

# Integrity and Authenticity of Digital Images by Digital Forensic Analysis of Metadata

Lizbardo Orellano-Benancio<sup>1</sup>, Ricardo Muñoz-Canales<sup>2</sup>, Paolo Rodriguez-Leon<sup>3</sup>, Enrique Lee Huamani<sup>4</sup>

<sup>1,2,3</sup>*Análisis Digital Forense (ADF-Lab), Análisis Digital Forense del Ministerio Público, Lima-Perú*

<sup>4</sup>*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima-Perú*

**Abstract**—During various court hearings, the thesis that every authentic digital file has precise metadata of its creation date was questioned. In this way, the problem was raised which indicates, if the metadata of a digital file (Image) whose label records the date of creation by the recording device of a digital image file are accurate and reliable. For this reason, during the forensic analysis carried out in this work, a record of the metadata of five (05) digital image files from known sources is shown and where their characteristics have been detailed, in addition a record of the metadata of the images used that were later manipulated with image editing software with which metadata comparisons were made to show the labels that suffered modifications in their content. Finally, the obtaining of HASH code with the SHA - 256 algorithm is shown, for digital assurance, of the edited and original files whose comparison allows observing the changes in the content at a binary level.

**Keywords**—Crime; Cybercrime; Digital Image; HASH; Metadata.

## I. INTRODUCTION

Today there are different types of crimes and cybercrimes, where part of the expert work is the forensic analysis of digital evidence. This type of evidence can be data, videos, images, audios, among others. The present experimental investigation focused on the forensic digital analysis of the metadata of five images, knowing the recording device. It is based on the thesis that the expert work of cases in which the validity as proof of digital evidence is discussed is made up of several specialized examinations. Thus, on the one hand, we have the expert work of forensic digital analysis of integrity and authenticity of images, where the first part of the analysis consists of obtaining the metadata and the hash code. On the other hand, the expert work of editing and manipulating digital images requires establishing the conditions of the equipment with which they were captured, so that the expert analysis carried out to date by the present members of the study shows, empirically, the following ways of recording such information:

- *Factory mode*: The configuration date is registered in the memory of the camera recorded during the factory process.
- *User configuration*: The user of the recording equipment configures the date and time according to the corresponding region and time zone; this is exclusive competence of the user.
- By updating or formatting the device's software, certain recording devices suffer alteration of the device's date and time.
- Internal battery change or deterioration of the battery, some image and video capture devices present the factory or user-set date stored in memory during power interruptions caused by internal or interchangeable battery change or when the battery is completely discharged.

The research approach is quantitative, descriptive in scope and experimental in design.

## II. BACKGROUND

In the context of a fiscal investigation, the analysis of the integrity and authenticity of an image is part of the expert work of digital forensic analysis. There are many research works related to this type of analysis and, in particular, to the analysis of metadata.

On the one hand, we find research in which protocols are disclosed to determine whether a digital image is intact and authentic. An example of this type of research is the one developed by Rodríguez-Santos, Delgado-Gutiérrez, Palacios-Luengas and Vasquez-Medina (2015), who present a methodology for detecting modifications in JPEG image files that includes, among other tests, digital assurance and metadata analysis [1]. Another example is the research by Zeng, Shi, Li, and Lu (2016), who present a methodology for image file authentication consisting of techniques based on the analysis of encoding metadata [2].

On the other hand, we find more specialized studies in which the focus is on metadata analysis. A case of this kind of studies is the one developed by Sandoval, Arenas, García and Hernández-Castro (2015), who show anomalies in the EXIF metadata of images created by mobile devices [3]. Another case is the study by Gangwar and Pathania (2018), who used metadata analysis with free computational tools to detect alterations by editing images [4].

The background information reviewed describes important concepts such as metadata, HASH code, modification date and others. It is vital that this terminology has no ambiguity or subjective connotation. With this in mind, forensic research circles, such as the Scientific Working Group on Digital Evidence (SWGDE), propose glossaries of these key concepts.

#### A. Metadata

Data, frequently embedded within a file, that describes a specific file or directory, which may include the locations where content is stored, dates and times, and application-specific information as well as permissions [5].

#### B. HASH code

An established mathematical calculation that generates a numerical value based on input data. This numerical value is called HASH or HASH value [5].

#### C. The date of the recording devices:

- File creation date: this is the date the file was "created". This date does not change whenever the file is opened, closed, saved or modified.
- File access date: this is the date on which the file was last accessed. An access can be a move, an open access or any other simple access.
- File modification date: this is the date on which there has been a change in the file content. For example, if the original content of a text file is modified, the file modification date will be updated.

#### D. The date of the recording devices:

The application of image science and domain expertise to discern whether a challenged image or video is an accurate representation of the original data according to some defined criteria and/or the determination of the original source of the image [6].

#### E. The date of the recording devices:

The process of altering the visual appearance of an image or specific features within an image that results in misrepresentation or misinterpretation [6].

#### F. Image content

Visual information within an image, such as subjects/objects, artifacts (due to compression and/or capture) and physical aspects of the image scene [6].

#### G. Image structure

Non-visual information about the image itself, such as the file type, file compression, metadata, or image source [6].

### III. METHODOLOGY

The methodology used is ISO 27037, which corresponds to the management of digital evidence related to the identification, collection and acquisition of digital evidence.

#### A. Recording Devices Used

On October 25, 2020, 5 images were captured with 5 different recording devices, one (01) video camera, one (01) digital camera, two (02) mobile phones and one (01) tablet; 3 of which had their internal date in factory mode. Table I shows the characteristics of the devices used while Figure 1 shows the recording devices.

**TABLE I**  
**Equipment Features Regarding Recording Devices Used**

Item	Appearance (in Time New Roman or Times)			
	Equipment	Brand	Model	Internal Date
1	Film camera	Sony	HDR-CX240	11/05/2037
2	Photographic Camera	Panasonic	DMC-FH5	01/01/2011
3	Cellphone	Alcatel	5059A	25/10/2020
4	Cellphone	Huawei	MED-LX9	25/10/2020
5	Tablet	Advance	RK30 SDK	01/01/2011



**Figure 1** The architecture of the software framework

### B. Collection and Extraction

The extraction method was Level 1 [12], which corresponds to manual procedures and photographic captures for the extraction of information, which included connection via computer equipment.

The files were extracted from each recording device as shown in Table II and verifying their creation date (see Figure 2).

**TABLE II**  
**FEATURES OF THE COLLECTION AND EXTRACTION EQUIPMENT**

Item	Equipment Employed		
	Equipment	Brand	Name of Captured Image
1	Film camera	Sony	DSC00091.JPG
2	Photographic Camera	Panasonic	P1020021.JPG
3	Cellphone	Alcatel	IMG_20201025_105238
4	Cellphone	Huawei	IMG_20201025_100435
5	Tablet	Advance	IMG_20110101_200214

Unidad USB (H:) > Metadatos

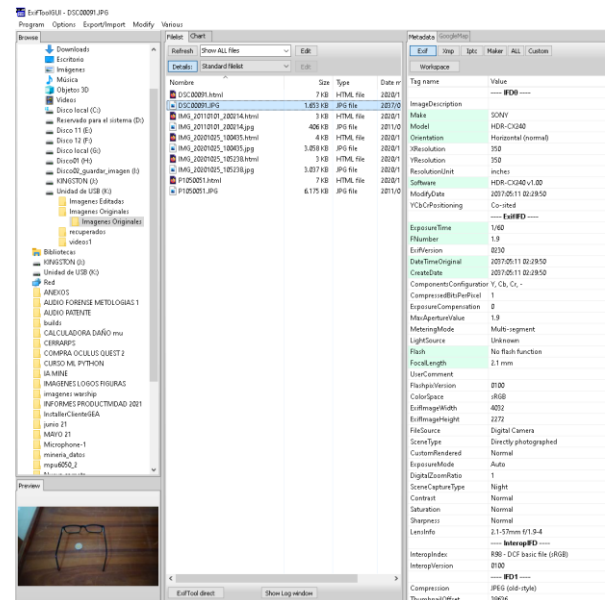
Nombre	Fecha de creación	Fecha de modificación	Tipo
#1050051	25/10/2020 11:58 AM	01/01/2011 12:00 AM	Archivo JPG
IMG_20201025_105238	25/10/2020 11:58 AM	25/10/2020 10:52 AM	Archivo JPG
IMG_20201025_100435	25/10/2020 11:58 AM	25/10/2020 10:04 AM	Archivo JPG
IMG_20110101_200214	25/10/2020 11:58 AM	01/01/2011 8:02 PM	Archivo JPG
DSC00091	25/10/2020 11:58 AM	11/05/2007 2:29 AM	Archivo JPG

**Figure 2** Digital image properties analyzed

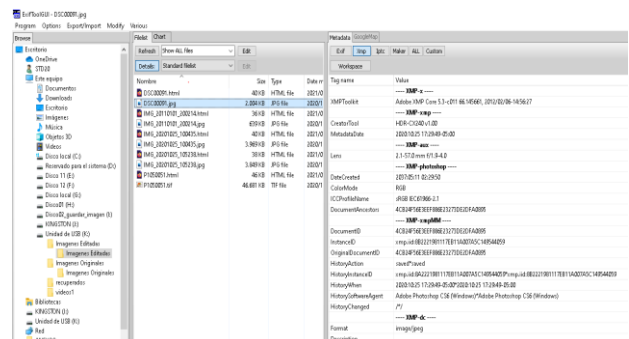
### C. Metadata Collection

As part of the experiment to be carried out in the process of image authentication [7], metadata examination is described exposing the main parameters obtained from the metadata of digital images with the aim of exposing the importance of this information and as an assistance to the research [8].

We proceeded to use the GUI-Exiftool software [7], [9], which is detailed in Figure 3 and Figure 4.



**Figure 3** GUI – Exiftool with Original Image EXIF Metadata



**Figure 4** GUI – Exiftool with Edited Image XMP profile metadata

#### D. Securing

The HashMyFile program is used to extract the HASH code that guarantees the integrity of the information Figure 5.



**Figure 5 Securing with code HASH SHA-256**

#### E. Metadata Analysis

The methodology used is based on the recommendations of the SWGDE (2018) for the analysis of image structure in the authentication process [6]. This methodology consists of the Image File Format Examination, Image Metadata Examination, Image File Packaging Examination, and Image Noise Examination. The image metadata examination consists in turn of the camera brand, model and serial number examination, the examination of the creation and modification date and time, the camera

configuration examination, the image resolution and size examination, Image Processing and History Examination, Original Filename Examination, and other.

Next, metadata were obtained in the form of tags and values which describe technical characteristics of the image capture device, characteristics of the photograph taken, and of the software used for its coding at the time it was generated. The comparison and analysis performed on the tags are presented below.

The metadata obtained referring to the brand, model, and software version of the capture equipment are detailed in the Table III.

**TABLE III  
TAGS AND CHARACTERISTIC VALUES OF  
ORIGINAL IMAGES 1**

Image Name	Equipment Employed			
	Brand	Model	Orientation	Software
1	Sony	HD R-CX240	Horizontal (normal)	HDR-CX240 v1.00
2	Panasonic	DM C-FH5	Horizontal (normal)	Ver.1.0
3	Alcatel	505 9A	Horizontal	Mediatek Camera Application
4	Huawei	ME D-LX9	Horizontal (normal)	MED-L09 10.1.0.134(C 25E8R2P2)
5	Advance	rk30sdk	Horizontal (normal)	-

Metadata referring to camera characteristics and date of creation can be found in Table IV.

## International Journal of Emerging Technology and Advanced Engineering

Website: www.ijetae.com (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 11, Issue 09, September 2021)

**TABLE IV**  
**TAGS AND CHARACTERISTIC VALUES OF ORIGINAL IMAGES 1**

Font Size	Metadata tags 2						
Image Name	Number	Date Time Original	Date of Creation	ISO	Flash	FocalLength	Time
DSC00091.JPG	1.9	2037:05:11 02:29:50	2037:05:11 02:29:50	-	No flash function	2.1 mm	1/60
P1050051.JPG	3.1	0000:00:00 00:00:00	0000:00:00 00:00:00	160	Auto, Fired	5.0 mm	1/60
IMG_20201025_105238	2.0	2020:10:25 10:52:39	2020:10:25 10:52:39	337	No Flash	3.5 mm	1/20
IMG_20201025_100435	1.8	2020:10:25 10:04:37	2020:10:25 10:04:37	809	No Flash	0.0 mm	1/25

Metadata referring to resolution and XMP profile in specific XMP Toolkit [11] tag of the digitized image in Table V.

**TABLE V**  
**Tags And Characteristic Values Of Original Images 3**

Image Name	Metadata tags 3		
	ExifImageWidth	ExifImageHeight	XMP Toolkit
DSC00091.JPG	Sony	HDR-CX240	Horizontal (normal)
P1050051.JPG	Panasonic	DMC-FH5	Horizontal (normal)
IMG_20201025_105238	Alcatel	5059A	Horizontal
IMG_20201025_100435	Huawei	MED-LX9	Horizontal (normal)
IMG_20110101_200214	Advance	rk30sdk	Horizontal (normal)

The information obtained as indicated above corresponds exclusively to non-manipulated digital photographic images, which are verified copies, by SHA 256 hash, obtained from the recording equipment sources.

We have included tags obtained from the metadata that XOMToolkit highlights and that are visible as file system characteristics in Windows systems when observing file properties.

Other tags that refer to the rest of the technical parameters of the captured image and capture camera have not been included in order to limit the analysis to the dates of file creation, resolution, XMP profile that are affected in Adobe Photoshop.

### *F. Editing and manipulation process*

Adobe Photoshop was used to edit and manipulate the files according to the following details:

For the image named "DSC00091.JPG", the first coin from the original image was cloned as shown in Figure 6.



**Figure 6 Cloning part of an image (coin)**

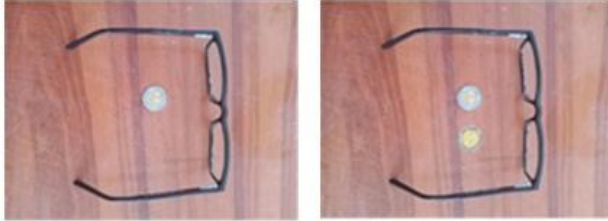
For the image with the name "IMG\_20110101\_200214.JPG", the coin was removed from the original image as follows as shown in Figure 7.



**Figure 7 Removing part of an image (Currency)**



For the image with the name "IMG\_20201025\_105238.JPG", part of the coin of the original image was enhanced as shown in Figure 8.



**Figure 8** Insertion of part of an image (Currency)

For the image with the name "P1050051.JPG", only the format change was made to "P1050051.TIF" as shown in Figure 9.



**Figure 9** Enhancement of part of an image (Currency)

For the image with the name "P1050051.JPG", only the format change was made to "P1050051.TIF" as shown in Figure 10.



**Figure 10** Change from JPG to TIF format

It is noted that the edits applied, did not alter the resolution of the original image, also the Adobe Photoshop software, with default settings, added metadata information in the XMP profile of the GUI-EXIFTTool.

#### IV. RESULTS

The results of the metadata and HASH codes of the edited and manipulated files are detailed:

##### *A. Editing and Manipulation*

The 5 images were edited and manipulated with Adobe Photoshop and the edited image metadata tags and values were obtained as shown in Table VI.

**TABLE VI**  
**TAGS AND VALUES OF EDITED IMAGE CHARACTERISTICS 1**

Image Name	Metadata tags				
	Brand	Model	Orientation	Software	Exposure Time
DSC00091.JPG	SONY	HDR-CX240	Horizontal (normal)	Adobe Photoshop CS6 (Windows)	1/60
P1050051.JPG	Panasonic	DMC-FH5	Horizontal (normal)	Adobe Photoshop CS6 (Windows)	1/60
IMG_20201025_105238	TCL	5059A	Horizontal	Adobe Photoshop CS6 (Windows)	1/20
IMG_20201025_100435	HUAWEI	MED-LX9	Horizontal (normal)	Adobe Photoshop CS6 (Windows)	1/25
IMG_20110101_200214	rk30sdk	rk30sdk	Horizontal (normal)	Adobe Photoshop CS6 (Windows)CS6 (Windows)	1/100

Metadata referring to resolution and XMP profile [11] of the edited digital image in Table VII.

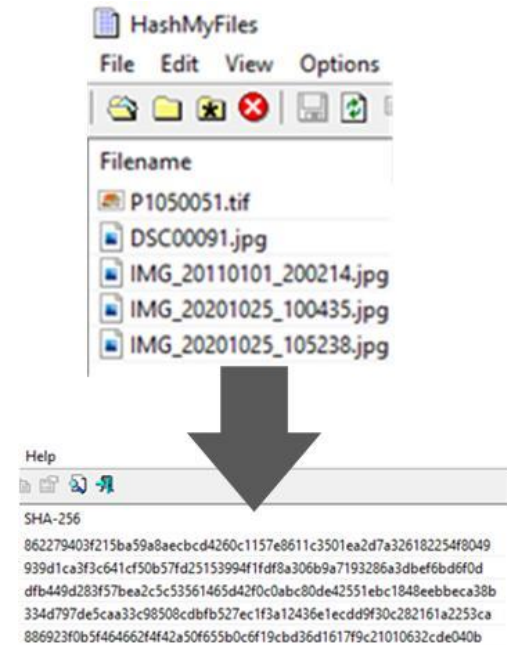
**TABLE VII**  
**TAGS AND VALUES OF EDITED IMAGE CHARACTERISTICS 2**

Image Name	Metadata tags 2		
	ExifImageWidth	ExifImageHeight	XMP Toolkit
DSC00091.JPG	4032	2272	Adobe XMP 5.0
P1050051.JPG	4608	3456	Adobe XMP 5.0
IMG_20201025_105238	3120	4160	Adobe XMP 5.0
IMG_20201025_100435	4160	3120	Adobe XMP 5.0
IMG_20110101_200214	1600	1200	Adobe XMP 5.0

As observed in Table VI and Table VII, there are variations in two groups of tags.

- *Software*: version of the software used for image processing.
- *XMP Toolkit*: Extended metadata profile used by Adobe Photoshop.

The remaining metadata of the digital images of the five images used, in the other parameters shown in Table VI and Table VII no changes were observed in the information except for the XMP profiles, it is detailed that the default configuration of the Adobe Photoshop software was used since it allows adding metadata to the images processed in its environment. In Figure 11 it is verified that the HASH codes of the edited images correspond to the HASH codes of the extracted and secured original files as detailed in Figure 5.



**Figure 11 HASH SHA-256 code of the edited and manipulated images**

## International Journal of Emerging Technology and Advanced Engineering

Website: [www.ijetae.com](http://www.ijetae.com) (E-ISSN 2250-2459, Scopus Indexed, ISO 9001:2008 Certified Journal, Volume 11, Issue 09, September 2021)

### V. CONCLUSIONS

From the five (05) extracted images it is concluded that the date of capture of the image depends on the date of configuration and/or state of the battery or internal battery that stores the configured information among which are the date and time, verifying that some recording devices of the present experiment had factory dates due to the discharge of the internal power supply.

Every device with digital image capture capability has as a property the assignment of a specific alphanumeric format for the assignment of the name of the recorded digital image file, concluding that it depends on the brand and model of the recording device.

The HASH codes and metadata will be affected depending on the editing or manipulation process they undergo, preserving in the metadata in some cases certain characteristics of the original file and certain characteristics of the editing or manipulation software among other data.

In the original digital image P1050051.JPG changed its HASH code when the format was changed to P1050051.TIF, which does not alter the context of the image, which could be part of a future research Figure 12.



**Figure 12 Original image and image with format change**

Likewise, it is important to emphasize that the original digital image IMG\_20201025\_105235.jpg was enhanced generating a new image with the same name Figure 13, verifying the variation of the HASH code, which could be part of a future investigation for image enhancement with the use of forensic or commercial software.



**Figure 13 Original image and enhanced image**

### REFERENCES

- [1] F. Rodríguez-Santos, G. Delgado-Gutiérrez, L. Palacios-Luengas y R. Vazquez-Medina, "Practical implementation of a methodology for digital images authentication using forensics techniques", *ACSIJ Advances in Computer Science: an International Journal*, Vol. 4, Issue 6, No.18, November 2015, ISSN : 2322-5157.
- [2] J. Zeng, S. Shi, Y. Li y Q. Lu, "Practical inspection workflow for digital image forensic authentication," 2016 4th International Symposium on Digital Forensic and Security (ISDFS), 2016, pp. 172-172, doi: 10.1109/ISDFS.2016.7473540.
- [3] A. Sandoval, D. Arenas, L. García y J. Hernández, "Analysis of errors in exif metadata on mobile devices", *Multimed Tools Appl* 74, 4735–4763 (2015). <https://doi.org/10.1007/s11042-013-1837-6>.
- [4] D. Gangwar, A. Pathania, "Authentication of Digital Image using Exif Metadata and Decoding Properties", *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, ISSN : 2456-3307, Volume 3, Issue 8, pp.335-341, November-December-2018. Available at doi: <https://doi.org/10.32628/CSEIT183815>.
- [5] Scientific Working Group on Digital Evidence, "SWGDE Digital & Multimedia Evidence Glossary", 2016, <https://drive.google.com/file/d/1ZZwOqgVOWo6qDeoJqv6VKafY2i1RJi2B/view>.
- [6] Scientific Working Group on Digital Evidence, "SWGDE Best Practices for Image Authentication", 2018, <https://drive.google.com/file/d/1Z0DsJMa6aDZIFJ9kRfOL8cow5VjhVT0t/view>.
- [7] ENFSI Best Practice Manual for Digital Image Authentication (2021) ENFSI-BPM-DI-02.
- [8] Galvan, F., &Battiato, S. IMAGE/VIDEO FORENSICS: THEORETICAL BACKGROUND, METHODS AND BEST PRACTICES Part three–Tools for operational scenarios.
- [9] Metodología para el Análisis Forense de las Evidencias Electrónicas, CA: AENOR, 2013.
- [10] Olivier, M. S. (2009). On metadata context in database forensics. *Digital Investigation*, 5(3-4), 115-123.
- [11] Mullan, P., Riess, C., &Freiling, F. (2019). Forensic source identification using JPEG image headers: The case of smartphones. *Digital Investigation*, 28, S68-S76.
- [12] J Reyes-Rodriguez, "Going deeper and deeper into cell phones" (NIST), 2019, <https://bit.ly/3mLucAS>