alterations of pixels.

Each image is being embedded to its maximum capacity. The peak signal-to-noise-ratio (PSNR) and mean-squared-error were calculated by comparing the watermarked image and original image. Four test images from different modalities and of different image sizes were selected for the study. The DICOM test images were obtained from 3rd-party software (i.e., OsiriX image navigation software[19]). For all the images used in the test, we did not encounter any image with a maximum pixel value greater than 65,527 which is a requirement for the watermarking scheme to handle overflow and underflow. Table 1 summarizes the performance results.

Embedding Capacity

The number of bits that can be embedded for the four test images ranges from 74,190 to 581,524 bits. For a larger image size, the maximum number of bits that can be embedded increases. For example, 581,524 bits of information could be embedded into XA image, which has the largest image size of 1,024 × 1,024 pixels. This was the largest embedding capacity of all four test cases. This was expected because more pixels were available for the hiding of information bits. Although the MR and CT image were of the same size, there was a difference in embedding capacity. This is mainly because the embedding technique is dependent on the pixel correlation of the image. Higher correlation (i.e., high similarity between pixel values) will result in higher embedding capacity.

Quality of Watermarked Medical Images (Imperceptibility Tests)

The PSNR values calculated for all images ranging between 34~35dB. Figure 4 shows that images embedded at maximum capacity are visually indistinguishable as the original images. It should be noted that a higher PSNR may not necessary translate to a better image quality. For example, a small distortion in a ROI may still result in a high PSNR but will have a significant impact on diagnosis results. Hence, it is important that original images are always restored. (Hence, the adoption of the reversible watermarking in our system.)

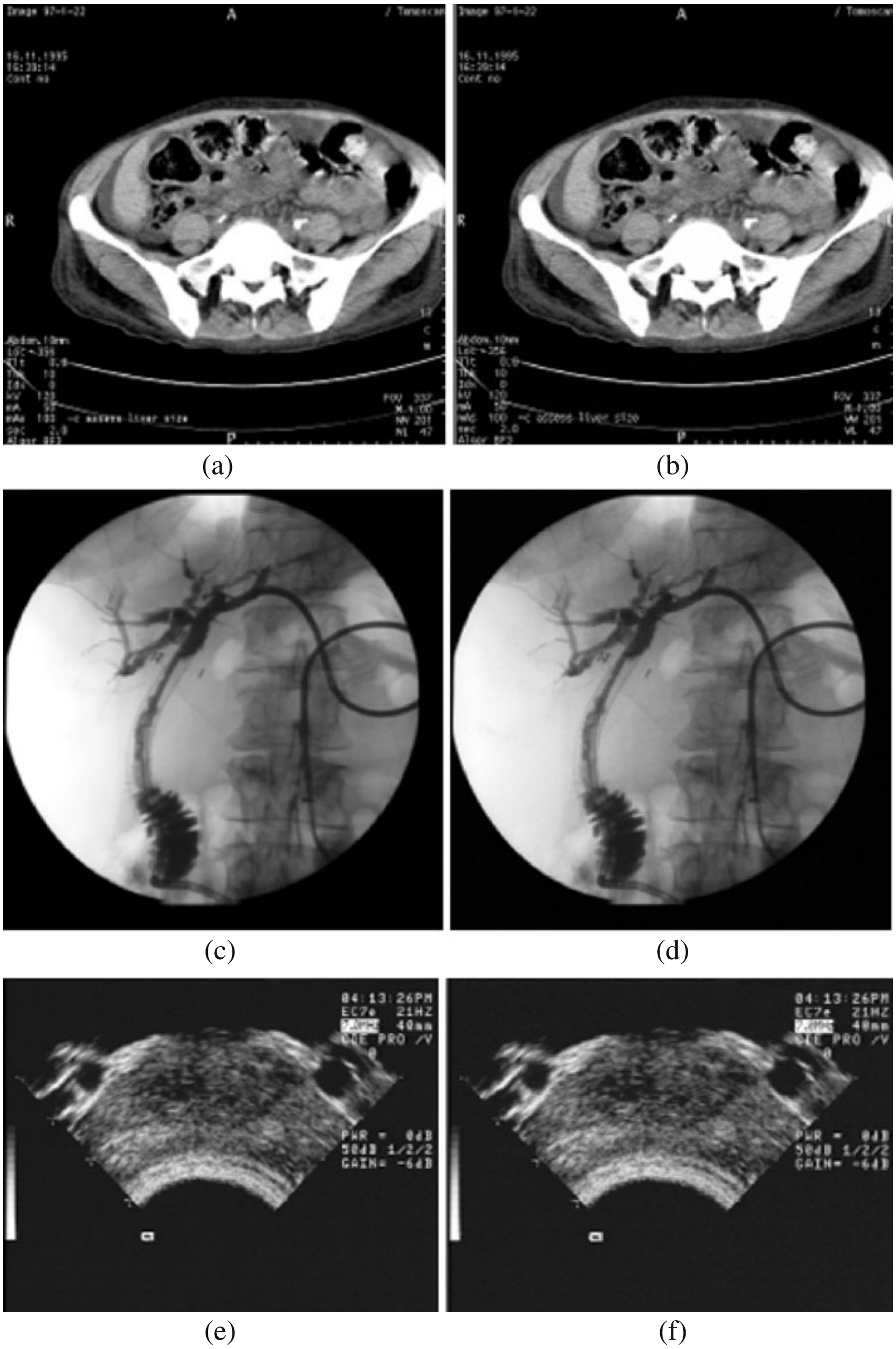
Run Time

Time taken for watermarking and dewatermarking process is an important factor to consider for

Table 1. Comparisons of Different Performance Measures on Different Medical Image Types

Image type	Image size	Bits per pixel	Maximum data that can be embedded (bits)	MSE	PSNR (dB)	Run time (s)
MRI	512 × 512	16	74190	21.5	34.8	5.48
XA	1024 × 1024	8	581524	20.7	35.0	20.7
US	480 × 640	16	79865	21.9	34.7	6.36
CT	512 × 512	8	127078	21.9	34.7	6.00

Terms used: *MRI* magnetic resonance imaging, *XA* X-ray angiography, ultrasound (US), Ultrasound (US), computed tomography (CT)



(a) (b)

(c) (d)

(e) (f)

Fig. 4. Original image (*left column*) compared with watermarked images (*right column*). a Original CT image b watermarked CT image, c original XA image, d Watermarked XA image, e original US image (f) Watermarked US image.

practical use in hospital system. It should not slow down the hospital's information system. The results showed that the time taken for the test images took an average of 9 s to process.

Robustness to Tampering

In order to demonstrate the tamper localization function in detecting forgery, counterfeited images were created by manually modifying the pixel values in the watermarked images using image processing software—ImageJ.[20] Figure 5 shows two examples of clinical relevant tampering. Figure 5b, e displays two samples of counterfeited images, and Figure 5c, f shows the corresponding images with tampered regions being localized by the tamper detection function. The localized tampered blocks were in encircled shaded boxes. Tampering was performed only at one location of the image in this test (as shown in Fig. 5). To test tampering at multiple locations, the watermarked images obtained from test images were being put through a separate systematic set of tampering. The tampering test includes tampering of a single pixel, a single block of size 8×8 and a spread of tampered blocks of size 8×8 at multiple locations in the image. Figures 6 shows representative results from systematic testing of tampering and localization capability of our scheme. Results show that the scheme is capable of detecting and localizing the various types of tampering, down to one pixel tampering and at multiple locations. The results also show that the tamper detection function was able to achieve a 100% detection rate.

DISCUSSION

This paper presented a fully reversible dual-layer watermarking scheme that has a tamper detection

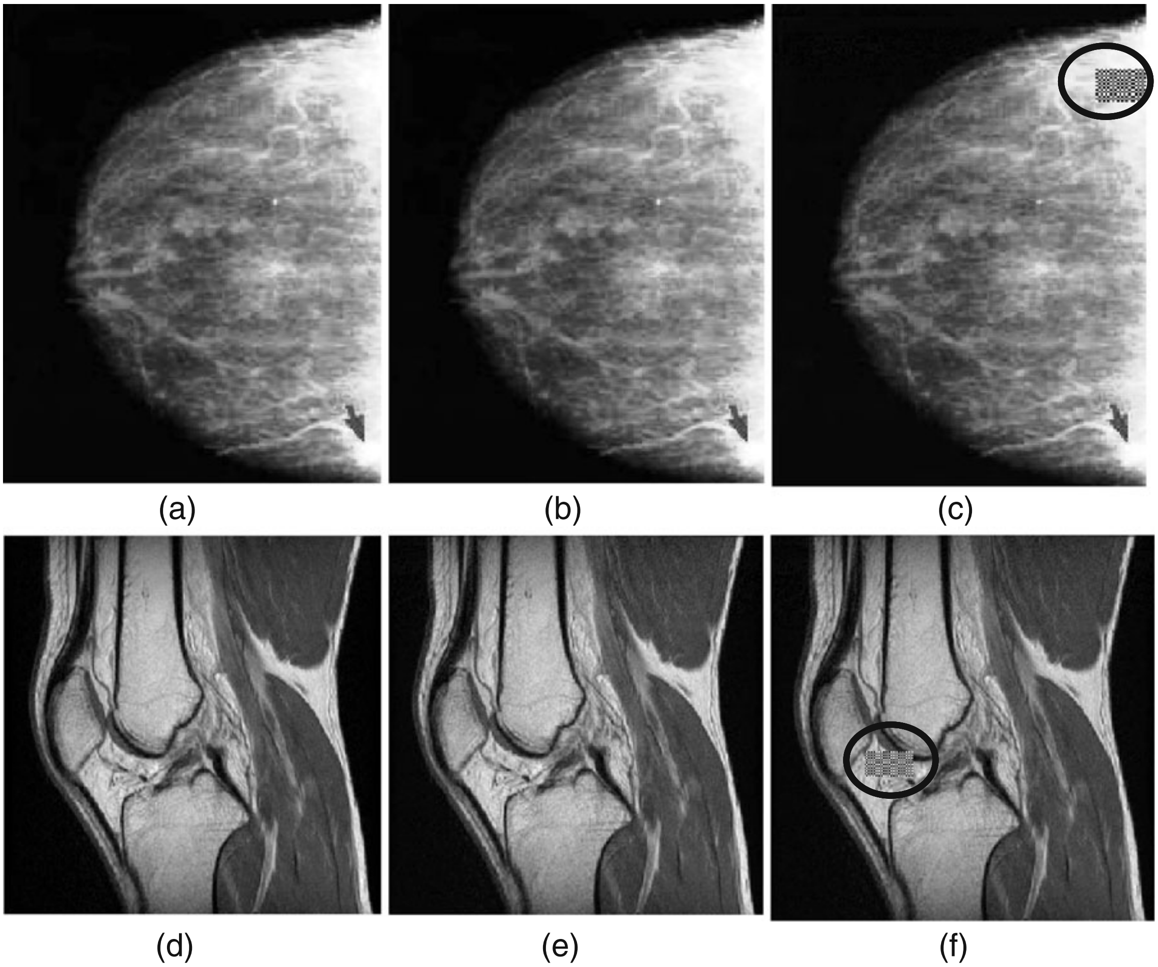


Fig. 5. Test images results using the tamper localization function. a Original mammogram image, b tampered mammogram by adding tumor like feature, c image displaying localization of tampering, d original MR knee image e tampered MR image by modifying femoral cartilage thickness, f image displaying localization of tampering.

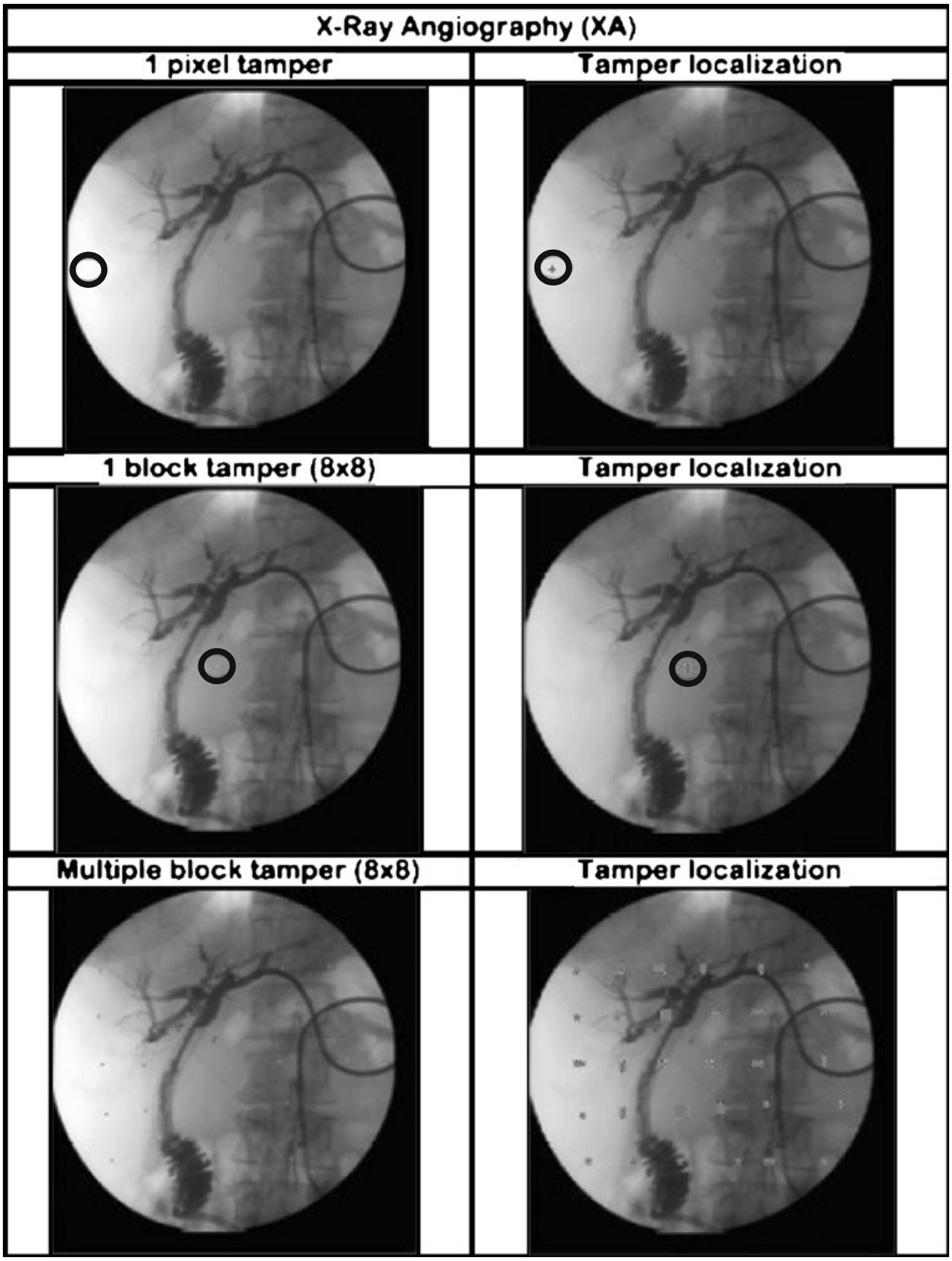


Fig. 6. Tamper detection results of tampered XA images.

and localization capability. The system was tested using sample CT, MR, US, and X-ray images. Results show that all the images were successfully watermarked and dewatermarked by the system. There are various types of watermarking schemes for medical images reported[3,7–14], and the schemes can be grouped in two main categories: non-reversible watermarking and reversible watermarking. For each of these watermarking techniques, RONI approach can be used. Fully reversible watermarking technique was used in our system, as opposed to non-reversible watermarking,[9] because our scheme requires original images to be used for diagnosis. This is an important clinical requirement. One important strength of watermarking is that the use of watermarks enables the images to be protected even when they leave the protection of the network.

The proposed watermarking scheme enables only authorized personnel to access the patient's medical images. The scheme ensures a high security of the random location signal which makes the hidden information fully secure to any intrusion or extraction attempts. This was achieved by using the public-key cryptographic watermarking technique in which only authorized personnel with the corresponding private key could decrypt the message hidden in the images and fully remove the watermarks embedded.

Our method which uses the public-key cryptography over digital signature has a number of advantages: (1) It is difficult to detect the encrypted random estimator signal (which is required for extraction of the source information) from the image. (2) Even if the encrypted random location signal could be retrieved, it is computationally infeasible to decrypt the random location signal without knowledge of the private key. Hence, we are able to ensure the integrity of the metadata and its authenticity. (3) There is no need to transmit the metadata together with the image since it is embedded into the image. This ensures that a hacker would not be able to easily separate the header, delete it, and create a new one.

We took a multi-layer watermarking approach to increase the data-hiding capacity, as opposed to single-layer watermarking schemes. This is possible because the watermarking technique is fully reversible. In our scheme, the first layer stores information related to the source and information used to check the integrity of the image and message. The second layer is designated for a tamper localization function. Because of the reversible nature of the scheme, it is possible to watermark over the two layers and subsequently retrieve information from the first layer by removing the second layer completely. This greatly increases the amount of data that can be embedded.

The reversible watermarking scheme enables a tamper detection function to be incorporated. Tamper detection is achieved by comparing the CRC of non-overlapping block embedded and the CRC of the corresponding block after dewatermarking. Using the tamper detection function, it is possible to determine where modifications were made to an image. Our results show that the tamper detection function was able to achieve a 100% detection rate, down to one pixel of tampering. In the current scheme, tampering is localized using a block of size 16×16. At present, DICOM uses digital signatures which can be used as a means to detect tampering. However, digital signatures can only detect the presence of tampering but cannot localize the tampering in the image. Using our scheme, it is possible to locate where tampering has been performed in the image. This added functionality will be very useful for doctors and legal professionals to detect any attempts that have been taken to tamper with the image or the watermark.

We recognized that the processing time needs to be significantly reduced before the scheme could be deployed for practical clinical use. Matlab which is not the optimum programming language to develop real-time software was used to rapidly develop the prototype program to test the concept. To reduce the time required to watermark an image, the scheme will be programmed using C++. Furthermore, our focus in this paper is on single image. Future work will involve modifying the scheme to handle multiple images or volume dataset so that the images can be watermarked in a reasonable time.

To avoid under and overflow, our scheme shifts pixel values of images by four gray levels values because many modalities produce images that do not utilize the full 16-bit range of pixel values. However, it is still possible that there might be images that utilize the entire 16-bit range of pixel values. For such images, shifting the pixels by four gray levels will not be possible. Hence, to avoid under and overflow for such cases, one possible method is to first check whether the pixel value is 0 or at the maximum allowable pixel value for the image. If the pixel selected has one of these two values, the pixel will not be modified. This method may not be optimum, but it will avoid under and overflow, and it is simple to implement. However, a small drawback to this method is that the embedding capacity will most likely be slightly reduced. In this study, we have focused on studying images with unsigned pixel values. DICOM also supports images with signed pixel values. The scheme proposed can be modified to handle signed pixel values by shifting the signed pixel values into the unsigned pixel range before watermarking and restoring the values back to the signed values after watermarking.

We have mainly tested our scheme based on the scenario in which there is a single recipient, and the recipient is known once the image is acquired by the sender. In certain scenario, the recipients of the images are not known beforehand, and the images will need to be archived first. In such scenario, a special archive public key could be used to first secure the images. The process could be handled by the archiving system. The corresponding private key should only be handled by authorized personnel who have the rights to retrieve the images from the archive. When the recipients were known, these images can then be dewatermarked using the archive private key and watermarked using the recipients' public keys before sending. This whole dewatermarking and dewatermarking process could be managed by the archival system. There could also be instances where the images need to be sent to multiple recipients. In such situations, two possible approaches could be used. In the first approach, multiple copies of the same image but watermarked using the public key of each recipient respectively will be sent. As a result, each recipient will only receive the images that have been watermarked using his/her public key. This first method can be considered as a one-to-one approach. A second approach is to use one common public and private key for multiple recipients. Using this approach, the recipients will most likely be grouped into a workgroup which will share the private key. This can be considered as a one-to-many approach. The second method involves managing workgroups and has an added difficulty in ensuring the security of the common private key. The first approach will provide better security because private keys are not shared.

Figure 7 depicts how the proposed scheme would fit into the overall picture archiving and communication system (PACS). This infrastructure is based on the image security system proposed by Cao et al.[21]. In the framework, a dedicated server called PACS Security Server is used to handle PACS-related security issues. This includes secure exchange of DICOM information over open networks and logging of security information (e.g., access rights and time stamping of the images) and as a key authority in charge of the storage and distribution of public keys. Referring to Figure 7, the watermarking system could be deployed at the sender's station. When a sender needs to send an image to a recipient in another hospital, a request for the public key will be sent to the PACS security server which keeps a log of the activity of every watermarking operation. The image will then be watermarked using the public key before sending to PACS security server. Once the image reaches the other hospital, the PACS security server at the other hospital will receive the image which will subsequently be sent to the recipient workstation. Before the recipient studies the image, the image will be fully dewatermarked using the recipient's private key.

CONCLUSION

Medical image security has become an important issue as images are communicated over open networks. However, there are no established techniques that can fully address these issues to be deployed in a hospital information system. In this paper, we described a reversible watermarking scheme which could be used to address the authentication and integrity problem of medical images. The tamper localization function together with the reversibility of the watermarks will make this scheme a well-suited

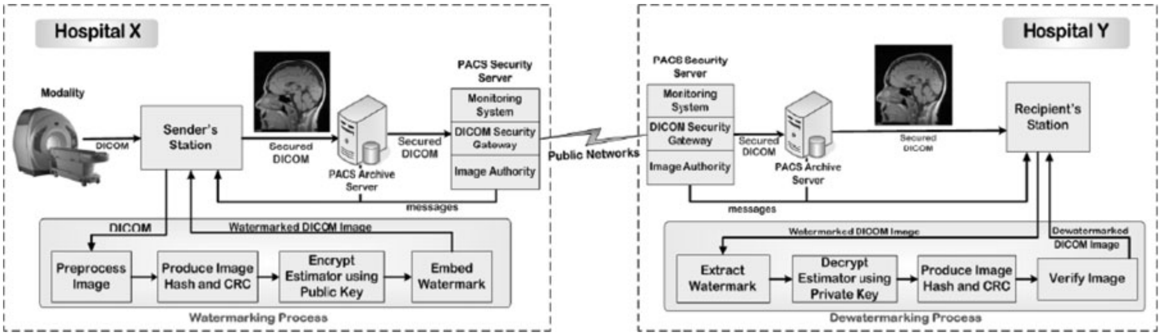


Fig. 7. An illustration of how the proposed scheme could fit into the overall PACS. This infrastructure is based on the image security system proposed by Cao et al.[21].

one for doctors as the scheme does not interfere with medical diagnosis.

REFERENCES

1. The Health Insurance Portability and Accountability Act (HIPAA), March 2009. [Online]. Available at: http://www.hhs.gov/ocr/ privacy/index.html

2. Digital Imaging and Communications in Medicine (DICOM), part 15: security profiles ed., National Electrical Manufacturers Association (NEMA), 2001, pS 3.15–2001

3. Coatrieux G, Maitre H, Sankur B, Rolland Y, Collorec R: Relevance of watermarking in medical imaging. Proc IEEE EMBS Information Technology Applications in Biomedicine. Arlington, VA 2000, pp 250–255

4. Zain JM, Fauzi AM, Aziz AA: Clinical evaluation of watermarked medical images. Proc EMBS 28th Annual International Conference of the IEEE 5459–5462, New York, USA, August 30–Sept. 3, 2006

5. Wang XY, Feng DG, Lai XJ, Yu HB: Collisions for hash functions MD4, MD5 HAVAL-128 and RIPEMD. Rump session of Crypto'04 and IACR Eprint archive, 2004

6. Kaminsky D: MD5 to be considered harmful someday *CryptologyePrint Archive*, 2004

7. Coatrieux G, Lamard M, Daccache W, Puentes W, Roux C: A low distortion and reversible watermark: application to angiographic images of the retina. Proc IEEE-EMBS Eng in Med Biol Soc 2224–2227, 2005

8. Miller ML, Cox IJ, Linnartz JPMG, and Kalker T: A review of watermarking principles and practices. Digital Signal Processing for Multimedia Systems, IEEE 461–485, 1999

9. Zhou XQ, Huang HK, Lou SL: Authenticity and integrity of digital mammography images. IEEE Trans Med Imag 20 (8):784–791, 2001

10. Macq B and Dewey F: Trusted headers for medical images. DFG VIII-DII Watermarking Workshop, Erlangen, Germany, 1999

11. Guo X, Zhuang TG: Lossless watermarking for verifying the integrity of medical images with tamper localization. J Digit Imaging 2008

12. Guo X, Zhuang TG: A region-based lossless watermarking scheme for enhancing security of medical data. J Digit Imaging 22(1):53–64, 2009

13. Wu JHK, Chang RF, Chen CJ, Wang CL, Kuo TH, Moon WK, Chen DR: Tamper detection and recovery for medical images using near-lossless information hiding technique. J Digit Imaging 21(1):59–76, 2008

14. Vleeschouwer CD, Delaigle J-F, Macq B: Circular interpretation of bijective transformations in lossless watermarking for media asset management. Multimedia, IEEE Transactions on 5(1):97–105, 2003

15. Fridrich J, Goljan M, Du R: Lossless data embedding—new paradigm in digital watermarking. EURASIP J Applied Signal Processing 2002(2):185–196, 2002

16. Diffie W, Hellman M: New directions in cryptography. IEEE Trans Inf Theory 22:644–654, 1976

17. Rivest RL, Shamir A, Adleman L: A method for obtaining digital signatures and public-key cryptosystems. Commun ACM 21:120–126, 1978

18. Rajatasereekul T, Kiettrisalpipop V: RSA encryption and decryption using matlab. ECE575 project, Oregon State University, 2002. [Online]. Available: http://islab.oregonstate.edu/koc/ece575/02Project/Kie+Raj/

19. Osirix Image Navigation Software Sample Datasets—available at: http://pubimage.hcuge.ch:8080/

20. ImageJ, Available at: http://rsbweb.nih.gov/ij/

21. Cao F, Huang HK, Zhou XQ: Medical image security in a HIPAA mandated PACS environment. Comput Med Imaging Graph 27:185–196, 2003