International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 8 Issue 3 March - 2024, Pages: 17-27

An Improved GCFIM Framework for Analyzing Digital Evidence Steganography

Noura Hamad[1] and Mahmoud Jazzar[2]

1College of Information Technology, Palestine Technical University Khadoori, Palestine
n.j.hamad2@students.ptuk.edu.ps
2College of Information Technology, Palestine Technical University Khadoori, Palestine
m.jazzar@ptuk.edu.ps

Abstract*— Criminals employ steganography, a form of anti-forensics, to conceal information within other files, complicating the retrieval of original evidence in digital crimes and impeding investigations. Digital forensic analysts must employ appropriate tools to uncover hidden messages. This research aims to detect concealed files in digital evidence using steganography analysis techniques. The study utilizes the Improved Generic Forensics Investigation Model framework, comprising seven stages: pre-process, collection and preservation, examination, analysis, reporting, presentation, and post-process. Tools such as FexImager for extracting forensic images from the digital evidence, Hiderman, and StegSpy were employed specifically for steganography analysis, while OSForensics and WinHex were utilized for forensic analysis. The results demonstrate the effectiveness of StegSpy and Hiderman in identifying the steganography in 18 files out of 20, and the OSForensics tool in detecting mismatched files. Furthermore, this experiment provides empirical evidence supporting the proficiency of the Improved Generic Forensics Investigation Model in steganography detection.*

Keywords— Anti-forensics techniques, Computer Forensics tools, General Computer Forensic Investigation Model, Steganography tools, Steganalysis.

1. INTRODUCTION

With the rapid growth of information technology and the internet, there has been an increase in the number of electronic crimes. Many criminals today profit from the use of information and communication technology in executing their offenses. Anti-forensic techniques are often used by digital criminals, making it even harder to discover digital evidence. Steganography is often recognized as an anti-forensic technology, It's the skill of writing in a way that's hidden or concealed. The purpose of steganography is to hide a message from a third party [1]. Criminals conceal their messages inside assets such as documents, audio, pictures, or videos. As they conceal the hidden file holding certain data, these assets act as the carrier[2].

Although the carrier file appears and functions normally, its information is corrupted. A significant distinction between cryptography and steganography is that Cryptography renders data indecipherable and incomprehensible, yet the ciphertext remains visible to human eyes [3]. In contrast, steganography enables the concealment of information in plain sight, utilizing a diverse range of secret information formats such as images, text, audio, video, and files. In other words, steganography conceals or hides information within carriers, concealing the fact that there is even a message there at all. While, cryptography conceals the contents of a message using algorithms such as RSA, AES, DES, and so on [4].

Steganalysis is to identify and detect suspected files, determine whether or not they have a payload encoded into them, and if possible recover that payload [5]. During investigation, steganographic information should be examined first. Steganography investigators should also be familiar with typical steganographic methods, software, tools, terminology, and websites. Steganography's procedure, software, and tactics will be easier to decipher with this information [6]. Forensic technologies were developed to discover criminals by gathering evidence based on a digital forensic model that was classified in general into stages: identification, collection, preservation, analysis, report, and presentation [7]. The purpose of this research is to detect and evaluate evidence collected from the suspect device and analyze various forms of files, including text, audio, pictures, and video formats, that have been concealed by criminals using steganography methods. An Improved General Computer Forensic Investigation Model framework will accomplish this goal. When it comes to obtaining relevant data for investigations, investigators require forensic software. Forensic software is sometimes multi-purpose, with the ability to carry out a variety of functions inside a single program. Computer forensic software enhances law enforcement's ability to access and evaluate digital evidence from suspicious computers. In this research, we will use computer forensic tools (FEXImager, OSForensics, WinHex), and steganography analysis tools (Hiderman, StegSpy).

2. Literature Review

Due to the significant increase in Internet usage, particularly in recent years, there has been a corresponding rise in concealment techniques and their prevalence. This poses a challenge for investigators seeking to uncover the truth. Consequently, numerous studies have been conducted in the realm of steganography and its detection methods to aid investigators in situations involving crimes. As per a study conducted by [8], one of the anti-forensic methods used by criminals to conceal information in other communications is steganography, which may create difficulties in the investigation and challenges in getting original evidence of the digital crime. For digital forensic analysts to be effective, they must be able to use appropriate techniques to locate and retrieve information that has been injected. Steganography methods were used in the study to examine the digital evidence that has been concealed. The static forensics approach is used to extract files that have been entered based on digital criminal case scenarios, and the five steps of the GCFI Model framework were used. They utilized Autopsy and WinHex as well as Hiderman and StegSpy. Based on the findings of a 20-file steganographic file insertion experiment, it seems that StegSpy and Hiderman are reliable tools for the steganographic examination of digital evidence in court. StegSpy has an 85% success rate in detecting hidden messages. The Hiderman tool was used to get steganographic messages out of 18 files, and the process of getting the messages out worked perfectly every time.

According to [9], As steganography becomes more widely employed in the digital world, there are several difficulties that computer forensic examiners should be aware of. A vast variety of tools and approaches are available, each with its strengths and disadvantages. Adaptations should be done gradually and using the most up-to-date information available. At first, the steganography overview gives the computer forensic examiner a general idea of what steganography is and how it works. It's possible that using tools to keep track of data changes would raise red flags, too. Steganography's methods and algorithms are dissected in detail to have a better grasp of how it works. There is a great deal of interest in the LSB approach since it is often employed in digital picture steganography and has just a little impact on the image's real color. This makes it almost impossible to see with the naked eye. This experiment's findings show that even if steganography tools A and B have similar qualities and techniques, using one of them for data extraction isn't conceivable because of the experiment's limitations. Forensic investigators must be aware of the type of steganography program installed, concealed, or erased from the victim's computer, as indicated by this study. This awareness is crucial because obtaining proof that the suspect uses a particular steganography application may raise doubts. These doubts become apparent during subsequent inquiries into locating concealed files on a computer. Nevertheless, experimental results demonstrate that investigators need to know what kind of steganography tool was used to extract the hidden information. According to [10], the paper presents a novel steganography technique utilizing justified text in PDF files. It employs Huffman coding to compress the secret message, embedding it within specific lines of the cover text. By replacing added spaces with normal spaces in host lines, the method maintains the file size and printability. Its inconspicuous nature safeguards against suspicion, preserving the text's originality without introducing grammatical errors. Compatible with various text sources, it outperforms other text steganography methods in average capacity percentage. Notably, it requires no advanced programming skills and supports all languages. Based on previous studies, in this research, we will repeat the experiment in [8] using the same tools of steganography, but we will try other computer forensic tools for analysis. This research aims to discover and examine evidence concealed by criminals utilizing steganography techniques in various file types such as text, audio, image, and video. The improved GCFIM framework and forensic tools are utilized to extract digital evidence data to submit it to the court.

3. METHODOLOGY

3.1 Case Stages

GCFIM consists of five stages including Pre-Processing, Acquisition & Preservation, Analysis, Presentation, and Post-Process [11], Still, it needs to be improved to enhance the steganography investigation process. Since the GCFIM Framework needed some development, we've made some additions to better illustrate the computer investigation process. Figure 1 shows the improved GCFIM framework's seven stages: pre-processing, collection and preservation, examination, analysis, reporting, presentation, and post-process stage. As shown in Figure 1 . Here are the explanations of every stage:

3.1.1- Pre-Processing stage:

This phase includes all of the preparatory work that must be done before the investigation and the formal gathering of data begins. Approval from the proper authorities and preparation for the utilized tools are among those activities.

3.1.2-Collection & Preservation stage

This stage included finding, gathering, storing, and preserving the electronic devices found at the crime scene which may contain data such as computers, smartphones, and USB drives. As a general rule, throughout this stage, the investigator should gather and protect any essential findings in preparation for the next step.

3.1.3- Examination Stage

Data acquisition and data extraction are the two main operations in this phase [12]. The kind of data source, such as a computer, smartphone, USB, or other electronic devices, must be considered while gathering data using data acquisition tools, software, and hardware. Each data source has a unique way of gathering information, so there are a variety of approaches. A digital forensic workstation's software and tools are used to conduct the data extraction stage. The obtained data is then used for further analysis and extraction of relevant evidence.

3.1.4-Analysis Stage

This phase is the heart and the main part of computer forensic investigation, which is what we're going to talk about. The analysis is done on the data to figure out where the crime came from and who did it, examination of the data acquired to determine the origin of the crime and the crime's purpose, and then the person who did it can be found.
The results of the analysis phase are recorded and given to the appropriate authorities.

3.1.5- Reporting Stage:

At this phase, a comprehensive report is made based on the analysis of the results. This report includes the methods used to get the outcomes, a clarification of how the tools and methods were chosen, and a statement about whether other methods need to be done. The goal of the report is to give factual information to the right people and help them figure out what happened. Also, the report should be written in a way that makes it easy for people who aren't experts in the digital forensic field to understand it. The verbal report tells the final results about whether or not there is proof to help the investigation. Also, the report explains any facts that aren't clear.

3.1.6-Presentation Stage:

Forensic analysts send the report they made in the previous step to the right authorities, and investigators to understand whether it is acceptable or not in court. This step is critical since the argument must not only be presented in a way that is easily understood by the person to whom it is being given but it must also be backed by appropriate and accepted evidence. This phase's major output is to either prove or reject the suspected illegal activities.

3.1.7- Post-Process Stage:

When the investigation is done correctly, this phase is about how to finish it off right. Digital or physical evidence needs to be given back to the rightful owner, or kept in a safe place if necessary. There should be a review of the investigative process so that the lesson can be learned and used to make the next investigation even better. As shown in Figure 1, Instead of proceeding sequentially from one phase to the next, the thing that makes this framework different from others is the opportunity to return to earlier stages must always be available[8]. We are dealing with scenarios that are always evolving in terms of crime scenes (both physical and digital), investigation tools employed, criminal instruments used, and investigators' degree of skill. As a result, it is very desirable to return to prior stages that we have completed, not only to remedy any shortcomings but also to learn new information.

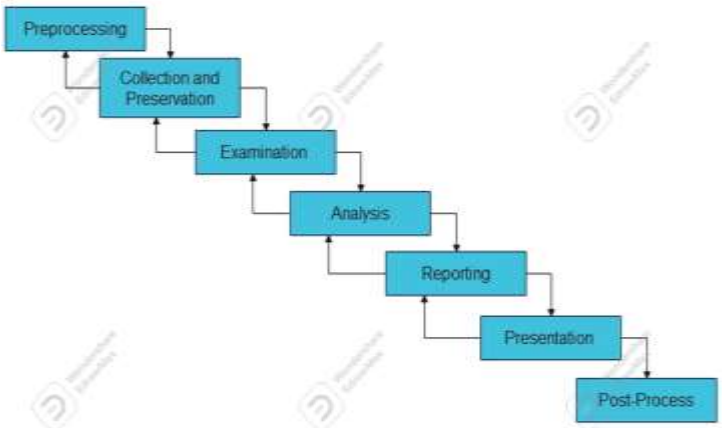

Figure 1: An Improved Generic Computer Forensic Investigation Framework for steganography analysis

Following the explanation of the stages of the enhanced GCFIM Framework in Figures 1 and 5, it is necessary to test the effectiveness of this Framework by applying it in a case scenario.

3.1 The Case Scenario

There are several types of steganography such as Image steganography, audio steganography, video steganography, and text steganography[13]. According to the basic model of steganography[14], as shown in Figure 2, the scenario of this case was conducted dependent on it.

Twenty different files were prepared in the formats of text, images, audio, and video files, with some of these files' attributes changed like the extension of files. Then a secret message(file) was hidden in most of the container files using Hiderman software. As shown in Figure 3, these files were finally stored on a USB to make a forensic image using the FEX Imager tool, to analyze this image, we use computer forensic tools like OSForensics and WinHex.

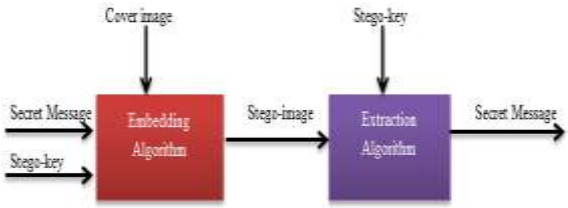

Figure2. The basic model of steganography [14]

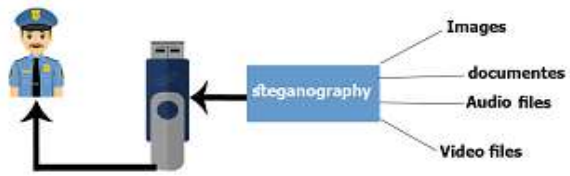

Figure 3. The Case Scenario

The scenario is executed using the Hiderman software. A text message is inserted into container files with many formats, including documents, movies, photos, and audio files, to produce a stego item, which is then saved on flash disk storage media. This case introduced files containing steganographic text. The processing time required to conceal files depends on the size of the file placed. The duration of the insertion operation is proportional to the file size. Figure 4, shows the use of the Hiderman tool to hide the text in every file to produce a stego item.

Figure 4. Hiding data using the Hiderman tool

4 . IMPLEMENTATION AND RESULTS

In this section, the investigator follows the steps in Figure 5. Initially, after receiving the investigation order, he prepares the official papers and proceeds to the crime scene to understand the circumstances of the crime. Subsequently, he collects and preserves the digital evidence from the crime scene. Next, he progresses to capturing a forensic image using the FEX Imager tool from the digital evidence, which in this scenario is a USB stick for analysis. He then utilizes this image for digital forensic analysis, employing tools such as OSForensics and WinHex. Following the analysis, he compiles a comprehensive report containing all the results obtained, which is intended to be presented in court for judgment of the suspect, either for acquittal or conviction.

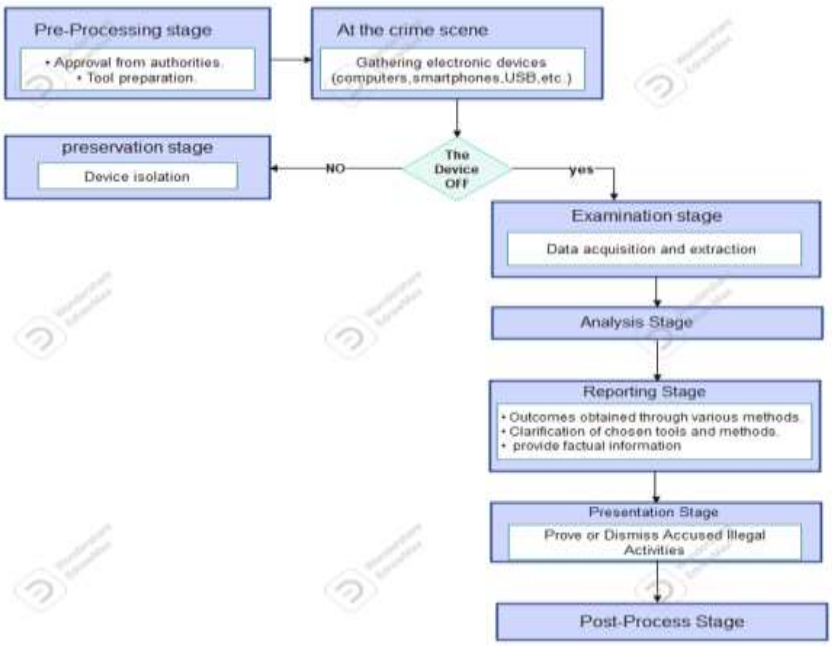

Figure 5. The improved GCFIM flowchart

A- Collection & Preservation

This is the first step in the process of identifying evidence in a crime scene, which is followed by the process of collecting and preserving the evidence's originality. The goal is to protect the evidence against changes in the physical form or data by preserving it in a secure place. The FEX Imager tool is used to gather data from physical evidence (USB) to produce a forensic image,as shown in Figure 6.

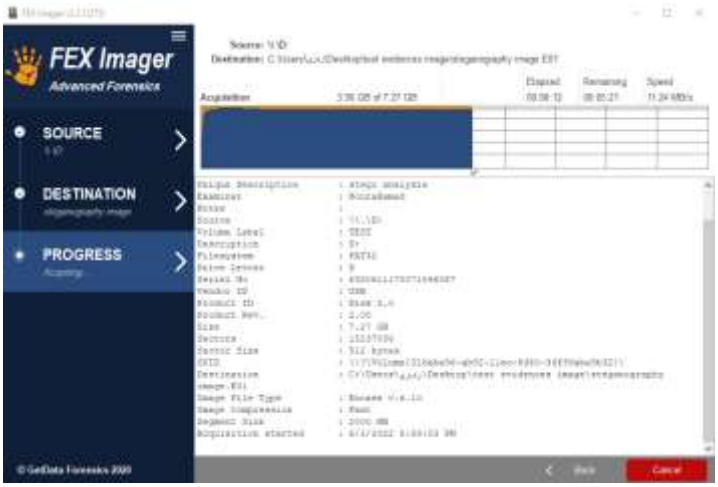

Figure 6. Making a forensic image using FEX Imager

To compare the MD5 and SHA1 hash values of the original evidence with the MD5 and SHA1 hash values of the forensic image that has been obtained, The FEX Imager tool is used for that purpose. The MD5 and SHA1 hash values of the original evidence file and image file are matched as shown in Figure 7.

Figure7. Testing MD5 and SHA values

B- Examination and analysis

After obtaining the forensic image, it will be examined to extract the evidence and analyze it using computer forensic tools such as OSForensics and WinHex. One of the most important and best programs used in computer digital forensics is the OSForensics tool. The forensic image has been added to extract data from it and analyze it. One of the most important features offered by the OSForensics tool is showing file properties, including mismatched files. Using this feature, all mismatched files appear as shown in Figure 8, so we can access metadata such as creation and modification dates for all files. Then, we will extract these files to analyze them using Stegspy and WinHex, as shown in Figure 9.

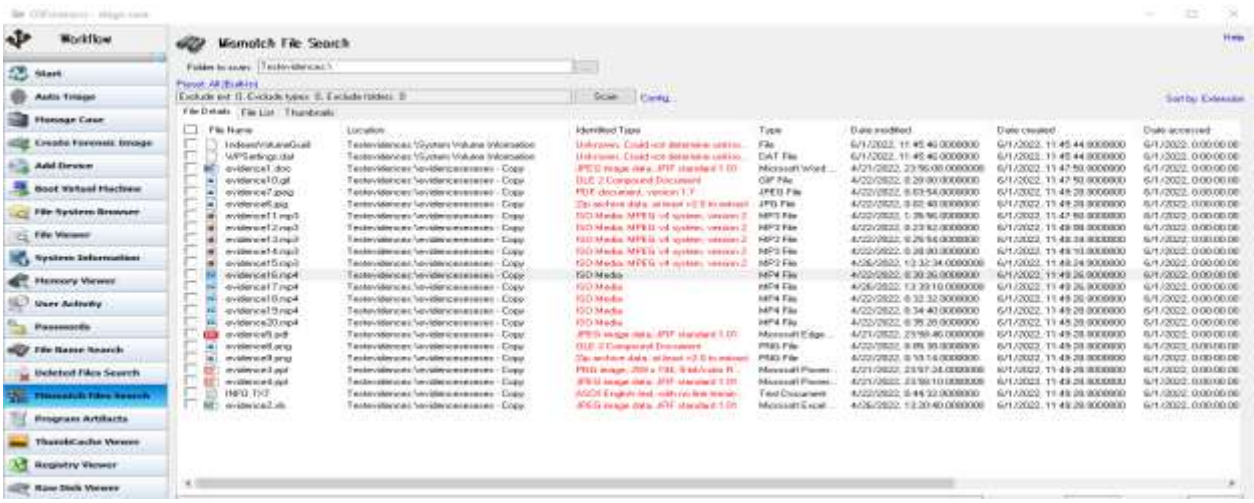

Figure8. Viewing all the mismatched files using the OSForensics tool.

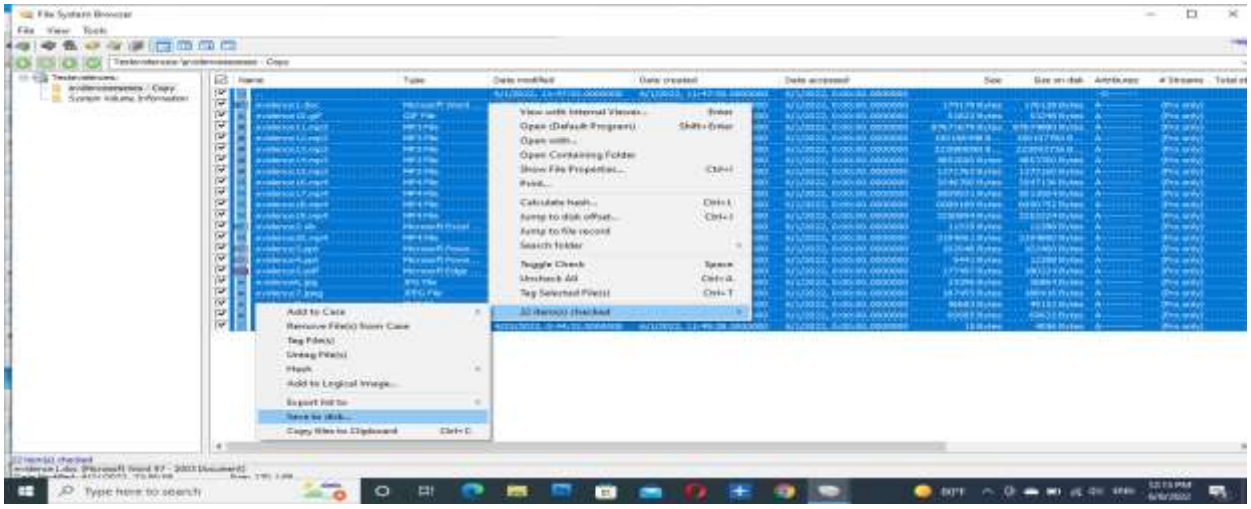

Figure9 . Files extraction using the OSForensics tool.

After extracting these files, we will use the Stegspy tool to examine if the files contain steganography or not. The detection of steganography shown in Figure 10.

Figure 10. Detecting steganography using Stegspy tool

Table 1 shows the secret file analysis findings. 18 out of 20 files tested contained steganographic data. It also shows that StegSpy was able to identify steganographic files that were hidden in different file formats and gave information about the marker values it found.

File type	filename	The original file format	found/not found	marker
image	Evidence1	.jpg	found	174177
	Evidence 2	.jpg	Not Found	--
	Evidence3	.png	found	102044
	Evidence4	jpg	found	9439
	Evidence5	.jpeg	found	177495
document	Evidence6	.xls	found	33394
	Evidence7	.Pdf	found	187453
	Evidence8	.ppt	found	46681
	Evidence9	.doc	found	69583
	Evidence10	.doc	found	51820
video	Evidence11	.Mp4	Not Found	--
	Evidence12	.Mp4	found	357491
	Evidence13	.Mp4	found	3250892
	Evidence14	.Mp4	found	4852598
	Evidence15	.Mp4	found	1371761
audio	Evidence16	.Mp3	found	3546698
	Evidence17	.Mp3	found	414758
	Evidence18	.Mp3	found	6089187
	Evidence19	.Mp3	found	574223
	Evidence20	.Mp3	found	3194059
document	info	.txt		--

Table 1. Steganography file analysis results

To complete the extraction procedure, we must input the key or password obtained from the contents of the data file. Every extracted file will be inserted into the Hiderman tool, with the password “Trial”, to extract the hidden messages from it, as shown in Figures 11 and 12.

Figure11. Extract the hidden files

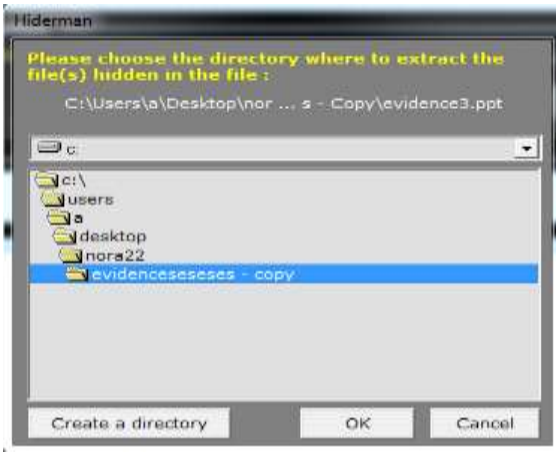

Figure12. Save the extracted hidden files

C - Analysis of the Extracted Hidden Files

After utilizing Stegspy and Hiderman to detect and extract concealed files from the suspicious files obtained from the suspect's USB drive, investigators successfully uncovered secret messages hidden within these files. where these messages were obscured using steganography techniques such as the Hiderman tool in this scenario. Subsequently, investigators subjected these messages to additional scrutiny using WinHex, a forensic tool. WinHex enabled the examination of hash values associated with each message file, providing investigators with the means to verify the integrity of the extracted data, as shown in Table 2.

Msg	Format	Hash (MD5)
Message 1	**txt**	**923CA44B00F09E9FAEF6D9E7C300EC6B**
Message 2	**txt**	**DD8A2211746DD2189DE0A12E34C980FA**
Message 3	**txt**	**9416C5929721FDB61B9CB0B5A1477159**
Message 4	**txt**	**67BFA40E9CFDFD78ED86B5BF000897E7**
Message 5	**txt**	**67BFA40E9CFDFD78ED86B5BF000897E7**
Message 6	**txt**	**C11044D3C0C7B68A9580DB4D48308BA9**
Message 7	**txt**	**3EF3084E3701B22E7386D1ED3137488E**
Message 8	**txt**	**9F28DC577A6BBBDE5681C58D384A0089**
Message 9	**txt**	**530DC3FE4763F88D8920B62F968F949E**
Message 10	**pdf**	**FEA581E35C5454F2094FEA3443BFD658**
Message 11	**txt**	**A93FDDC028D6BC734E117EDF3BF2B048**
Message 12	**txt**	**9277CF828943EB70B3CEA31890008AD0**
Message 13	**pub**	**E9CC5187300F2AC11DA0A8DF9C51BEA6**
Message 14	**txt**	**6871EE354ADF9B378CBD48499325A40F**
Message 15	**ppt**	**9E0E0EA0BADC8AB1209C80305C4FDCC9**
Message 16	**txt**	**CCB66F0C23EE20F5AC32B026046F53B2**
Message 17	**Pdf**	**9225BF0FEF27332E4995B5ED7B558366**
Message 18	**txt**	**BE0C7180AE43409967E9D728E2E466EF**

Table 2: The hidden files hash value.

Using the WinHex tool, investigators can examine the header and footer of every file. If any text appears in the ASCII part after the end of the files, it may suggest the presence of hidden or embedded files within the container. This indicates that additional data is stored beyond what is initially visible, potentially indicating the use of steganography or other concealment techniques, as shown in Table 3, Figures 13 and 14.

File	Header	Footer
JPG	**FF D8 FF**	**FF D9**
PNG	**89 50 4e 47**	**49 45 4e 44 42 60 82**
Docx	**50 4B 03 04 14 00 06 00**	**50 4B 05 06**
PDF	**25 50 44 46**	**25 25 45 4F 46**

Table3. Files Header and Footer

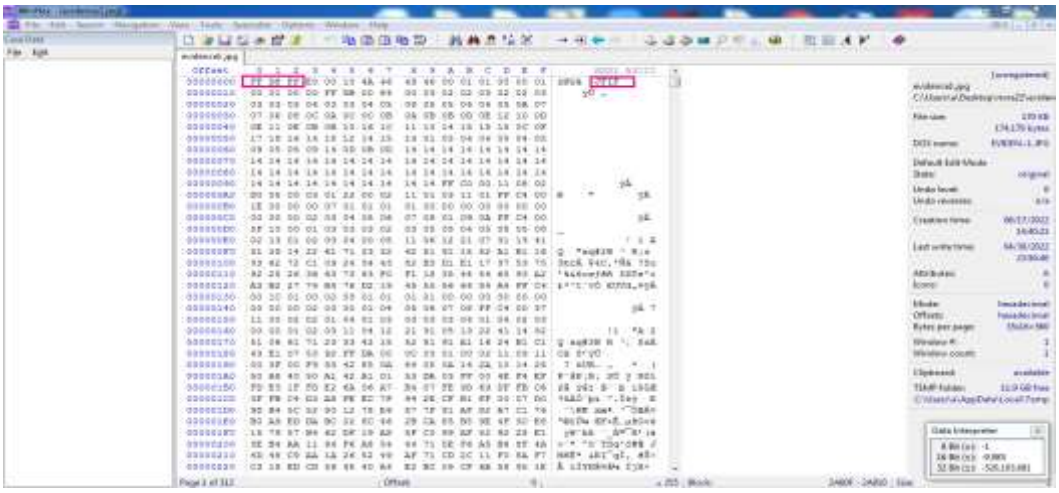

Figure 13. The header of the JPG file.

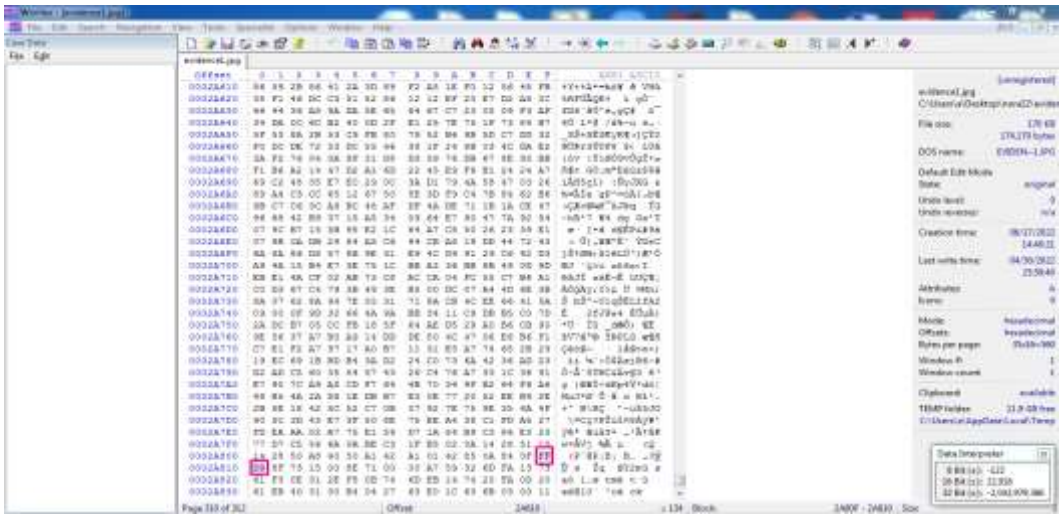

Figure14. The footer of the JPG file

D – Reporting and presentation

After completing the collection, extraction, and analysis of the obtained information, the investigator compiles a detailed report for submission and presentation in court. The report must be clear, concise, and unbiased, catering to both technical and non-technical audiences, with the investigator prepared to testify and provide clarification in court.

5- CONCLUSION

As steganography becomes more widely employed in the digital world, several difficulties need to be addressed in computer forensic analysis. There are various tools and approaches, each with its own set of strengths and shortcomings. Continuous adaptation and the integration of more recent methodologies are essential. Investigators must be proficient in analyzing anti-forensic techniques such as steganography. This study employs a static forensics approach, utilizing the improved Generic Computer Forensic Investigation Model framework. The Hiderman, Stegspy, OSForensics, and Winhex tools complemented each other effectively in detecting and extracting the hidden files. However, identifying evidence that a suspect used a specific steganography tool may yield uncertain results, as the use of steganography itself can obscure the presence of concealed files.

As demonstrated in the experiment, investigators need a deep understanding of steganography tools to successfully extract hidden information. For future work, exploring different file types and utilizing alternative steganography and forensic tools will be pursued to enhance detection capabilities.

6 - REFERENCES

[1] An Overview of Steganography for the Computer Forensics Examiner. Steganography for the Computer Forensics Examiner. (n.d.). Retrieved May 23, 2022, from https://www.garykessler.net/library/fsc_stego.html.

[2] J. N. Cheltha C, M. Rakhra, R. Kumar and H. Walia, "A Review on Data hiding using Steganography and Cryptography," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2021, pp. 1-4, doi: 10.1109/ICRITO51393.2021.9596531.

[3] Subramanian, N., Elharrouss, O., Al-Maadeed, S., & Bouridane, A. (2021). Image steganography: A review of the recent advances. IEEE Access, 9, 23409-23423.

[4] Almuhammadi, S., & Al-Shaaby, A. (2017). A Survey on Recent Approaches Combining Cryptography and Steganography. 63–74. https://doi.org/10.5121/csit.2017.70306.

[5] Yari, I. A., & Zargari, S. (2018). An Overview and Computer Forensic Challenges in Image Steganography. Proceedings - 2017 IEEE International Conference on Internet of Things, IEEE Green Computing and Communications, IEEE Cyber, Physical and Social Computing, IEEE Smart Data, IThings-GreenCom-CPSCom-SmartData 2017, 2018-January (February), 360–364. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.60

[6] Panhalkar, T. (2020, July 20). Detecting steganography. Infosavvy Security and IT Management Training. Retrieved April 15, 2022, from https://info-savvy.com/detecting-steganography

[7] Hamad, N., & Eleyan, D. (2022). Digital Forensics Tools Used in Cybercrime Investigation-Comparative Analysis. 113–127. https://doi.org/https://doi.org/10.37896/JXAT14.04/314909

[8] Hajar Akbar, M., Ahmad, U., Yogyakarta, D., & Sunardi, I. (2020). Analysis of Steganographic on Digital Evidence using General Computer Forensic Investigation Model Framework. In IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 11, Issue 11). www.ijacsa.thesai.org

[9] Yari, I. A., & Zargari, S. (2017, June). An overview and computer forensic challenges in image steganography. In 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (pp. 360-364). IEEE.

[10] Khosravi, B., Khosravi, B., Khosravi, B., & Nazarkardeh, K. (2019). A new method for pdf steganography in justified texts. Journal of information security and applications, 45, 61-70. https://doi.org/10.1016/j.jisa.2019.01.003.

[11] Yusoff, Y., Ismail, R., & Hassan, Z. (2011). Common Phases of Computer Forensics Investigation Models. International Journal of Computer Science and Information Technology, 3(3), 17–31. https://doi.org/10.5121/ijcsit.2011.3302.

[12] Khweiled, R., & Jazzar, M. (2021). An Improved Framework For Cyberbullying Investigation Process on WhatsApp application. J. Xi'an Univ. Archit. Technol., 13(9), 238-246. DOI: 10.37896/JXAT13.9/313822.

[13] Rahman, S., Uddin, J., Zakarya, M., Hussain, H., Khan, A. A., Ahmed, A., & Haleem, M. (2023). A Comprehensive Study of Digital Image Steganographic Techniques. IEEE Access, 11(January), 6770–6791. https://doi.org/10.1109/ACCESS.2023.3237393

[14] Emam, M. M., Aly, A. A., & Omara, F. A. (2016). An Improved Image Steganography Method Based on LSB Technique with Random Pixel Selection. IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 7). Retrieved from www.ijacsa.thesai.org