

# AN IMPLEMENTING OF ZIGZAG PATTERN IN NUMBERING WATERMARKING BITS FOR HIGH DETECTION ACCURACY OF TAMPERS IN DOCUMENT

Nur Alya Afikah Usop

Faculty of Computing, Universiti Malaysia Pahang,  
26600 Pekan, Pahang, Malaysia  
[alyausop@gmail.com](mailto:alyausop@gmail.com)

Syifak Izhar Hisham

Faculty of Computing, Universiti Malaysia Pahang,  
26600 Pekan, Pahang, Malaysia  
[syifakizhar@ump.edu.my](mailto:syifakizhar@ump.edu.my)

‘Aqilah Abd Ghani

Faculty of Computing, Universiti Malaysia Pahang,  
26600 Pekan, Pahang, Malaysia  
[Aqilahghani.edu@gmail.com](mailto:Aqilahghani.edu@gmail.com)

Jasni Mohd Zain

Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Kompleks Al-Khawarizmi,  
Universiti Teknologi MARA,  
40450, Shah Alam, Selangor, Malaysia  
[jasni@tmsk.uitm.edu.my](mailto:jasni@tmsk.uitm.edu.my)

## Abstract

Digital form nowadays has been easily manipulated by many corrupters especially in formal business. PDF document is one of the main digital forms that need to be accessed when dealing with offer letters, scanned documents, digital receipt, etc. Thus, PDF is one of the critical data that needs to be protected from any manipulations of its originality. An experiment is conducted to develop an improved zigzag-LSB watermarking scheme for all characters of text documents in PDF format which proved to have the best embedding system of numbering. The techniques used special numbering manners by spreading the numbered original data as far as possible from the original locations. The schemes embedded authentication data in least significant bits (LSB). These methods have been proven to be effective since it uses an image-based approach with a numbering pattern that is robust to deletion, image splicing, and copy attacks. Results show three types of attacks in PDF documentation managed to detect tamper location after being tampered.

**Keywords:** PDF watermarking; text watermarking; delete attack; image splicing attack; copy attack.

## 1. Introduction

Nowadays documented information has been widely used in most formal and official business. It helps a lot in various sectors such as study institutions, companies, banks, and business transactions which need a configured form to be accessed online especially in the emailing process. The provided information in every documentation required a concrete privacy and confidentiality since it may lead to interruption from intruders who want to eliminate the original issued information. Back in the past years, we all knew that scammers had freely ruined

many native documents, most focusing on document falsification. Hence, research on text security has recently been active.

Due to the problems, watermarking has been one of the best methods in order to keep the data safe from any malicious attacks. This method would assist in documenting authenticity and recovery.

The generic watermarking process includes embedding and extracting operations. Embedding process of digital watermarking is known when the insertion of watermark is keyed into the host signal followed by an algorithm while the extraction process in digital watermarking occurs when watermark is recovered from the watermarked signal using the host signal and the secret key. The classification of the embedding and extracting process can be divided into spatial domain and transform domain [Liew and Zain (2009)]. The method of embedding in the spatial domain never discovered a bad distortion towards the covered image. However, compared to transform domain, the robustness is low besides it rarely survives from attacks. Transform domain can overcome possible compression and is more robust against geometric transformation such as rotation, scaling, translation and cropping. The disadvantage of the transform domain is the time consumed for computation would be longer [Song et al. (2010)].

The advantages of fragile watermarking which usually aim to be exterminated and become undetected after an image had successfully modified in any mode, we can assume that image is safe from any tamper if fragile watermark is detected correctly due to the watermark embedding in an image [Zain and Clarke (2005)]. The appropriateness of detection in fragile watermarking could be identified based on free of alteration or tampering in an image due to the watermark embedding [Liew (2011)].

Therefore, a focus on tamper detection by fragile watermarking is proposed. This scheme is an intensification which has been previously conducted on medical images [Hisham (2016)]. The watermark embedding process is applied using the domain of spatial which is the Least Significant Bits (LSB) of image pixels. The LSB is identified to be imperceptible since it is safe from any significant distortions on original images [Hisham et al. (2013)]. Parity check approach and comparisons between average intensities approach are used for authentication. Here we introduced a digital watermarking scheme which is more effective in recovery bits distribution for all characters of text documents in PDF format.

## 2. Materials and Method

The temptation starts after undergoing a watermarked process and being ready to be published or used in official documentation and business. The attacks usually will be deletion attacks, copy move forgery/ image splicing attacks, insertion attacks, combine attacks, and copy attacks. The experiment will be focused on online PDF documents and scanned PDF documents.

### 2.1. Conversion Process

Before starting the watermark embedding, the PDF document should firstly be converted into image format which is BMP. The software used in conversion is MATLAB software. The reason for choosing this software is because it would save the file with protected original data.

### 2.2. Embedding Process

The process started by numbering in a zigzag pattern. The zigzag pattern can be seen in Figure 1. The process started at the upper center of the document since most important and privacy information was provided at the section. The equation used in mapped each of the blocks is;

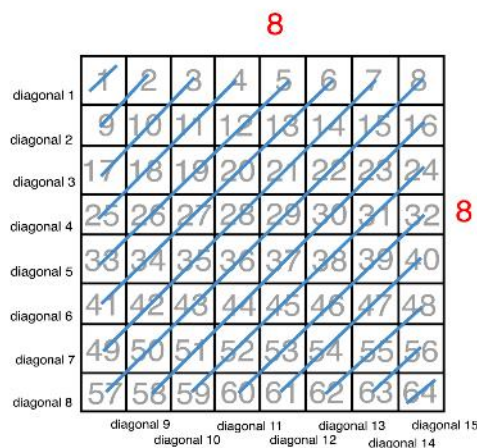


Fig. 1. Diagonals of Zigzag SCAN.

$$B = [(k \times o) \bmod Nb] + 1 \quad (1)$$

B: watermarked block  
Nb: Number blocks  
k: secret key

$$o = s(A) \quad (2)$$

Where o refers to; (zigzag SCAN formula)  
A: array of number blocks  
s: polar coordinates

Referring to the equation above, B is the watermarked block, Nb is the number of blocks, k is the secret key, while o and s are the polar coordinates with real numbers. The processed of cover image should be in diagonal size since it starts from the left in order to mapped all blocks. The algorithm below explains how zigzag-SCAN numbering manner occurred in image blocks [Afifah and Zain (2007)].

- (1) Set the block size, C. In this simulation, we set it as  $8 \times 8$ . Calculate blocks per row and blocks per column based on the width and the height of the image, referring to equation (1).
- (2) Set the key number, k.
- (3) Calculate the coordinate for each block to get the number of columns and the number of rows.
- (4) Starting point is selected. The zigzag SCAN path is set as in Figure 1 below.
- (5) After all image blocks are numbered in zigzag SCAN, each block is mapped using equation 1.

Diagonal 1: 1  
Diagonal 2: 2, 9  
Diagonal 3: 3, 10, 17  
Diagonal 4: 4, 11, 18, 25  
Diagonal 5: 5, 12, 19, 26, 33  
Diagonal 6: 6, 13, 20, 27, 34, 41  
Diagonal 7: 7, 14, 21, 28, 35, 42, 49  
Diagonal 8: 8, 15, 22, 29, 36, 43, 50, 57  
Diagonal 9: 16, 23, 30, 37, 44, 51, 58  
Diagonal 10: 24, 31, 38, 45, 52, 59  
Diagonal 11: 32, 39, 46, 53, 60  
Diagonal 12: 40, 47, 54, 61  
Diagonal 13: 48, 55, 62  
Diagonal 14: 56, 63  
Diagonal 15: 64

The following algorithm describes how the 3-tuple watermark of each sub-block was generated and embedded, as adapted and improved from [Hisham (2016)] \*Original image set as C.

- (6) Set the LSB of each pixel within the block of C to zero.
- (7) Calculate the average intensity of the block, AvgC and each of its sub-blocks, AvgCs, respectively.
- (8) Generate the authentication watermark, v, of each sub-block. V is 1 if the AvgCs is bigger than AvgC or 0 if otherwise [Zain and Fauzi (2006)].

$v = 0$  if  $AvgCs < AvgC$   
 $v = 1$  if  $AvgCs > AvgC$

- (9) Generate the parity check bit, p, of each sub-block. P is 1 if the parity number is odd, and 0 if otherwise.

$p = 0$  if parity number is even  
 $p = 1$  if parity number is odd

The watermarked data of an MRI image can be seen in Figure 2. The embedded data consists of v, p, and r of the watermark data.

### 2.3. Detection Process

In the detection process, two characteristics that need to be measured are parity bits and intensity average comparisons. The document needs to be firstly embedded directly onto the blocks in order to localize the tampering area. When blocks are being tampered locally, the changes in pixel intensities and average intensities can be seen. The exactness of the detection can be measured by intensity comparisons if parity bits fail to detect. There were two levels in this process;

#### Level 1

Perform the following procedures for each  $4 \times 4$ -pixel sub-block  $Z_s$  within the block  $B_r$ :

- (1) Firstly,  $v$  and  $p$  should be extracted from  $Z_s$ .
- (2) Calculate the average intensity for each sub-block  $Z_s$ , represented as  $\text{avg } Z_s$ , by setting the LSBs of each pixel within each  $Z_s$  to zero.
- (3) If  $\text{avg } Z_s \geq \text{avg } Z$ , put the algebraic relation  $v'=1$ ; else, set it to 0.  $P_s$  represents the total number of 1s in  $\text{avg } Z_s$ .
- (4)  $Z_s$ ' parity check bit,  $p'$ , should be set to 1. Set it to 0 if  $P_s$  is even.
- (5)  $P'$  should be compared to  $p$ , and  $v'$  should be compared to  $v$ . Label  $Z_s$  as tampered and complete the detection for  $Z_s$  if they are not equal; else, mark it as valid.

#### Level 2

- (1) Determine the block number of block  $C$ , which includes the intensity feature of block  $Z$ .
- (2) Locate Block  $C$ .
- (3) Assume block  $Z$  is valid and complete the test if block  $C$  is declared tampered.
- (4) Execute the following steps if block  $C$  is correct:
- (5) Extract the LSBs from each pixel in the associated block within block  $C$  to get the 7-bit intensity of each  $Z_s$ , padding one zero to the end to get an 8-bit result.
- (6) If the values are different from  $\text{avg } Z_s$ , label  $Z$  as tampered.

### 3. Results and Discussion

The tested samples are focusing on RGB type, grayscale type, and bitmap format. There were PDF documents which were basically the size of A4. Referring to the algorithm, the embedding scheme with zigzag-SCAN pattern produced good results in terms of elapsed times, PSNR values, tamper blocks, and detected tampered regions in the documents.

PSNR is explained as one of the metrics in order to identify deterioration of embedded documents between original documents [Hisham (2016)]. Based on the results of the tested document, most of all documents had achieved up to 55 dB of PSNR. Values which are over 36 dB in PSNR are considered to be acceptable in terms of deterioration [Huynh-Thu and Ghanbari (2008)].

Accuracy is another metric to compare the exactness of detection which refers to the noise or blue dots that are viewed in the documents. The less noise/ blue dots, the higher the accuracy can be achieved. In this observation, three blue dots will be considered as 1 percent (%) of inaccuracy. The accuracy cannot be measured on the whole one document since the sizes are large.

The document size is included as one of the metrics in order to compare whether the size of a document will affect results in accuracy or not. The size of the tested documents is collected after PDF format has been converted into BMP format. We can assume that the larger size of the document can deduct the accuracy percentage due to greater pixels involved.

Below shows the online document which is built up in pixels. The embedded documents have an average elapsed time (sec.) which is in between 40-45 minutes. Based on the results, the PSNR value is considered to be acceptable since all types of documents below managed to achieve more than 55 dB.

#### 3.1. Online Document

Table 1 below shows the online document which is built up in pixels. The embedded documents have an average elapsed time (sec.) which is in between 39- 60 minutes. Based on the results, the PSNR value is considered to be acceptable since all types of documents below managed to achieve more than 55 dB while the average elapsed time is measured as not a computational time saver due to the large amount of time consumed to be completed. Table 1 shows the comparisons between original online documents and watermarked online documents in BMP format.

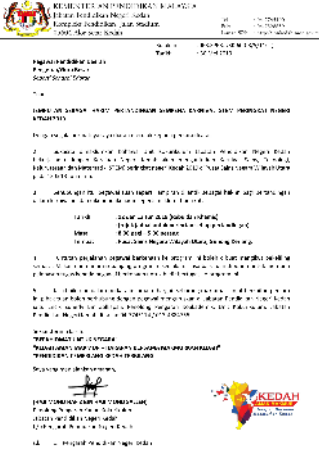

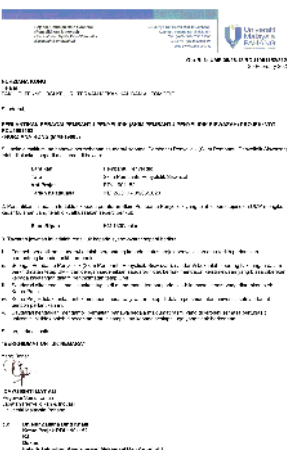
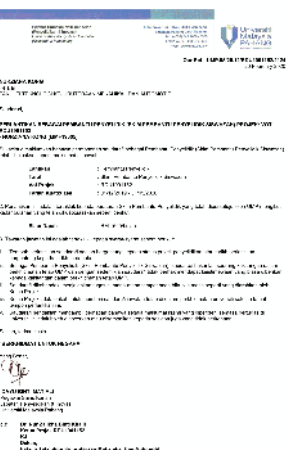
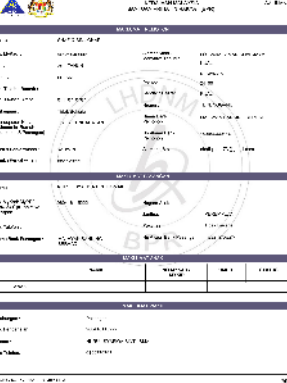
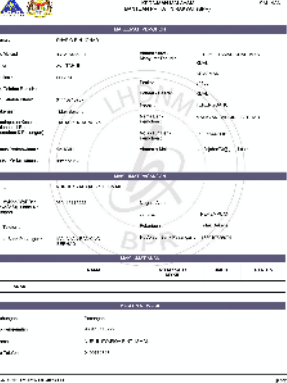
Original document	Watermarked document	Document Size (MB)	PSNR (dB)
<p><b>Invitation Letter</b></p> 		24.9 MB	62.5466
<p><b>Offer Letter</b></p> 		24.8 MB	62.5420
<p><b>Application Form</b></p> 		24.0 MB	62.0113

Table 1. The comparisons between original online documents and watermarked online documents in BMP format.

Below shows the graph of document size (MB) for each three types of document versus PSNR value. Referring to Figure 2, we can estimate that PSNR values are acceptable which are above than 55dB.

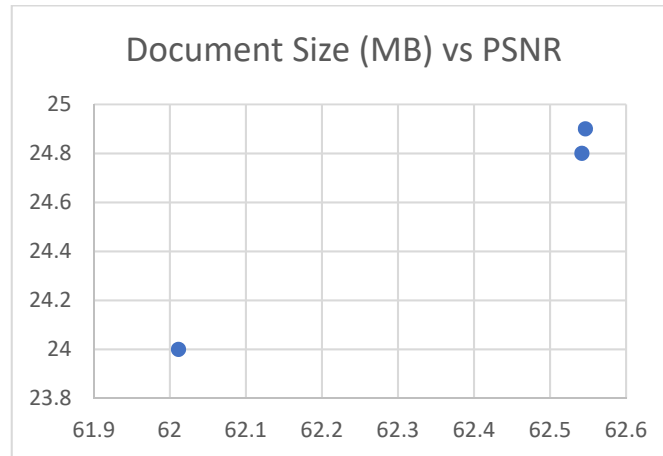


Fig. 2. Document size against PSNR values.

Table 2 shows the three (3) types of online documents which are usually used in formal dealing/ business. Referring to the table below, all attacks are able to be detected by watermarking scheme in BMP format. Table 2 shows the results and comparison between three (3) types of attacks which usually used in formal dealing/ business with three (3) types of different documents.

Sample Data/ Attacks	Invitation Letter	Offer Letter	Application Letter
Deletion attack	√	√	√
Image splicing/ Replacement attack	√	√	√
Copy attack	√	√	√

Table 2. The comparisons between three (3) types of attacks with three (3) types of different documents.

Figure 3 below shows the graph of document size (MB) against embedding elapsed time (sec.). As we can see the larger the size of the document, the more time taken to complete the embedding process.

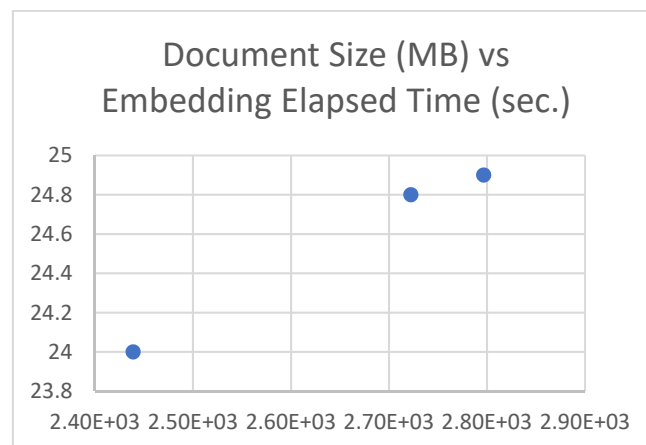


Fig. 3. Document Size (MB) vs Elapsed Time (sec.) for Embedding.

The graph in Figure 4 below shows the document size against detecting elapsed time (sec.). As we can see, the time taken for each document is not too far but consumes a large elapsed time to complete the detection process.

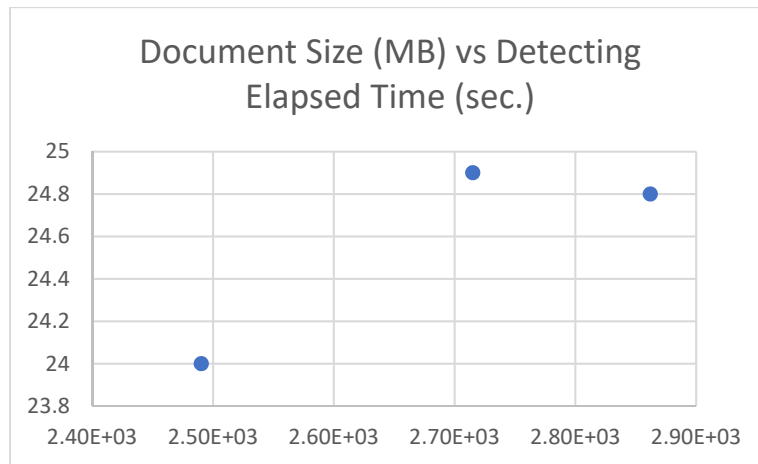


Fig. 4. Document size (MB) against detecting elapsed time (sec.).

### 3.1.1. Deletion Attack

The delete attack usually happens in most document scamming where the original content is being deleted. For example, intruders try to blank the original info in a certificate or offer letter, and it may continue with a combine attack. The delete attack usually happens in most document scamming where the original content is being deleted. For example, intruders try to blank the original info in a certificate or offer letter, and it may continue with a combine attack.

The Table 3 below compares accuracy and document size which range from 20 MB to 24 MB. The documents have different sizes to each other while the accuracy is considered to be good since all of the results show more than 80 percent (%). Results of the deletion attack in BMP Zigzag-LSB viewed the blue region with few noises around. Noises would not be the main issues since every region has different size of cropping due to the big size of file format in the document. Table 3 shows the results of deletion attack for online document in BMP format.

Tampered Image	Tampered Detection	Accuracy	Document Size (MB)
<p>Invitation Letters\-</p>		91%	24.9 MB
<p>Offer Letter</p>		83%	24.8 MB
<p>Application Form</p>		94%	24.0 MB

Table 3. The results of deletion attack for online documents in BMP format.

The graph in Figure 5 shows the size of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of deletion attacks on the documents has high accuracy.

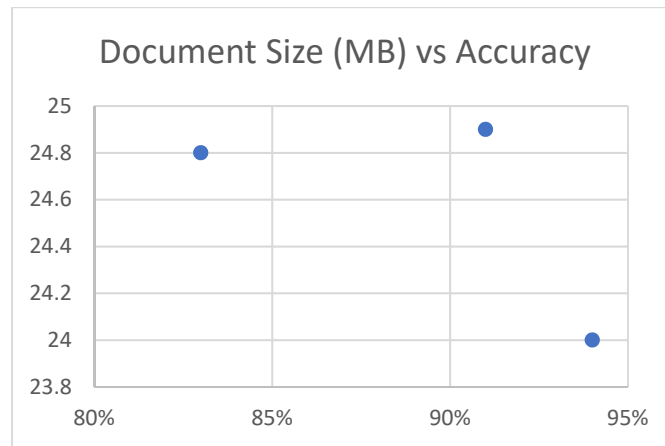


Fig. 5. Document Size (MB) vs Accuracy.

### 3.1.2. Image Splicing/ Insertion Attack

Image splicing is one of replacement attacks where another source is replaced or added into the document. As an example, intruders try to change the original logo to a fake one. Table 4 below compares accuracy and document size which range from 20 MB to 24 MB. The documents have different sizes to each other but the accuracy results have a closed range in percentage. Based on the results, we can conclude that online documents of BMP format in the Zigzag-LSB watermarking scheme are able to detect image splicing/ insertion attacks based on their high accuracy which are above 85% since they are able to fully spot the logo. Blue dots around are noises where it would not be the main issue since every region has different size of cropping due to the big size of file format in the document. Table 4 shows the results of image splicing/ insertion attack for online document in BMP format.

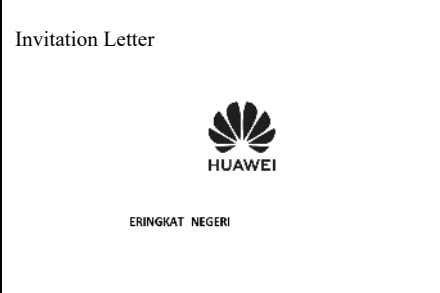



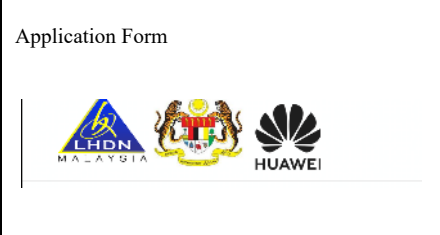

Tampered Image	Tampered Detection	Accuracy	Document Size (MB)
<p>Invitation Letter</p> 		92%	24.9 MB
<p>Offer Letter</p> 		88%	24.8 MB
<p>Application Form</p> 		93%	24.0 MB

Table 4. The results of image splicing/ insertion attack for online documents in BMP format.



The graph in Figure 6 shows the size of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of image splicing/ insertion attacks on the documents has high accuracy.

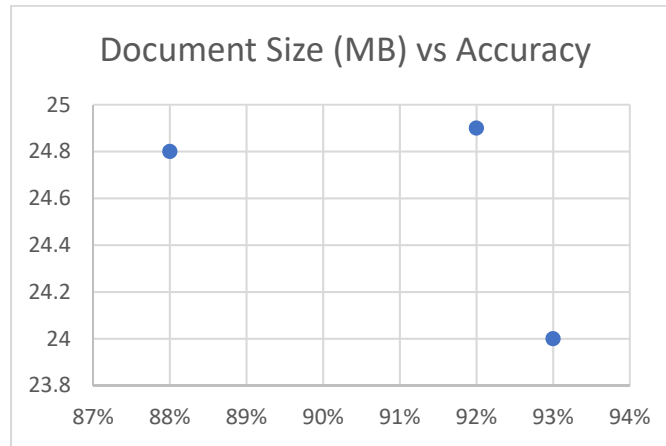




Fig. 6. Document Size (MB) vs Accuracy.

### 3.1.3. Copy Attack

The test for copy attack here is by copying the whole document and pasting it into another new document. Table 5 below compares accuracy and document size which range from 20 MB to 24 MB. The documents have different sizes to each other but the accuracy results have a closed range in percentage (%). Based on the results, we can conclude that online documents of BMP format in the Zigzag-LSB watermarking scheme are able to detect copy attacks based on their high accuracy which are more than 90% due to little noises and considered to be fully detected since it localizes the whole document. However, the elapsed time taken can be categorized as not a computational time-saver due to the large amount of time consumed to complete a document. Table 5 shows the results of copy attack for online document in BMP format. combine attack is a combination between delete and insert tamper in the document.

Tampered Image	Tampered Detection	Accuracy	Elapsed Time (sec.)
<p>Invitation Letter</p> 		93%	6.0691e+03
<p>Offer Letter</p>		93%	2.7029e+03



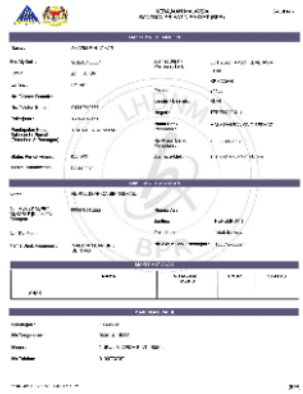
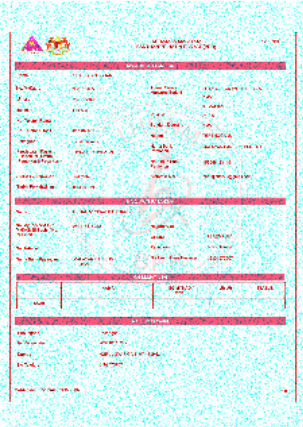
			
<p>Application Form</p> 		<p>95%</p>	<p>2.7269e+03</p>

Table 5. The results of copy attack for online documents in BMP format.

The graph in Figure 7 shows the size of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of copy attacks on the documents has high accuracy.

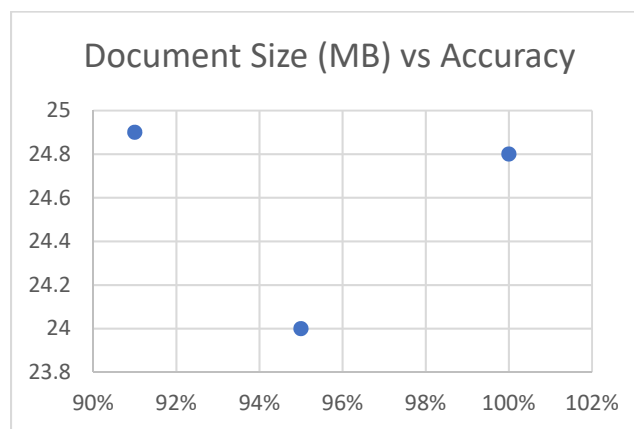
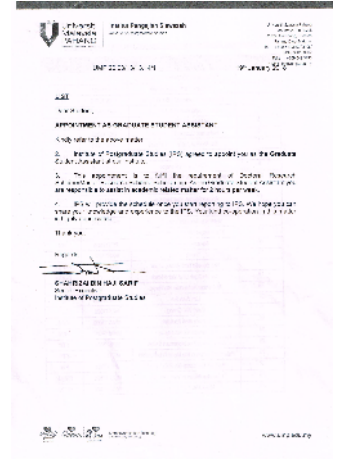



Fig. 7. Document Size (MB) vs Accuracy.

### 3.2. Scanned Document

Table 6 below shows the scanned document which is built up in vectors. The embedded documents have an average elapsed time (sec.) which is in between 53- 1 hours. Based on the results, the PSNR value is considered to be acceptable since all types of documents below managed to achieve more than 55 dB and the average elapsed time is measured as not a computational time-saver due to more than 42 minutes to be completed. Table 6 shows the comparisons between original scanned documents and watermarked scanned documents in BMP format.

Original document	Watermarked document	Document Size (MB)	PSNR (dB)
<p>Certificate</p> 		23.6 MB	61.5944
<p>Appointment Letter</p> 		23.6 MB	61.6900
<p>Offer Letter</p>		24.8 MB	64.1322

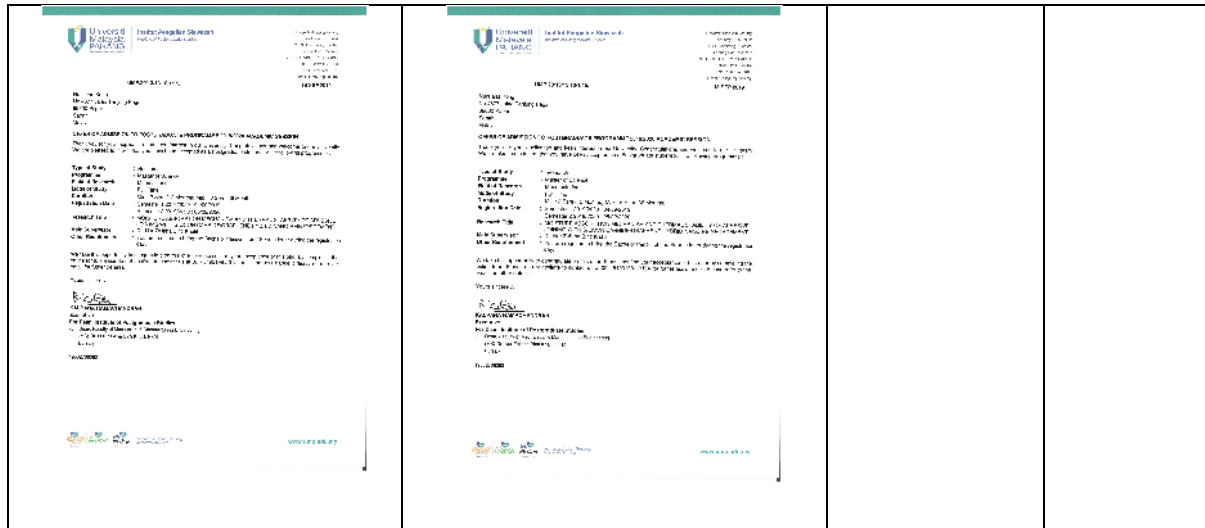


Table 6. The comparisons between original scanned documents and watermarked scanned documents in BMP format.

The graph in Figure 8 below shows the document size (MB) for each three types of documents versus PSNR value. Referring to the graph, we can estimate that PSNR values are acceptable which are above than 55dB.

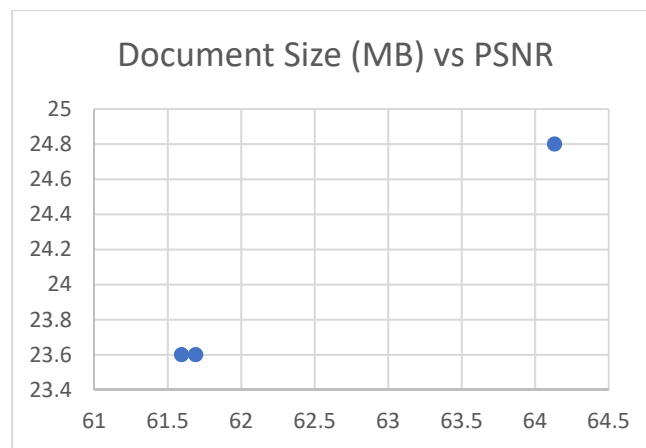


Fig. 8. Document size against PSNR values.

Table 7 shows the three (3) types of scanned document files which are usually used in formal dealing/business with possible attacks. Referring to the table below, all attacks are able to be detected by watermarking scheme in BMP format. Table 7 shows the results and comparison between three (3) types of attacks which usually aimed in formal dealing/ business.

Sample Data/ Attacks	Certificate	Appointment Letter	Offer Letter
Deletion attack	√	√	√
Image splicing/ Replacement attack	√	√	√
Copy attack	√	√	√

Table 7. The comparisons between original scanned documents and watermarked scanned documents in BMP format.

Table 8 shows the average of elapsed time of online documents which are in between 40 to 50 minutes. Hence, the online documents are measured as not a computational time saver due to their ability of completing processes which is in the range of 40- 50 minutes per document. Table 8 shows the results of computation times between three (3) types of scanned document files which usually used in formal dealing/ business.

Computation Time (sec.)	Certificate	Appointment Letter	Offer Letter
Computing Time (elapsed time, sec.) for embedding	3.5788e+03	3.2840e+03	3.1959e+03
Computing Time (elapsed time, sec.) for detection	3.5673e+03	2.6243e+03	2.6982e+03

Table 8. The results of computation times between three (3) types of scanned document files in BMP format.

The graph in Figure 9 below shows the document size against embedding elapsed time (sec.). As we can see, the elapsed time taken are in the same range, and since the size of the documents are quite large, the more time needed to complete the process.

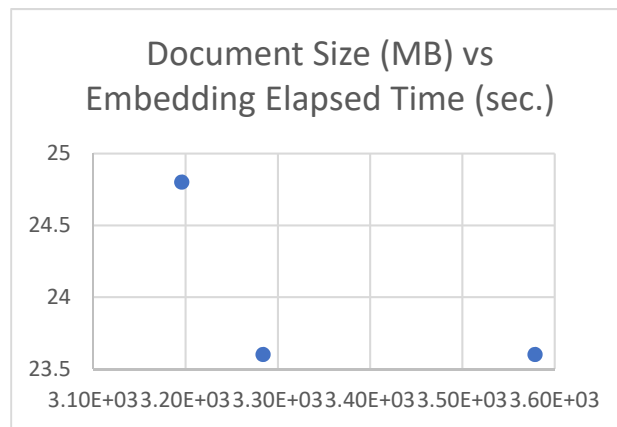


Fig. 9. Document Size (MB) vs Elapsed Time (sec.) for Embedding.

Figure 10 below shows the document size against detecting elapsed time (sec.). As we can see, the elapsed time taken are in the same range, and since the size of the documents are quite large, the more time needed to complete the process.

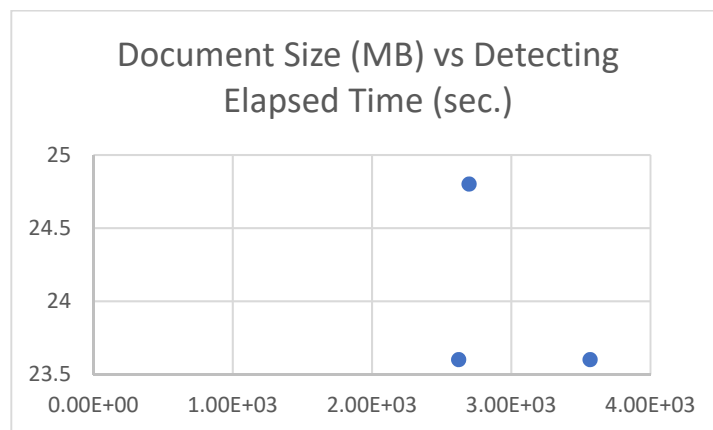


Fig. 10. Document size (MB) against detecting elapsed time (sec.).

### 3.2.1. Deletion Attack

The delete attack usually happens in most document scamming where the original content is being deleted. For example, intruders try to blank the original info in a certificate or offer letter, and it may continue with a combine attack. Table 9 below compares accuracy and document size which range from 23 MB to 24 MB. The documents have different sizes to each other while the accuracy is considered to be good since all of the results show more than 74 percent (%). Results of the deletion attack in BMP Zigzag-LSB viewed the blue region with few noises

around. Noises would not be the main issues since every region has different size of cropping due to the big size of file format in the document. Table 9 shows the results of deletion attack for scanned document in BMP format.

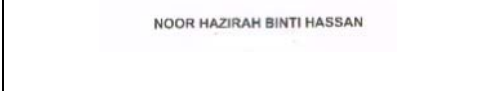

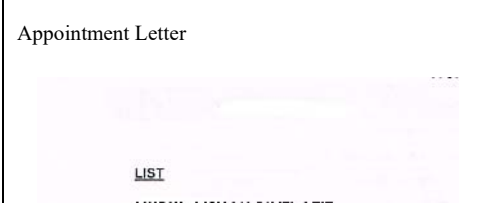


Tampered Image	Tampered Detection	Accuracy	Document Size (MB)
<p>Certificate</p> 		75%	23.6 MB
<p>Appointment Letter</p> 		89%	23.6 MB
<p>Offer Letter</p> <p><b>Type of Study</b> : Research <b>Programme</b> : Master of Science <b>Field of Research</b> : Manufacturing</p> <p><b>Duration</b> : Min : 2 Sem (12 Months) <b>Registration Date</b> : Semester 1 2019/2020 Semester 2 2019/2020</p>		100%	24.8 MB

Table 9. The results of deletion attack for scanned documents in BMP format.

The graph in Figure 11 shows the size of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of deletion attacks on the documents has high accuracy.

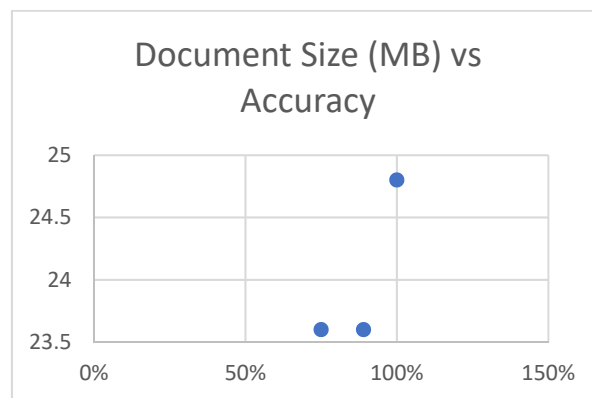


Fig. 11. Document Size (MB) vs Accuracy.

### 3.2.2. Image splicing/ insertion attack

Image splicing is one of replacement attacks where another source is replaced or added into the document. As an example, intruders try to change the original logo to a fake one. Table 10 below compares accuracy and document size which range from 20 MB to 24 MB. The documents have different sizes to each other but the accuracy results have a closed range in percentage (%). Based on the results, we can conclude that scanned documents of BMP format in Zigzag-LSB watermarking scheme are able to detect image splicing/ insertion attack based on their high

accuracy which are above 88% since they are able to fully spot the logo. Blue dots around are noises where it would not be the main issue since every region has different size of cropping due to the big size of file format in the document. Table 10 shows the results of image splicing/ insertion attack for scanned document in BMP format.





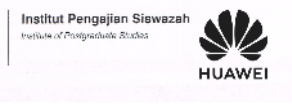

Tampered Image	Tampered Detection	Accuracy	Document Size (MB)
<p>Certificate</p> 		89%	23.6 MB
<p>Appointment Letter</p> 		92%	23.6 MB
<p>Offer Letter</p> 		95%	24.8 MB

Table 10. The results of image splicing/ insertion attack for scanned documents in BMP format.

The graph in Figure 12 shows the size of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of image splicing/ insertion attacks on the documents has high accuracy.

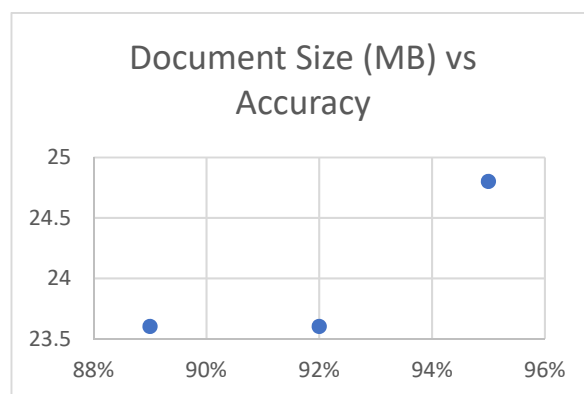


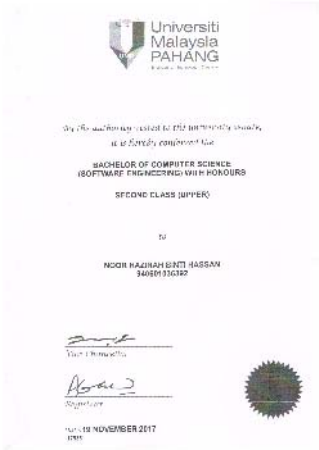


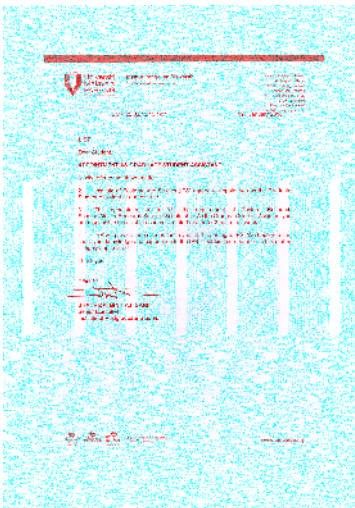
Fig. 12. Document Size (MB) vs Accuracy.

### 3.2.3. Copy Attack

The test for copy attack here is by copying the whole document and pasting it into another new document. Table 11 below is comparing accuracy and document size which range from 20 MB to 24 MB. The documents have different sizes to each other but the accuracy results have a close range in percentage (%). Based on the results,



we can conclude that scanned documents of BMP format in Zigzag-LSB watermarking scheme are able to detect copy attacks based on their high accuracy which are more than 94% due to little noise and considered to be fully detected since it localizes the whole document. However, the elapsed time taken can be categorized as not a computational time-saver due to the large amount of time consumed to complete a document. Table 11 shows the results of copy attack for scanned documents in BMP format.

Tampered Image	Tampered Detection	Accuracy	Elapsed Time (sec.)
<p>Certificate</p> 		95%	2.7323e+03
<p>Appointment Letter</p> 		95%	2.6828e+03
<p>Offer Letter</p>		95%	6.3969e+03



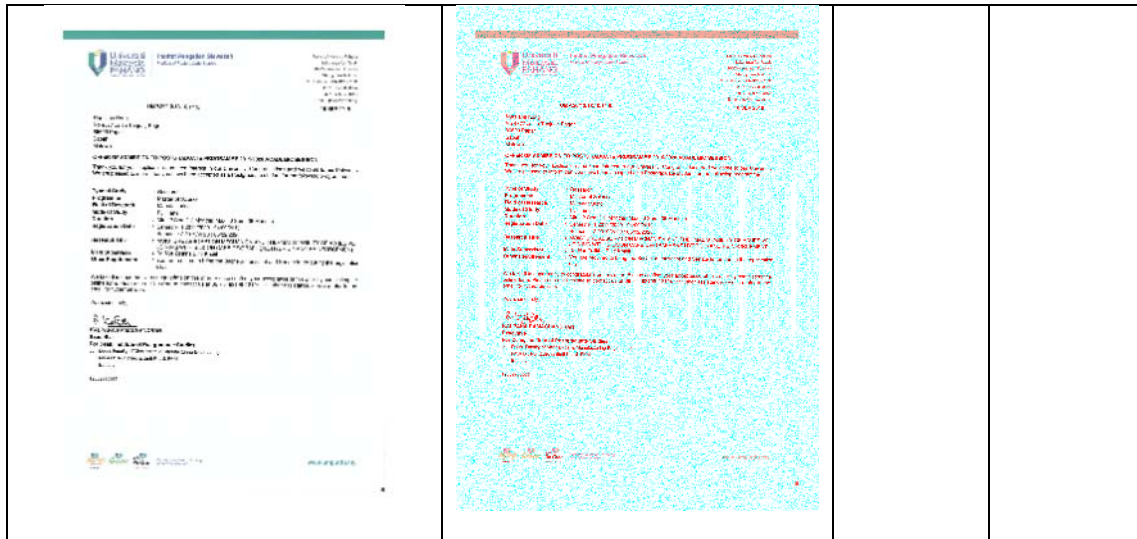


Table 11. The results of copy attack for scanned documents in BMP format.

The graph in Figure 13 shows sizes of the document (MB) against accuracy (%). Based on the graph below, we can assume that the accuracy for each of the tested documents are in the same range. The detection of copy attacks on the documents has high accuracy.

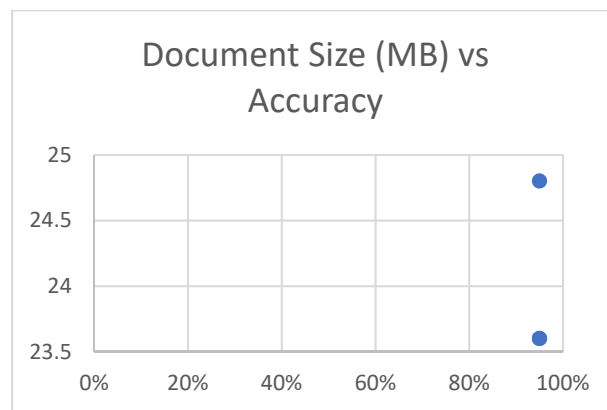


Fig. 13. Document Size (MB) vs Accuracy.

#### 4. Conclusions

The proposed scheme is to verify the authenticity of different types of documents which are usually used in most formal businesses and deals. The watermark is fragile due to the ability to detect the modifications made in the documents with different attacks. Based on the results, the scheme managed to detect all three types of attacks which have high possibility towards PDF; deletion attack, image splicing attacks, and copy attack. This method proved the ability to embed and detect the whole part of the document. Zigzag-LSB watermarking scheme proved that it has the best performance in terms of quality and user-friendly level.

#### 5. Future Research

Even though overall results are good, in terms of future research, the study will be focused on the computational time which needs to be improved due to the longer time taken to be done.

#### 6. Acknowledgement

This research works is supported by an FRGS-RACER grant entitled ‘A New Embedding Algorithm using Hilbert-Peano Pattern in Enhancing Authentication Systems for Textual Documents (RDU192623)’ RACER/1/2019/ICT04/UMP//2 supported by the Ministry of Higher Education, and a PGRS grant (PGRS200369) supported by Universiti Malaysia Pahang.

## 7. Conflict of Interest

The authors declare that there is no conflict of interest in this paper.

## References

- [1] Afifah, N.; Zain, J. (2007): Using spiral scan technique for medical image watermarking with tamper detection and recovery, in: Proc. in National Conference on Software Engineering and Computer System, pp. 20–21.
- [2] Dikanev, P.; Vybornova, Y. (2021, September): Method for Protection of PDF Documents against Counterfeiting using Semi-Fragile Watermarking. In 2021 International Conference on Information Technology and Nanotechnology (ITNT) (pp. 1-4). IEEE.
- [3] Hisham, S. I. (2013): A Quick Glance at Digital Watermarking in Medical Images, Biomedical Engineering Research Journal Vol. 2, Issue 2, pp. 79-87.
- [4] Hisham, S. I. (2016): Enhanced LSB Watermarking Methods based on Scanning Patterns for Authentication of Medical Images, ph.d. thesis, Universiti Malaysia Pahang.
- [5] Huynh-Thu, Q.; Ghanbari, M. (2008): Scope of validity of PSNR in image/video quality assessment. Electronics letters, 44(13), 800-801.
- [6] 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), pp. 20–22.
- [7] Liew, S.C. (2011): Tamper Localization and Recovery Watermarking Schemes for Medical Images in PACS, Doctor of Phil. Thesis, Universiti Malaysia Pahang, Malaysia.
- [8] Liew S. C.; (2013): Tamper Localization and Lossless Recovery Watermarking Scheme with ROI Segmentation and Multilevel Authentication, Journal of Digital Imaging Volume 26, Issue 2, pp 316-325.
- [9] Liew, S. C.; (2013): Tamper localization and loss- less recovery watermarking scheme with roi segmentation and multilevel authentication, in: Journal of digital imaging, pp. 26(2), 316–325.
- [10] Saini, L. K.; Shrivasta, V. (2014): A survey of Digital Watermarking Techniques and its Applications, International journal of Computer Science Trends and Technology (IJCST). Volume 2, Issue (3).
- [11] Song, C.; (2010): Analysis of digital image watermark attacks, in: 2010 7th IEEE Consumer Communications and Networking Conference, IEEE, pp. 1–5.
- [12] Zain, J.; Clarke, M. (2005). Security in telemedicine: issues in watermarking medical images, Sciences of Electronic, Technologies of Information and Telecommunications, Tunisia.
- [13] Zain, J. M.; Fauzi, A. R. (2006, September) Medical image watermarking with tamper detection and recovery. In 2006 International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 3270-3273). IEEE.

## Authors Profile



### Nur Alya Afikah binti Usop

She has been graduated from Universiti Malaysia Sarawak (UNIMAS) in 2019 with bachelor degree of Computational Science. Currently furthering Master degree in Universiti Malaysia Pahang (UMP) in Research of Multimedia which focusing on image processing security. She actively involved in conferences and research presentation, she had done two reviewed research papers, ongoing paper journal and finishing her master study.



### Syifak Izhar Hisham

She believes that she has strong background in researching, academic writing and presenting in various viva, speech conference, and product competition sessions. Her research concentrations include Data Security, Digital Watermarking, Image Processing, E-Learning, Multimedia Technology, and Images Authentication. She has experiences in lecturing and teaching with 5 years experiences in various institutions and various fields.



### Jasni Mohamad Zain

Jasni Mohamad Zain received her Bachelor degree in Computer Science from University of Liverpool, England, UK in 1989; and PhD from Brunel University, West London, UK in 2005. She starts her career as a tutor in 1997 at University of Technology Malaysia (UTM). She currently holds the post as Director of Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), UiTM Shah Alam, Malaysia. She was the Deputy Director (Cybertechnology) Research Nexus UiTM and the Head of Advanced Analytics and Engineering Centre (AAEC). She was the Dean of Faculty of Computer Systems & Software Engineering, University Malaysia Pahang for 8 years. She has been actively presenting papers and keynote address in national and international conferences. Her research interests include Data Mining, digital watermarking, image processing and network security. She has graduated 15 PhDs and 6 Masters by research under her

supervision and published more than 100 refereed articles. She has a patent file for digital watermarking (PI 2008047).



**‘Aqilah Abd Ghani**

'Aqilah Abd Ghani received her Bachelor degree in Graphic and Multimedia Technology from University Malaysia Pahang, Malaysia in 2017. She started to learn official computing courses at Pahang Matriculation College (KMPh) and continue her study at University Malaysia Pahang (UMP) until now. Currently, she doing a Master of Research in Image & Signal Processing at Faculty pf Computing, UMP. Actively involved in conferences and research presentation, she had done two Scopus indexed research papers, ongoing paper journal and finishing her master study. She also aims to pursue her academic journey as a Postdoctoral, Ph.D. researcher one day.