

5-1-2022

Error Level Analysis Technique for Identifying JPEG Block Unique Signature for Digital Forensic Analysis

Nor Amira Nor Azhan
Universiti Teknologi Malaysia

Richard Adeyemi Ikuesan
Zayed University, richard.ikuesan@zu.ac.ae

Shukor Abd Razak
Universiti Teknologi Malaysia

Victor R. Kebande
Blekinge Tekniska Högskola

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Computer Sciences Commons](#)

Recommended Citation

Azhan, Nor Amira Nor; Ikuesan, Richard Adeyemi; Razak, Shukor Abd; and Kebande, Victor R., "Error Level Analysis Technique for Identifying JPEG Block Unique Signature for Digital Forensic Analysis" (2022). *All Works*. 5139.

<https://zuscholars.zu.ac.ae/works/5139>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact scholars@zu.ac.ae.

Article

Error Level Analysis Technique for Identifying JPEG Block Unique Signature for Digital Forensic Analysis

Nor Amira Nor Azhan ¹, Richard Adeyemi Ikuesan ² , Shukor Abd Razak ¹  and Victor R. Kebande ^{3,*} 

¹ School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Skudai 81310, Malaysia; noramira.norazhan@gmail.com (N.A.N.A.); shukorar@utm.my (S.A.R.)

² Department of Computing & Applied Technology, College of Technological Innovation, Zayed University, Abu Dhabi P.O. Box 144534, United Arab Emirates; richard.ikuesan@zu.ac.ae

³ Department of Computer Science (DIDA), Blekinge Institute of Technology, 37179 Karlskrona, Sweden

* Correspondence: victor.kebande@bth.se

Abstract: The popularity of unique image compression features of image files opens an interesting research analysis process, given that several digital forensics cases are related to diverse file types. Of interest has been fragmented file carving and recovery which forms a major aspect of digital forensics research on JPEG files. Whilst there exist several challenges, this paper focuses on the challenge of determining the co-existence of JPEG fragments within various file fragment types. Existing works have exhibited a high false-positive rate, therefore rendering the need for manual validation. This study develops a technique that can identify the unique signature of JPEG 8 × 8 blocks using the Error Level Analysis technique, implemented in MATLAB. The experimental result that was conducted with 21 images of JFIF format with 1008 blocks shows the efficacy of the proposed technique. Specifically, the initial results from the experiment show that JPEG 8 × 8 blocks have unique characteristics which can be leveraged for digital forensics. An investigator could, therefore, search for the unique characteristics to identify a JPEG fragment during a digital investigation process.

Keywords: digital forensics; file fragment identification; JPEG fragment; file-carving; error level analysis; JPEG signature



Citation: Azhan, N.A.N.; Ikuesan, R.A.; Razak, S.A.; Kebande, V.R. Error Level Analysis Technique for Identifying JPEG Block Unique Signature for Digital Forensic Analysis. *Electronics* **2022**, *11*, 1468. <https://doi.org/10.3390/electronics11091468>

Academic Editor: Byung Cheol Song

Received: 3 February 2022

Accepted: 29 April 2022

Published: 3 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The field of Digital Forensics (DF) has faced a lot of growth in the recent past because of increased digital crimes and diverse forensic analysis strategies that have attempted to uncover events (usually posthumously) that occur in digital media. This can only be achieved through the use of scientifically verifiable methodologies [1–4]. Discounting that, diverse attention has been shifted to digital forensic analysis, which is a growing body of research that has an intent of not only understanding the threats to digital media [5] but also to provide mitigation strategies. Furthermore, research in multimedia forensics, which is a branch of digital forensics, presents a comprehensive study of the facets of multimedia and data from an investigation perspective. In essence, multimedia forensics utilizes scientific methods that aid in the extraction of media-related facts from digital devices [6] that can be used in forensic hypothesis formation for purposes of litigation. Joint Photographic Expert Group (JPEG) is a file format that can be considered to be the most adopted multimedia data due to its compression capabilities and its possibility of being commercialized. The JPEG compression level can further be termed as the quality of the image, which is theoretically device-specific and unique for each JPEG file since different devices may produce different image quality. More specifically, the JPEG format permits the customization (manipulation) of the image quality in different instances.

The process of recovering potential digital evidence and conducting forensic analysis from media-related devices is often considered the principal focus of multimedia forensics [7–10]. Before the analysis is conducted, images need to be extracted from the storage

devices. Basically, in this context, analysis is a step towards justifying how a potential incident could have happened. This allows one to give a detailed description of each context, including links and relationships, and explanations [11–15]. Precisely, this helps to build strong arguments for multimedia forensics given that files can be deleted by suspects. In some specific situations, it may force files to be fragmented [16], where a special technique is often required to recover such media evidence, JPEG for instance. Generally, two available methods considered in the recovery process include file recovery and file carving. The traditional approach of file recovery works based on the information provided by the file system, while file carving works using the internal structure of the file to be recovered [17]. In such cases, a priori knowledge of the file type is usually known either through the file signature, header information, or other metadata on the media. These methods are contingent on the assumption that the file structure is contiguously stored in the allotted disk space. However, such an assumption does not always hold for JPEG files, as it permits fragmentation and a lossy compression process.

Therefore, it is challenging to deal with fragmented files in contrast to contiguous files stored in disk blocks [18]. Thus, a digital forensic technique of file carving is used to “carve” (copy) [19] bytes of data from the disk image, regardless of the type of file system. Existing studies have mainly focused on developing and enhancing an automated carver while some are focused on separating the process using a SmartCarver [20]. With an automated carver, the carving process is conducted directly through raw disk images by applying traditional carving which analyzes the JPEG header, footer, and some relevant markers. An example of the JPEG marker that is mainly used in carving includes the application-specific APP data marker, which is also referred to as thumbnail image/s [21–24]. The data or blocks between fragments that possess such characteristics will then be examined. Moreover, the main issue arises when non-JPEG blocks are decoded as JPEG fragments, which can be seen among the earliest carving approach called Bifragment Gap Carving (BGC) [18]. A Sequential Hypotheses Testing [25] is further proposed to enhance BGC by detecting fragmentation points. However, forward and reverse testing is needed if the carving result is found inconclusive. To avoid false identification of JPEG fragments, existing automated carving processes would specifically include a JPEG validator that is capable of minimizing the rate of false-positive carvings. Generally, a libjpeg decoder is used to validate the JPEG file, which further requires manual validation [18,23,26–29].

Nonetheless, the SmartCarver technique has proven to overcome the challenges of fragmented files by separating the carving process with three major steps: Pre-processing, collating, and reassembling. Existing works based on the second step of the SmartCarver technique (collation), perform poorly in accurately identifying compressed JPEG fragments, which also affects the fragmentation reassembly process. In a collation step, all unknown fragments are examined as to whether they can classify JPEG fragments. Therefore, a group of JPEG fragments (can be more than one JPEG file) can be prepared for further reassembling process. This can be performed using file fragment classification, to identify the known and unknown data fragments. Three categories of classification have been summarized for JPEG fragments classification: The signature-based approach, statistical approach, and Artificial Intelligence (AI) approach [30]. While the signature-based approach is based on a JPEG header, footer, and markers, a statistical approach is based on entropy information and byte frequency analysis. The results of JPEG classification are largely unsuccessful through the use of an AI approach, as different approaches use different datasets for training and testing. The existing classification of common file types has shown a high degree of accuracy, however, inconsistent whenever fragments contain compressed data [31]. A false-positive rate of 86% was observed when JPEG fragments are classified among 10 different file types [32], while a 96.8% rate was observed within 23 file types [33]. A study in [16] asserts that a detailed study and the analysis of a JPEG internal file structure could help in reducing the false-positive rate. To address the low accuracy rate of JPEG fragment identification, and the high false-positive rate, the current study thus attempts to define an alternative method for the identification of JPEG fragments by applying the error level

analysis to highlight the unique signature of JPEG 8×8 blocks. Thus, the scope of this study is limited to an approach that attempts to provide a baseline for accurate identification and detection, in line with the assertions from [16,18]. In attempting to accurately carve a JPEG file among a series of compressed JPEG files, a forensic examiner could consider the approach presented in this study as a step towards a reliable carving process. In this regard, this is the first study, to the knowledge of the authors, that provides a reliable basis for the identification of compressed JPEG file fragments using an inherent marker that is capable of further revealing JPEG file modification.

The remainder of the paper is organized as follows: background and related works are discussed in Section 2 while the proposed error level analysis technique for identifying JPEG block unique signature for purposes of forensic analysis is discussed in Section 3. Thereafter, results and discussions are discussed in Section 4 with a conclusion and a mention of future work given in Section 5.

2. Background and Related Literature

2.1. Existing Literature

Among all multimedia file types, the JPEG file format is the most common file type encountered during the investigation. This can be attributed to its unique compression format as widely adopted in digital cameras [34]. While files, such as Microsoft Word and other text-based files, can be easily carved based on their content, JPEG is relatively difficult due to its compression technique. It requires specific skills and techniques to deal with this type of file. Generally, three major challenges to JPEG image file forensics are issues of file headers, carving file fragments (as well as the reconstruction of a complete file), and the handling of unknown files [35]. Basic carving methods deal with non-fragmented files while advanced carving focuses more on fragmentation issues. A synopsis of the two concepts of the file carving technique is presented in Table 1.

Table 1. Carving concepts [16].

Basic Carving	Advanced Carving
<ul style="list-style-type: none"> • The file is not fragmented. • The file is not compressed. • The beginning of the file is not overwritten. 	<ul style="list-style-type: none"> • The file is sequentially fragmented. • The file is non-sequentially fragmented (out of order). • Missing fragments.

Current studies emphasize the fragmentation scenario which can also be applied to a non-fragmented situation. Existing carving methods make use of header and footer to carve JPEG files and fragments [17]. In this way, the starting point and ending point of the file fragments can be identified easily before proceeding to handle the remainder of the JPEG content. Furthermore, JPEG fragments are identified based on JPEG markers, analysis of byte frequency [36], and based on fragments with 'FF 00' byte.

Thumbnail images are useful in the carving process and also have a specific JPEG marker. Several carving methods make use of thumbnail images to carve JPEG files or fragments. An automated tool called myKarve [21] is developed to carve JPEGs and their thumbnails when the JPEGs are linearly fragmented. The limitation is that the tool can only identify non-JPEG fragments, such as PDF, Word, and Excel files. Further improvement was made on myKarve with a tool called PattrecCarve [22]. The PattrecCarve tool performs carving based on hex patterns and was subsequently developed to successfully carve thumbnail/s or embedded JPEG files stored in contiguous fragments. A more recent work scaled up the thumbnail image to validate JPEG fragments [23]. However, the presence of thumbnail image/s that are still intact in a disk will only make the aforementioned methods useful in JPEG carving.

Other carving techniques use object validation as one step to carving JPEG fragments. Object validation works by determining whether the given bytes belong to a known JPEG content [18]. Thus, all JPEG markers are used as references to validate files (or fragments).

The earliest work was applied in [18] where its approach is limited to only two fragments. Generally, the first step of this approach is searching for the header and footer of JPEG files. Then, the data between those markers are validated using libjpeg, a modified version of the JPEG decompressor. The decompressor decompresses or decodes any available data based on the first fragment containing the JPEG header. A false-positive result will require manual validation from the user. Our proposed method attempts to overcome the manual validation process since a compressed file, such as a PDF, can also be validated as JPEG data by using an object validator. Other validators, DERRIC [27], and EXCAVATOR [28], are developed to meet high-quality data analysis tools for data carving. In those methods, the carving is performed by adapting the method in [18] where files are usually fragmented into two fragments.

Usually, unknown files are handled by dividing fragments based on disk block size and a machine learning algorithm is used to classify fragments [35]. All possible disk block sizes of 256 B, 512 B, 1 kB, 2 kB, 4 kB, 8 kB and 16 kB are tested [37]. Furthermore, a fragment size of 100 B has also been used [38]. Machine learning approaches that have been adapted include k-means, Fisher’s Linear Discriminant, Support Vector Machine, Artificial Neural Network, and Decision Tree. Several statistical features are extracted from fragments to determine the characteristics of each fragment, such as average, standard deviation, Kurtosis, and entropy. However, to determine the best approach, different compositions and the size of datasets are used for the training and testing phase [39]. Generally, the false-positive rate can also affect the classification accuracy, especially for high entropy file fragments, such as ZIP and PDF files [33]. One fundamental assumption often associated with the file carving is that the file is not corrupted, and it has not been over-written in the digital media. These assumptions also hold in this study.

2.2. JPEG Compression

Image compression aims to reduce the amount of image file data to be stored and transmitted [40]. It can be in the form of a lossless or lossy compression technique. The lossless compression technique does not remove any data from the original file while the lossy compression technique discards information to achieve a better compression quality. JPEG uses a lossy image compression method. Figure 1 shows how JPEG compression and decompression can be performed.

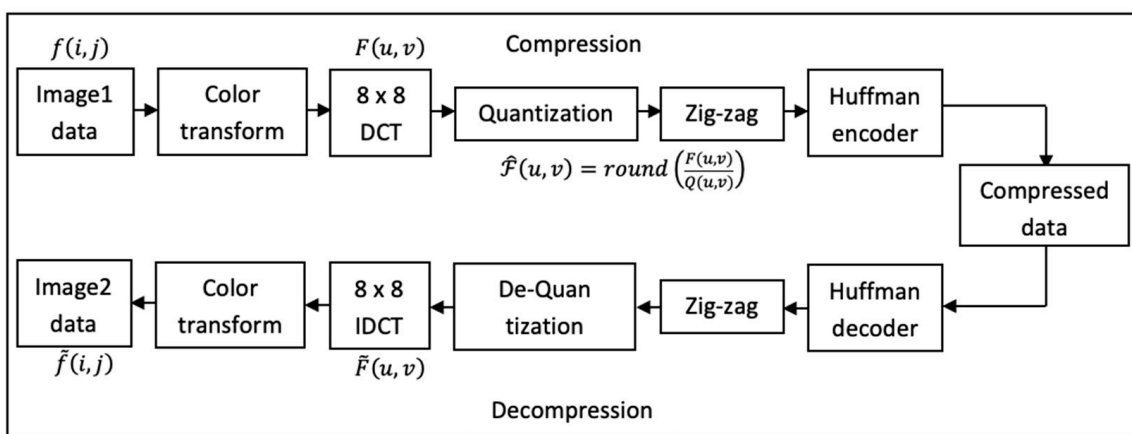


Figure 1. JPEG compression–decompression process.

There are five steps of compression as well as the decompression of JPEG files. The original RGB pixel values are first converted to a luminance/chrominance space (YCbCr). Then, the image is divided into blocks of 8 × 8 pixels. Each pixel is subtracted from 128 for the level shift, thus subjecting all pixels to be within the range of –128 to 127. Then, two-dimensional DCT is applied to each block to obtain transform coefficients in the frequency domain. The resulting coefficients are quantized by dividing with the values defined in the

quantization table. This is where the main compression of JPEG files takes place. The third step put the quantized transform coefficients in zig-zag order, in a sequence of low to high frequencies. As a result, more zeroes are likely to occur at the end of the sequence which is then followed by the Huffman encoder process. The decoding process is the reverse of the encoding process. Error-values of the original image (8×8 blocks) can be analyzed with the uncompressed version of the JPEG by taking the difference of Image1 and Image2 of the same image data, as given in expression (1):

$$DIFF(i, j) = f(i, j) - \tilde{f}(i, j) \tag{1}$$

The end process of the compression steps often produces compressed JPEG data called bitstreams. Carving is conducted by dealing with this bitstream of data. Carving with file types, such as text and documents, seems easier [41] while it is challenging when dealing with compressed data as with JPEG. It requires specific skills and knowledge to extract all valuable data to carve bitstreams of JPEG compressed data, with no chance of viewing the real JPEG values instead of the non-real values. Thus, our technique is fully implemented based on these bitstream patterns.

The observable errors of JPEG files arise when files are rounded, quantized, and errors are truncated [42]. Furthermore, the DCT process induces round-off errors. More errors are also added when the output of DCT is divided by the quantization coefficient using values stored in quantization tables. Different JPEG tools use different quantization table data, such as GIMP and Adobe Photoshop. In addition, different digital devices, such as digital cameras and smartphones, use customized quantization tables for the creation of JPEG files [43].

2.3. Error Level Analysis

Error Level Analysis (ELA) is widely used in image forensics research. More recently, the ELA approach is used in detecting image tampering and modification [44]. This technique interprets the error pattern by examining the difference between original images and a modified version of the same image. It works by comparing pixels in the original image with pixels in the modified image.

The amount of error in ELA is based on 8×8 blocks. JPEG images can be explained with two conditions through ELA:

- A JPEG is said to be original if all 8×8 blocks have a similar error pattern. Therefore, the 8×8 pixel block can be said to have attained local minima.
- A JPEG is said to be manipulated if any 8×8 block has a higher error pattern and an 8×8 pixel block is not at its local minima.

ELA process can be carried out by resaving a given image using a specific compression quality level, then computing and observing the differences between the compression levels as illustrated in the expression given in Equation (2):

$$\begin{aligned}
 & \overbrace{I_{A0}(i, j) - I_{B1}(i, j)}^{\text{Resavings} \quad \text{Recompres}} = ELA_1 \\
 & I_{A1}(i, j) - I_{B2}(i, j) = ELA_2 \\
 & I_{A2}(i, j) - I_{B3}(i, j) = ELA_3 \\
 & \quad \quad \quad \vdots \\
 & \quad \quad \quad \vdots \\
 & I_{An}(i, j) - I_{Bn}(i, j) = ELA_n
 \end{aligned} \tag{2}$$

The above expressions showed how ELA works. In the expression, 'I' refers to a JPEG image. Taking one example of ELA calculation, where A_n denotes a JPEG resaved for n times of 75% quality setting, and B_n denotes a JPEG recompressed for n times of 95% quality setting. Thus, the above ELA calculation can be said as "using ELA of 95%". In

other words, it is to examine what will be the error value if the JPEG is recompressed with 95% compression quality. The amount of compression error, $ELA_{1,2,\dots,n}$, will be decreased as JPEG is resaved for $1, 2, \dots, n$ times. As a result of resaving times, each 8×8 JPEG block will slowly reach its local minima and will become darker.

On the other hand, a study in [44] also demonstrated another way of detecting manipulation by approximating JPEG quality for the 8×8 block. However, this approach fails whenever a JPEG is compressed using Adobe Photoshop because the values of approximation are far different compared to the quality setting in that tool. The proposed method is, however, immune to this approximation. This was validated using a test image that was compressed and decompressed using Adobe Photoshop. Figure 2 shows the result of ELA for one image based on an ELA of 95%. Noticed the presence of an added object (highlighted in the circle) in the JPEG during the first and second resaved. This is where ELA is applicable in detecting image tampering and modification. The dataset used in this study is available online here.

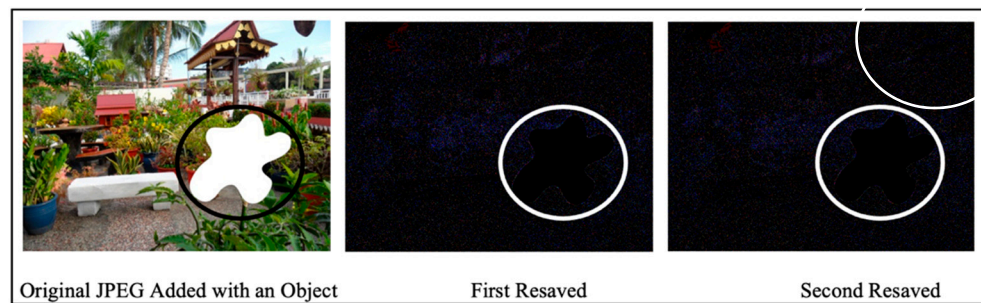


Figure 2. An example of an image depicting ELA results.

3. Proposed Algorithm

The proposed algorithm is inspired by techniques used in detecting image manipulation. ELA works by resaving JPEGs with a particular compression level, while analysis is performed by observing the rate of ELA change relative to the resave count. The flowchart of the algorithm adopted for this study is further shown in Figure 3.

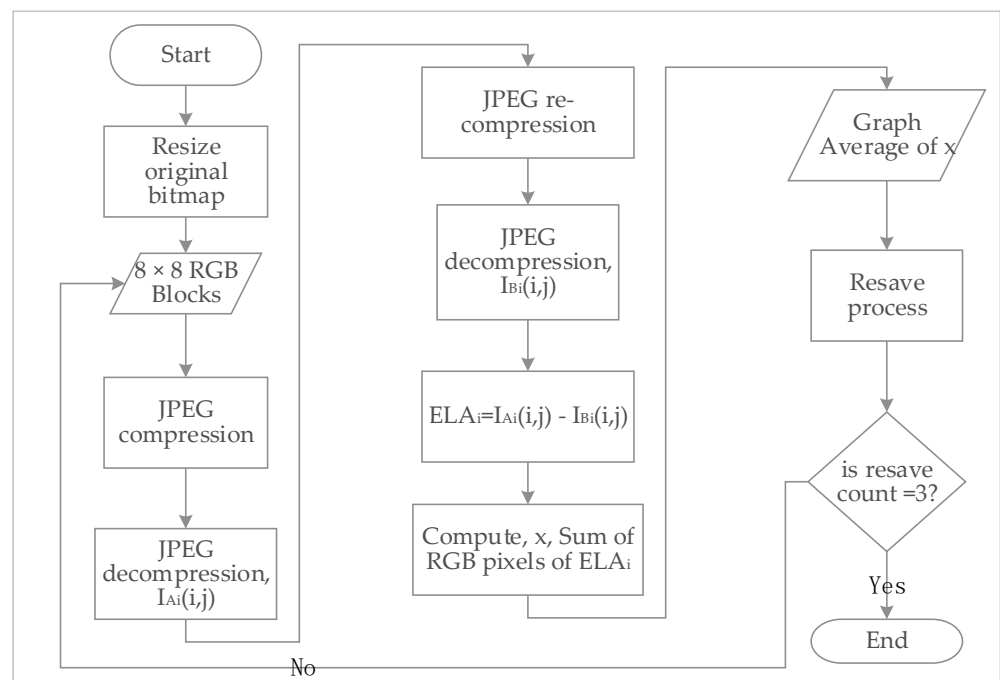


Figure 3. Flow of proposed method.

An ELA in the image forensics area makes use of JPEG pixels to observe the error level. Whereas in the proposed method, an uncompressed version of a JPEG in the format of a Bitmap image is used to examine the error level. As shown in Figure 3, we observed three phases of the ELA error rate (three repetitions in a flow chart), which depicts the number of times the JPEG file is resaved as shown in Figure 4. For every resaved JPEG, the error is calculated by examining the differences (*DIFF*) between the original image and the uncompressed image. To further simplify the results, error analysis based on average calculation is computed. The JPEG signature of 8×8 blocks is finalized based on graph representation. The summary of the three repetitions in the flow chart (in Figure 3) is further illustrated in Figure 4.

1. Run first compression–decompression and calculate rate of ELA,

$$DIFF_1(i_1, j_1) = f(i_1, j_1) - \tilde{f}(i_1, j_1)$$
2. Run second compression–decompression and calculate rate of ELA,

$$DIFF_2(i_2, j_2) = f(i_2, j_2) - \tilde{f}(i_2, j_2)$$
3. Run third compression–decompression and calculate rate of ELA,

$$DIFF_3(i_3, j_3) = f(i_3, j_3) - \tilde{f}(i_3, j_3)$$

Figure 4. Simplified algorithm of ELA calculation.

4. Results and Discussions

An experiment was conducted using 21 JPEG files of JFIF format, each comprised of 48, 8×8 blocks: a total of 1008 blocks. The original images are resized for the sake of seeking the initial results of the experiment. All images are from a personal collection to confirm there is no modification of the images. Otherwise, ELA will not work. The resizing of images and compression–decompression are performed using Adobe Photoshop. Using Adobe Photoshop, 75% of JPEG resaving, and 95% of the compression level were performed. The observation and analysis showed that different tools give different rounding errors of decompression results and Adobe Photoshop produces the best results. The ELA error rates were then calculated using Matlab R2013a. Observations were then carried out based on a graph creation process by using an ELA of 95%. For each run, the average difference for each 8×8 JPEG block was observed. The blocks were set up as shown in Table 2. The JPEGs were resized to 64×48 pixels. The setup position was only for experiment references, not a JPEG baseline sequence.

Table 2. JPEG block location setup for experiment.

Block 1	Block 7	Block 13	Block 19	Block 25	Block 31	Block 37	Block 43
Block 2	Block 8	Block 14	Block 20	Block 26	Block 32	Block 38	Block 44
Block 3	Block 9	Block 15	Block 21	Block 27	Block 33	Block 39	Block 45
Block 4	Block 10	Block 16	Block 22	Block 28	Block 34	Block 40	Block 46
Block 5	Block 11	Block 17	Block 23	Block 29	Block 35	Block 41	Block 47
Block 6	Block 12	Block 18	Block 24	Block 30	Block 36	Block 42	Block 48

The ELA calculation was performed for three runs and was obtained based on the following steps:

1. $I_{A0}(i, j) - I_{B1}(i, j) = ELA_1$
 2. $I_{A1}(i, j) - I_{B2}(i, j) = ELA_2$
 3. $I_{A2}(i, j) - I_{B3}(i, j) = ELA_3$
- (3)

By assuming the JPEG file (in this case is I_{A0}) is stored in a disk is using 75% compression quality-based JPEG DHT metadata, this image is recompressed using 95% quality (in this case it is I_{B1}) and the difference of both files is declared as ELA_1 . In the second step, I_{A0} is further resaved by using the same compression level which is 75%, recompressed with

Table 3. Cont.

Block	Re-Save	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
42	1st	0.19	0.95	0.56	1.92	1.83	1.50	1.83	2.08	1.94	1.91	0.70	0.80	1.61	1.31	1.98	0.70	0.98	0.89	0.19	1.78	1.61
	2nd	0.19	0.95	0.56	1.91	1.83	1.50	1.83	2.08	1.94	1.91	0.70	0.80	1.61	1.31	1.98	0.70	0.98	0.89	1.08	1.78	1.61
	3rd	0.19	0.95	0.56	1.91	1.83	1.50	1.83	2.08	1.94	1.91	0.70	0.80	1.61	1.31	1.98	0.70	0.98	0.89	1.22	1.78	1.61
43	1st	1.72	1.30	0.75	0.42	0.75	1.42	0.92	1.70	1.25	1.25	0.84	1.13	0.94	0.73	0.44	0.75	2.03	1.56	1.22	1.02	1.44
	2nd	1.84	1.14	0.75	0.42	0.75	1.42	0.75	1.70	1.25	1.25	0.84	1.08	0.94	0.73	0.44	0.77	2.03	1.56	1.22	1.02	1.44
	3rd	1.58	1.14	0.75	0.42	0.75	1.42	0.75	1.70	1.25	1.25	0.84	1.08	0.94	0.73	0.44	0.77	2.03	1.56	1.22	1.02	1.44
44	1st	1.48	2.88	1.27	1.53	1.42	1.17	2.11	1.38	0.69	1.38	1.28	2.36	1.78	2.30	0.23	0.69	1.25	1.55	0.42	0.81	1.97
	2nd	1.48	2.30	1.27	1.53	1.42	1.17	1.80	1.38	0.69	1.38	1.28	2.36	1.78	2.30	0.23	0.69	1.25	1.55	0.42	0.81	1.97
	3rd	1.48	2.42	1.27	1.53	1.42	1.17	1.73	1.38	0.69	1.38	1.28	2.36	1.78	2.30	0.23	0.69	1.25	1.55	0.42	0.81	1.97
45	1st	1.09	1.97	0.52	2.31	1.17	0.98	0.84	1.91	1.39	2.23	2.13	2.38	0.95	2.42	1.47	0.28	1.94	1.53	0.63	0.84	1.78
	2nd	1.09	1.97	0.52	2.11	1.17	0.98	0.84	1.91	1.39	2.23	2.13	2.38	0.95	2.42	1.47	0.28	1.94	1.53	0.64	0.84	1.78
	3rd	1.09	1.97	0.52	2.17	1.17	0.98	0.84	1.91	1.39	2.23	2.13	2.38	0.95	2.42	1.47	0.28	1.94	1.53	0.67	0.84	1.78
46	1st	1.25	1.34	0.52	1.48	1.81	1.86	2.02	2.39	0.70	1.66	0.69	2.69	1.89	2.38	0.97	0.67	1.20	0.72	0.14	2.08	1.48
	2nd	1.25	1.34	0.52	1.48	1.81	1.63	2.02	2.39	0.70	1.66	0.69	2.69	1.89	2.56	0.97	0.67	1.20	0.72	0.14	1.86	1.48
	3rd	1.25	1.34	0.52	1.48	1.81	1.92	2.02	2.39	0.70	1.66	0.69	2.69	1.89	2.31	0.97	0.67	1.20	0.72	0.14	1.81	1.48
47	1st	1.25	1.58	0.42	1.88	0.72	0.95	0.89	1.72	1.64	1.89	1.30	2.19	0.94	0.39	1.36	2.06	1.31	0.86	0.89	1.45	1.56
	2nd	1.25	1.58	0.28	1.83	0.72	1.06	0.89	1.72	1.64	1.89	1.30	2.19	0.94	0.39	1.45	2.00	1.22	0.86	1.22	1.45	1.56
	3rd	1.25	1.58	0.28	2.27	0.72	0.94	0.89	1.72	1.64	1.89	1.30	2.19	0.94	0.39	1.45	2.17	1.22	0.86	1.03	1.45	1.56
48	1st	0.92	0.66	0.56	1.48	1.13	0.94	2.00	1.80	2.08	1.39	1.55	1.41	1.13	1.47	1.86	1.16	0.72	0.73	0.73	0.78	0.89
	2nd	0.92	0.66	0.56	1.59	1.13	0.94	1.89	1.80	2.08	1.39	1.55	1.41	0.73	1.47	1.86	0.92	0.72	0.73	0.73	0.78	0.89
	3rd	0.92	0.66	0.56	1.66	1.13	0.94	1.86	1.80	2.08	1.39	1.55	1.41	0.73	1.47	1.86	0.92	0.72	0.73	0.73	0.78	0.89

Observed ELA differential.

The highlighted sequences in Table 3 illustrate the component that belongs to the particular blocks that have changed in the ELA values for each file-resave process, with 18% of the total of 1008 blocks. However, the number of a particular block is not constant for each JPEG as different JPEGs consist of various pixel values and colors. The unchanged values depict a block that was not affected by the compression process. It was observed that the lower the quality of the image, the lesser the value of the ELA. For instance, a closer observation of Table 3 reveals that a reduction of value can be observed from 95% quality to 75% quality. This thus shows a linear relationship between the quality of the image and the corresponding ELA value. In some instances, the value of the corresponding ELA remains the same within the modified range (quality). However, the corresponding values are still greater than the unmodified ELA values. Furthermore, for some blocks of the image, the value of the ELA remains unchanged irrespective of the quality of modification. The graph in Figure 5 shows the pattern that occurred during the first re-compressed $DIFF1(i1, j1)$, while Figure 6 depicts the pattern during the second re-compressed $DIFF2(i2, j2)$. Furthermore, Figure 7 shows the pattern of the third re-compressed $DIFF3(i3, j3)$.

From the observation in Figures 5–7, for each re-compression cycle, the value of ELA is below 4. However, the scatter plot further shows that the ELA values for most of the images range between 0 and 3. On further examination of the scatter plots, values above 3 were generated from images img4, img13, img14, and img15, respectively. To fully examine the probable cause of these outliers, further evaluation of the range of ELA values for those four images was carried out, by modifying the quality of JPEG re-savings. This process involves decreasing the quality value of the image during resaving, 70% of compression quality in this case. ELA calculation is then carried out for the three runs, and a slight modification to the steps is as follows:

$$\begin{aligned}
 &1. I_{A1}(i, j) - I_{B2}(i, j) = ELA_1 \\
 &2. I_{A2}(i, j) - I_{B3}(i, j) = ELA_2 \\
 &3. I_{A3}(i, j) - I_{B4}(i, j) = ELA_3
 \end{aligned} \tag{4}$$

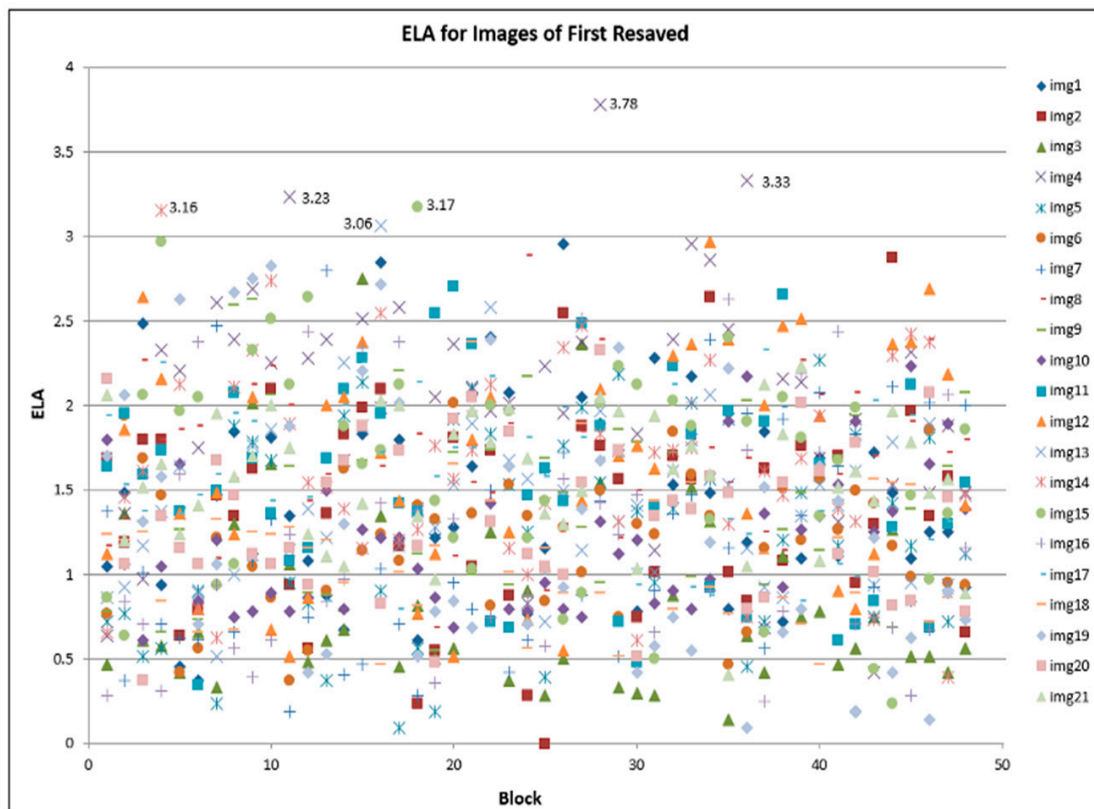


Figure 5. ELA for images of first resaved.

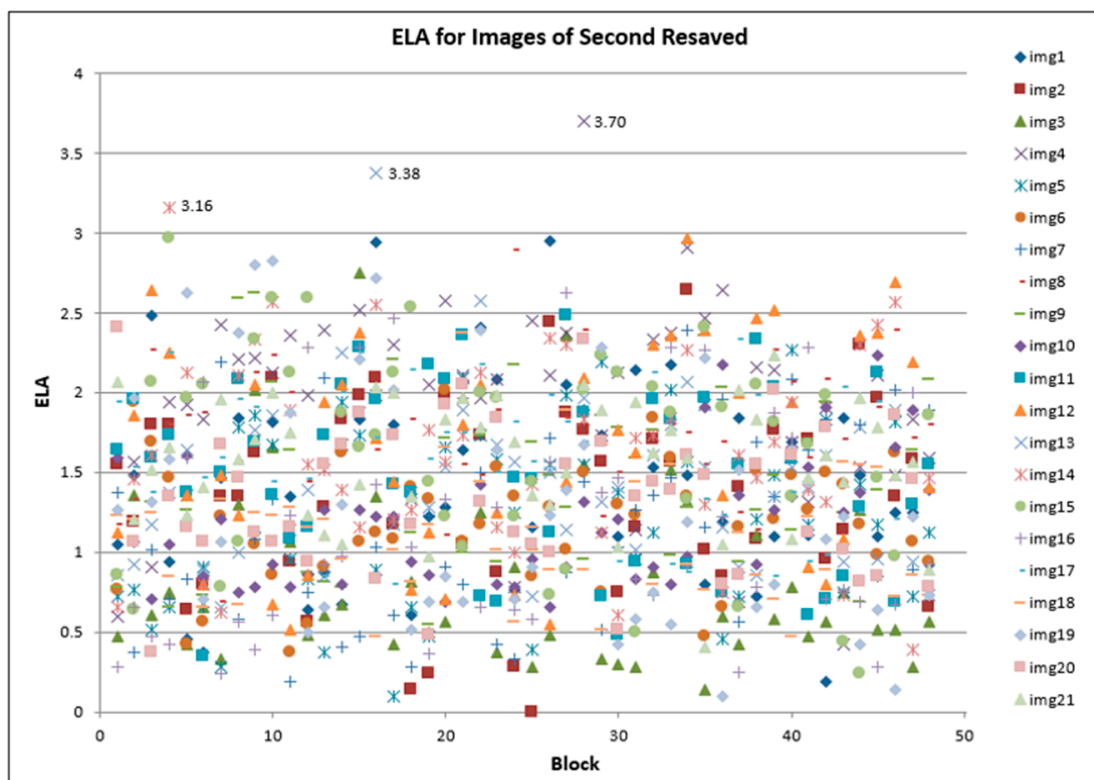


Figure 6. ELA for images of second resaved.

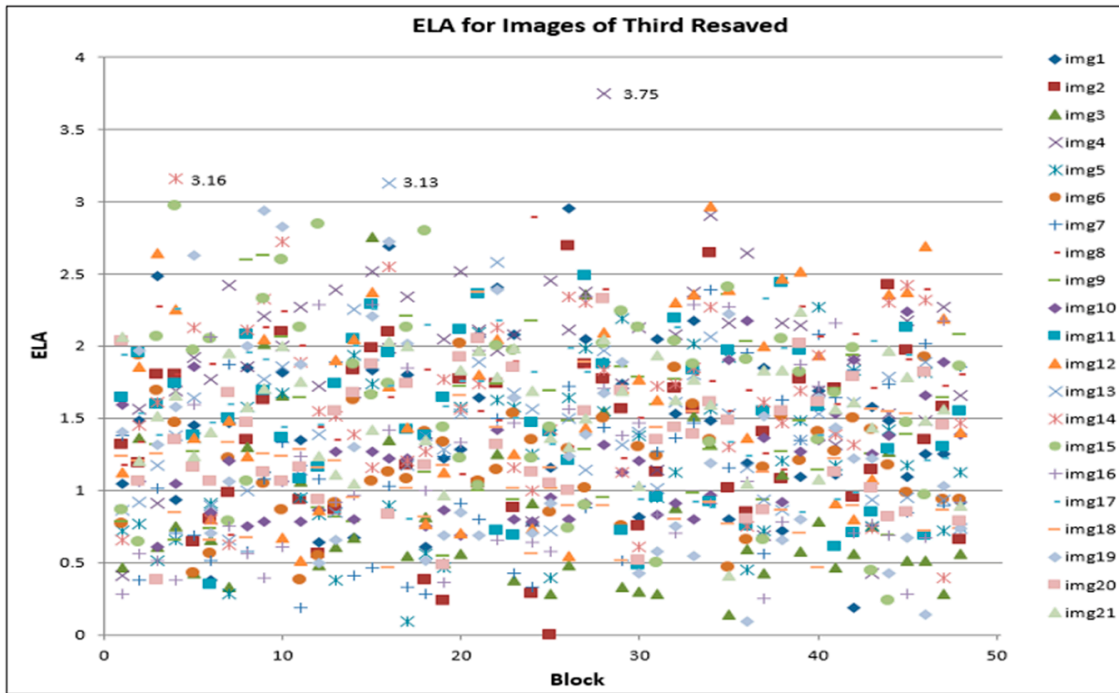


Figure 7. ELA for images of third resaved.

In the previous steps, I_{A0} originally stored in a disk with 75% quality compression quality is re-compressed with 95% quality to observe ELA of 95% quality. Compared to the above-stated steps, I_{A0} of 75% quality compression quality in a disk is resaved again with 70% quality to obtain I_{A1} . I_{A1} is then re-compressed with 95% quality to produce ELA_1 . In the second step, the I_{A1} is resaved with 70% quality to reach I_{A2} . I_{A2} is then re-compressed with 95% quality to produce ELA_2 . The same step is repeated for the third step. The result of this process is shown in Table 4.

Table 4. ELA for images img4, img13, img14 and img15.

Block	Re-Save	Image			
		Img4	Img13	Img14	Img15
1	1st	0.97	0.81	1.16	0.52
	2nd	0.75	0.81	1.16	0.52
	3rd	0.75	0.81	1.16	0.52
2	1st	1.20	1.03	1.64	1.47
	2nd	1.28	1.03	1.64	1.47
	3rd	1.28	1.03	1.64	1.47
3	1st	1.23	0.91	1.28	1.84
	2nd	1.23	0.91	1.28	1.84
	3rd	1.23	0.91	1.28	1.84
4	1st	1.47	1.36	1.94	1.30
	2nd	1.23	1.36	1.67	1.30
	3rd	1.23	1.36	1.67	1.30
5	1st	1.94	0.98	1.55	1.77
	2nd	1.77	0.98	1.55	1.77
	3rd	1.77	0.98	1.55	1.77

Table 4. Cont.

Block	Re-Save	Image			
		Img4	Img13	Img14	Img15
6	1st	1.38	0.86	0.63	1.67
	2nd	1.38	0.86	0.63	1.67
	3rd	1.38	0.86	0.63	1.67
7	1st	2.08	1.02	0.98	1.00
	2nd	2.08	1.02	0.98	0.75
	3rd	2.08	1.02	0.98	0.75
8	1st	1.61	1.22	1.67	1.30
	2nd	1.36	1.22	1.67	1.30
	3rd	1.36	1.22	1.67	1.30
9	1st	1.58	1.22	1.22	1.17
	2nd	1.58	1.22	1.22	1.17
	3rd	1.58	1.22	1.22	1.17
10	1st	2.16	1.55	2.41	1.27
	2nd	2.16	1.55	2.41	1.27
	3rd	2.16	1.55	2.41	1.27
11	1st	2.92	0.95	1.73	1.66
	2nd	2.55	0.95	1.73	1.66
	3rd	2.42	0.95	1.73	1.66
12	1st	1.78	1.50	1.36	2.41
	2nd	1.78	1.50	1.36	2.19
	3rd	1.78	1.50	1.36	2.19
13	1st	1.42	0.84	1.80	1.61
	2nd	1.42	0.84	1.80	1.61
	3rd	1.42	0.84	1.80	1.61
14	1st	1.45	1.77	1.25	1.86
	2nd	1.45	1.77	1.25	1.86
	3rd	1.45	1.77	1.25	1.86
15	1st	1.36	1.80	1.33	1.80
	2nd	1.36	1.80	1.33	1.80
	3rd	1.36	1.80	1.33	1.80
16	1st	1.95	2.16	1.30	1.22
	2nd	1.95	2.36	1.30	1.22
	3rd	1.95	2.36	1.30	1.22
17	1st	1.64	0.91	1.31	1.95
	2nd	1.64	0.91	1.31	1.95
	3rd	1.64	0.91	1.31	1.95
18	1st	1.70	1.59	1.28	2.06
	2nd	1.70	1.59	1.28	2.06
	3rd	1.70	1.59	1.28	2.06

Table 4. Cont.

Block	Re-Save	Image			
		Img4	Img13	Img14	Img15
19	1st	1.25	1.23	1.28	1.34
	2nd	1.25	1.23	1.28	1.34
	3rd	1.25	1.23	1.28	1.34
20	1st	2.52	1.44	1.83	1.08
	2nd	1.86	1.44	1.83	1.08
	3rd	1.70	1.44	1.83	1.08
21	1st	1.44	1.84	1.20	1.16
	2nd	1.44	1.84	1.20	1.16
	3rd	1.44	1.84	1.20	1.16
22	1st	1.02	1.55	1.64	1.53
	2nd	1.02	1.55	1.64	1.53
	3rd	1.02	1.55	1.64	1.53
23	1st	2.13	1.83	1.61	1.45
	2nd	2.13	1.83	1.61	1.45
	3rd	2.13	1.83	1.61	1.45
24	1st	0.94	1.03	0.78	0.88
	2nd	0.94	1.03	0.78	0.88
	3rd	0.94	1.03	0.78	0.88
25	1st	1.67	0.75	0.47	1.28
	2nd	1.67	0.75	0.47	1.28
	3rd	1.67	0.75	0.47	1.28
26	1st	1.73	1.02	1.30	1.06
	2nd	1.73	1.02	1.30	1.06
	3rd	1.73	1.02	1.30	1.06
27	1st	1.83	0.81	1.91	0.98
	2nd	1.83	0.81	1.91	0.98
	3rd	1.83	0.81	1.91	0.98
28	1st	2.59	2.17	1.25	1.36
	2nd	2.59	2.17	1.25	1.36
	3rd	2.59	2.17	1.25	1.36
29	1st	1.61	0.98	1.02	1.50
	2nd	1.61	0.98	1.02	1.50
	3rd	1.61	0.98	1.02	1.50
30	1st	1.31	1.13	0.78	1.30
	2nd	1.36	1.13	0.78	1.30
	3rd	1.36	1.13	0.78	1.30
31	1st	0.78	0.88	1.17	0.38
	2nd	0.78	0.88	1.17	0.38
	3rd	0.78	0.88	1.17	0.38

Table 4. Cont.

Block	Re-Save	Image			
		Img4	Img13	Img14	Img15
32	1st	2.16	1.55	1.20	2.09
	2nd	1.77	1.55	1.20	2.09
	3rd	1.77	1.55	1.20	2.09
33	1st	2.27	1.73	1.75	1.20
	2nd	2.22	1.73	1.75	1.20
	3rd	2.22	1.73	1.75	1.20
34	1st	2.09	1.00	1.39	1.73
	2nd	2.09	1.00	1.39	1.73
	3rd	2.09	1.00	1.39	1.73
35	1st	2.11	1.30	1.22	1.47
	2nd	2.16	1.30	1.22	1.47
	3rd	1.92	1.30	1.22	1.47
36	1st	1.61	1.58	0.83	2.19
	2nd	1.61	1.58	0.83	2.19
	3rd	1.61	1.58	0.83	2.19
37	1st	0.89	1.09	1.53	0.33
	2nd	0.89	1.05	1.53	0.33
	3rd	0.89	0.95	1.53	0.33
38	1st	2.11	1.39	1.13	1.06
	2nd	2.11	1.39	1.13	1.06
	3rd	2.11	1.39	1.13	1.06
39	1st	1.63	0.92	1.56	1.56
	2nd	1.63	0.92	1.56	1.56
	3rd	1.63	0.92	1.56	1.56
40	1st	1.44	1.02	1.36	1.44
	2nd	1.44	1.02	1.36	1.44
	3rd	1.44	1.02	1.36	1.44
41	1st	1.30	1.38	1.27	1.31
	2nd	1.30	1.38	1.27	1.31
	3rd	1.30	1.38	1.27	1.31
42	1st	1.22	1.39	1.36	1.19
	2nd	1.22	1.39	1.36	1.19
	3rd	1.22	1.39	1.36	1.19
43	1st	0.84	0.95	0.52	0.61
	2nd	0.84	0.95	0.52	0.61
	3rd	0.84	0.95	0.52	0.61
44	1st	0.92	1.77	1.67	0.47
	2nd	0.92	1.77	1.67	0.47
	3rd	0.92	1.77	1.67	0.47

Table 4. Cont.

Block	Re-Save	Image			
		Img4	Img13	Img14	Img15
45	1st	1.89	0.34	1.50	1.56
	2nd	1.58	0.34	1.50	1.56
	3rd	1.28	0.34	1.50	1.56
46	1st	1.17	1.36	1.30	1.08
	2nd	1.17	1.36	1.30	1.08
	3rd	1.17	1.36	1.30	1.08
47	1st	1.98	1.22	0.64	0.92
	2nd	1.72	1.22	0.64	0.92
	3rd	1.80	1.22	0.64	0.92
48	1st	1.33	0.70	1.05	1.47
	2nd	1.33	0.70	1.05	1.47
	3rd	1.33	0.70	1.05	1.47

Similar patterns were observed between Tables 3 and 4. However, a scatter plot of the combined images (by applying the modified process to all the images), shown in Figures 8–10 respectively, deviates from the observed pattern in Figures 5–7. Whilst the ELA threshold for the first sets of the re-saved processes (as presented in the expression in Equation (2)) had values greater than 3, the current process (as presented in the expression in Equation (3)) generates values that are below ELA of 3.

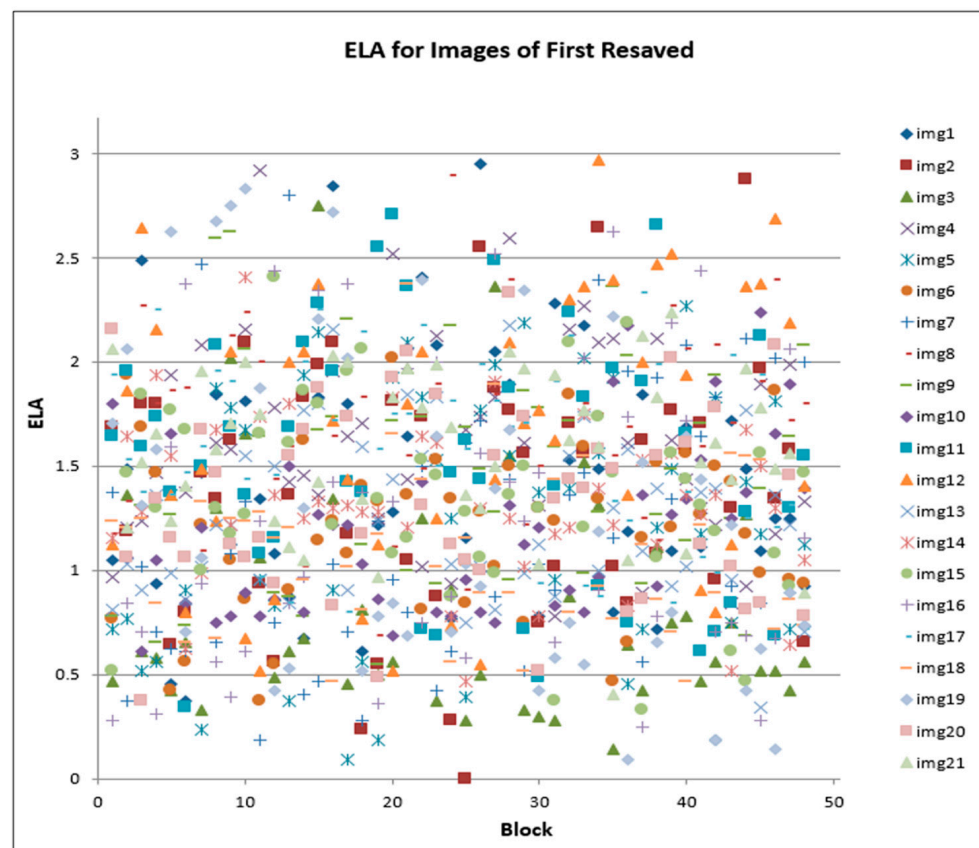


Figure 8. ELA for images of first resaved.

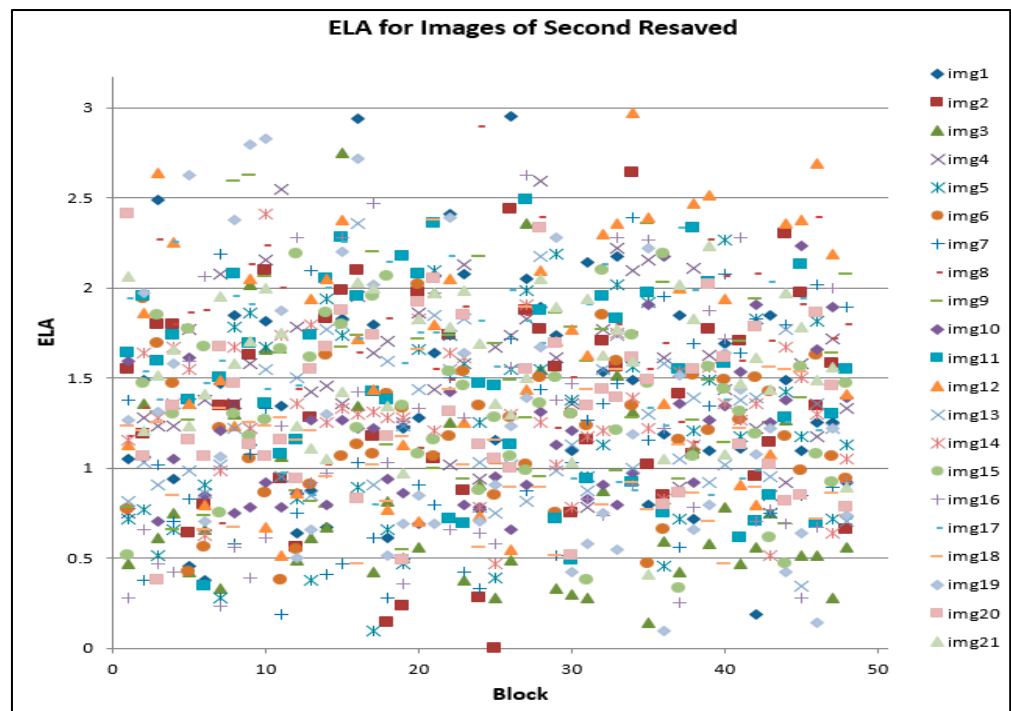


Figure 9. ELA for images of second resaved.

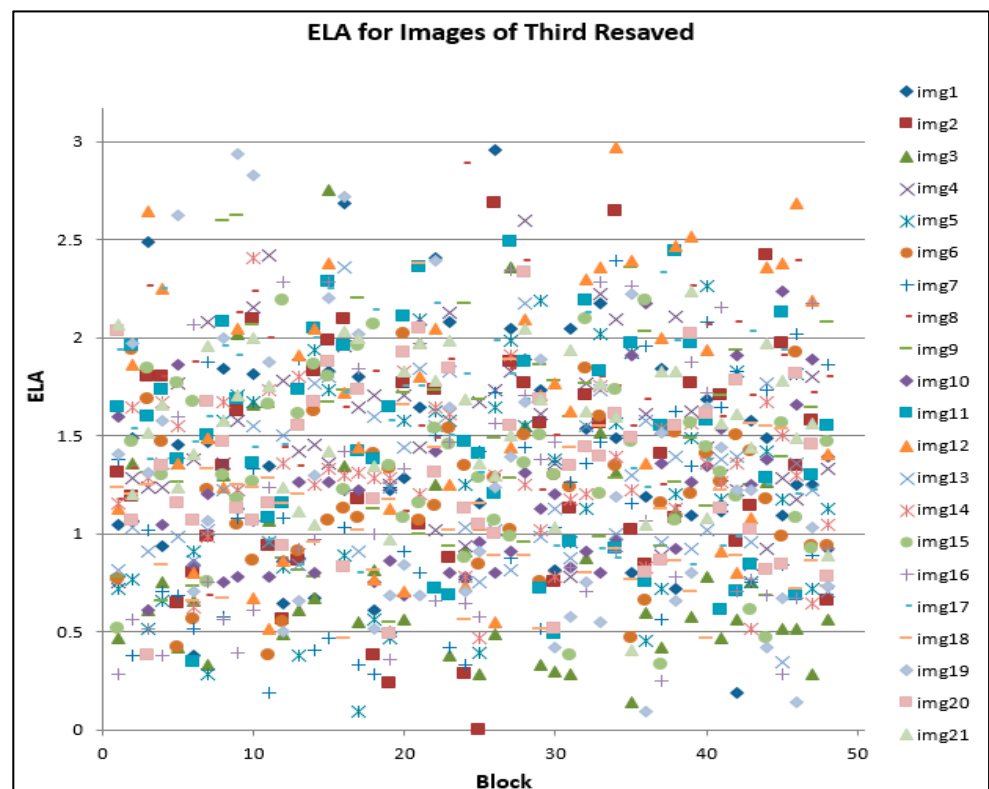


Figure 10. ELA for images of third resaved.

Therefore, in each resaved phase (first, second and third resaved) associated with each JPEG, it was observed, as shown in Figures 8–10, that the ELA values are below 3.0. A closer examination of each block of each image shows that there is a small change whenever images are resaved three times. In image forensics, this is where ELA can be used

to detect whether images are modified or not. There will be a significant degree of changes if images have been modified for some purpose. Furthermore, to ascertain if a given JPEG file has been modified, a forensic investigator can compare the ELA of the given JPEG file to the original image. Further study will, however, adapt this benefit towards developing a novel carving algorithm that can be utilized for the subsequent stages of JPEG image-file forensic analysis. This process could be specifically useful for identifying modified images during the investigation. A forensic examiner equipped with this knowledge will be able to, in addition to other knowledge, identify falsified images even when the metadata and the image header seems to be correct. This approach, therefore, offers signature-based identification metrics which cannot be easily altered [45,46]. In a preliminary observation, the study has observed ELA values that are greater than 3, tested on non-JPEG file (PDF file) fragments with JPEG fragments.

Timely and effective forensic investigation is paramount in any setting and owing to the complexity and structure of the compressed JPEG files, the author emphasizes not only focusing on the fragments but also using the file fragments as a way of linking the potential crime to a suspect. While the authors admit that pre-processing these data may have some implications due to the error rate, our approach has significantly outlined that it could be useful when implemented for purposes of digital forensic analysis. This in many instances could find traces from JPEG files that could, in most cases, defeat anti-forensic techniques. While this study advocates that this approach could only exist in a reactive approach, we also opine to the need for incorporating some forensic readiness [47,48] mechanisms which could allow automated signature identification in a proactive approach. This is because an image forger usually exercises an objective of consistent alteration in a camouflaging manner to be undetectable and defeat forensic investigation tools. Further study would be needed to confirm this observation and provide reliable proof of JPEG identification. Notably, leveraging error level analysis to identify the unique signatures could also be applied in diverse forensic investigation mechanisms, for example, in an IoT environment during end-to-end communication, where data traveling as plaintext or data that are altered when a cryptosystem is broken, over the networks. Given that wireless sensor networks are more vulnerable to these kinds of attacks, applying forensic analysis in these environments would be ideal in extracting artifacts that can be used to create key forensic hypotheses [49–51]. This forensic hypothesis would duly follow an investigation cycle that is deemed scientific based on acceptable processes and standards for the artifacts to be admissible in case of a security incident.

5. Comparison with Existing Techniques

The experiment that has been conducted in this study was mainly focused on highlighting how the suggested ELA technique is leveraged in identifying JPEG block signatures for purpose of digital forensics. Section 4 highlighted systematic steps that showed that 21 images have been used to realize the exercise where unique characteristics have been identified. It is worth noting that this study mainly capitalized on the drawbacks of existing works that have exhibited a high false-positive rate, therefore rendering the need for manual validation. While this study develops a technique that can identify the unique signature of JPEG 8×8 blocks using the Error Level Analysis technique, the authors have noted or drawn comparison with existing works that closely match the proposed, as is shown in Table 5. Researchers in [52] have suggested an approach that utilizes ELA in image forensics; however, the authors note that that study was more generalized. Moreover, researchers in [53] use lossy compression using ELA; however, this has a limitation where the color drops below 256. In addition, other studies in [54–57] have identified photo forensics algorithms using ELA, ELA for semi-automatic wavelet soft-thresholding, forgery identification using forensic tools, and face-swap image exposure on deep learning based ELA with a number of limitations inclined towards weakness with the addition of noise, generalized study, and difficulty in identifying identification when deep learning training models are used.

Table 5. Comparison with other studies.

REF	Focus	Limitation
[52]	ELA for image forensics	Study is generalized
[53]	Image forensics using lossy compression using ELA	Cannot be applied across non-lossy compression, such as PNG or where color drops below 256
[54]	Photo forensics algorithm using ELA	Inclined only towards lossy compression techniques
[55]	ELA for semi-automatic wavelet soft-thresholding	Study is weakened with produced noise
[56]	Forgery identification using forensic tools	Study is generalized and applies ELA, metadata analysis, JPEG luminance
[57]	Face-swap image exposure on deep learning based ELA	Deep learning training model cannot explicitly explain the principle of identification as opposed to ELA

The advantage of the proposed approach in this paper is drawn based on the comparisons that have been highlighted from other studies as follows:

- Alteration of images can significantly be reduced with a lower degree of false-positives where the features extracted from the images are subjected to ELA, and this can help to build a forensic hypothesis.
- The experiment that has been conducted in this study has shown that this approach is effective.
- Our approach has utilized a simple dataset which in the context of this study overcomes the intensive need for rigorous training while pointing out specific features, which from a digital forensic perspective may save an investigator time.

The experimental results that have been shown in Section 4 shows that the efficiency rate of identification of this approach can be improved if embedded with other techniques; however, this study has potentially been positioned as a build-up owing to the fact that it has put across a proof of concept based on the identified problem.

6. Future Directions

Research on image quality assessment can be significantly leveraged for multimedia forensics. Evaluation metrics, such as the peak-signal-to-noise ratio (PSNR), the structural similarity index measure (SSIM), root mean square error (RMSE), and the feature similarity indexing method (FSIM), can be used to evaluate as well as discriminate the structural content of an image. For instance, studies in [49–51] have observed that the SSIM is more sensitive to JPEG compression relative to the PSNR. The JPEG image compression process introduces distinguishable structural distortion which can be measured by these metrics. By using the absolute errors computation capability of the PSNR and RMSE, coupled with the perception and saliency-based errors provided by the SSIM and FSIM, future works will focus on exploiting these features in addition to the ELA for forensic analysis. Using these, a forensic investigator can examine the luminance, structural and contrast properties of a given image. Furthermore, the authors intend to validate the result of these metrics on a larger scale by identifying ELA, SSIM, FSIM, RMSE, and PSNR values for larger JPEG file sizes, with more blocks, which can precisely be used to identify JPEG fragments in the carving process. This output will then be used to develop an intelligent system for carving compressed JPEG files. This is currently a largely missing component for digital forensic examiners.

7. Conclusions

A multitude of digital forensic challenges show that there is still a lack of suitable techniques that can facilitate electronic discovery, especially for JPEG file formats. This creates the need of developing suitable techniques that can solve this perennial challenge. In this paper, we have introduced a new technique of identifying a unique JPEG block

signature using error level analysis, which is a significant process during forensic analysis. This has been achieved by observing the stages of JPEG compression level and identifying JPEG 8×8 blocks based on ELA values that have a range of 0 to 3.0.

Author Contributions: Conceptualization, N.A.N.A., R.A.I., S.A.R. and V.R.K.; methodology, N.A.N.A. and R.A.I.; software, N.A.N.A., R.A.I., S.A.R. and V.R.K.; validation, N.A.N.A., R.A.I., S.A.R. and V.R.K.; formal analysis, N.A.N.A., R.A.I., S.A.R. and V.R.K.; investigation, N.A.N.A., R.A.I., S.A.R. and V.R.K.; resources, N.A.N.A., R.A.I. and S.A.R.; data curation, N.A.N.A., R.A.I. and S.A.R.; writing—original draft preparation, N.A.N.A., R.A.I. and S.A.R.; writing—review and editing, N.A.N.A., R.A.I. and S.A.R.; visualization, N.A.N.A., R.A.I. and S.A.R.; supervision, R.A.I. and S.A.R.; project administration, N.A.N.A., R.A.I. and S.A.R.; funding acquisition, V.R.K. All authors have read and agreed to the published version of the manuscript.

Funding: APC was funded by Blekinge Institute of Technology, BTH, Sweden.

Data Availability Statement: The dataset used in this study is available online here.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adeyemi, I.R.; Razak, S.A.; Zainal, A.; Azhan, N.A.N. A Digital Forensic Investigation Model for Insider Misuse. In *Advances in Computational Science, Engineering and Information Technology, AISC*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 225, pp. 293–305.
2. Ikuesan, A.R.; Venter, H.S. Digital forensic readiness framework based on behavioral-biometrics for user attribution. In Proceedings of the 2017 IEEE Conference on Application, Information and Network Security (AINS), Miri, Malaysia, 13–14 November 2017; Volume 2018-Janua, pp. 54–59.
3. Adeyemi, I.R.; Razak, S.A.; Salleh, M.; Venter, H.S. Leveraging human thinking style for user attribution in digital forensic process. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2017**, *7*, 198–206.
4. Mohlala, M.; Ikuesan, A.R.; Venter, H.S. User attribution based on keystroke dynamics in digital forensic readiness process. In Proceedings of the 2017 IEEE Conference on Application, Information and Network Security (AINS), Miri, Malaysia, 13–14 November 2017; Volume 2018-Janua, pp. 1–6.
5. Adeyemi, I.R.; Razak, S.A.; Azhan, N.A.N. A Review of Current Research in Network Forensic Analysis. *Int. J. Digit. Crime Forensics* **2013**, *5*, 1–26. [[CrossRef](#)]
6. Piva, A. An Overview on Image Forensics. *ISRN Signal Process.* **2013**, *2013*, 496701. [[CrossRef](#)]
7. Yasinsac, A.; Erbacher, R.F.; Marks, D.G.; Pollitt, M.M.; Sommer, P.M. Computer forensics education. *IEEE Secur. Priv.* **2003**, *1*, 15–23. [[CrossRef](#)]
8. Abdullah, M.T.; Mahmud, R.; Ghani, A.A.A.; Abdullah, M.Z.; Sultan, A.B.M. Advances in Computer Forensics. *Int. J. Comput. Sci. Netw. Secur.* **2008**, *8*, 215–219.
9. Singh, A.; Ikuesan, A.R.; Venter, H.S. Digital Forensic Readiness Framework for Ransomware Investigation. In *International Conference on Digital Forensics and Cyber Crime, Proceedings of the 10th International EAI Conference (ICDF2C 2018), New Orleans, LA, USA, 10–12 September 2018*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 259, p. 259.
10. Makura, S.M.; Venter, H.S.; Ikuesan, R.A.; Kebande, V.R.; Karie, N.M. Proactive Forensics: Keystroke Logging from the Cloud as Potential Digital Evidence for Forensic Readiness Purposes. In Proceedings of the 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT 2020), Doha, Qatar, 2–5 February 2020; pp. 200–205.
11. Kebande, V.R.; Ikuesan, R.A.; Karie, N.M.; Alawadi, S.; Choo, K.-K.R.; Al-Dhaqm, A. Quantifying the need for supervised machine learning in conducting live forensic analysis of emergent configurations (ECO) in IoT environments. *Forensic Sci. Int. Rep.* **2020**, *2*, 100122. [[CrossRef](#)]
12. Kebande, V.R.; Karie, N.M.; Venter, H.S. Cloud-Centric Framework for isolating Big data as forensic evidence from IoT infrastructures. In Proceedings of the 2017 1st International Conference on Next Generation Computing Applications (NextComp), Mauritius, 19–21 July 2017; pp. 54–60.
13. Karie, N.M.; Kebande, V.R.; Venter, H.S. Diverging deep learning cognitive computing techniques into cyber forensics. *Forensic Sci. Int. Synerg.* **2019**, *1*, 61–67. [[CrossRef](#)]
14. Adeyemi, I.R.; Abd Razak, S.; Salleh, M. Understanding Online Behavior: Exploring the Probability of Online Personality Trait Using Supervised Machine-Learning Approach. *Front. ICT* **2016**, *3*, 8. [[CrossRef](#)]
15. Kebande, V.R.; Karie, N.M.; Michael, A.; Malapane, S.; Kigwana, I.; Venter, H.S.; Wario, R.D. Towards an integrated digital forensic investigation framework for an IoT-based ecosystem. In Proceedings of the 2018 IEEE International Conference on Smart Internet of Things (SmartIoT), Xi'an, China, 17–19 August 2018; pp. 93–98.
16. Simone, M.P. *Data Carving Concepts*; SANS Institute: Bethesda, MD, USA, 2009; p. 27.
17. Beek, C. *Introduction to File Carving*; White Paper; McAfee Foundstone Professional Services: Mission Viejo, CA, USA, 2011.
18. Garfinkel, S.L. Carving Contiguous and Fragmented Files with Fast Object Validation. *Digit. Investig.* **2007**, *4*, 2–12. [[CrossRef](#)]

19. Richard III, G.G.; Roussev, V. Scalpel: A Frugal, High Performance File Carver. In Proceedings of the Digital Forensic Research Conference DFRWS 2005, New Orleans, LA, USA, 17–19 August 2005; pp. 1–10.
20. Pal, A.; Memon, N. The Evolution of File Carving. *IEEE Signal Process. Mag.* **2009**, *26*, 59–71. [[CrossRef](#)]
21. Deris, M.M.; Mohamad, K.M. Carving JPEG Images and Thumbnails Using Image Pattern Matching. In Proceedings of the 2011 IEEE Symposium on Computers & Informatics, Kuala Lumpur, Malaysia, 20–23 March 2011; pp. 78–83.
22. Abdullah, N.A.; Ibrahim, R.; Mohamad, K.M. Carving Thumbnail/s and Embedded JPEG Files Using Image Pattern Matching. *J. Softw. Eng. Appl.* **2013**, *6*, 62–66. [[CrossRef](#)]
23. Birmingham, B.; Farrugia, R.A.; Vella, M. Using Thumbnail Affinity for Fragmentation Point Detection of JPEG Files. In Proceedings of the IEEE EUROCON 2017—17th International Conference on Smart Technologies, Ohrid, Macedonia, 6–8 July 2017.
24. Guo, H.; Xu, M. A Method for Recovering JPEG Files Based on Thumbnail. In Proceedings of the 2011 International Conference on Control, Automation and Systems Engineering (CASE), Singapore, 30–31 July 2011.
25. Pal, A.; Sencar, H.T.; Memon, N. Detecting file fragmentation point using sequential hypothesis testing. *Digit. Investig.* **2008**, *5*, 2–13. [[CrossRef](#)]
26. Cohen, M.I. Advanced JPEG carving. In Proceedings of the e-Forensics'08: 1st International ICST Conference on Forensic Applications and Techniques in Telecommunications, Information and Multimedia, Adelaide, Australia, 21–23 January 2008; Volume 1, pp. 1–6.
27. van den Bos, J.; van der Storm, T. Bringing Domain-Specific Languages to Digital Forensics. In Proceedings of the 2011 33rd International Conference on Software Engineering (ICSE), Honolulu, HI, USA, 21–28 May 2011.
28. van den Bos, J.; van der Storm, T. Domain-Specific Optimization in Digital Forensics. In *International Conference on Theory and Practice of Model Transformations, Proceedings of the 5th International Conference (ICMT 2012), Prague, Czech Republic, 28–29 May 2012*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 121–136.
29. De Bock, J.; De Smet, P. JPGCarve: An Advanced Tool for Automated Recovery of Fragmented JPEG Files. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 19–34. [[CrossRef](#)]
30. Poisel, R.; Rybnicek, M.; Tjoa, S. Taxonomy of Data Fragment Classification Techniques. In *International Conference on Digital Forensics and Cyber Crime, Proceedings of the Fifth International Conference (ICDF2C 2013), Moscow, Russia, 26–27 September 2013*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 67–85.
31. Garfinkel, S.L. Digital forensics research: The next 10 years. *Digit. Investig.* **2010**, *7*, S64–S73. [[CrossRef](#)]
32. Veenman, C.J. Statistical disk cluster classification for file carving. In Proceedings of the Third International Symposium on Information Assurance and Security, Manchester, UK, 29–31 August 2007; pp. 393–398.
33. Fitzgerald, S.; Mathews, G.; Morris, C.; Zhulyn, O. Using NLP techniques for file fragment classification. *Digit. Investig.* **2012**, *9*, 44–49. [[CrossRef](#)]
34. Alshammary, E.; Hadi, A. Reviewing and Evaluating Existing File Carving Techniques for JPEG Files. In Proceedings of the 2016 Cybersecurity and Cyberforensics Conference (CCC), Amman, Jordan, 2–4 August 2016; pp. 55–59.
35. Warlock, Digital forensics: File Carving. InfoSec Publication. Available online: <https://resources.infosecinstitute.com/topic/file-carving/> (accessed on 1 February 2022).
36. Kadir, N.F.A.; Abd Razak, S.; Chizari, H. Identification of fragmented JPEG files in the absence of file systems. In Proceedings of the 2015 IEEE Conference on Open Systems (ICOS), Melaka, Malaysia, 24–26 August 2015; pp. 1–6.
37. Gopal, S.; Yang, Y.; Salomatin, K.; Carbonell, J. Statistical Learning for File-Type Identification. In Proceedings of the 2011 10th International Conference on Machine Learning and Applications and Workshops, Honolulu, HI, USA, 18–21 December 2011.
38. Ahmed, I.; Lhee, K.; Shin, H.; Hong, M. Fast Content-based File Type Identification. In *International Conference on Digital Forensics, Proceedings of the 7th IFIP WG 11.9 International Conference on Digital Forensics, Orlando, FL, USA, 31 January–2 February 2011*; Springer: Berlin/Heidelberg, Germany; pp. 65–75.
39. Roussev, V.; Garfinkel, S.L. File Fragment Classification-The Case for Specialized Approaches. In Proceedings of the 2009 Fourth International IEEE Workshop on Systematic Approaches to Digital Forensic Engineering, Berkeley, CA, USA, 21–21 May 2009; pp. 3–14.
40. Shaban Al-Ani, M.; Awad, F.H. The Jpeg Image Compression Algorithm. *Int. J. Adv. Eng. Technol.* **2013**, *6*, 1055–1062.
41. Alherbawi, N.; Shukur, Z.; Sulaiman, R. A Survey on Data Carving in Digital Forensics. *Asian J. Inf. Technol.* **2016**, *15*, 5137–5144.
42. ChandraSekhar, C.; Ramesh, C. A Novel Compression Technique for JPEG Error Analysis and for Digital Image Applications. *Int. J. Latest Trends Comput.* **2012**, *3*, 84–89.
43. Azhan, N.; Abd Razak, S.; Adeyemi, I.R. Analysis of DQT and DHT in JPEG Files. *Int. J. Inf. Technol. Comput. Sci. (IJITCS)* **2013**, *10*, 1–11.
44. Krawetz, N. A Picture 's Worth . . . Version 2 Table of Contents. *Solutions* **2008**, 1–43.
45. Ikuesan, A.R.; Venter, H.S. Digital behavioral-fingerprint for user attribution in digital forensics: Are we there yet? *Digit. Investig.* **2019**, *30*, 73–89. [[CrossRef](#)]
46. Ikuesan, A.R.; Salleh, M.; Venter, H.S.; Razak, S.A.; Furnell, S.M. A heuristics for HTTP traffic identification in measuring user dissimilarity. *Hum.-Intell. Syst. Integr.* **2020**, *2*, 17–28. [[CrossRef](#)]
47. Kebande, V.R.; Venter, H.S. On digital forensic readiness in the cloud using a distributed agent-based solution: Issues and challenges. *Aust. J. Forensic Sci.* **2018**, *50*, 209–238. [[CrossRef](#)]

48. Kebande, V.R.; Venter, H.S. A comparative analysis of digital forensic readiness models using CFRaaS as a baseline. *Wiley Interdiscip. Rev. Forensic Sci.* **2019**, *1*, e1350. [\[CrossRef\]](#)
49. Setiadi, D.R.I.M. PSNR vs SSIM: Imperceptibility quality assessment for image steganography. *Multimedia Tools Appl.* **2021**, *80*, 8423–8444. [\[CrossRef\]](#)
50. Horé, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
51. Sara, U.; Akter, M.; Uddin, M.S. Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. *J. Comput. Commun.* **2019**, *07*, 8–18. [\[CrossRef\]](#)
52. Abd Warif, N.B.; Idris, M.Y.I.; Wahab, A.W.A.; Salleh, R. An evaluation of Error Level Analysis in image forensics. In Proceedings of the 2015 5th IEEE International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 10–11 August 2015; pp. 23–28.
53. Cha, S.; Kang, U.; Choi, E. The image forensics analysis of jpeg image manipulation (lightning talk). In Proceedings of the 2018 International Conference on Software Security and Assurance (ICSSA), Seoul, Korea, 26–27 July 2018; pp. 82–85.
54. Gunawan, T.S.; Hanafiah, S.A.M.; Kartiwi, M.; Ismail, N.; Za'bah, N.F.; Nordin, A.N. Development of photo forensics algorithm by detecting photoshop manipulation using error level analysis. *Indones. J. Electr. Eng. Comput. Sci.* **2017**, *7*, 131–137. [\[CrossRef\]](#)
55. Jeronymo, D.C.; Borges, Y.C.C.; Coelho, L.D.S. Image forgery detection by semi-automatic wavelet soft-Thresholding with error level analysis. *Expert Syst. Appl.* **2017**, *85*, 348–356. [\[CrossRef\]](#)
56. Parveen, A.; Khan, Z.H.; Ahmad, S.N. Identification of the forged images using image forensic tools. In *Communication and Computing Systems, Proceedings of the 2nd International Conference on Communication and Computing Systems (ICCCS 2018), Gurgaon, India, 1–2 December 2018*; CRC Press: Boca Raton, FL, USA, 2018; p. 39.
57. Zhang, W.; Zhao, C. Exposing Face-Swap Images Based on Deep Learning and ELA Detection. *Proceedings* **2019**, *46*, 29.