2014

Stegacryption of DICOM Metadata

Jonathan Blackledge
Technological University Dublin, jonathan.blackledge@tudublin.ie

A. Al-Rawi
University of Bahrain, abdulrahman.alrawi@gmail.com

Recommended Citation

Blackledge, J., Al-Rawi, A.: Stegacrypion of DICOM Metadata. IET ISSC 2014, University of Limerick, Ireland, June 26-27, 2014.

Stegacryption of DICOM Metadata

J. Blackledge* **and A. Al-Rawi****

**School of Electrical*
and Electronic Engineering
Dublin Institute of Technology, Ireland

*** College of Applied Studies*
University of Bahrain
Kingdon of Bahrain

E-mail: *jonathan.blackledge@dit.ie **aalrawi@uob.edu.bh

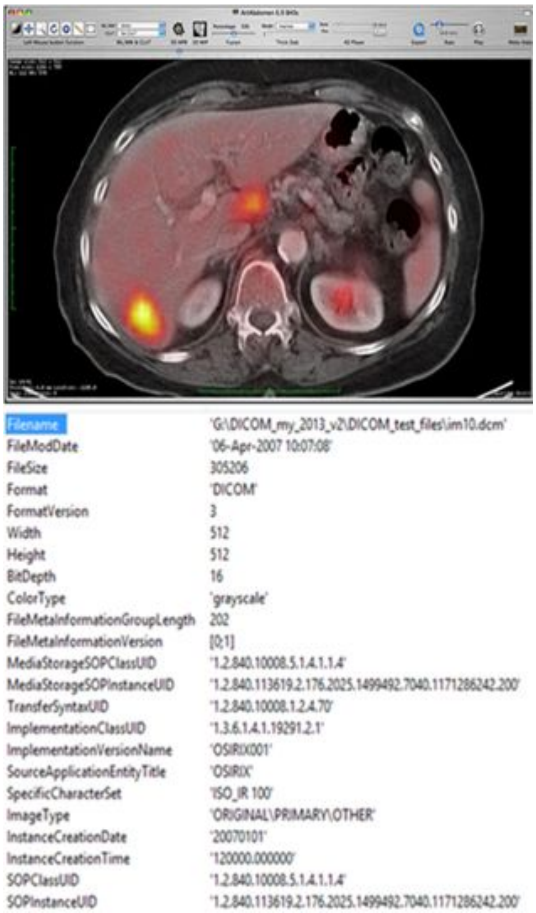
Abstract — **Digital Imaging and Communications in Medicine (DICOM) files are an international data standard for storing, distributing and processing medical images of all types. DICOM files include a header file containing Metadata on details which may include information on the patient. This often inhibits the free distribution of DICOM files due to issues relating to the confidentiality of data on identifiable living people, thereby limiting the potential for other radiologists to provide a diagnosis, for example, through distribution of the data over the Internet. This problem is a current limiting condition with regard to the development of Tele-medical imaging. Thus in this paper we consider a method of encrypting and embedding (or Stegacrypting) DICOM Metadata into the DICOM image, thereby providing a solution to a problem that currently inhibits the distribution of medical images using a file type that is an established international standard. The proposed method removes or 'anonymises' the private data, encrypt it and then embeds it into the DICOM image in an imperceptible way. The specific algorithm developed retains the private data attached to a DICOM image even when the image is converted into a standard image file format.**

Keywords — **Coding and Encryption, Information Hiding, Medical Image Processing, Digital Imaging and Communications in Medicine**

I Introduction

The Digital Imaging and Communications in Medicine or DICOM format is an international standard for visualising and and processing medical images. A number of 'DICOM viewers' are now available for interpreting and processing medical images such as OsiriX [1]. The increasing use of such facilities means that DICOM data is and will continue to become an increasingly important and essential aspect of Tele-medical imaging world wide for empowering mobile health using medical informatics technologies such as MedoPad [2] and tele-medical imaging systems such as MoletestUK [3], for example. In this regard, one of the principal problems associated with the world wide distribution of DICOM data over the internet is the confidentiality of the patient information which is held in the Metadata associated with a DICOM image by default. Thus the problem is to find a solution to the distribution of DICOM files that conforms to data protection legislation which prevents the processing and analysis of data on identifiable living people and thereby includes patient data associated with a DICOM file.

Figuire 1 show a typical example of a DICOM image visualised using the OsiriX imaging software and an example of a section of the Metadata associated with such images. In the context of this figure, we consider a method of encrypting the Metadata and hiding the resulting ciphertext in the corresponding image so that the image can then be distributed as a DICOM file or otherwise with a guarantee of patient confidentiality. The method of both encrypting the data and using Steganographic methods of hiding it in an image is known as *Stegacryption* and in this paper we provide an algorithm that uses an integer wavelet transform whose output is used to embedded a binary ciphertext.



Filename	'G:\DICOM_my_2013_v2\DICOM_test_files\im10.dcm'
FileModDate	'06-Apr-2007 10:07:08'
FileSize	305206
Format	'DICOM'
FormatVersion	3
Width	512
Height	512
BitDepth	16
ColorType	'grayscale'
FileMetaInformationGroupLength	202
FileMetaInformationVersion	[0;1]
MediaStorageSOPClassUID	'1.2.840.10008.5.1.4.1.1.4'
MediaStorageSOPInstanceUID	'1.2.840.113619.2.176.2025.1499492.7040.1171286242.200'
TransferSyntaxUID	'1.2.840.10008.1.2.4.70'
ImplementationClassUID	'1.3.6.1.4.1.19291.2.1'
ImplementationVersionName	'OSIRIX001'
SourceApplicationEntityTitle	'OSIRIX'
SpecificCharacterSet	'ISO_IR 100'
ImageType	'ORIGINAL\PRIMARY\OTHER'
InstanceCreationDate	'20070101'
InstanceCreationTime	'120000.000000'
SOPClassUID	'1.2.840.10008.5.1.4.1.1.4'
SOPInstanceUID	'1.2.840.113619.2.176.2025.1499492.7040.1171286242.200'

Fig. 1: Example DICOM image viewed using OsiriX [1] (above) and some example DICOM Metadata (below).

II Digital Imaging and Communications in Medicine

Digital Imaging and Communications in Medicine (DICOM) is a standard for handling, storing, and transmitting information in medical imaging. DICOM files can be exchanged between two entities that are capable of receiving image and patient data in DICOM format. The National Electrical Manufacturers Association (NEMA) holds the copyright to this standard [4] and is the international standard for medical images and related information (ISO 12052). DICOM defines the formats for medical images that can be exchanged with the data and quality necessary for clinical use and is implemented in almost every radiology, cardiology imaging, and radiotherapy device (X-ray, CT, MRI, ultrasound, etc.), and, increasingly, in devices in other medical domains such as ophthalmology and dentistry [5].

DICOM was introduced in 1993 after some ten years of standards development from the early 1980s when only manufacturers of CT or MR imaging devices could decode the images that the early machines generated. It differs from some, but not all, data formats in that it groups information into data sets. Thus, for example, a file of an X-Ray image actually contains the patient ID, Name, etc. within the file, so that the image can never be separated from this information by mistake. This is similar to the way that image formats such as JPEG can also have embedded tags to identify and describe the image.

DICOM has an information model which differentiates it from other standards used in the medical industries sector. The model is based on information objects which include definitions on the information to be exchanged. Each image type, and therefore information object, has specific characteristics. A CT image, for example, requires different descriptors in the image header compared to an ultrasound image or an ophthalmology image. These templates are identified by unique identifiers which are registered by the National Electrical Manufacturers Association (NEMA), the DICOM standard facilitator. Information objects are also known as part of the Service Object Pair (SOP) Classes. An example of a SOP Class is the CT Storage SOP Class, which allows CT images to be exchanged [6].

The DICOM standard contains a number of major enhancements to previous versions of the ACR-NEMA Standard including the following [7]:

1. It is applicable to a networked environment, whereas the ACR-NEMA Standard was applicable in a point-to-point environment only.

2. It is applicable to an off-line media environment, while the ACR-NEMA Standard did not specify a file format or choice of physical media or logical files ystem.

3. It specifies how devices claiming conformance to the Standard react to commands and data exchanged in addition to specifying levels of conformance.

4. It is structured as a multi-part document.

5. It introduces explicit Information Objects not only for images and graphics but also for waveforms, reports, printing, etc.

6. It specifies an established technique for uniquely identifying any Information Object.

In the following section, we review some of the principal encrypted or otherwise information hiding techniques that have been specifically designed for medical images including DICOM data.

III Medical Image Information Hiding

A number of techniques have been proposed for both encrypting and hiding data relating to the medical imaging field using both spatial and transform based techniques. For example, in [8], a method for imperceptibly embedding patient information in an associated medical images is considered. The patient's name and ID are converted into contiguous binary streams and each stream encoded using arithmetic coding. The encoded information is then embedded into the image pixels using a basic Least Significant Bits approach and sent to the receiver. The receiver requires the decoding software to decode the extracted data and regenerate the patients personal information. A medical image watermarking scheme based on histogram modification and block division differences is considered in [9] and a block-based approach coupled with a histogram shift between the local minimum and maximum frequencies is considered in [10] and [15]. Specific organ segmentation based approaches are considered in [11] for CT imaging, an LSB modification scheme that detects and recovers image tampering using a Region-of-Interest approach is considered in [12], [13] and [14] and a method for distortion-free, reversible and fragile medical image watermarking is given in [16]. Other watermarking schemes that focus on applications in medical imaging include those that combine lossless compression and encryption [17] including blind watermarking assuming a DICOM format [18], [19] and [20].

Application of the Discrete Cosine Transform for watermarking medical images and high capacity multiple watermarking methods are reported in [21] and [22], respectively. An approach that utilises the wavelet transform is considered in [23] in which a dual-tree wavelet transform with Bivariate Shrinkage is used. The Dual-Tree Complex Wavelet Transform (DT-CWT) uses a dual tree composed of a discrete and complex wavelet transform which enhances the robustness of the method and overcomes DWT drawbacks such as poor directionality, shift sensitivity and absence of phase information. However, Bivariate Shrinkage, a method for image thresholding, yields high performance with regard to de-noising images utilising the statistical dependence between wavelet coefficients and their parent. The embedding method starts by computing the wavelet transform using the DT-CWT and selecting the appropriate subbands. The watermark is 'managed' by repetition (using a key) for increased robustness and the data is embedded into the wavelet coefficients with an ability to balance robustness and fidelity. The extraction process is performed by estimating the appropriate coefficients from the CWT transform of the stego image and resorting the watermark bits. An adaptive data hiding method using integer wavelet transform coefficients is proposed in [24] using the adaptive data hiding algorithm presented in [25]. The original medical image histogram is modified to overcome the underflow/overflow problem and a 4-level integer transform undertaken. The multiple embedding watermark process starts by first deciding upon the hiding capacity followed by embedding the data into the LH1, the EPR data into HL2 and LH3, the index watermark in HL3 and LH3 and finally, the IAC (Image Authentication Code) data in HL4 and LH4. Data extraction is achieved using the same method after assessment of the embedding data length, watermark embedding (which is based on the 'edge coefficients') being based on the absolute values.

Having researched the relevant literature, it is clear that there are no methods currently available for Stegacrypting DICOM images and no international standard has been developed to-date. In this context, the following section considers a new algorithm that has been prototyped using a MATLAB programming environment based on extending the approach used in [24] .

IV DICOM Information Hiding

The method proposed in this section aims to protect the private information associated with a DICOM image from unauthorised access. The DICOM standard embeds the confidential data into the DICOM header. Here, we embed the encrypted confidential information into the DICOM image itself and remove it from the DICOM 'Object'. This provides a way of protecting the private data from unauthorised personnel while keeping that data accessible to authorised users even when the DICOM object is converted into an ordinary image format. This is of particular value with regard to the distribution of DICOM images between radiologists as current practices restrict this activity due to the confidential nature of the patient information that accompanies a DICOM image. In turn, this restriction limits the open access approach associated with 'best practice' in terms of research and development in medical image analysis and the implementation of new medical image processing algorithms for specific medical conditions, diagnostic requirements and training.

Figure 2 illustrates the proposed DICOM information embedding and extraction algorithms.

The principal algorithms associated with the proposed DICOM Metadata information hiding methods are summarised as follows

a) Algorithm I: DICOM Encryption and Embedding Algorithm

Step 1: Read the DICOM image data as *img* and the DICOM Metadata as *info*.

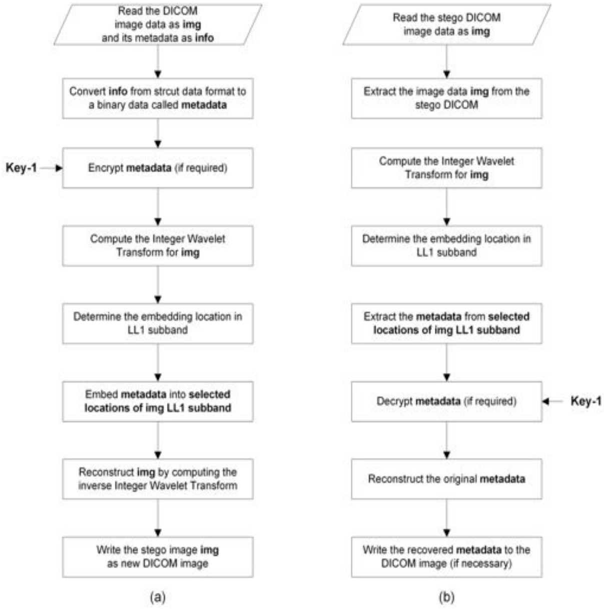

Read the DICOM image data as img and its metadata as info
Convert info from strcut data format to a binary data called metadata
Key-1
Encrypt metadata (if required)
Compute the Integer Wavelet Transform for img
Determine the embedding location in LL1 subband
Embed metadata into selected locations of img LL1 subband
Reconstruct img by computing the inverse Integer Wavelet Transform
Write the stego image img as new DICOM image
(a)
Read the stego DICOM image data as img
Extract the image data img from the stego DICOM
Compute the Integer Wavelet Transform for img
Determine the embedding location in LL1 subband
Extract the metadata from selected locations of img LL1 subband
Decrypt metadata (if required)
Key-1
Reconstruct the original metadata
Write the recovered metadata to the DICOM image (if necessary)
(b)

Fig. 2: DICOM information hiding. (a) Embedding algorithm; (b) Extraction algorithm.

Step 2: Specify the confidential attributes in the *info* structure based on the Confidentiality Profile Attributes listed in [26].

Step 3: Encrypt the *info* confidential attributes and convert them into a binary stream.

Step 4: Apply an Integer Wavelet Transform to the DICOM image data *img* to obtain the LL1, LH1, HL1 and HH1 coefficients.

Step 5: Determine the LSBs to be used as the embedding location(s) in the LL sub-band of the Wavelet transform.

Step 6: Determine the DICOM image data *img* LSB to be used as the embedding location(s).

Step 7: Embed the encrypted binary stream into the specified embedding location.

Step 8: Remove the confidential attributes content from the *info* structure.

Step 9: Reconstruct the original image data *img* by applying the Inverse Integer Wavelet Transform.

Step 10: Write the modified *img* and *info* into a new stego-DICOM file.

The following points should be noted:

- Any commercial or otherwise encryption method can be applied that generates a binary ciphertext of the Matedata.
- If the DICOM image is to be saved as an ordinary image, the *img* pixel values must be quantised to the intended image pixel colour range (8-bis, 16-bits, etc.). The original *img* range (minimum and maximum values) must be also stored.
- The LL sub-band of the transformed medical image is selected for embedding the watermark because it is more robust to attacks such as low-pass filtering. However, changing the LL coefficients causes more perceptual distortion to the DICOM image if many pixels are altered. This issue can be solved by compressing the watermark data (before encryption).

b) Algorithm II: DICOM Extraction and Decryption Algorithm

Step 1: Read the stego-DICOM image data as *img* and the DICOM Metadata as *info.*

Step 2: Specify the confidential attributes in the *info* structure based on the Confidentiality Profile Attributes listed in [26].

Step 3: Apply the Integer Wavelet Transform to the DICOM image data *img* to obtain the LL1, LH1, HL1 and HH1 coefficients.

Step 4: Determine the DICOM image data *img* LSBs used in the embedding process from the LL sub-band of the Wavelet transform, and then extract the hidden data.

Step 5: Decrypt the extracted bits using the same encryption key to recover the confidential attribute values.

Step 6: Re-write the extracted values to the confidential attributes in the *info* structure.

Step 7: Write the modified *info* to a DICOM file.

The following point should be noted: if the image to be read in Step 1 is not DICOM, the stored DICOM image data range (minimum and maximum values) must be used to extract the hidden information.

Figure 3 illustrates the perceptual quality of a typical modified medical image using the prototype m-code for implementing the Algorithms I and II as given in [27].

c) Example Results

A set of 10 DICOM images of different sizes are examined in 4. The original and stego medical images are shown in order to illustrate the perceptual quality of the modified images. Moreover, the MSE (Mean Square Error) and PSNR (Peak Signal-to-Noise Ratio) values are listed as subjective measures for the Stegocrypted images.

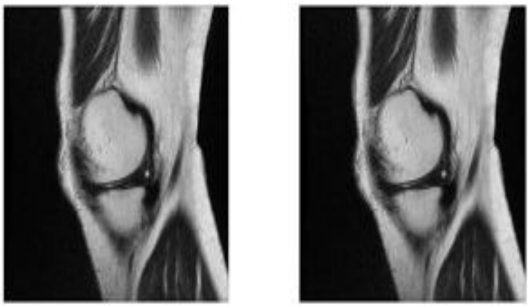
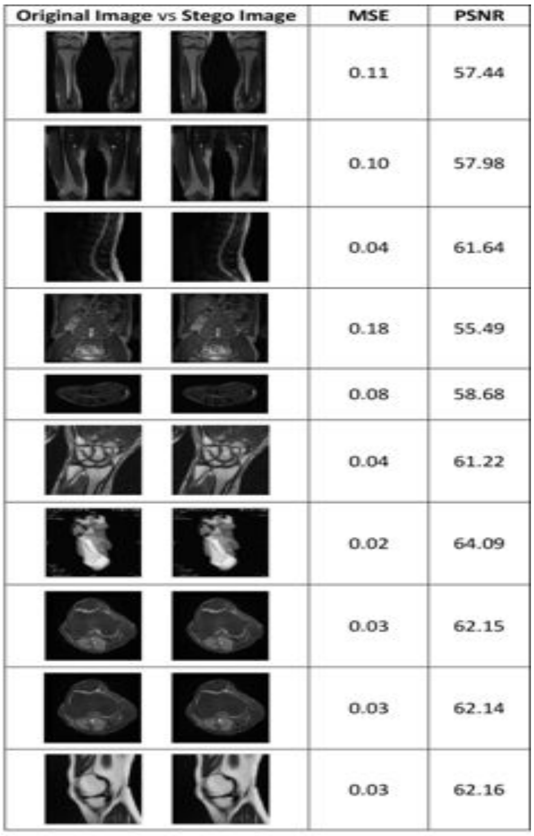


Fig. 3: DICOM watermarking method. Left: Original DICOM image, Right: Stego DICOM image.



Original Image vs Stego Image	MSE	PSNR
	0.11	57.44
	0.10	57.98
	0.04	61.64
	0.18	55.49
	0.08	58.68
	0.04	61.22
	0.02	64.09
	0.03	62.15
	0.03	62.14
	0.03	62.16

Fig. 4: DICOM Stegocryption results. From left to right: Original image, Stegocrypted image, MSE and PSNR

V Summary and Conclusions

The algorithms presented in this paper are an attempt to solve a problem in the area of Telemedicine that has arisen from need to maintain patient confidentiality. This is a relatively common problem in the field of Health Informatics and the e-Health Services technology. Health Informatics is the appropriate and innovative application of concepts and technologies to improve health care and health which may be subdivided into two principal categories: (i) Tele-Health which is related to direct (video conferencing) or indirect (website delivery) of health information or health care to a recipient; (ii) e-Health which encompasses products, systems and services, including tools for health authorities and professionals and personalised health systems for patients and citizens. In principle, the solution presented in this paper applies to both categories especially with regard to information relating to medical images which is predicated on the use of DICOM data. For example, with regard to the interpretation of medical images, DICOM data should ideally be available to distribute to radiologists world-wide who can provide a diagnosis from which a final decision can be made based on a voting algorithm. Irrespective of weather this is achieved through visual inspection using an iPAD [2], for example, and/or image processing using a system such as OririX [1], the patient confidentiality problem needs to be solved on a routine basis and it is in this context that the results presented in this paper are given.

Acknowledgments

Jonathan Blackledge is supported by the Science Foundation Ireland Stokes Professorship Programme.

References

[1] OsiriX, `http://www.osirix-viewer.com/`

[2] MedoPad, `http://www.medopad.com/medopad_Ltd/Medopad.html`

[3] MoletestUK, `http://moletestuk.com/` and `http://en.wikipedia.org/wiki/Moletest`

[4] DICOM, `http://en.wikipedia.org/wiki/Dicom`.

[5] About DICOM, `http://medical.nema.org/Dicom/about-DICOM.html`, The Association of Electrical Equipment and Medical Imaging Manufacturers (NEMA).

[6] J. M. Blackledge, M. D. Blackledge and J. N. Courtney, *Non-Gaussian Anisotropic Diffusion for Medical Image Processing using the OsiriX DICOM*, International Society for Advanced Science and Technology (ISAST) - Transaction on Computing and Intelligent Systems, vol: 4, issue: 1, pages: 24 - 31, 2012.

[7] The DICOM standard, `http://medical.nema.org/Dicom/about-DICOM.html`, The Association of Electrical Equipment and Medical Imaging Manufacturers (NEMA).

[8] V. N. Kumar, M. Rochan, S. Hariharan, K. Rajamani, *Data Hiding Scheme for Medical Images Using Lossless Code for Mobile HIMS*, Communication Systems and Networks (COMSNETS), 2011 Third International Conference on , vol., no., pp.1,4, 4-8 Jan. 2011.

[9] A. Lavanya, V. Natarajan, *Data Hiding Using Histogram Modification of Difference in Medical Images Based on Block Division*, Recent Trends In Information Technology (ICRTIT), 2012 International Conference on , vol., no., pp.141,144, 19-21 April 2012.

[10] M. Fallahpour, D. Megias, M. Ghanbari, *High Capacity, Reversible Data Hiding in Medical Images*, Image Processing (ICIP), 2009 16th IEEE International Conference on , vol., no., pp.4241,4244, 7-10 Nov. 2009.

[11] S. K. Lee, S. J. Lim, Y. H. Suh, Y. S. Ho, *Lossless Data Hiding for Medical Images with Patient Information*, Image Processing, 2007. ICIP 2007. IEEE International Conference on , vol.3, no., pp.III - 253,III - 256, Sept. 16 2007-Oct. 19 2007.

[12] B. W. R. Agung, Adiwijaya, F. P. Permana, *Medical Image Watermarking with Tamper Detection and Recovery using Reversible Watermarking with LSB Modification and Run Length Encoding (RLE) Compression*, Communication, Networks and Satellite (ComNetSat), 2012 IEEE International Conference on , vol., no., pp.167,171, 12-14 July 2012.

[13] J. M. Zain, A. R. M. Fauzi, *Medical Image Watermarking with Tamper Detection and Recovery*, Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE , vol., no., pp.3270,3273, Aug. 30 2006-Sept. 3 2006.

[14] S. C. Liew, S. Y Liew, J. M. Zain, *Reversible Medical Image Watermarking For Tamper Detection And Recovery With Run Length Encoding Compression*, World Academy of Science, Engineering and Technology, Issue 50, p799, February 2011.

[15] M. Fallahpour, D. Megias, M. Ghanbari, *Reversible and High-Capacity Data Hiding in Medical Images*, Image Processing, IET , vol.5, no.2, pp.190,197, March 2011.

[16] DN. V. harwadkar, B. B. Amberker, Supriya, P.B. Panchannavar, *Reversible Fragile Medical Image watermarking with Zero Distortion*, Computer and Communication Technology (ICCCT), 2010 International Conference on , vol., no., pp.248,254, 17-19 Sept. 2010.

[17] M. K. Kundu, S. Das, *Lossless ROI Medical Image Watermarking Technique with Enhanced Security and High Payload Embedding*, International Conference on Pattern Recognition, 1457-1460, 2010.

[18] R. C. Raul, F. U. Claudia, G. J. Trinidad-Bias, *Data Hiding Scheme for Medical Images*, Electronics, Communications and Computers, 2007. CONIELECOMP '07. 17th International Conference on , vol., no., pp.32,32, 26-28 Feb. 2007.

[19] L. R. Knudsen, W. Meier, B. Preneel, V. Rijmen, S. Verdoolaege, *Analysis Methods for (Alleged) RC4*, Lecture Notes in Computer Science Vol. 1514, 1998, pp 327-341, Springer Verlag, 1998.

[20] Y. Wang and A. Pearmain, *Blind Image data Hiding Based on Self Reference*, Pattern Recognition Letters, Vol. 2, No. 15,pp. 1681-1689, November 2004.

[21] C. Dong, J. Li, Y. Chen, *A DWT-DCT Based Robust Multiple Watermarks for Medical Image*, Photonics and Optoelectronics (SOPO), 2012 Symposium on , vol., no., pp.1,4, 21-23 May 2012.

[22] B. M. Irany, X. C. Guo, D. Hatzinakos, *A high Capacity Reversible Multiple Watermarking Scheme for Medical Images*, Digital Signal Processing (DSP), 2011 17th International Conference on , vol., no., pp.1,6, 6-8 July 2011.

[23] R. M. Kongo, L. Masmoudi, N. Idrissi, N. Hassanain, M. Cherkaoui, A. Roukhe, *A Medical Image Watermarking Scheme Based on Dual-Tree Wavelet Transform*, Innovative Computing Technology (INTECH), 2012 Second International Conference on , vol., no., pp.144,152, 18-20 Sept. 2012.

[24] N. A. Memon,S. A M Gilani, *Adaptive Data Hiding Scheme for Medical Images Using Integer Wavelet Transform*, Emerging Technologies, 2009. ICET 2009. International Conference on , vol., no., pp.221,224, 19-20 Oct. 2009

[25] B. L. Lai, L. W. Chang, *Adaptive Data Hiding for Images Based on Harr Discrete Wavelet Transform*, Lecture Notes in Computer Science, Vol. 4319, pp 1085-1093, Springer Verlag 2006.

[26] The DICOM Supplement 55: Attribute Level Confidentiality (including De-identification), `ftp://medical.nema.org/medical/dicom/final/sup55_ft.pdf`, The Association of Electrical Equipment and Medical Imaging Manufacturers (NEMA).

[27] `http://eleceng.dit.ie.jblackledge/DICOM.zip`