# Technical Framework for File Carving Tools

**[1]Ziad Saif Alrobieh,[2]Ali Mohammed Abdullah Ali Raqpan**

*[1]Department of Communication & Computer Engineering, Alsaeed Faculty for Engineering & Information Technology, Taiz University, Taiz, Republic of Yemen;*

*[2]Department of Information Technology, Alsaeed Faculty for Engineering & Information Technology, Taiz University, Taiz, Republic of Yemen;*

ziadrh@yahoo.com ; ali1raqpan@gmail.com

## ABSTRACT

Governments and societies seek to combat crime of all kinds, including electronic crimes, which are where computing has a fundamental corner in it, such as if the evidence is inside the computer or the crime has been committed with it, and for this several techniques have emerged to deal with electronic evidence, including File Carving technology, which is a technology that works to restore files without relying on the structure of the file system and this is what distinguishes them from traditional restoration. This technology has several tools, each tool depends on different algorithms as these algorithms are applied to the file depending on several requirements, including header, footer, or the contents of the file , And on this diversity and difference in algorithms and methods has led to the variation of each tool from the other, and this difference can be known by performing measurements experiments on tools, through a set of dataset to obtain several factors, namely the speed of recovery operations, recall, precision and The number of files retrieved. properly show the disadvantages of these tools, such as fragmentation has been working on this address defects developed ways through which the restoration, such as Smart Carving, and this technology is still under work and employment and development.

**Keywords:** File Carving, Digital Evidence, Fragmented Files, Metadata, File Carving Algorithms, File Carving Methodologies, File Carving Tools.

## 1. Introduction

The science of Digital Forensic has become one of the most advanced sciences in the field of research and innovation[2], it depends on techniques that are constantly changing as a result of the increase in cybercrime, among these technologies is the mechanism of recovery, which has significantly evolved from the traditional method to a set of innovative methods that do not depend on the file system such as File Carving Technique , despite the strength and importance of this technology, it faced problems and difficulties, but on the other hand, work continued to devise solutions to these problems, whether technical or Programmatic. The tools for this technology developed after it was the first tool Foremost that was invented by the US Department of Defense, and thus it is considered the core of File Carving technology. During emergence of the information revolution, which was initially aimed at making the world a small village, but with the development of means of communication and the Internet and the development of storage methods, types of crimes appeared, which are electronic crimes, which made governments and organizations work to combat them, the first work was to seek to develop electronic criminal investigation science (Forensics Science)[3], The beginning of this development was in the year 1984 when the US Federal Bureau of Investigation intervened and worked to establish the Computer Analysis and Response Team (CART), and work began to adopt the approach to developing this field, and several techniques appeared to combat cyber-crime that work to prove the crime with evidence[1]. At the beginning of the nineties of the twentieth century a tendency appeared to create a mechanism for calculating the electronic evidence, which was the first for Daubert, with the passage of time a technology emerged that works to retrieve files without relying on the metadata of the operating system and the file system, which is File Carving technique got rid of One of the traditional restoration restrictions, which appeared at the end of the last century, and those interested in this field sought to develop it With this technology, the first event was held in 2006 under the auspices of DFRWS then followed in 2007, the aim was to develop tools and mechanisms for this technology, which faced several problems that were worked on to confront and solve them. New methods were invented, for example what was revealed by both Pal & Memon, which is known as (Smart Carving), from this acceleration in development, many experiments and measurements emerged on these tools to know the performance and efficiency of these tools. Perhaps the most prominent of these studies is what S.J.J.Kloet did in 2007 to find out how Evaluating these tools ,how they work, and obtaining the results of evaluating the tools.Many studies and research have continued to this date in this field, and the wheel of development is still continuing.The reminder of this paper is organized as follows. Section 2 File Carving Technology ,Section 3 Methodologies of Measures, Section 4 Experiments and Results and Section 5 Conclusion.

## 2. File Carving Technology

### 2.1 Recovery and File Carving:

Traditional Recovery: It is a file system-based mechanism to recover deleted files, and most file systems do not change the physical file location during the deletion mechanism. Therefore, it appears that the defect of this mechanism is that in the event that the file system structure is disrupted or deleted, the files cannot be restored and cannot deal with the Data-Set that is not from the file system structure. Therefore, it cannot be used in electronic investigations and the need for other techniques arose.

File Carving Technology: It is one of the file recovery techniques that are used in electronic investigations depending on the file structure and content without relying on the file system, and files are recovered from the unallocated space [6]. File carving is a difficult and complex process, so it can be used for the following problems:

1. Upon deletion, when the file is deleted, the file system indexes become unmarked to the file's content, and in this case the content remains unchanged until the clusters that contained this content are overwritten by another file.
2. When the files are in storage media or devices and their file system is undefined, here the importance of this technology emerges as it is independent and does not depend on the file system.
3. With hidden files, which are files that are not shown to the file system.
4. File Carving Technology identifies and retrieves files from the original data that have been deleted or destroyed in the file system, memory, or data resulting from the "Swap-Space" process.

## 2.2 File Carving Algorithms:

In file carving technology, String Matching Algorithm is used because the file system metadata does not give us a description of the location of each file within file system[5]. Its purpose is to find a clear pattern. The string matching algorithms have two main goals, namely:

1- Reducing the number of symbol comparisons.
2- Reducing the time required in case analysis.

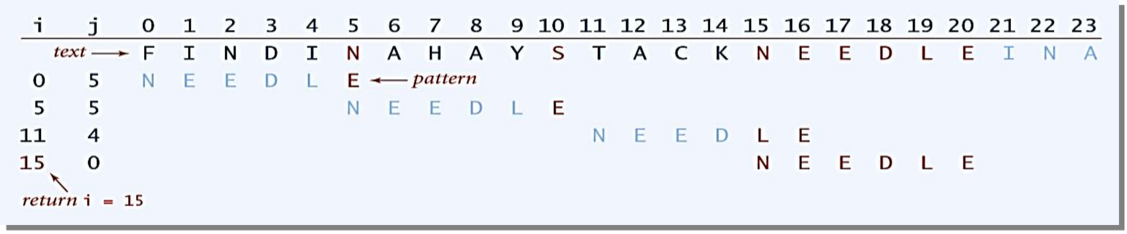
Most algorithms have two stages. The first stage is the preprocessing stage of a group of patterns and the second stage is the search for samples. There are several applications of string matching algorithms such as search engines, computer security, bio-information, and DNA analyzes, in addition to their use in file carving technology of two types:

1- Fast: It is designed to search for one text from the sample at a time, and it may also be known as “Single”.
2- Multiple patterns: They search for a group of samples simultaneously, which is faster than the “Fast” pattern.

Several algorithms fall under the fast chain matching, including:

- The Brute Force Algorithm: It is one of the oldest and simplest string matching algorithms and may be known as "Naive". The idea of this algorithm is based on a comparison between the sample "Pattern" and the "Text" string, where the letter is taken from "Pattern" and compared with a character from "Text" and when Matching these two letters moves to the next letter in both "Pattern" and "Text". If the match does not happen, there is a cursor moving to the next letter in "Text" and there is another cursor in "Pattern" that returns to the beginning, and we compare again and continue the process until all "Pattern" characters match with a part of "Text" or access End of "Text", which is why "Time-Complexity" is equal to O (NM).
- Boyer-Moore Algorithm: The Boyer-Moore algorithm has two basic approaches, which are the detection of "bad character heuristic" and the discovery of "good suffix heuristic". This algorithm is contrary to most pattern comparison algorithms as it performs matching from the last letter in the pattern, i.e. from right to left. It is based on the "sub-linear" concept, meaning that it does not scan every symbol in the text. As for the bad character detection method, it works to match the pattern with the string with which the patterns want to be identified. If the pattern does not match the text string, we determine the location of the mismatch, and the symbol in this place is called the bad character, so we displace this pattern until the match occurs or the pattern skips the place of the mismatch. This method has "Time-Complexity" equal to O (n / m) where (m) is the length of the pattern in the best case, but in the worst case it is O (nm) and occurs when the string of text matches the pattern. As for the method of good suffixes, two patterns are matched, for

example A, B, since A contains a branched string (a). This string is compared with a branched string of B and to be (b), and the patterns are shifted until a match is achieved with(b), or (a) series of B antecedents' matches string (a), or B skips string (a).Scalpel uses the "Boyer-Moore" algorithm  to define the header and footer of the data.



| i | j | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
|  | *text* → | F | I | N | D | I | N | A | H | A | Y | S | T | A | C | K | N | E | E | D | L | E | I | N | A |
| 0 | 5 | N | E | E | D | L | E ← *pattern* |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 | 5 |  |  |  |  |  | N | E | E | D | L | E |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 11 | 4 |  |  |  |  |  |  |  |  |  |  |  | N | E | E | D | L | E |  |  |  |  |  |  |  |
| 15 | 0 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | N | E | E | D | L | E |  |  |  |

*return* i = 15

Figure 2.1 Boyer-Moore Algorithm

- The "Knuth-Morris-Pratt" Algorithm: It is a faster algorithm than "Brute Force" and was invented in 1975 to improve it by eliminating duplicate comparison. Where we first compare the first letter of "Pattern" with the first letter of "Text", and if they are different, we go to the next site of "Text", but if we compare three letters of "Text" with "Pattern" and they are correct, but the letter differs The fourth here in this case we have compared the previously identical letters so we go to the next site of "Text" which has a similarity of these letters that we matched, so the operations are less and the algorithm is faster, which is the linear time-O (N + M) algorithm.
- Raita Algorithm: This algorithm searches for the sample in the text string by comparing each "Character" symbol from the sample to the given text, but the mechanism is as follows:

1- The part of the text "string" is defined in the form of a window and its size is the size of the "Pattern" sample.

2- The sample is compared with the window from the text by taking the last symbol from the sample with the last symbol from the window.

3- If a match is made, we move to the first symbol of the sample and compare it with the first symbol from the window.
4- If the matching is also done, the middle symbol from the sample is compared with the middle symbol of the window. Steps (1) to (4) are known in the "Preprocessing" stage.
5- If the previous steps are successfully completed, the actual comparison process will start, which is from the second symbol to the last symbol, but on one symbol each time.
6- If no matching occurs, a process of shifting the window occurs until it finds the last element in it that matches the last element of the sample, and so the algorithm continues until the text string ends.
7- The amount of displacement for a mismatch is known as "Bad Character Shift".

It is considered the development and improvement of the performance of the “Boyer-Moore” algorithm and the “Hors-pool” algorithm. This algorithm was published in 1991 by “Timo Raita”, and for the “Preprocessing” phase it is O (m) and for the search phase it is “Time Complexity” O (mn), Where (m) is the sample size, and (n) is the textual series size. There are other algorithms such as the "Berry-Ravindran" algorithm and the "Karp-Rabin" algorithm, which were invented in 1987 and are different from the rest of the algorithms as they use "Hashing" techniques, the "Hors-pool" algorithm, and the "Quick Search" algorithm. "Shift-OR" algorithm, "The Smith" algorithm, and these algorithms are mostly not used in file carving tools, although they are similar to the previous algorithms, but they may be used in the future. Several algorithms fall under the multi-pattern string matching, including:

- Commeritz-Walter Algorithm: It is a combination of the “Boyer-Moore” algorithm and the “Aho-Corasick” algorithm. In the “Preprocessing” stage, the “Aho-Corasick” algorithm is used, but in a slightly different way, and in the research phase, the “Boyer-Moore” algorithm is used.

- Wu-Manber Algorithm: It is an upgrade from the "Boyer-Moore" algorithm.
- Aho-Corasick Algorithm :It resembles a tree, but with links between the internal nodes, and it consists of two parts, as the first part matches the finite samples, which were formed from a group of "string" texts called "pre-processing", and in the second part the group of texts is applied as inputs for samples of matching, when the match occurs, a signal is made.This algorithm is an extension and development of the "Knuth-Morris-Pratt" algorithm. Its idea is based on the creation of a "Finite Automaton" that uses a set of words. It basically collects all samples in a "tree" and then turns into a "non-deterministic automaton (NFA)" Then it turned into "deterministic automaton (DFA)".The upgraded version of scalpel now uses the "Aho-Corasick" algorithm with multiple modes, which makes it much faster in prospecting (carving).

### 2.3 File Carving Methodologies:

There are several methods and mechanisms used by the file carving tools when excavating, and these mechanisms include:

- ❖ Header Based method: there are more than one method that differs in terms of the mechanism and is similar in terms of head use [7], including:
  - Header-Footer method: It is used to search for files with a specific header, which is one of the signs and symbols that distinguish the beginning of each file, as well as the footer that distinguishes the end of the file, and it has several defects, most notably:
    1- The difference in the number of bytes that contain the header and the footer, some of them may be long and others may be short, such as (.jpeg) images where the header and footer consist of (2-bytes), and this difference

may lead to the appearance of False Positive during the recovery process.

2- This method may not be compatible with fragmented files, that is, it may link a file header to the footer of another file.
3- This method may not be able to recover some files because it does not have a fixed header and footer, such as text files and "HTML".

- Header / Maximum Size method: a method that uses the header of the file and then takes the highest possible value for the file space. This method is a practical with some file formats such as (JPEG, MP3) because these formats may have some fragments attached to them at the end of the file, and this method has the same disadvantages of the "Header-Footer" method, in addition to:

  1- You may recover a file size larger than the original file size.
  2- You may recover a small part of the original file, and this applies to connected files that may be restored incorrectly if they are larger than (Maximum Size).

- Header / Embedded Length method: depends on a specific level of knowledge of the internal file structure for all file types.

The header and the footer is every file that contains the values or what are known as magic numbers, which are in the first byte and the last byte, and through it the work of restoring all the blocks between them, but not all files have a footer, determines the end of the file (EOF) to be difficult. The term Header: it is a unique number. It may also be known as Signatures. These numbers are like a fingerprint as each type of file has a special number that is at the beginning of the file structure. Some files do not have a footer such as MS-OFFICE, others do not have a header or even a footer, such as some text files (.txt) and some e-mail messages. These files are self-defined and do not need a header or footer.

Table 2.1 file types, their extension, header and footer

| File Type | Extension | Header | Footer |
|---|---|---|---|
| Mpeg4 | .mp4 | 00 00 00 18 66 74 79 70 6D 70 34 32 | - |
| WinRAR ZIP Archive | .zip | PK\x03\x04 | \x3c\xac |
| Icon File | .ico | \x00\x00\x01\x00 | - |
| PDF | .pdf | 25 50 44 46 | %EOF |
| Image GIF | .gif | \x47\x49\x46\x38\x37\x61 | \x00\x3b |
| Image PNG | .png | \x89\x50\x4E\x47 | - |
| MIDI File | .mid | \x4D\x54\x68\x46 | - |
| Hyper Text Mark-Up Links | .htm | <html> | </html> |
| Image JPEG | .jpg | \xff\xd8\xff\xe0\x00\x10 | \xff\xd9 |
| PST | .pst | !BDN | - |
| Text Documents | .txt | Begin\040PGP | - |
| WAV | .wav | 52494646 (RIFF) | - |

❖ File Structure Method: In this method, it is necessary to understand the contents of the file to build a general knowledge of the internal file structure, It is a method that helps collect the files that have been fragmented, meaning that it is a solution to the problem of fragmentation, and it is similar to the "Header / Embedded Length" method, but the latter did not solve the problem of fragmentation, and this method may be called "Deep-Carving" or the name "Semantic Carving" [8], it is used by the Scalpel & Photorec tools, it depends on the use of the inner layer of the file such as the header and footer as well as Size information. If the file is not fragmented and the File Structure Data is intact, then using this method gives excellent results, and it may help to know the fragmentation occurring or any defect. As a result of using this method, Known False Positive or False Negative may appear as a result of using this method.

- Fragment-Recovery Carving method: it works to reconfigure the parts to the original file, and this method may be called Split Carving.
- Bi-fragment Gap Carving method: It works on determining the data between two parts of the file that has the fragmentation, by doing a set of tests to determine it. It may be considered an improvement to the Split Carving method. This method may be used rarely for image and video files, but it has disadvantages, namely:
    - It is incapable of dealing with nonlinear fragmentation, as it dictates that the fragmentation be equally in the file.
    - It does not work with files that are split into more than two parts.
- Block-Based Carving method: It is a method that analyzes the input in the form of a block and then another block (block-by-block) to determine whether the block is part of the file or an additive to the file, and it may also be known as Content-Based [9] such as Loose-Structure for example (MBOX, HTML, XML).

  Since the storage media, such as the hard disk, the data is stored on Sectors paths of these discs, and thus the partitioning occurs on the boundaries of these paths. Therefore, this method depends on examining each data block to see that it belongs to this part of the file to be examined.

  There is a flaw in this technique and how to find a mechanism to calculate the values that distinguish between the data blocks that belong to the file and that do not belong.
- Repacking Carving method: It modifies the extracted data by adding a new header or footer or adding any other information. This method has been used by the tool (Garfinkel's ZIP carver).
- File Validation Method: It works to verify that the recovered data is intact and identical to the data to be restored, including:

    - The statistical method (Statically Carving): It is a method for analyzing the input graphically or statistically such as

Entropy to determine if the input is part of the extracted file.

- By inspecting and optimizing each block or cluster file [10].

## 2.4 Smart Carving:

Smart Carving techniques, which is one of the modern methods in this field, which also helped in solving the problem of fragmented files through three main processes that take place before restoration, namely pre-processing, collecting, and then recombining the file (Reassembly) [11].

1- Pre-processing: It is the first process where it is concerned with working on decrypting files if they are encrypted and decompressing compressed files in order to facilitate dealing with files directly. In this process (Allocated and Unallocated) clusters are determined, and upon completion of these operations All three, the file is ready to move to the next process.

2- Collecting and Comparison: Similar clusters are classified according to the File Signature factor, as each file depending on its format, has this unique factor. Here, string matching algorithms can be relied upon, or entropy can be taken as it helps to reveal the probability of the file type.

3- Reassembly: the parts are properly arranged, then joined and combined with some, such as the original file, and then the output of this assembly is checked whether it is correct or not.

## 2.5 File Carving Operations:

They are operations that evidence or file to be restored goes through and explain the integration between the human element and the automatic methods[12].

- Preprocessing: it is extracting information about the file by specifying the beginning and end of the file and the length of the file size, as well as specifying all the sectors in which the parts of the file are located.
- Assembly: by generating a copy of the file temporarily, taking all sectors and arranging them in a sequential order and is acceptable.
- Discriminator: Ensure the correctness of the results and avoid mistakes.

These operations are carried out in a series, depending on the machine or electronic devices, and there are operations in which the human element directly interferes (Acquire Image, Verification, Investigation) and this integration between the machine and the human element achieves the success of file carving process.

## 2.6 File Carving Tools:

File carving tools have the best performance in electronic criminal investigation tools for several factors, including[4]:

1- The percentage of files that are able to recover them.
2- Accuracy and reliability of the outputs of these tools.
3- The speed of processing carried out by the carving tools.

File carving tools, especially open source, can work on any operating system even if the metadata for that system is destructive, and for this reason, the controversy between open source and closed source software remains constant on the level of the support mechanism, security and philosophy of structure and reliability.

Table 2.2 (types of tools and their compatibility with operating systems)

| Tool name | Licenses | Operating System | Modifiable | Interface |
|---|---|---|---|---|
| Scalpel | Open source | Windows, Linux, Mac-OS X | Yes | Script |
| Encase | Proprietary | Windows, Mac-OS X | No | Interfaces |
| Foremost | Open source | Linux | No | Script |
| FTK | Proprietary | Windows | Yes | Interfaces |
| Autopsy | Proprietary | Windows, Linux, Mac-OS X | No | Interfaces |

| | | | | |
|---|---|---|---|---|
| Photorec | Open source | Windows, Linux, Mac-OS X | Yes | Script<br>& interfaces |
| Adroit | Proprietary | Windows, Mac-OS X | No | interfaces |
| Bulk Extractor | Open source | Windows, Linux, Mac-OS X | No | Script<br>& interfaces |
| EVTxtract | Open source | Windows, Linux | No | Script |
| Forensics Exploral | Proprietary | Windows | Yes | Interfaces |
| Defraser | Proprietary | Windows | No | Script<br>& interfaces |

## 3. Methodologies of Measures:

### 3.1 Methods for Testing File Carving Tools:

1- Procedures and process tests: It is the work of analyzing and testing the tools and comparing the results with previous tests and knowing the differences between each test and concluding the reasons for the difference.

2- Performance measures: It is the ability of the tools to properly carving files[13].

$$\text{Carving Recall } (Cr) = (All-Sfn-Ufn) / All) \qquad (3.6)$$

$$\text{Supported Recall } (Sr) = (Sp-Sfn) / Sp \qquad (3.7)$$

$$\text{Carving Precision } (Cp) = tp / (tp + Ufp + (0.5 * Kfp)) \qquad (3.8)$$

$$\text{F-measure } (Fm) = 1 / (\alpha * (1 / Cp) + (1- \alpha) * (1 / Cr)) \qquad (3.9)$$

- All: The total number of files in Data Set.
- Supported File (Sp): The total number of files type in "Data Set" that are available and supported by the tool.
- True Positive (tp): the file that has been properly recovered from Data Set. To find out the (tp) values, we calculate the value of MD5 Hash in the tool for data types and compare it with MD5 Hash in Data Set Documentation.
- False Positive (fp): The recovered file is but not intact and it is of two types:
    - Known (Kfp): are files that are recovered but improperly. They are defined and supported by the carving tool.
    - Unknown (Ufp): are files that are recovered but incorrectly and are not defined or supported by the carving tool.
- False Negative (fn): It is part of the files that have not been recovered and is divided into two types:
    - Supported (Sfn): It is the part of the file that was not recovered through the tool and it is supported by the tool.
    - Unsupported (Ufn): This is a part of the file that was not recovered by the tool and it is not supported by the tool.
- Alpha (α): It is a factor used to denote the importance of Recall and Precision and it may be equal to (0.5).

3- Data set: This is done by making a copied image from storage media such as the hard disk or storage memory through the Imaging process. This process is done through a set of tools or text commands. Each Data Set is distinguished by its name and type, as well as MD5 Hash Value and File Location. There are types of data sets available on websites that are available for scientific measurement processes, including "11-carve-fat.dd" .

### 3.2 Image Storage Media:

The image identical to the digital evidence (Digital Forensics Imaging): It is a term for the process by which it creates an image identical to the medium in which the evidence is stored for the examination and analysis process of the investigation so that there is no change to the contents of the original evidence. There are several types of storage media images that are used in electronic criminal investigations, including (dd), which is known as (Split Raw Images), as well as (Expert Witness Format -EWF) and also WHX File (.whx) and other types, and most of file carving tools, especially open source they were built to work on the type (dd). Among the most popular tools for creating images corresponding to storage media are: Encase , Pro-Discover , FTK Imager ,Klennet-Imager ,Guymager . When creating an imaging, the tool must be stable and not make any changes to the storage media, without errors, and to prevent Bit-Stream Duplicate. FTK Imaging: It is an evidence acquisition tool that is used to make a preview of the evidence and make a copy of the storage media.(dd) image: is an image identical to the storage media in terms of size and type of the storage medium and this type has problems, the most prominent of which is if it deals with large data with a size of more than a terabyte, and because of this this type of image has been divided into a group of Segments and with this segments are easy to handle and easy to store.(EWF) image: It has become the most used now, as the image is divided into a series of sections (E01, E02, E03, ……). it is used by the tools FTK and Encase.

## 4. Experiments and Results:

### 4.1 Self-Experiments:

Experiments that were applied to images of various storage media that were created for the purpose of the study as well as to the storage medium itself in order to know how to deal with different images as well as how each tool deals with the storage medium and the goal was to highlight the effectiveness of the tools on the images of different storage media. Device: Dell Latitude E6430.Operating System Version:

Windows 10 Pro, Kali GNU / Linux Rolling (Version 3.26.2).Processor: Intel Core i5-3340M, 2.70 (GHz).RAM: 8 (GB).System Type: 64- (Bit).

Table 4.١ (Tools of Experiments and Results )

| Tool | Kind of Media | Carving Time | Total Number of Files | Files Restored | True Positive | Recall | Precision | $F_{score}$ |
|---|---|---|---|---|---|---|---|---|
| Autopsy 4.14.0 | Media Image EWF | 160s | ٢٥٠ | ١٦٠ | ٩٠ | ٠.٦٤ | ٠.٥٦٢٥ | ٠.٥٩٨٧٥٣ |
| | | Results Estimates | | | | Mediocre | Mediocre | Mediocre |
| | Media Image DD | 385s | ٢٥٠ | ١٥٩ | ٩٣ | ٠.٦٣٦ | ٠.٥٨٤٩ | ٠.٦٠٩٣٨١ |
| | | Results Estimates | | | | Mediocre | Mediocre | Mediocre |
| | Media Image WHX File | 99s | ٢٥٠ | ٠ | ٠ | 0 | - | - |
| | | Results Estimates | | | | Bad | - | - |
| | Raw Image | 211s | ٢٥٠ | ١٥٩ | ٩٣ | ٠.٦٣٦ | ٠.٥٨٤٩ | ٠.٦٠٩٣٨١ |
| | | Results Estimates | | | | Mediocre | Mediocre | Mediocre |
| | USB Flash 963 (MB) | 344s | ٢٥٠ | ١٥٩ | ٩٣ | ٠.٦٣٦ | ٠.٥٨٤٩ | ٠.٦٠٩٣٨١ |
| | | Results Estimates | | | | Mediocre | Mediocre | Mediocre |
| photorec_ win 7.0 | Media Image EWF | tool did not recognize an image | | | | | | |
| | Media Image DD | tool did not recognize an image | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Media Image WHX File | tool did not recognize an image | | | | | | |
| | Raw Image | 11s | ٢٥٠ | ١٤٩ | ١٤٧ | 0.596 | 0.986577 | 0.743092 |
| | | Results Estimates | | | | Mediocre | Almost Perfect | Mediocre |
| | USB Flash 963 (MB) | 10s | ٢٥٠ | ١٥٠ | ١٤٩ | 0.6 | 0.993333 | 0.748117 |
| | | Results Estimates | | | | Mediocre | Almost Perfect | Good |
| foremost version 1.5.7 | Media Image EWF | 44s | ٢٥٠ | 3 | 3 | 0.012 | 1 | 0.023715 |
| | | Results Estimates | | | | Bad | Excellent | Bad |
| | Media Image DD | 59s | ٢٥٠ | 168 | 45 | 0.672 | 0.267857 | 0.023715 |
| | | Results Estimates | | | | Mediocre | Bad | Bad |
| | Media Image WHX File | 53s | ٢٥٠ | 1 | 1 | 0.672 | 1 | 0.007968 |
| | | Results Estimates | | | | Bad | Excellent | Bad |
| | Raw Image | 55s | ٢٥٠ | 163 | 43 | 0.652 | 0.263804 | 0.375627 |
| | | Results Estimates | | | | Mediocre | Bad | Bad |
| | USB Flash 963 (MB) | 65s | ٢٥٠ | 173 | 65 | 0.692 | 0.375723 | 0.487018 |
| | | Results Estimates | | | | Mediocre | Bad | Bad |
| photorec 7.0 | Media Image EWF | tool did not recognize an image | | | | | | |
| | Media | ٢٤٨s | ٢٥٠ | 146 | 144 | 0.584 | 0.986303 | 0.733618 |
| | | Results Estimates | | | | Medioc | Almost | Good |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Image DD | | | | | re | Perfect | |
| | Media Image WHX File | 11s | ٢٥٠ | 1 | 1 | 0.004 | 1 | 0.007968 |
| | | Results Estimates | | | | Bad | Excellent | Bad |
| | Raw Image | 295s | ٢٥٠ | 146 | 144 | 0.584 | 0.986303 | 0.733618 |
| | | Results Estimates | | | | Mediocre | Almost Perfect | Good |
| | USB Flash 963 (MB) | 298s | ٢٥٠ | 146 | 144 | 0.584 | 0.986303 | 0.733618 |
| | | Results Estimates | | | | Mediocre | Almost Perfect | Good |
| scalpel-2.0 | Media Image EWF | 65s | ٢٥٠ | 55 | 46 | 0.22 | 0.836364 | 0.348365 |
| | | Results Estimates | | | | Bad | Good | Bad |
| | Media Image DD | 83s | ٢٥٠ | 169 | 87 | 0.676 | 0.514793 | 0.584485 |
| | | Results Estimates | | | | Mediocre | Mediocre | Mediocre |
| | Media Image WHX File | 61s | ٢٥٠ | 7 | 5 | 0.028 | 0.02 | 0.023333 |
| | | Results Estimates | | | | Bad | Bad | Bad |
| | Raw Image | 80s | ٢٥٠ | 183 | 71 | 0.732 | 0.284 | 0.409228 |
| | | Results Estimates | | | | Mediocre | Bad | Bad |
| | USB Flash 963 (MB) | 88s | ٢٥٠ | 186 | 71 | 0.744 | 0.284 | 0.411081 |
| | | Results Estimates | | | | Mediocre | Bad | Bad |

## 4.2 Comparative Experiences:

It is to carry out an experiment on "Data Set" for which previous experiences have occurred and compare the values of results with previous values and build results through comparison to find out the essence of the difference. We have relied on "11-carve-fat.dd", as an experiment was conducted on it by "Kloet" in In the year 2007 as in Table (1.4), then experiments were conducted on it by "Thomas Laurenson" in the year 2013 and compared with the results of "Kloet" as in Table (2.4), as it depended on the changes that appeared in the measurement parameters, and we are in the process of this study we followed The same approach in terms of measuring mechanisms and determining results, which was in Table (3.4).

Table 4.٢ (Kloet 2007 results)

| Tool | Performance Measurements | | |
|---|---|---|---|
| | Recall | Precision | Fscore |
| Scalpel | 1 | 0.003 | 0.01 |
| | Excellent | Bad | Bad |
| Foremost | 0.933 | 0.786 | 0.85 |
| | Very Good | Good | Very Good |
| Photorec | 0.931 | 0.857 | 0.89 |
| | Very Good | Very Good | Very Good |

Table 4.٣ (Thomas Laurenson 2013 results)

| Tool | Performance Measurements | | |
|---|---|---|---|
| | Recall | Precision | Fscore |
| Scalpel | 0.860 ↓ | 0.917↑ | 0.854↑ |
| | Very Good | Very Good | Very Good |
| Foremost | 0.708↓ | 1↑ | 0.829↓ |
| | Mediocre | Excellent | Good |
| Photorec | 0.933↑ | 1↑ | 0966↑ |

| | | | |
|---|---|---|---|
| | Very Good | Excellent | Almost Perfect |

Table 4.٤ (Results of the current study 2020)

| Tool | Performance Measurements | | |
|---|---|---|---|
| | Recall | Precision | Fscore |
| Scalpel | 0.860↓- | 0.917↑- | 0.854↑- |
| | Very Good | Very Good | Very Good |
| Foremost | 0.933-↑ | 1↑- | 0.965↑↑ |
| | Very Good | Excellent | Almost Perfect |
| Photorec | ١↑↑ | 0.9375↑↓ | 0.968↑↑ |
| | Excellent | Very Good | Almost Perfect |

## 5. Conclusion:

This chapter comes to clarify the outcome of the study from several applied and research aspects in a way that summarizes the study in specific lines that talk about the essence of the topic, as this chapter deals with the conclusions created by the study and the difficulties faced by the study as well as the future work related to the study and what the study owners aspire to. This technology has helped many governments and organizations that rely on electronic investigations to solve cases, and its tools have become the most used tools as they give high-accuracy results when retrieving evidence, as well as a result of its continuous development and work to solve the problems of this technology, as well as inventing new mechanisms, and it is useful to know that File Carving technology, when mechanisms that do not depend on metadata have been adopted, have become more flexible and effective than traditional restoration, since each operating system has its own metadata, and a special file system, which leads to the inability to recover if the system itself is deleted, Several algorithms have been improved and developed, such as String Matching Algorithm, which in turn helped innovate many new carving methodologies, and with all these changes the need for evaluation mechanisms for the technology and its tools emerged, so it adopted a methodology based on three factors: (Recall-Precision-Fscore) in addition to additional factors, such as the time of execution of the restoration process or the speed of the tool in execution, etc., but in this study we added a new factor, which is the type of image in the storage medium, where there are tools that are not There is a type of image and it supports others to a lesser and higher degree. The importance of the images in this process is that they keep a copy of the evidence and work on it in order not to harm the original evidence.

## REFERENCES

[1]. B. V. Prasanthi "Cyber Forensic Tools: A Review" , International Journal of Engineering Trends and Technology (IJETT) – Volume-41 , ISSN: 2231-5381, Number-5 - November 2016, https://www.ijettjournal.org.

[2]. B. Carrier, "Open Source Digital Forensics Tools the Legal Argument", carrier@cerias.purdue.edu .

[3]. A. Sivaprasad, Prof. S. Jangale, "A Complete Study on Tools & Techniques for Digital Forensic Analysis", Information Technology V.E.S.I.T Mumbai, India, abi.lecturer@gmail.com, smitajangale@yahoo.com.

[4]. T. Laurenson, "Performance Analysis of File Carving Tools", 28th Security and Privacy Protection in Information Processing Systems (SEC), Jul 2013, Auckland, New Zealand, pp.419-433,10.1007/978-3-642-39218-4_31, hal-0146384.

[5]. K. Hulin, "Digital Forensics III - File Carving", Department of Computer Science The University of Texas at Dallas, September 23rd, 2011.

[6]. A. Pal and N. Memon, "The Evolution of File Carving [The benefits and problems of forensics recovery]", IEEE SIGNAL PROCESSING MAGAZINE, 1053-5888/09/$25.00©2009IEEE, MARCH 2009.

[7]. Alexey V. Gubin, "Klennet Storage Software", 2017 – 2019.

[8]. L. Miguel Pereira Constantino Romano, “File carving in practice”, Universidade do Minho Escola de Engenharia Departamento de Inform´atica, October2015.

[9]. S. Kamal Ahmad Khalid, R. Raad Ali, K. Malik Mohamed, S. Jamel, “A REVIEW OF DIGITAL FORENSICS METHODS FOR JPEG FILE CARVING”, Journal of Theoretical and Applied Information Technology 15th September 2018, Vol.96. No 17, ISSN: 1992-8645, www.jatit.org .

[10]. Cor J. Veenman, “Statistical Disk Cluster Classification for File Carving”, Conference Paper: August 2007, DOI: 10.1109/ISIAS.2007.4299805.

[11]. A. Nur Elmi Abdi, “MP4-Karver: CARVING OF CORRUPTED MP4 VIDEOS USING ASMD REPAIRING TECHNIQUE”, Faculty of Computer Science and Information Technology Universiti Tun Hussein Onn Malaysia, July 2016.

[12]. G. Richard, V. Roussev, “Scalpel: A Frugal, High Performance File Carver”, The Digital Forensic Research Conference DFRWS 2005 USA New Orleans, LA (Aug 17th – 19th), http:/dfrws.org.

[13]. S.J.J. Kloet, “Measuring and Improving the Quality of File Carving Methods”, Eindhoven University of Technology - Department of Mathematics and Computer Science, Almere, October 29, 2007.