

# File Carving: Analyzing Data Retrieval in Digital Forensics

Purna Chandra Sethi,

Assistant Professor

Department of Computer Science,

Rama Devi Women's University, Bhubaneswar, Odisha, India

Email: pcsethi [AT] rdwu.ac.in

**Abstract—** In the current scenario, mostly the data are stored in digital media. Managing the storage and security of huge volume of data is emerging as a significant challenge for data science researchers and engineers. As data is considered as more costly and powerful than anything else, so during damage or loss of data thousands of dollars are being invested for data recovery. File carving is a technique used for data recovery from the file without the any contextual information when the storage media is formatted or file system got damaged. In this study, we have tried to describe the various types of file carving techniques and the tools used for file carving, along with their limitations and the categories of files which are supported along with the scenario for such recovery.

**Keywords-** Digital Forensic; File Carving; Data Recovery; File Carving Tools and Techniques.

## I. INTRODUCTION

Digital forensic analysis techniques are used to investigate and analyse digital devices, data, and networks to uncover evidence for legal proceedings. These techniques are crucial for identifying and understanding digital crimes, such as cyber-attacks, data breaches, intellectual property theft, and other cybercrimes. Some of the common digital forensic techniques includes: disk imaging, file carving, metadata analysis, timeline analysis, keyword search, network forensic, memory forensic, hashing and integrity verification, data recovery, data visualization, etc. Each of these techniques is used for different types of file recovery applications.

Disk imaging is a fundamental practice in digital forensics that entails the creation of a bit-by-bit duplicate, often referred to as a "disk image" or "forensic image," of a digital storage medium like a hard drive, solid-state drive (SSD), or USB drive. This replica encompasses all the data and metadata found on the original device, encompassing deleted, concealed, and unallocated data. Disk imaging is crucial for preserving the integrity of digital evidence and conducting forensic investigations. Some of the commonly used disks imaging techniques are: physical imaging, local imaging, live imaging, dead imaging, remote imaging, sparse imaging, etc. The different disk imaging techniques can be implemented using various forensic imaging tools and software solutions designed

for digital forensic investigation. The condition of the storage media, the requirements of the investigation, and the legal considerations are the major parameters during forensic analysis. The most popular file carving techniques are depicted in table-1.

Table-1: Popular File Carving Techniques

| Sl. No. | Technique               | Description   |
|---------|-------------------------|---|
| 1       | Header/Footer Carving   | In this method, file carving tools search for known file signatures or headers that indicate the beginning of a file. Once a header is found, the tool scans the raw data for the corresponding footer or end marker to determine the file's length and extract the entire file. This method is effective for recovering files with well-defined file signatures, such as JPEG images (which typically start with the bytes FF D8) or ZIP archives. |
| 2       | Entropy-Based Carving   | Entropy-based file carving methods analyze the entropy (randomness) of the data to identify potential file boundaries. Files often exhibit higher entropy compared to unallocated or free space, making it possible to detect file boundaries based on changes in entropy levels. This method is useful for recovering files that may not have distinct headers or signatures.  |
| 3       | File Structure Analysis | Some file carving tools leverage knowledge of file structures and formats to reconstruct files based on their internal organization. Through the analysis of standard file types such as documents, images, videos, and archives, these utilities can discern file fragments and reconstruct files even in scenarios where their headers or   |

|   |                              |  |
|---|------------------------------|--|
|   |                              | signatures are missing or damaged.   |
| 4 | Fragmented File Carving      | Fragmented file carving methods focus on reconstructing files that are fragmented across multiple non-contiguous sectors or clusters on the storage media. These methods use techniques such as file system analysis, cluster chaining, or file system journal parsing to identify and assemble fragmented file fragments into complete files.               |
| 5 | Signature Extension          | Some file carving tools allow users to define custom file signatures or extensions to search for specific file types that may not be supported by default. By specifying custom signatures or extensions, forensic analysts can tailor the file carving process to target specific types of files or data structures.  |
| 6 | Content-Aware Carving        | Content-aware file carving methods analyse the content and context of data fragments to identify potential file boundaries and reconstruct files based on their internal structure and relationships. These methods may use heuristics, machine learning algorithms, or pattern recognition techniques to recognize file contents and extract relevant data. |
| 7 | File Fragment Classification | File carving tools may classify recovered file fragments according to their file type, format, or content to prioritize reconstruction efforts and identify incomplete or corrupted files. Classification algorithms help forensic analysts differentiate between valid file fragments and irrelevant data   |

|   |                 |  |
|---|-----------------|--|
|   |                 | formats and allows users to define custom file signatures for carving specific file types. Scalpel is highly configurable and can be adapted to different forensic scenarios.  |
| 2 | PhotoRec        | PhotoRec is a free and open-source file carving tool that specializes in recovering multimedia files, including photos, videos, and audio files, from digital storage devices. It supports a broad range of file systems and files formats and can recover files from damaged or formatted partitions. PhotoRec is part of the TestDisk suite of data recovery tools.  |
| 3 | Foremost        | Foremost file carving is a command-line tool designed for extracting files according to the headers, footers, and data structures. It provisions several file types, including images, documents, archives, and executables, and can recover files from disk images, raw partitions, and other storage media. Foremost is highly configurable and can be customized to target specific file types.   |
| 4 | Scalpel-Next    | Scalpel-Next is an enhanced version of the Scalpel file carving tool with additional features and improvements. It offers better performance, increased file type coverage, and improved accuracy in file recovery. Scalpel-Next supports both file carving and file system analysis and are suitable for use in digital forensic investigations.  |
| 5 | R-Studio        | It is a commercial data recovery software suite, incorporates file carving functionalities within its feature set. It boasts support for a diverse array of file systems and types, facilitating the recovery of files from damaged, corrupted, or deleted partitions. R-Studio is renowned for its user-friendly interface and offers advanced options for file recovery and reconstruction.  |
| 6 | Autopsy         | Autopsy, an open-source digital forensic platform, incorporates file carving functionalities within its suite of features. It offers a graphical user interface (GUI) that facilitates forensic investigations and the analysis of disk images, file systems, and data artifacts. Autopsy seamlessly integrates a variety of forensic tools and techniques, including file carving, timeline analysis, keyword searching, and metadata extraction. |
| 7 | EnCase Forensic | EnCase Forensic is a commercial digital forensic software suite that offers  |

These file carving methods can be implemented in standalone file carving tools or integrated into comprehensive digital forensic software suites. Forensic analysts typically use a combination of these methods, along with manual verification and validation, to recover files and data from storage media during digital investigations. Several tools are available for implementing file carving in forensic investigations. Some widely used tools for file carving implementation are described in table-2.

Table-2: Popular File Carving Tools

| Sl. No. | Name of the Tool | Description   |
|---------|------------------|---|
| 1       | Scalpel          | Scalpel is an open-source file carving tool designed for efficient and precise file recovery from various storage media. It supports a wide range of file types and |

|  |  |  |
|--|--|--|
|  |  | comprehensive file carving facilities for extracting files contents from storage media. It provides advanced features for disk imaging, file system analysis, evidence acquisition, and data recovery. EnCase Forensic is widely used by law enforcement agencies, government organizations, and corporate security teams for forensic investigations. |
|--|--|--|

These tools offer varying degrees of functionality, performance, and ease of use, depending on the specific requirements of the forensic investigation. Forensic analysts may choose the most suitable tool based on factors such as file type support, platform compatibility, scalability, and budget constraints. Additionally, it's essential to consider factors such as documentation, support, and community resources when selecting a file carving tool for implementation in digital forensic investigations.

The organization of the remaining sections of this paper is: section-2 contains the brief overview of the literature survey in this research field. Section-3 focuses on the comparison analysis of different file carving techniques based on various parameter consideration following the section-4 showing the conclusion and the future scope for research in this field.

## II. LITERATURE SURVEY

The data recovery is becoming a major concern in case of loss or damage of storage device. The file carving technique typically finds specific header/footer signatures and data structures in files. But file carving has been around for a long time facing the problem of recovering fragmented files, files are important targets in digital forensics. (doc, hwp, xls, etc.) are relatively easy to fragment, so it is very important to propose a solution for recovery. Various carving techniques and tools are continuously being developed to overcome this, and various researches and techniques are being carried out to verify the functionality. This paper addresses some file carving techniques and its parameters for validation.

In [1], the researchers proposed a new technique called Hadoop Distributed File System (HDFS) which was used for carving of file system that is huge in size. The researchers present a framework that integrates various steps to introduce a novel file carving technique aimed at recovering the maximum number of file pieces affected by 10% corruption, while also ensuring successful file carving. In [2], the researchers conducted a review of file carving techniques, emphasizing four key aspects: the availability of realistic datasets for tool testing, object validation in fragmented data storage environments, content-based validation, and its implications for digital investigation practices, as well as semantic validation to mitigate false positive rates.

In [3], the researchers made a testing of six file carving tools and made performance comparison using a new dataset for recovery and reassembly of the fragmented files resulting

to maximize the file recovery and minimize the invalid file output. In [4], the researchers conducted a thorough review to provide a comprehensive understanding of the current state-of-the-art in the field. Additionally, the paper introduces a novel file carving ontology aimed at facilitating knowledge base recovery of data. The review cites seventy research papers in this field to compile research contributions and identify future scope for advancements.

In [5], the researchers offer an overview of the evolution of file carving, highlighting its advantages and challenges in forensic recovery. The paper compares traditional data recovery methods with techniques like file carving, file systems and fragmentation analysis, FAT32, file allocation, deletion, and recovery, as well as NTFS, appending/editing files, and file structure-based carving. In [6], the researchers described a method for detecting the point of fragmentation in a file to aid in data recovery. They utilized sequential hypothesis testing to compare adjacent pairs of blocks from the beginning of the file until the fragmentation point was identified, aiming to minimize errors in detecting the fragmentation point.

In [7], the researchers conducted a survey of files retrieved from over 350 hard drives to reconstruct fragmented files for forensic purposes. They assert that accurate file carving is attained through a multi-tier decision-making process, which requires swift validation by candidate strings extracted from the media drive. The study includes validation procedures for three file types: JPEG, Microsoft OLE (MSOLE), and ZIP formats, and outlines the approach for utilizing high-speed validators to reconstruct fragmented data.

During the fragmentation of the files, the files could be disconnected and their order could be disrupted due to which the straight forward carving fails. So in [8], the researchers introduced a classifying clustering technique to address the challenges posed by file fragmentation, which can lead to disconnection and disordering of files, thus impeding straightforward carving processes. The proposed technique leverages the statistics of clusters and incorporates a set of characteristic features and statistical patterns. These are used to train a supervised classification model capable of identifying a variety of relevant file types. The approach also restricts the number of neighboring clusters to enhance the classification performance.

In [9], the researcher contends that the file carving process relies on the file format and assumes that the data are stored at the block-level device. However, this process can generate numerous false-positive files, which may be either invalid or nonexistent, during recovery from the target device. The author argues that in real-world scenarios involving large volumes of data, a substantial amount of false positive data is produced, introducing uncertainty to the carved data. To address this issue, the researcher proposes an in-place approach to file carving, which reconstructs file system metadata without duplicating the file contents. This approach aims to reduce significant storage requirements and turnaround time.

In [10], the researchers conducted a performance comparison of the file carving process between PhotoRec and Foremost, focusing on three parameters: the number of returned files, file validation, and processing speed. The validation of the process was conducted using the SHA1 (Secure Hash Algorithm-1). The findings indicate that the PhotoRec method demonstrates superior performance, delivering faster processing speeds compared to the Foremost method.

In [11], the researchers address the challenges encountered during the file carving process and propose algorithms for automatically generating file fingerprints from a set of input files. Subsequently, these fingerprints are used to recognize unknown file types, rather than relying on associated metadata. The recognition is facilitated by three algorithms: byte-by-frequency analysis, byte frequency cross-correlation analysis, and file header/trailer analysis. The study concludes that the test accuracy varies between 23% and 96% across different iterations.

### III. PREPARE YOUR PAPER BEFORE STYLING

In this digital forensic era when thousands of terabytes of data is getting digitized daily, the file management as well as recovery is becoming a major concern. Many digital forensic tools are used for data recovery during loss or damage of file contents as well as storage device. Out of all the techniques, file carving techniques has a special area due to the accuracy as well as simplicity in the implementation. Different factors are considered during the performance analysis of different carving techniques. The factors associated are: accuracy, speed of implementation, robustness, supported file types, ease of use, resource requirements, validation mechanisms, community adoption, etc. The accuracy of a file carving technique refers to its ability to correctly identify and reconstruct files from raw data. Techniques that minimize false positives (incorrectly identified files) and false negatives (missed files) are considered more accurate. Speed refers to how quickly a file carving technique can process data and recover files. Faster techniques are preferred, especially in forensic investigations where time is critical. Robustness refers to a technique's ability to handle various types of file corruption, fragmentation, and storage media formats. Techniques that are more resilient to these challenges are considered more robust. Supported FileTypes deals with some file carving techniques may specialize in certain types of files or formats. Evaluating which file types a technique supports can help determine its suitability for specific applications. The ease of use of a file carving technique includes factors such as user-friendliness, accessibility, and the availability of documentation and support resources. Resource requirements, such as memory usage and processing power, can vary between file carving techniques. Techniques that are more resource-efficient may be preferable in certain situations. Validation mechanisms ensure the integrity and authenticity of recovered files. Techniques that incorporate robust validation methods, such as checksum verification or file signature analysis, inspire greater confidence in the recovered data. The community adoption represents the adoption and usage of a file carving technique

within the forensic community can indicate its reliability and effectiveness. The techniques with a larger user base and community support may offer more resources and expertise for troubleshooting and improvement. By considering these criteria and conducting comparative evaluations, researchers and practitioners can make informed decisions about which file carving techniques are best suited to their specific requirements and objectives.

The comprehensive comparison analysis of different carving techniques are depicted in table-3.

## IV. CONCLUSION AND FUTURE SCOPE

### 4.1. CONCLUSION

File carving techniques are essential in digital forensics and data recovery, enabling the extraction of files from raw data without relying on file system metadata. Several file carving methods are used for different types of file applications and each technique has its advantages and limitations. Signature-based carving is effective for well-known file types with distinct signatures. However, it may struggle with fragmented or corrupted files and might miss files with non-standard signatures. Content-based carving analyzes the content of data blocks to reconstruct files. This approach is more resilient to fragmentation and corruption compared to signature-based methods. However, it requires more computational resources and may produce false positives. Structure-based file carving techniques involve identifying and reconstructing files based on their internal structure or organization. These techniques often rely on knowledge of file formats and their specific structural characteristics. The major advantages of this technique are it is precise, resilience to fragmentation, support for complex file format, reduced false positive, etc. However, it has dependency on the file format knowledge, complexity and computational overhead, vulnerable to data corruption and limited with scalability. Advanced file carving techniques implemented by combining multiple carving techniques can improve overall effectiveness and reliability. Hybrid approaches often integrate signature-based, header/footer, and content-based methods to overcome individual limitations and enhance file recovery rates. Machine learning techniques show promise in enhancing file carving accuracy, especially in complex scenarios involving fragmented, compressed, or encrypted data. However, challenges remain in training models with diverse datasets and ensuring generalization to new file types and formats.

The effectiveness of file carving techniques depends on factors such as file system complexity, data fragmentation, corruption levels, and the diversity of file types present. While each technique has its strengths and weaknesses, a combination of methods or advancements in machine learning-based approaches holds potential for improving file recovery capabilities in digital forensics and data recovery applications.

**4.2. FUTURE SCOPE**

File carving is one of the demanding research fields due to its high demand for digital data recovery during the loss or damage of the file contents. The future scope of research field encompasses several key areas, including: enhanced fragmentation handling, increased file format support, improved robustness to data corruption, efficient handling of encrypted and compressed data, integration of artificial

intelligence and machine learning, automation and scalability, forensic data visualization and analysis, privacy and security considerations, real-time and forensic incident response, interdisciplinary collaboration, etc. Exploration of the key features for faster data recovery along with secured data processing and related fields opens many research areas to advance the state-of-the-art in file carving techniques and address emerging challenges in data analysis and recovery.

Table-3: Comprehensive Comparison Analysis of Different File Carving Techniques

| Sl. No.                              | File Carving Techniques        | Feature Description  | Steps Involved   | Tools Used                       | Feature Description and Performance Analysis   |
|--------------------------------------|--------------------------------|--|--|----------------------------------|--|
| <b>Basic File carving Techniques</b> |                                |  |  |                                  |  |
| 1.                                   | <b>Signature-based Carving</b> | This is commonly referred to by various names, depending on the specific approach or tool being used.                    | <ol style="list-style-type: none"> <li>1. Identification of Signatures</li> <li>2. Scanning the Raw Data</li> <li>3. Extraction of File Fragments</li> <li>4. Reconstruction of Files</li> <li>5. Validation and Verification</li> </ol> | a. Header/Footer Carving         | This technique involves searching for known file signatures or headers that indicate the beginning and end of a file. It is one of the most traditional forms of signature-based carving.                                    |
|                                      |                                |  |  | b. File Signature Analysis       | This approach focuses on analysing the signatures or magic numbers of files to identify file types and boundaries within the raw data.   |
|                                      |                                |  |  | c. File Carving by File Headers  | Some tools or techniques specifically target file headers to identify file types and extract data fragments based on header information.   |
|                                      |                                |  |  | d. Magic Number Carving          | This technique also known as files signatures or file markers, are specific byte sequences that indicate the file type or format. Magic number carving involves searching for these sequences to identify and extract files. |
|                                      |                                |  |  | e. Pattern Matching Carving      | It is used to search for specific byte sequences or patterns within the raw data that indicate the presence of files or file boundaries.   |
|                                      |                                |  |  | f. File Fragment Identification  | In this approach, signatures are used not only to identify file boundaries but also to identify and reconstruct fragmented file fragments within the raw data.   |
| 2.                                   | <b>Content-based Carving</b>   | This is also known as content-aware carving or content carving, is a file carving technique used in digital forensics to | <ol style="list-style-type: none"> <li>1. Raw Data Analysis</li> <li>2. Data Fragment Identification</li> <li>3. Fragment Assembly</li> <li>4. File Reconstruction</li> <li>5. Validation and Verification</li> </ol>                    | a. Fragmentation Analysis        | It involves analysing the fragmentation patterns of data to identify and reconstruct file fragments.   |
|                                      |                                |  |  | b. Entropy-Based Carving         | This technique analyses the entropy (randomness) of the data to identify potential file boundaries and fragments based on changes in entropy levels.   |
|                                      |                                |  |  | c. File Carving by Data Patterns | It specifically target data patterns or signatures within the raw data to identify and extract files.  |

|    |                         |   |  |   |   |
|----|-------------------------|---|--|---|---|
|    |                         | recover files based on their content rather than relying solely on file system metadata or file headers.  | <ol style="list-style-type: none"> <li>6. Content Analysis</li> <li>7. Error Handling</li> <li>8. Documentation and Reporting</li> </ol>   | <ol style="list-style-type: none"> <li>d. Data Chunking</li> </ol>                          | This approach involves dividing the raw data into smaller chunks or blocks and analysing the content of each chunk to identify potential file fragments.  |
|    |                         |   |  | <ol style="list-style-type: none"> <li>e. Content Validation Carving</li> </ol>             | This technique often include validation mechanisms to verify the integrity and authenticity of recovered files based on their content.  |
|    |                         |   |  | <ol style="list-style-type: none"> <li>f. Machine Learning-Based Carving</li> </ol>         | It is an advanced content-based carving technique utilizes machine learning algorithms to analyse the content of data fragments and identify potential file boundaries or file types.   |
| 3. | Structure-based carving | Structure-based carving techniques in digital forensics rely on knowledge of file structures and formats to reconstruct files based on their internal organization. | <ol style="list-style-type: none"> <li>1. File Structure Analysis</li> <li>2. Header/Footer Identification</li> <li>3. Fragment Assembly</li> <li>4. File Reconstruction</li> <li>5. Metadata Extraction</li> <li>6. Validation and Verification</li> <li>7. Error Handling</li> <li>8. Documentation and Reporting</li> </ol> | <ol style="list-style-type: none"> <li>a. File Structure Analysis</li> </ol>                | This technique involves analysing the internal structure of common file types, such as documents, images, videos, and archives, to identify file fragments and reconstruct files based on their internal organization.          |
|    |                         |   |  | <ol style="list-style-type: none"> <li>b. File System Metadata Parsing</li> </ol>           | This may involve parsing file system metadata, such as file allocation tables (FAT), master file tables (MFT), or inode structures, to reconstruct files based on their metadata entries.                                       |
|    |                         |   |  | <ol style="list-style-type: none"> <li>c. Directory Structure Reconstruction</li> </ol>     | This technique focuses on reconstructing directory structures, including directories, subdirectories, and file names, to identify relationships between files and directories and recover files based on their directory paths. |
|    |                         |   |  | <ol style="list-style-type: none"> <li>d. File Header/Footer Analysis</li> </ol>            | This technique analyse file headers and footers to identify file formats and structures, allowing for the reconstruction of files even if their headers or signatures are missing or damaged.                                   |
|    |                         |   |  | <ol style="list-style-type: none"> <li>e. Fragmented File Recovery</li> </ol>               | This technique involve identifying and assembling fragmented file segments scattered across multiple non-contiguous sectors or clusters on the storage media into complete files.   |
|    |                         |   |  | <ol style="list-style-type: none"> <li>f. Metadata Extraction and Reconstruction</li> </ol> | This approach involves extracting file metadata, such as timestamps, file attributes, and file system information, and using this metadata to reconstruct files or infer file relationships.                                    |

| <b>Advanced File Carving Techniques</b> |                            |  |                            |  |  |
|---|----------------------------|--|----------------------------|--|--|
| 4.                                      | Graph theory-based carving | Thistechniques leverage concepts and algorithms from graph theory to analyse and reconstruct files from digital storage media. While there may not be widely recognized names specifically dedicated to this approach, such techniques can be developed and implemented using graph-based methodologies. | 1. Graph Representation    | a. Graph-based File Reconstruction     | This approach involves representing file fragments and their relationships as graphs, where nodes represent data blocks and edges represent connections between blocks based on file structure or content. |
|   |                            |  | 2. Node Identification     | b. Graph Analysis for File Carving     | Graph theory is utilized to analyse the structure and content of file fragments, enabling the identification and reconstruction of files based on graph-based algorithms.                                  |
|   |                            |  | 3. Cluster Analysis        | c. Graph Mining for File Recovery      | This technique is applied to extract patterns and relationships from file fragments, facilitating the reconstruction of files from raw data.   |
|   |                            |  | 4. Fragment Assembly       | d. Graph-based Fragment Assembly       | Graph theory principles are used to assemble fragmented file segments into complete files by analysing the connections and dependencies between data blocks.   |
|   |                            |  | 5. File Reconstruction     | e. Graph-based Metadata Reconstruction | Graph-based methods are employed to reconstruct file metadata, such as file attributes and relationships, aiding in the identification and recovery of files from digital storage media.                   |
| 6. Validation and Verification          |                            |  |                            |  |  |
| 7. Error Handling                       |                            |  |                            |  |  |
| 8. Documentation and Reporting          |                            |  |                            |  |  |
| 5.                                      | Bi-fragment Gap Carving    | This technique focuses on identifying and reconstructing file fragments with gaps between them.  | 1. Raw Data Analysis       | a. Gap-aware Fragment Carving          | This technique emphasizes on identifying and handling file fragments with gaps between them during the carving process.  |
|   |                            |  | 2. Gap Identification      | b. Bi-fragment Reconstruction Method   | This technique' highlights the unique approach to reconstructing file fragments that are separated by gaps, ensuring more accurate and complete file recovery.   |
|   |                            |  | 3. Fragment Identification | c. Fragment Gap Bridging Technique     | This emphasizes the capability to bridge the gaps between bi-fragments during the carving process to reconstruct files effectively.  |
|   |                            |  | 4. Fragment Assembly       | d. Gap-tolerant File Carving Algorithm | This technique provides the tolerance for gaps between file fragments, ensuring robust file reconstruction even in the presence of fragmented data.  |
|   |                            |  | 5. File Reconstruction     | e. Bi-fragment Gap Recovery Method     | This emphasizes the technique for recovering file fragments with gaps between them, enabling comprehensive data recovery from digital storage media.   |
| 6. Validation and Verification          |                            |  |                            |  |  |
| 7. Error Handling                       |                            |  |                            |  |  |
| 8. Documentation and Reporting          |                            |  |                            |  |  |
| 6.                                      | Smart carving              | This typically involve the use of advanced algorithms, heuristics, or machine learning approaches to improve the efficiency and  | 1. Data Pre-processing     | a. Intelligent File Carving            | This emphasizes the use of intelligent algorithms or techniques to enhance the file carving process, leading to more efficient and accurate results  |
|   |                            |  | 2. Feature Extraction      | b. Adaptive Carving Methodology        | This suggests that the carving technique adapts its approach based on the characteristics of the data being analysed, resulting in smarter and more effective file recovery.                               |
|   |                            |  | 3. Algorithm Selection     | c. Context-aware                       | This highlights ability to consider  |
| 4. Training (if applicable)             |                            |  |                            |  |  |
| 5. File Fragment Identification         |                            |  |                            |  |  |

|  |  |  |                                      |  |
|--|--|--|--------------------------------------|--|
|  | accuracy of file carving, may be referred to by various names. | 6. Fragment Assembly<br>7. File Reconstruction<br>8. Validation and Verification<br>9. Error Handling<br>10. Documentation and Reporting | Carving Technique                    | contextual information, such as file system structures or data patterns, to improve the accuracy of file carving.  |
|  |  |  | d. Machine Learning-assisted Carving | This indicates that the carving technique incorporates machine learning algorithms to learn from data and improve its performance over time, leading to smarter file recovery.       |
|  |  |  | e. Heuristic-driven Carving Approach | This suggests that the carving technique utilizes heuristics or rules of thumb to guide the file recovery process, enabling more intelligent and efficient carving.                  |
|  |  |  | f. Predictive File Carving Algorithm | This implies that the carving technique can predict file boundaries or structures based on data patterns or other indicators, resulting in more accurate and targeted file recovery. |

### REFERENCES

- [1] E. Alshammari, G. Al-Naymat and A. Hadi, "A New Technique for File Carving on Hadoop Ecosystem," *2017 International Conference on New Trends in Computing Sciences (ICTCS)*, Amman, Jordan, 2017, pp. 72-77, doi: 10.1109/ICTCS.2017.16.
- [2] Nadeem Alherbawi, Zarina Shukur, Rossilawati Sulaiman, "Systematic Literature Review on Data Carving in Digital Forensic," *Procedia Technology*, Volume 11, 2013, Pages 86-92, ISSN 2212-0173, <https://doi.org/10.1016/j.protcy.2013.12.165>.
- [3] Laurenson, T. (2013). "Performance Analysis of File Carving Tools." In: Janczewski, L.J., Wolfe, H.B., Sheno, S. (eds) *Security and Privacy Protection in Information Processing Systems. SEC 2013. IFIP Advances in Information and Communication Technology*, vol 405. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-39218-4\\_31](https://doi.org/10.1007/978-3-642-39218-4_31)
- [4] R. Poisel and S. Tjoa, "A Comprehensive Literature Review of File Carving," *2013 International Conference on Availability, Reliability and Security*, Regensburg, Germany, 2013, pp. 475-484, doi: 10.1109/ARES.2013.62.
- [5] A. Pal and N. D. Memon, "The evolution of file carving," *IEEE Signal Processing Magazine*, vol. March 2009, pp. 59– 71, March 2009.
- [6] A. Pal, H. T. Sencar, and N. D. Memon, "Detecting file fragmentation point using sequential hypothesis testing," *Digital Investigation*, vol. 5, no. Supplement 1, pp. S2 – S13, 2008.
- [7] S. L. Garfinkel, "Carving contiguous and fragmented files with fast object validation," *Digital Investigation*, vol. 4, no. Supplement 1, pp.2 – 12, 2007.
- [8] C. J. Veenman, "Statistical disk cluster classification for file carving," in *Proc. of Int. Symposium on Information Assurance and Security (IAS'07)*, vol. 0. Manchester, UK: IEEE CS, 2007, pp. 393–398.
- [9] G. G. Richard, V. Roussev, and L. Marziale, "In-place file carving," in *Proc. of IFIP Int. Conf. on Digital Forensics*, 2007, pp. 217–230.
- [10] Nurhayati and N. Fikri, "The analysis of file carving process using PhotoRec and Foremost," *2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT)*, Kuta Bali, Indonesia, 2017, pp. 1-6, doi: 10.1109/CAIPT.2017.8320663.
- [11] M. McDaniel and M. H. Heydari, "Content based file type detection algorithms," in *Proc. of the 36th Annual Hawaii Int. Conf. on System Sciences (HICSS'03) - Track 9*. Washington, DC, USA: IEEE CS, 2003, p. 332.1.