

A FRAMEWORK OF ENHANCED  
AUTHENTICATION FOR PDF TEXTUAL  
DOCUMENTS USING ZIGZAG-LSB  
EMBEDDING ALGORITHM

NUR ALYA AFIKAH BINTI USOP

MASTER OF SCIENCE

UNIVERSITI MALAYSIA PAHANG  
AL-SULTAN ABDULLAH

## THESIS DECLARATION LETTER (OPTIONAL)

Librarian,  
Universiti Malaysia Pahang Al-Sultan Abdullah,  
Lebuh Persiaran Tun Khalil Yaakob,  
26300, Gambang, Kuantan, Pahang.

Dear Sir,

### CLASSIFICATION OF THESIS AS RESTRICTED

Please be informed that the following thesis is classified as RESTRICTED for a period of three (3) years from the date of this letter. The reasons for this classification are as listed below.

Author's Name

Thesis Title

Reasons	(i)
	(ii)
	(iii)

Thank you.

Yours faithfully,

---

(Supervisor's Signature)

Date:

Stamp:

Note: This letter should be written by the supervisor and addressed to the Librarian, *Universiti Malaysia Pahang Al-Sultan Abdullah* with its copy attached to the thesis.



## SUPERVISOR'S DECLARATION

I/We\* hereby declare that I/We\* have checked this thesis/project\* and in my/our\* opinion, this thesis/project\* is adequate in terms of scope and quality for the award of the degree of \*Doctor of Philosophy/ Master of Science.

---

(Supervisor's Signature)

Full Name : TS. DR. SYIFAK BINTI IZHAR HISHAM  
Position : SUPERVISOR  
Date : 17 DECEMBER 2023

---

(Co-supervisor's Signature)

Full Name : PROFESOR MADYA TS. DR. ROHANI BINTI ABU BAKAR  
Position : CO-SUPERVISOR  
Date : 24 JANUARY 2024



## STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Universiti Malaysia Pahang Al-Sultan Abdullah or any other institutions.

---

(Student's Signature)

Full Name : NUR ALYA AFIKAH BINTI USOP

ID Number : MCM 19001

Date : 26 NOVEMBER 2023

A FRAMEWORK OF ENHANCED AUTHENTICATION FOR PDF TEXTUAL  
DOCUMENTS USING ZIGZAG-LSB EMBEDDING ALGORITHM

NUR ALYA AFIKAH BINTI USOP

Thesis submitted in fulfillment of the requirements  
for the award of the degree of  
Doctor of Philosophy

Faculty of Computing

UNIVERSITI MALAYSIA PAHANG AL-SULTAN ABDULLAH

JANUARY 2024

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my primary supervisor, Ts. Dr. Syifak binti Izhar Hisham, for her invaluable guidance, unwavering support, and insightful feedback throughout the entire research process. Her expertise and dedication have been instrumental in shaping the direction and quality of this thesis. I truly value the unwavering support she has provided since the first day of my Master's program. I am also deeply thankful to my secondary supervisor, Profesor Madya Ts. Dr. Rohani Binti Abu Bakar, and my internal examiner, Profesor Madya Ts. Dr. Ferda Ernawan, for their constructive criticism, thoughtful suggestions, and encouragement that greatly contributed to the refinement of this work.

I extend my sincere appreciation to my lab-mate, particularly 'Aqilah Mohd Ghani and Noor Hazirah binti Hassan, for their invaluable assistance in resolving challenges and their encouragement that propelled me towards the successful completion of my Master's research. Their support has been instrumental in overcoming obstacles and maintaining motivation throughout this academic journey. I would like to express my gratitude to all the staff members of the Faculty of Computing at UMP who have provided assistance in various ways, making my Master's journey both smooth and meaningful.

My sincere thanks go to my family; mum, dad, and eldest sister for their unwavering encouragement and understanding throughout this challenging academic journey. Their support has been my pillar of strength. I am grateful to my beloved companion and my friends for their patience, motivations, and understanding along this Master journey. Lastly, I want to express my gratitude to everyone who willingly contributed their time and insights to this research. Their involvement was crucial to the successful completion of this research. This thesis represents the collective effort and support of many individuals, and I am deeply appreciative of their contributions. I pray for the best for everyone.

## ABSTRAK

Tahap keselamatan untuk dokumen adalah penting kerana banyak prosedur dan proses formal pada masa kini memerlukan dokumen digital sebagai bukti atau pengesahan proses, seperti resit perbankan internet, surat perlantikan rasmi, transkrip akademik dan sijil, dokumen yang dipindai, dan dokumen komersial untuk pemasaran dalam talian. Fenomena penggunaan meluas dokumen digital dalam e-mel, laman sosial, laman web e-dagang, dan banyak laman web yang berkaitan dengan kerajaan adalah satu trend untuk mencapai audiens, menjimatkan masa pemrosesan dan sumber kertas, serta mengurangkan kos pos. Kemajuan pesat teknologi maklumat dalam multimedia digital bermakna kandungan asal dokumen boleh dengan mudah dimanipulasi, diedit, atau ditamper oleh pihak lain. Oleh itu, satu prosedur watermark digital diperlukan untuk mengesan sebarang perubahan dalam maklumat asal dalam dokumen. Suatu kerangka kerja penambahbaikan pengesahan untuk dokumen teks PDF menggunakan algoritma sembunyi Zigzag-lsb dicadangkan untuk melaksanakan teknik skema watermarking baru yang boleh digunakan dalam saiz dokumen secara umum, untuk mengkaji sistem penomboran semasa untuk mengamankan pengesahan dokumen digital, dan menilai skema watermarking untuk dokumen digital dari segi kadar ketepatan, serta menganalisis prestasinya. Bit paling tidak signifikan (LSB) digunakan oleh teknik-teknik tersebut untuk menyulitkan data pengesahan. Dengan menyebarkan data asal yang bernombor sejauh mungkin dari lokasi asal, strategi-strategi itu menggunakan sistem penomboran khusus. Proses yang serupa digunakan dalam fasa penciptaan dan penanaman watermark. Keberkesanan algoritma itu telah dibuktikan kerana ia menggunakan pendekatan berdasarkan imej selepas penukaran antara dokumen dan imej dengan corak penomboran yang rapuh terhadap serangan padam, gantian, sisipan, gabungan, dan salinan. Sebagai kesimpulan, corak penomboran baru telah terbukti lebih cekap berbanding corak penomboran sedia ada dalam watermarking kerana hasilnya menunjukkan prestasi terbaik dari segi kadar ketepatan.

## ABSTRACT

The security level for documents is necessary since a lot of formal procedures and processes nowadays need digital documents as proof or process verification, such as internet banking receipts, official letters of appointment, academic transcripts and certificates, scanned documents, and commercial documents for online marketing. The phenomenon of widespread usage of digital documents in emails, social networking sites, e-commerce websites, and many government-linked websites is a trend for reaching the audience, saving processing time and paper resources, as well as cutting the cost of postage. The rapid advancement of information technology in digital multimedia means that the original content of documents can be easily manipulated, edited, or tampered with by others. Hence, a digital watermarking procedure is needed to detect any changes in of the original information in document becomes necessary. A framework of enhanced authentication for PDF textual documents using Zigzag-lsb embedding algorithm is proposed in order to apply a new technique of watermarking scheme which can be used in general size of document, to study the current numbering of watermarking system to secure the authentication of digital documents, and evaluate the watermarking scheme for digital documents in terms of accuracy rate, and analyse the performance. Least significant bits (LSB) were used by the techniques to encrypt authentication data. By dispersing the numbered original data as far away from the original locations as feasible, the strategies made use of specific numbering systems. Similar processes are used in the generation and embedding phases of a watermark. The effectiveness of the algorithm has been established since it uses an image-based approach after conversion between document and images with a numbering pattern that is fragile to deletion, replacement, insertion, combine, and copy attack. As a conclusion, the new numbering pattern have demonstrated to be more competent compared to the existing numbering pattern in watermarking since the results have shown the best performance in terms of accuracy rate.



## TABLE OF CONTENT

<b>DECLARATION</b>	
<b>TITLE PAGE</b>	
<b>ACKNOWLEDGEMENTS</b>	<b>ii</b>
<b>ABSTRAK</b>	<b>iii</b>
<b>ABSTRACT</b>	<b>iv</b>
<b>TABLE OF CONTENT</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>x</b>
<b>LIST OF FIGURES</b>	<b>xii</b>
<b>LIST OF SYMBOLS</b>	<b>xiv</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xv</b>
<b>LIST OF APPENDICES</b>	<b>xvi</b>
<b>CHAPTER 1 INTRODUCTION</b>	<b>17</b>
1.1 Background of Research	17
1.2 Problem Statement	19
1.3 Research Questions	19
1.4 Research Objectives	20
1.5 The Final Outcome	20
1.6 Research Scope	20
1.7 Thesis Outline	21
1.8 Conclusion	21
<b>CHAPTER 2 LITERATURE REVIEW</b>	<b>22</b>
2.1 Introduction	22
2.2 General Concepts	23
2.3 Requirement for Digital Watermarking Methods	27

2.3.1	Embedding/ extracting Domain	27
2.3.2	Availability of Reference Data	30
2.3.3	Fragile Watermarking	31
2.3.4	Document Type	35
2.3.5	Image	36
2.3.6	Vector Data	37
2.3.7	Raster Data	37
2.4	Technique Classification of Digital Watermarking	38
2.4.1	Image Based Approach	38
2.5	Characteristics of Digital Watermarking	39
2.5.1	Imperceptibility	39
2.5.2	Security	40
2.5.3	Robustness	41
2.5.4	Capacity	42
2.6	Properties of Digital Watermarking	42
2.6.1	Effectively	42
2.6.2	Image Fidelity	43
2.6.3	Payload Size	43
2.6.4	False Positive	43
2.7	Possible Attacks to Digital Watermarking	43
2.7.1	Deletion Attack	43
2.7.2	Insertion Attack	44
2.7.3	Combined Attack	45
2.7.4	Copy Attack	45
2.7.5	Replacement Attack	45
2.8	Numbering Pattern	45

2.8.1	Scanning	45
2.9	Performance Measurement Methods	48
2.9.1	Mean Square Error (MSE)	48
2.9.2	Structural Similarity Index Measure (SSIM)	49
2.9.3	Peak Signal Noise Ratio (PSNR)	50
2.9.4	Accuracy Rate	50
2.9.5	Detection Range	51
2.9.6	Tamper Localisation Measurement	51
2.10	Revision on Existing Fragile Watermarking in Image	52
2.10.1	Conclusion of Fragile Watermarking in Image	58
<b>CHAPTER 3 RESEARCH METHODOLOGY</b>		<b>60</b>
3.1	Introduction	60
3.2	Research Strategy and Development	60
3.2.1	Initial Planning	61
3.2.2	Analysis	62
3.2.3	Design and Implementation	62
3.2.4	Spiral Pattern	64
3.2.5	Zigzag SCAN Pattern	74
3.3	Research Dataset	81
3.4	Research Instrument	83
3.5	Research Evaluation	83
3.6	Conclusion	84
<b>CHAPTER 4 RESULTS AND DISCUSSION</b>		<b>85</b>
4.1	Introduction	85
4.2	Table of Results	85

4.3	Spiral-LSB Numbering Technique	85
4.3.1	Watermark Embedding of 8x8 Block size Documents	86
4.3.2	Watermark Embedding of 4x4 Block size Documents	88
4.3.3	Watermark Detection of 8x8 Block Size Document	89
4.3.4	Watermark Detection of 4x4 Block Size Document	95
4.3.5	Results and Discussion	96
4.3.6	Limitations of Spiral-LSB	99
4.4	Zigzag-LSB Numbering Technique	99
4.4.1	Watermark Embedding of 8x8 Block size Document	100
4.4.2	Watermark Embedding of 4x4 Block size Document	102
4.4.3	Watermark Detection of 8x8 Block size Document	104
4.4.4	Watermark Detection of 4x4 Block size Document	109
4.4.5	Results and Discussion	110
4.4.6	Limitations of Zigzag-LSB	116
4.5	Tamper Localisation Results	117
4.6	Conclusion and Contributions	121
	<b>CHAPTER 5 CONCLUSION</b>	<b>125</b>
5.1	Introduction	125
5.2	Summary	125
5.2.1	Advantages of block-based watermarking	126
5.3	Conclusion and Limitation	127
5.4	Suggestion and Further Research	128
	<b>REFERENCES</b>	<b>130</b>
	<b>APPENDICES</b>	<b>139</b>

## LIST OF TABLES

Table 2. 1	The comparison between spatial domain and transform domain (Hisham et al., 2016)	29
Table 2. 2	Example of LSB (adapted from (Sutaone, 2008))	34
Table 2. 3	Existing scheme of fragile watermarking in Image	56
Table 3. 1	The example of sample datasets in experiment	82
Table 4. 1	Comparison between the original data and watermarked data with document size and PSNR value using 8x8 Block Size	86
Table 4. 2	Comparison between the original data and watermarked data with document size and PSNR value using 4x4 Block Size	88
Table 4. 3	The obtained results and comparison were conducted among five (5) types of online data commonly utilized in formal dealings or business transactions using 8x8 block size	94
Table 4. 4	Comparison between the original data and tampered data using 4x4 block size	95
Table 4. 5	Comparison of elapsed time (sec.) between 8x8 Block Size and 4x4 Block Size Data	97
Table 4. 6	Performance measurement for detection in 8x8 Block Size Data	98
Table 4. 7	The original data and watermarked data with size and PSNR values using 8x8 block size	100
Table 4. 8	Comparison between the original data and watermarked data with document size, PSNR value and elapsed time using 4x4 Block size	103
Table 4. 9	The findings and comparisons among the five (5) types of online documents commonly employed in formal business dealings using 8x8 block size	109
Table 4. 10	Comparison between the original data and watermarked data with document size, detection accuracy, and detection range using 4x4 block size	110
Table 4. 11	Performance measurement for detection in 8x8 block size data	112
Table 4. 12	Comparison for detection in examination slip using 8x8 block size and 4x4 block size data	114
Table 4. 13	Comparison between original and watermarked data with the PSNR value produced by Spiral-LSB and Zigzag-LSB method	115
Table 4. 14	Comparison with other numbering pattern watermarking methods	116
Table 4. 15	Comparison between the tampering rate and elapsed time in Spiral-LSB and Zigzag-LSB technique based on possible attacks using 8x8 block size	117
Table 4. 16	Result of deteted tampered region after being tested with possible attacks	118

Table 4. 17	Table Comparisons of tampering localisations between possible attacks	120
Table 5. 1	The summary of contributions	123

## LIST OF FIGURES

Figure 2. 1	General concepts of watermarking	23
Figure 2. 2	The visible watermark on cover image	24
Figure 2. 3	The embedder and extractor in watermarking system	26
Figure 2. 4	Generic concept for digital watermarking technique	26
Figure 2. 5	Least Significant Bits (LSB)	33
Figure 2. 6	Scenario of insertion attack	44
Figure 2. 7	Few types of scanning; (a) Raster, (b) Spiral, (c) Zeta, (d) Diagonal, (e) Hilbert.	46
Figure 2. 8	The zigzag SCAN (diagonal scanning) pattern	47
Figure 2. 9	The zigzag SCAN (diagonal scanning) pattern	48
Figure 3. 1	Research work of Spiral-LSB	61
Figure 3. 2	Conversion of document into BMP image format	65
Figure 3. 3	The proposed scheme of spiral numbering, mapping, generation and embedding of watermark	65
Figure 3. 4	The example of spiral numbering pattern in centre 5x5 block	67
Figure 3. 5	The 8x8 pixels of block, B with sub blocks, Bs	68
Figure 3. 6	The conversion of document into image using MATLAB	71
Figure 3. 7	The list of attacks used in the document and being converted back into image	73
Figure 3. 8	The detection phase where attacked image is being tested to detect attacks	74
Figure 3. 9	Conversion of document type of PDF into BMP image format	76
Figure 3. 10	The proposed scheme of numbering, mapping, generation and embedding of watermark	76
Figure 3. 11	Diagonals of Zigzag SCAN	79
Figure 4. 1	(From left) (a) Tampered data with deletion attack; (b) detected tampered data	90
Figure 4. 2	(From left) (a) Tampered data with replacement attack; (b) detected tampered data	91
Figure 4. 3	(From left) (a) Tampered data with insertion attack; (b) detected tampered data	92
Figure 4. 4	(From left) (a) Tampered data with combined attack; (b) detected tampered data	93
Figure 4. 5	(From left) (a) Tampered data with copy attack; (b) detected tampered data with original document size; (c) detected tampered data with resized document	94

Figure 4. 6	PSNR value for 20 datasets in 8x8 block sizes	96
Figure 4. 7	PSNR value for 20 datasets in 4x4 block sizes	97
Figure 4. 8	(From left) (a) Tampered data with deletion attack; (b) detected tampered data	104
Figure 4. 9	(From left) (a) Tampered data with replacement attack; (b) detected tampered data	105
Figure 4. 10	(From left) (a) Tampered data with insertion attack; (b) detected tampered data	106
Figure 4. 11	(From left) (a) Tampered data with combined attack; (b) detected tampered data	107
Figure 4. 12	(From left) (a) Tampered data with copy attack; (b) detected tampered data	108
Figure 4. 13	PSNR value for 20 datasets in 8x8 block size	111



## REFERENCES

- 101GIS. (2021). *Vector data vs Raster Data: Which one should I choose?* 101gis. <https://101gis.com/vector-data-vs-raster-data/>
- Akter, A., & Ullah, M. A. (2014). DIGITAL WATERMARKING WITH A NEW ALGORITHM. In *IJRET: International Journal of Research in Engineering and Technology*. <http://www.ijret.org>
- Aminuddin, A., & Ernawan, F. (2023). AuSR3: A new block mapping technique for image authentication and self-recovery to avoid the tamper coincidence problem. *Journal of King Saud University - Computer and Information Sciences*, 35(9). <https://doi.org/10.1016/j.jksuci.2023.101755>
- Bashardoost, M., Mohd Rahim, M. S., Saba, T., & Rehman, A. (2017). Replacement Attack: A New Zero Text Watermarking Attack. *3D Research*, 8(1). <https://doi.org/10.1007/s13319-017-0118-y>
- Bhalerao, S., Ansari, I. A., & Kumar, A. (2021). A secure image watermarking for tamper detection and localization. *Journal of Ambient Intelligence and Humanized Computing*, 12(1), 1057–1068. <https://doi.org/10.1007/s12652-020-02135-3>
- Bourbakis, N. G. (1997). IMAGE DATA COMPRESSION-ENCRYPTION USING G-SCAN PATTERNS. *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation Vol.2*, 1117–1120.
- Brassil, J. T., Low, S., Maxemchuk, N. F., & O’Gorman, L. (1995). Electronic Marking and Identification Techniques to Discourage Document Copying. *IEEE Journal on Selected Areas in Communications*, 13(8), 1495–1504. <https://doi.org/10.1109/49.464718>
- Bravo-Solorio, S., & Nandi, A. K. (2011). Secure fragile watermarking method for image authentication with improved tampering localisation and self-recovery capabilities. *Signal Processing*, 91(4), 728–739. <https://doi.org/10.1016/J.SIGPRO.2010.07.019>

- Cayre, F., Fontaine, C., & Furon, T. (2005). Watermarking security: Theory and practice. In *IEEE Transactions on Signal Processing* (Vol. 53, Issue 10 II, pp. 3976–3987). <https://doi.org/10.1109/TSP.2005.855418>
- Chakraborty, S., Jalal, A. S., & Bhatnagar, C. (2017). LSB based non blind predictive edge adaptive image steganography. *Multimedia Tools and Applications*, 76(6), 7973–7987. <https://doi.org/10.1007/s11042-016-3449-4>
- Chao, H., & Fan, J. (2004). Layout and Content Extraction for PDF Documents. In *LNCS* (Vol. 3163).
- Concolato, C., Schmitz, P., Association for Computing Machinery. Special Interest Group on Hypertext, H. and W., Association for Computing Machinery. Special Interest Group on Systems Documentation, Adobe Systems, Hewlett-Packard Company, Ecole nationale supérieure des télécommunications (France), Association for Computing Machinery, & ACM Digital Library. (2012). *DocEng 2012 : proceedings of the 2012 ACM Symposium on Document Engineering : September 4-7, 2012, Paris, France*.
- Di Martino, F., & Sessa, S. (2019). Fragile watermarking tamper detection via bilinear fuzzy relation equations. *Journal of Ambient Intelligence and Humanized Computing*, 10(5), 2041–2061. <https://doi.org/10.1007/s12652-018-0806-3>
- Dittmann, J., Wohlmacher, P., & Nahrstedt, K. (2001). Using cryptographic and watermarking algorithms. *IEEE Multimedia*, 8(4), 54–65. <https://doi.org/10.1109/93.959103>
- Durvey, M., & Satyarthi, D. (2014). A Review Paper on Digital Watermarking. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(4), 99–105.
- Elbasi, E. (2022). A non-blind Watermarking Technique using Flexible Scaling Factor in Wavelet Transform. *2022 45th International Conference on Telecommunications and Signal Processing, TSP 2022*, 150–155. <https://doi.org/10.1109/TSP55681.2022.9851257>

- Ernawan, F., Aminuddin, A., Nincarean, D., Razak, M. F. A., & Firdaus, A. (2022). Three Layer Authentications with a Spiral Block Mapping to Prove Authenticity in Medical Images. *International Journal of Advanced Computer Science and Applications*, 13(4), 211–223. <https://doi.org/10.14569/IJACSA.2022.0130425>
- Gemelli, A., Vivoli, E., & Marinai, S. (2022). Graph Neural Networks and Representation Embedding for Table Extraction in PDF Documents; Graph Neural Networks and Representation Embedding for Table Extraction in PDF Documents. *2022 26th International Conference on Pattern Recognition (ICPR)*. <https://doi.org/10.1109/ICPR56361.2022.9956590>
- Giri, K. J., Peer, M. A., & Nagabhushan, P. (2014). A channel wise color image watermarking scheme based on discrete wavelet transformation. *2014 International Conference on Computing for Sustainable Global Development, INDIACom 2014*, 758–762. <https://doi.org/10.1109/IndiaCom.2014.6828064>
- Goos, G., Hartmanis, J., Van, J., Board, L. E., Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Zurich, E., Mitchell, J. C., Naor, M., Nierstrasz, O., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., & Weikum, G. (2005). LNCS 3710 - Digital Watermarking. In *Lecture Notes in*.
- Goyal, L. M. K. V. A. P. A. P. P. (2014). A Robust Method for Integrity Protection of Digital Data in Text Document Watermarking. *IJIRST-International Journal for Innovative Research in Science & Technology*, 1(6). [www.ijirst.org](http://www.ijirst.org)
- Hartung, F., & Kutter, M. (1999). Multimedia watermarking techniques. *Proceedings of the IEEE*, 87(7), 1079–1107. <https://doi.org/10.1109/5.771066>
- Hisham, S. I. (2016). *ENHANCED LSB WATERMARKING METHODS BASED ON SCANNING PATTERNS FOR AUTHENTICATION OF MEDICAL IMAGES SYIFAK IZHAR HISHAM*.
- Hisham, S. I., Muhammad, A. N., Badshah, G., Johari, N. H., & Mohamad Zain, J. (2017). Numbering with spiral pattern to prove authenticity and integrity in medical images. *Pattern Analysis and Applications*, 20(4), 1129–1144. <https://doi.org/10.1007/s10044-016-0552-0>

- Hisham, S. I., Zain, J. M., Arshad, N. W., & Chuin, L. S. (2015a). HILBERT-LSB-C as Authentication System for Color Medical Images. *2015 4th International Conference on Software Engineering and Computer Systems (ICSECS)*, 15–20.
- Hisham, S. I., Zain, J. M., Arshad, N. W., & Chuin, L. S. (2015b). HILBERT-LSB-C as Authentication System for Color Medical Images. *4th International Conference on Software Engineering and Computer Systems (ICSECS)*, 15–20.
- Huang, H.-C., Hang, H.-M., & Pan, J.-S. (2004). *An Introduction to Watermarking Techniques*. www.worldscientific.com
- Iqbal, M. M. (2019). A Robust Digital Watermarking Algorithm for Text Document Copyright Protection based on Feature Coding. *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 1940–1945.
- Jalil, Z., & Mirza, A. M. (2009). A review of digital watermarking techniques for text documents. *2009 International Conference on Information and Multimedia Technology, ICIMT 2009*, 230–234. <https://doi.org/10.1109/ICIMT.2009.11>
- Jalil, Z., & Mirza, A. M. (2010). Text watermarking using combined image-plus-text watermark. *2nd International Workshop on Education Technology and Computer Science, ETCS 2010, 1*, 11–14. <https://doi.org/10.1109/ETCS.2010.494>
- Jaseena, K. U., & John, A. (2011). An Invisible Zero Watermarking Algorithm using Combined Image and Text for Protecting Text Documents. / *International Journal on Computer Science and Engineering (IJCSE)*, 3(6), 2265–2272.
- Jassim, F. (2013). *Increasing Compression Ratio in PNG Images by k-Modulus Method for Image Transformation*. 1–10.
- Kang, H., & Iwamura, K. (2014). *Image Protection System with Steganography and Authentication*. <https://doi.org/10.1109/IIH-MSP.2014.123>
- Kantner, L., Shroyer, R., & Rosenbaum, S. (2002). Structured Heuristic Evaluation of Online Documentation. *IEEE International Professional Communication Conference*, 331–342.

- Khadam, U., Iqbal, M. M., Azam, M. A., Khalid, S., Rho, S., & Chilamkurti, N. (2019). *Digital Watermarking Technique for Text Document Protection Using Data Mining Analysis*.  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8713871>
- Kim, Y.-W., & Oh, I.-S. (2004). *Watermarking text document images using edge direction histograms*. <https://doi.org/10.1016/j.patrec.2004.04.002>
- Liew, S.-C., & Mohd Zain, J. (2011). Tamper Localization and Lossless Recovery Watermarking Scheme. *Software Engineering and Computer Systems: Second International Conference, ICSECS, 179*, 555–566.
- Liew, S.-C., & Zain, J. M. (2009). A Review of Medical Image Watermarking and Its Implementations. *In Proceedings of Malaysian Technical Universities Conference on Engineering and Technology (MUCEET2009)*, 20–22.  
<https://www.researchgate.net/publication/282610645>
- Lin, H., Yang, S., & Xu, L. (2011). Watermark algorithm for color image authentication and restoration. *Proceedings of 2011 International Conference on Electronic and Mechanical Engineering and Information Technology, EMEIT 2011, 6*, 2773–2776. <https://doi.org/10.1109/EMEIT.2011.6023677>
- Lin, X., Gao, L., Tang, Z., Baker, J., & Sorge, V. (2014). Mathematical formula identification and performance evaluation in PDF documents. *International Journal on Document Analysis and Recognition, 17*(3), 239–255.  
<https://doi.org/10.1007/s10032-013-0216-1>
- Mehta, S., Prabhakaran, B., Nallusamy, R., & Newton, D. (2016). *mPDF: Framework for Watermarking PDF Files using Image Watermarking Algorithms*.  
<http://arxiv.org/abs/1610.02443>
- Mishra, A., Rajpal, A., & Bala, R. (2018). Bi-directional extreme learning machine for semi-blind watermarking of compressed images. *Journal of Information Security and Applications, 38*, 71–84. <https://doi.org/10.1016/j.jisa.2017.11.008>
- Morkel, T., Eloff, J. H. P., & Olivier, M. S. (2005). An overview of image steganography. *Proceedings of the Fifth Annual ...*, 1(2), 1–11.

- Muhamad, A., & Zain, J. M. (2007). *Using Spiral Scan Technique for Medical Image Watermarking with Tamper Detection and Recovery*.
- Qin, C., Chang, C. C., & Chen, P. Y. (2012). Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism. *Signal Processing*, 92(4), 1137–1150. <https://doi.org/10.1016/j.sigpro.2011.11.013>
- Qin, C., Ji, P., Zhang, X., Dong, J., & Wang, J. (2017). Fragile image watermarking with pixel-wise recovery based on overlapping embedding strategy. *Signal Processing*, 138, 280–293. <https://doi.org/10.1016/j.sigpro.2017.03.033>
- Ramakrishnan, C., Patnia, A., Hovy, E., & Burns, G. A. (2012). *Layout-aware text extraction from full-text PDF of scientific articles*. <http://code.google.com/p/lapdf/text/>.
- Saini, L. K., Shrivastava, V., Tech, M., & Scholar, R. (2014). A Survey of Digital Watermarking Techniques and its Applications. *International Journal of Computer Science Trends and Technology*, 2. [www.ijcstjournal.org](http://www.ijcstjournal.org)
- Salch, M. (2019). *Which Graphic File Format is Best: Vector and Raster Images - Tell Your Tale Marketing & Design*. <https://tellyourtale.com/graphic-design/which-graphic-file-format-is-best-vector-and-raster-images/>
- Santhi, V., & Arulmozhivarman, P. (2013). Hadamard transform based adaptive visible/invisible watermarking scheme for digital images. *Journal of Information Security and Applications*, 18(4), 167–179. <https://doi.org/10.1016/j.istr.2013.01.001>
- Setiadi, D. R. I. M. (2020). PSNR vs SSIM: imperceptibility quality assessment for image steganography. *Multimedia Tools and Applications*, 80(6), 8423–8444. <https://doi.org/10.1007/s11042-020-10035-z>
- Shahreza, M. S. (2005). An Improved Method for Steganography on Mobile Phone. *WSEAS Transactions on Systems*, 4(7), 955–957. <http://mohammad.shirali.ir>

- Shamimi Kamaruddin, N., Kamsin, A., Por, Y., & Rahman, H. (2018). *A Review of Text Watermarking: Theory, Methods, and Applications*.  
<https://doi.org/10.1109/ACCESS.2018.2796585>
- Sharma, G., & Coumou, D. J. (2006). Watermark synchronization: Perspectives and a new paradigm. *2006 IEEE Conference on Information Sciences and Systems, CISS 2006 - Proceedings*, 1182–1187. <https://doi.org/10.1109/CISS.2006.286644>
- Shukla, B., Khatri, S. K., Kapur, P. K., Amity University, Amity University. Amity Institute of Information Technology, Computer Society of India., Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, & Institute of Electrical and Electronics Engineers. (2016). *2016 5th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions) : September 7-9, 2016, venue, Amity University Uttar Pradesh, Noida, India*.
- Singh, B., & Sharma, M. K. (2021). Tamper detection technique for document images using zero watermarking in wavelet domain. *Computers and Electrical Engineering*, 89. <https://doi.org/10.1016/j.compeleceng.2020.106925>
- Singh, P., & Chadha, R. S. (2013a). A Survey of Digital Watermarking Techniques, Applications and Attacks. *Certified International Journal of Engineering and Innovative Technology (IJEIT)*, 2(9), 165–175.  
<https://www.researchgate.net/publication/342344131>
- Singh, P., & Chadha, R. S. (2013b). Review to Digital Watermarking and a Novel Approach to Position the Watermark in the Digital Image. *International Journal of Engineering and Advanced Technology (IJEAT)*, 2(4), 2249–8958.
- Song, C., Sudirman, S., Merabti, M., & Jones, D. L. (2010). Analysis of Digital Image Watermark Attacks. *2010 7th IEEE Consumer Communications and Networking Conference*, 1–5.  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5421631>
- Stev. (2021). *What Does it Mean to Scan a Document? | Scan To Computer*.  
<https://scantocomputer.com/what-does-it-mean-to-scan-a-document/>

- Sutaone, M. S. (2008). Image Based Steganography Using LSB Insertion Technique. *IET International Conference on Wireless, Mobile and Multimedia Networks*, 146–151.
- Suthaharan, S. (2004). Fragile image watermarking using a gradient image for improved localization and security. *Pattern Recognition Letters*, 25(16), 1893–1903. <https://doi.org/10.1016/j.patrec.2004.08.017>
- Thanki, R., Borra, S., Dwivedi, V., & Borisagar, K. (2017). A RONI Based Visible Watermarking Approach for Medical Image Authentication. *Journal of Medical Systems*, 41(9). <https://doi.org/10.1007/s10916-017-0795-3>
- Thanki, R., Kothari, A., & Trivedi, D. (2019). Hybrid and blind watermarking scheme in DCuT – RDWT domain. *Journal of Information Security and Applications*, 46, 231–249. <https://doi.org/10.1016/j.jisa.2019.03.017>
- Varshney, Y. (2017). Attacks on Digital Watermarks: Classification, Implications, Benchmarks. *International Journal on Emerging Technologies (Special Issue NCETST-2017)*, 8(1), 229–235. [www.researchtrend.net](http://www.researchtrend.net)
- Vidhya, R., Brindha, M., & Gounden, N. A. (2020). Analysis of zig-zag scan based modified feedback convolution algorithm against differential attacks and its application to image encryption. *Applied Intelligence*, 50(10), 3101–3124. <https://doi.org/10.1007/s10489-020-01697-1>
- Wang, C., Zhang, H., & Zhou, X. (2018). A self-recovery fragile image watermarking with variable watermark capacity. *Applied Sciences (Switzerland)*, 8(4). <https://doi.org/10.3390/app8040548>
- Wolfram Mathworld. (2013). *Wolfram MathWorld: The Web's Most Extensive Mathematics Resource*. <https://mathworld.wolfram.com/>
- Wong, P. W., & Memon, N. (2001). Secret and public key image watermarking schemes for image authentication and ownership verification. *IEEE Transactions on Image Processing*, 10(10), 1593–1601. <https://doi.org/10.1109/83.951543>



- Yang, H., & Kot, A. C. (2004). Text document authentication by integrating inter character and word spaces watermarking. *2004 IEEE International Conference on Multimedia and Expo (ICME)*, 2, 955–958.  
<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1394360>
- Yu, M., Wang, J., Jiang, G., Peng, Z., Shao, F., & Luo, T. (2015). New fragile watermarking method for stereo image authentication with localization and recovery. *AEU - International Journal of Electronics and Communications*, 69(1), 361–370. <https://doi.org/10.1016/j.aeue.2014.10.006>
- Zain, J. M., & Fauzi, A. R. M. (2006). Medical image watermarking with tamper detection and recovery. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 3270–3273.  
<https://doi.org/10.1109/IEMBS.2006.260767>
- Zhang, X., & Wang, S. (2008). Fragile watermarking with error-free restoration capability. *IEEE Transactions on Multimedia*, 10(8), 1490–1499.  
<https://doi.org/10.1109/TMM.2008.2007334>
- Zhou, X., Zhao, W., Wang, S., & Peng, R. (2009). *A semi-fragile watermarking scheme for content authentication of chinese text documents; A semi-fragile watermarking scheme for content authentication of chinese text documents.*  
<https://doi.org/10.1109/ICCSIT.2009.5234911>