Enhancing Information Hiding and Segmentation for Medical Images using Novel Steganography and Clustering Fusion Techniques

by

Hayat Shahir Al-Dmour

A dissertation submitted in fulfillment of the requirements for the degree

Doctor of Philosophy

School of Biomedical Engineering

Faculty of Engineering and Information Technology

University of Technology Sydney

January 2018

Certificate of Original Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Hayat Shahir Al-Dmour
Monday 14th August, 2017

Abstract

In recent years, there has been rapid development in digital medical imaging. The continuous development of medical imaging is expected to make further contributions to healthcare systems, where the increased use of medical imaging in a variety of clinical settings has played an important role in improving health services. The main objective of the research presented in this thesis is to investigate digital image steganography and segmentation in order to offer a systematic way for designing and developing them, with a particular concentration on medical imaging security and magnetic resonance (MR) brain image segmentation.

The first objective presents digital steganography, which refers to the science of concealing important information in digital media such as text, image, audio and video. The importance of this science comes from the fact that if the message is visible, then the attack is highly possible. So, the purpose of digital image steganography is to hide the existence of the secret message from a third party that is unauthorized to see it. The second objective presents digital segmentation, which aims to divide the image into meaningful and non-overlapping regions. The segmentation process is considered an essential process in many important biomedical applications, such as tumour detection, quantitative tissue analysis and computer-integrated surgery.

A major requirement for any steganography method is to minimize the changes that are introduced to the cover image by the data embedding process without compromising the embedding capacity. The main aim of this research is to propose techniques that achieve a high level of capacity, imperceptibility and security. In other words, the proposed methods attempt to reduce the degradation of the stego image to the level that makes the introduced changes not noticeable to the Human Visual System (HVS). Since the HVS is less sensitive to changes in sharp regions of images compared to uniform regions, many researchers have attempted to identify edge pixels and embed the secret message in them in order to enhance imperceptibility and increase the embedding capacity by varying the number of embedded bits per pixel based on edges' strength. However, the identification of edges in steganography systems is usually faced with some challenges that are mainly related to changes that are caused by the embedding process, which lead to having slight difference between the edges of the cover (original) image and the stego image (output of the embedding process). In addition to proposing a method that attempts to resolve this issue, we incorporate coding theory to help in reducing modifications caused by the embedding process.

In medical image security systems, information security schemes are used to conceal coded Electronic Patient Records (EPRs) into medical images. This will help to protect the EPRs' confidentiality without affecting the image quality and particularly the Region of Interest (ROI), which is essential for diagnosis. A method that converts EPR data into ciphertext using private symmetric encryption method is proposed. A simple edge detection method has been developed to embed the confidential information in edge pixels, which will lead to an improved stego image quality. To increase the efficiency, two message coding mechanisms have been utilized to enhance the ± 1 steganography. The first one, which is based on Hamming code, is simple and fast, while the other which is known as the Syndrome Trellis Code (STC), is more sophisticated as it attempts to find a stego image that is close to the cover image through minimizing the embedding impact. The proposed steganography algorithm embeds the secret data bits into the Region of Non Interest (RONI), where due to its importance; the ROI is preserved from modifications.

In order to enhance the performance of clustering-based medical image segmentation, an efficient fully-automatic brain tissue segmentation algorithm based on a clustering fusion technique is presented. In the training phase of this algorithm, the pixel intensity value is scaled to enhance the contrast of the image. The brain image pixels that have similar intensity values are then grouped into objects using a superpixel algorithm. Then, three clustering techniques are utilized to segment each object. For each clustering technique, a neural network (NN) model is fed with features extracted from the image objects and is trained using the labels produced by that clustering technique. In the testing phase, a pre-processing step that includes scaling and resizing of the brain image is applied before the superpixel algorithm partitions the image into multiple objects (similar to the training phase). The three trained neural network models are then used to predict the respective class of each object and the obtained classes are combined using majority voting.

The performance of all proposed methods have been tested and evaluated on different datasets using different criteria such embedding rate, mean square error (MSE), peak signal-to-noise ratio (PSNR), weighted peak signal-to-noise ratio (wPSNR), embedding efficiency, jaccard similarity (JS), dice similarity coefficient (DSC), root mean square error (RMSE), accuracy, sensitivity and specificity. Also, the effectiveness of the proposed steganography algorithm is proven using one of the efficient steganalysis techniques. The obtained results showed that our proposed methods outperform some of the well-established methods in the literature.

Acknowledgments

First and foremost, I would like to express my gratitude to Allah (Glorified and Exalted is He) for blessing me with endurance and fortitude to smoothly accomplish this dissertation.

During this research period of four years, there have been many people who have walked alongside me with their guidance, support and inspiration. I am much grateful to my principle supervisor **Dr. Ahmed Al-Ani** for incessant guidance and encouragement. Discussions with him helped me to understand my potential besides producing this high-quality dissertation. Without him, the road towards the completion of this research would have been difficult.

Thanks would be a simple word to applaud the sacrifice of my parents, especially my deceased father. He has been a pillar of support to shape my values, perseverance and in large as a good human. Extended thanks to my sisters and brothers who also supported me in each step of these four years.

I express my profound gratitude to all my colleagues and friends within and outside the premises of UTS, in particular **Obaid Aamir** and **Karthick Thiyagarajan**, for their friendship and support throughout my degree.

Lastly but not the least, I endorse the Mutah University for their commitment to support my Ph.D. candidature with scholarships. This enabled me to realise my long lasting dream into reality.

Dedication

Every challenging work needs self-efforts as well as guidance and support of parents. This thesis is dedicated to the memory of my father, ***Shaher Al-Dmour****, who passed away before I completed my degree. I wish that he could be with me to share the success of my graduation with a Doctor of Philosophy degree.*

This thesis is also dedicated to my mother, ***Khadija Al-Dmour****. This dissertation stands as a testimony for her endless support, prayers, love and beyond to overcome my hardships to complete my degree.*

To my beloved sisters and brothers, for their support and patience throughout these stressful years.

Abbreviations

PSNR	Peak Signal-to-Noise Ratio
BPNN	Back Propagation Neural Network
CSF	Cerbuspinal Fluid
DCT	Discrete Cosine Transform
DHHS	Department of Health and Human Services
DICOM	Digital Imaging and Communication In Medicine
DSC	Dice Similarity Coefficient
DWT	Discrete Wavelet Transform
ECC	Error Correction Code
EPR	Electronic Patient Record
FCM	Fuzzy C-means
GM	Gray Matter
HAS	Human Auditory System
HIPAA	Health Insurance Portability and Accountability Act
HVS	Human Visual System
ID	Identity Card
IDCT	Inverse Discrete Cosine Transform
ISP	Internet Service Provider

IWT	Integer Wavelet Transform
JPEG	Joint photographic expert group
JS	Jaccard Similarity
KLD	Kullback–Leibler Divergence
LSB	Least Significant Bit
LSBM	Least Significant Bit Matching
MIS	Medical Information System
MSB	Most Significant Bit
MSE	Mean Square Error
MRI	Magnetic Resonance Image
PACS	Picture Archiving and Communication System
PMM	Pixel Mapping Method
PoV	Pair of Values
PRNG	Pseudo Random Number Generator
PSNR	Peak Signal-to-Noise Ratio
PVD	Pixel Value Difference
RLC	Run Length Coding
RMSE	Root Mean Square Error
ROI	Region of Interest
RONI	Region of Non-Interest
SOM	Self-organized Map
SSIM	Structural Similarity Index
STC	Syndrome Trellis Code
TBPC	Tree-based Parity Check
TPVD	Tri-pixel Value Differencing
VoIP	Voice over Internet Protocol
WM	White Matter
wPSNR	weighted Peak Signal-to-Noise Ratio

Contents

List of Figures

List of Tables