A Review on the Application of Artificial Intelligence and Automation in Digital Forensics

MTA Deen[1#] and B Hettige[2]

[1]Department of Computer Science, General Sir John Kotelawala Defence University, Sri Lanka
[2]Department of Computer Engineering, General Sir John Kotelawala Defence University, Sri Lanka

[#] 37-cs-0007@kdu.ac.lk

Abstract*: As a branch of forensic science, Digital Forensics is concerned with identifying, acquiring, processing, analysing, and reporting on digital data. For law enforcement investigators, Digital Forensics support is crucial since electronic evidence is present in almost all criminal activities. An array of electronic evidence can be gathered from a variety of sources, including computers, smartphones, remote storage, unmanned aerial systems, and shipborne equipment. The main objective in Digital Forensic is to extract data from electronic evidence, process it into actionable intelligence and present the findings for prosecution. The success, efficiency, and efficacy of a typical forensic inquiry are significantly influenced by the knowledge and prior experience of the investigator or any security agent. The outcomes of a digital investigation will be more effective and efficient if the power of intelligence in the available computer resources is utilized. In modern computer science, Artificial Intelligence (AI) is a well-established field that can often provide a means of solving computationally complex or large problems in a realistic timeframe. The influence of AI on several fields in modern society and its achievements throughout time suggest that it can help with a variety of challenging Digital Forensics investigative issues. This review paper outlines various methods of evaluating, optimizing and standardizing applications of artificial intelligence and Automation models used in digital forensics.*

Keywords: *Digital Forensic, Artificial Intelligence, Automation, Machine Learning, Intelligent Forensics*

1. Introduction

Most of us heavily rely on digital devices and the Internet to operate and improve our quality of life and/or businesses as a result of the development of technology. We rely on these tools and technologies to process, store, and transfer data, which causes a large volume of data to be created, gathered, and shared electronically. In recent years the occurrence of malicious cyber activity has increased as many individuals and international cooperate sector depends on digital infrastructure of the highest level and Information and Communication Technologies (ICT). Data breachers, information leakage, information security breaches, hacking, malware and ransomware attacks, phishing scams and botnets are some of the malicious activities that have caused severe loss for companies and individuals. As a result, companies have resorted to cyber-crime prevention and detection.

There has been extensive exploration of computational intelligence techniques in a variety of domains. Digital Forensic is such domain. Digital forensics is also referred to as computer and network forensics. Using science, Digital Forensics involves identifying, collecting, examining, and analysing data, while protecting its integrity and maintaining a strict chain of custody (Kent, Chevalier, Grance & Dang, 2006). Simply Digital Forensics can be defined as the process of identifying, preserving, analysing, and documenting digital evidence that is to be presented as evidence in court of law when required. Digital Forensics traces its roots back to 1970 when engineers recovered the only copy of a database that was deleted unintentionally (Garfinkel, 2010). Digital Forensics Investigators are individuals who follow evidence and solve crimes through digital means. Digital Forensics Investigators role is to recover deleted files, cracking passwords, finding the initial source of a security breach and after collecting the evidence is then analysed, stored, and translated to make it presentable before the court of law (Digital Forensics, 2022). Security experts, law enforcement agencies investigating cybercrime and Digital Forensics investigators face new challenges as a result of today's massive amounts of data, heterogeneous technologies, borderless networks and complex modern hardware/software frameworks. From both ethical and technological perspective, modern Digital Forensics face obstacles.

Artificial Intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. Machine Learning, Deep Learning, Expert systems, natural language processing, and speech recognition are some applications of AI. Automation is the combination of modern hardware and software to carry out a task or process with zero or minimal human invention. In the context of AI, large volumes of labelled training data are ingested, analysed for correlations and patterns, and predictions are made based on the patterns found in the data. Automation and Artificial Intelligence are two modern computing fields that work together. The idea of AI is to simulate the cognitive and reasoning processes of the human brain to help streamline and/or automate laborious procedures. Massive development efforts are being made to create AI powered software, applications, operating omatea wide range of processes, operations, and activities. The main aim by integrating AI and Automation is to achieve efficiency, accuracy, and cost-reduction. Machine Learning (ML) is a section of AI. When a computer system or algorithm can process a large amount of data and ultimately can draw predictions and conclusions it is referred to as Machine Learning (Choy et al.,2018). Machin Learning analyses historical data using regression models and classification to form future predictions. Intelligent Automation (IA) is a term used to refer to technology solutions that make use of AI, Automation, and machine learning (Jarrett & Choo, 2021).

Digital forensics is a complex and evolving field where AI and Automation is making significant headway. AI, ML and Automation are considered as emerging application in Digital Forensics. It is also recorded that the US Federal and State Law Enforcement Agencies have been investigating the potential applications of technology that is powered by Artificial Intelligence to improve the effectiveness of Digital forensics (Jarrett & Choo, 2021). This will result in increasing the accuracy of Digital Forensic investigations. The field of Digital Forensics is one that is taking on greater significance in computing and frequently necessitates the thoughtful study of vast quantities of complex data. Therefore, it would appear that using AI is the best strategy to address many of the issues that Digital Forensics is currently facing. AI can be used to help identify hidden trends in the collected evidence as AI has no bounds to how much data it can process and analyse.

2. Methodology

In this section of the document the methodology and approach followed in conducting the review is discussed. For the purpose of composing this review, a systematic stategy was used, in which the topic, angle, goals, and title were selected initially. Research articles, documents and resources were explored and analysed. These were then used to develop a deeper knowledge of the ideas, concepts, and technology and then further studied and reviewed thoroughly, to find the data and information that will be most useful and acceptable for this investigation. To do this study, articles between 2000 to 2022 were found via Google Scholar and other research archives using keywords related to Digital Forensic, Automation and Artificial Intelligence. The criteria followed when selecting research articles for the study are as follows:

- English-language writings that are clear
- considering research conducted during the 2000 to 2022.
- Accessibility of the entire document or article

Factors such as technology used, drawbacks of the technology, user feedback, advantages, methods followed, case studies and etc were thoroughly reviewed, analysed, and compared during this study.

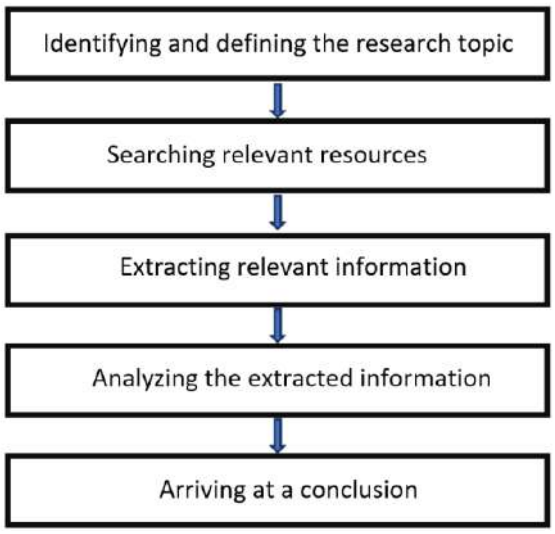
Identifying and defining the research topic
Searching relevant resources
Extracting relevant information
Analyzing the extracted information
Arriving at a conclusion

Figure 1. Methodology for systematic literature

Finally, conclusions were reached by using the data and expertise obtained through reading, analysing, and constructing these research works.

I. Intelligent Forensics

Intelligent forensics is an interdisciplinary method that uses resources in a more intelligent way while utilizing technological advancements to solve a case. A variety of technologies and techniques from Artificial Intelligence, computational modelling, and social network analysis are included in intelligent forensics, which helps to focus digital investigations and cut down on the time spent looking for digital evidence. Intelligent forensics can be used both pro-actively, prior to an incident, and reactively, following an occurrence. The proactive application of intelligent forensics aims to spot threats before an incident occurs. Intelligent Forensics is currently being used by secret/military services and law enforcement agencies. Techniques like social network analysis (SNA) and Artificial Intelligence (AI) are used in Intelligent Forensics. In order to deal with the complexity of huge data sources of digital evidence, there are a number of viable intelligent forensic solutions. The solutions centre on either condensing the scope of the investigation, accelerating the investigation instruments, or utilizing intelligent forensics. instead of using queries to find data like in conventional digital forensics, intelligent forensics uses improved methods and approaches. (Irons & Lallie, 2014)

II. Application of Ai and Automation In Digital Forensic

Law enforcement organizations have been able to pinpoint important trends in a variety of crimes thanks to the impact of Intelligent Automation in Digital Forensics (Reiber, 2018). The prospects for using Artificial Intelligence in computer forensics are identified in this section, with the discussion concentrating on the ways in which using AI and Automation might improve computer forensics investigations.

A. Representation of knowledge

A concept in AI systems is representation of knowledge and ontology. The information we want to be able to reason about is what we call as representation of knowledge, and ontology is how we formally structure that knowledge representation so that we can reason about it. It is noteworthy that representation of knowledge can be characteristics of item in the domain, how these can be processed and how the processes are applied.

In the article 'The use of Artificial Intelligence in Digital Forensics: An Introduction' by Dr Faye Mitchell it is stated that AI has the greatest potential to impact Digital Forensics since it can offer expertise to aid in the standardization of the representation of knowledge and information in the field. Even the most fundamental activities in Digital Forensics, such the interchange of image information amongst forensic imaging tools (Turner, 2005), are made more difficult by this absence of standards. This indicates that Digital Forensics lags behind generally recognized best practices. Clear advantages would result from the development of an international domain ontology for digital forensics. Set up a formal framework for communicating about digital evidence, the ability to build a large, usable case repository (Duce, Mitchell & Turner, 2007). Such a case repository would include examples of Digital Forensic investigations with known attributes and outcomes. This has proven to be incredibly helpful in other AI fields and could be effective for teaching Digital Forensic professionals as well as testing the performance of specialists, whether they be humans or AI systems. A standardized ontology may be extremely helpful in developing a uniform, reusable body of background knowledge that AI systems could exploit.

B. Artificial Intelligence and Automation

As discussed above Artificial Intelligence and Automation are making significant headway in the field of Digital Forensics. The "Evidence Analysis" stage of Digital Forensics has been highlighted by the researchers as having a high relevance for AI. Digital forensic tools with AI capabilities process objective data, analyse it, and then create strong prospective hypotheses that can be used as evidence in a court of law (Costantini et all. 2019).

1) Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation: The research carried out by Xiao, Li and Xu in 2019 proposes an Artificial Intelligence based approach to conduct through video-based evidence analysis and data extraction. In order to analyse low-quality footage, the researchers suggest a forensic video analysis framework that uses an effective video/image enhancement algorithm. For use in Digital Forensic investigations, a Closed-Circuit Television (CCTV) footage quality improvement technique based on Contrast Limited Adaptive Histogram Equalization (CLAHE) is developed. A deep-learning-based object detection and tracking system is suggested to aid in the video-based forensic investigation by detecting and identifying suspected suspects and tools from film (Xiao, Li & Xu, 2019).

2) Automated Forensic Examiner (AFE): Fahdi, Clarke & Furnell proposes an Automated Forensic Examiner that aims to solve the sorting and identification challenge in a case using Artificial Intelligence. The techniques used here are technical competency measure, dynamic criminal knowledge base and visualization to give the Digital Forensics investigator a comprehensive understanding of the case. To find evidence, the proposed method employs an iterative method. Next, it performs associative mapping to connected occurrences, and as a result, the system produces an evidence trail. Automated Evidence Profiler (AEP) and Self Organizing Maps (SOMs) are the core components of the AFE. The growing disparity between the quantity and scope of cases requiring forensic investigation and time requirements has been successfully addressed by the suggested strategy. This solution is an effective way to address the current problems. This solution comes to the conclusion that normal Digital Forensic procedures can be automated using methods like SOM and AEP, making the entire process more cost-effective and efficient. (Fahdi, Clarke & Furnell, 2013)

3) AI Framework: In the article 'Towards an Artificial Intelligence Framework by Actively Defend Cyberspace' by Masombuka, Grobler & Watson the techniques and methodologies for applying AI in Digital Forensics investigations and the motivation of an active defence framework is highlighted. The AI framework addresses advanced threats and emphasises on proactive measures, real time detection, active monitoring and mitigation of key threats. The article also explores innovative strategies, like the use of Artificial Intelligence (AI) systems with the ability to learn, adapt, and analyse data in real time to detect user behaviour, would help defend the cyberspace. According to the research done the proposed framework is aimed on strengthening the security backbone of an organization and illustrates the importance of combing AI and cybersecurity. The framework being presented also aims to provide the groundwork for future studies and research on the significance of defending cyberspace through AI. (Masombuka, Grobler & Bruce Watson, 2018).

4) Network Intrusion Detection: Intrusion Detection System (IDS) is used to detect cyber-attacks or malicious activity. Artificial Intelligence is often regarded as the better method for modifying and creating IDS and plays a crucial role in detecting intrusions. Neural network algorithms are a novel Artificial Intelligence method that can be used to solve difficulties in the present day.

In the study carried out by Kanimozhi and Jacob a system is proposed to detect a specific type of botnet attacks that poses a severs threat to banking and financial sectors. Artificial Intelligence is used to develop the suggested system using a realistic cyber defence dataset (CSE-CIC-IDS2018), The proposed Artificial Intelligence-based intrusion detection system for classifying botnet attacks is strong, precise, and accurate. It is recorded with a performance accuracy of 99.97%. The newly proposed system can be used for real-time network traffic data analysis as well as for traditional network traffic analysis and cyber-physical system traffic analysis. This system can also be enhanced to reorganize other types of attacks. Kanimozhi & Jacob, 2019)

C. Pattern Recognition

Pattern recognition is a subset of AI that excels at identifying particular types or groups of data in an inquiry. Patten recognition includes image recognition and recognizing a pattern in a disk image that would suggest it is a sound file, or a pattern in an email message that suggests SPAM. Many of the strategies make heavy use of probabilistic reasoning, statistics, or both. Understanding how the human perceptual system functions is necessary for the more sophisticated and precise sorts of image recognition that may be used to locate particular types of pictures (Mitchell. 2014). Machine Learning, Artificial Neural Nets or decisions trees techniques help in Pattern Recognition.

A study carried out by Fahdi, Clarke & Furnell apply unsupervised pattern recognition to identify notable artefacts utilizing Self-Organising Map (SOM). It can be considered that SOM as a supportive method for visually interpreting and analysing data produced by computer forensic tool (Fei et al, 2005) & (Fei at all, 2006). This study highlights that the application of SOM to identify notable artefacts works well in one case but poor identification in another case. As a conclusion the study states that their experimental results display a good level of performance. The findings imply that SOM can be utilized to benefit forensic investigators, such as by aiding in the visualization of artifacts and speeding up human analytical processes. However, it has never been attempted to utilize SOM to analyse a forensic image with the sole goal of identifying significant files using metadata retrieved at the file system and application levels. (Fahdi, Clarke & Furnell, 2016)

C. Neural Network

Studies have also proved that Neural Networks can be trained to identify appropriate and inappropriate behaviour and even model the behaviour patterns of different users so that it would be feasible to alert unusual use pattern to the currently logged in user. To identify exceptions and uncover patterns of behaviour, employ machine learning and data mining approaches. It is possible to create systems that continuously learn and enhance system performance in addition to big data analytics and high-speed computing platforms in order to stay up with shifting trends in the computer forensics industry. (Irons & Lallie, 2014)

3. Discussion

The impact of Artificial Intelligence and Automation on Digital Forensics is significant. The overall efficiency and quick recognition of patterns and other evidence is phenomenal. The efficiency and speed have enabled forensic professionals to produce leads and resolve cases with less effort, expense, and time.

Applying Artificial Intelligence and Automation in Digital Forensics also rises a few challenges.

Since Intelligent Automation enabled tools are still in development and may not always produce accurate, complete, or robust information necessary for forensic cases, the accuracy of the forensic outcome is somewhat dependent on the abilities of the human investigator (James & Gladshev, 2013). It is required to either hire highly skilled investigators or give the investigator extensive training and skill development to get around this issue.

Another problem is the use of various, sophisticated media formats, which the existing AI systems may find difficult to gather or analyse (Fahdi, Clarke & Furnell, 2013).

If any crucial information is absent from a knowledge representation, the forensic outcome may be impacted and may lead to incorrect conclusions. In pattern recognition algorithms, there is also a chance of producing a significant number of false positives and false negatives.

4. Conclusion

From the seed of a concept, cultivated by daring pioneers, developed and expanded by professionals, to its present condition, Digital Forensics has bloomed in less than thirty years. To provide the field with a strong basis for the future, many people have dedicated their time, experience, and enthusiasm. Recent trends suggest that cybercrime and the use of digital investigations are occurring in a changing and larger-scale context. There is a need to examine more effective and efficient processes and procedures in digital investigations in order to handle cybercrime in terms of identifying, gathering, recovering, analysing, and recording. Because anti-forensic strategies will keep getting more advanced, new technologies are required for better Digital Forensics. Although the application of Artificial Intelligence and Automation in Digital Forensics is still in its infancy, it has a lot to offer the field.

This review paper identifies potential opportunities provided by applying Artificial Intelligence and Automation principles and procedures to Digital Forensics and, as a result, applying intelligent techniques to digital investigation and address the issues of the larger and more complex domains where cybercrimes occur. This study identified and analysed a number of applications and frameworks for combining Artificial Intelligence and Automation with Digital Forensics. The study also reveals the potential impacts and challenges of incorporating Artificial Intelligence and Automation in Digital Forensics.

Cost reduction, improved efficiency and speed of forensic investigations, accurate data and information processing and increased probability of solving higher number of cases in limited amounts of time is the impact of Artificial Intelligence and Automation in Digital Forensics. We can see that that these different ways have simplified the life of a digital forensic investigator. Additionally, I believe that there is a need to increase public awareness of the use, applicability, and effects of Automation and Artificial Intelligence in the field of Digital Forensics. The general public, including individuals, groups, manufacturers, law enforcement and cyber security experts, needs to be made aware of these benefits.

References

Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, Geis JR, Pandharipande PV, Brink JA, Dreyer KJ. Current Applications and Future Impact of Machine Learning in Radiology. Radiology. 2018 Aug;288(2):318-328. doi: 10.1148/radiol.2018171820. Epub 2018 Jun 26. PMID: 29944078; PMCID: PMC6542626.

Costantini, S., De Gasperis, G. & Olivieri, R. Digital forensics and investigations meet artificial intelligence. Ann Math Artif Intell 86, 193–229 (2019). https://doi.org/10.1007/s10472-019-09632-y

D. A. Duce, F. R. Mitchell and P. Turner, 'Digital Forensics: Challenges and Opportunities', in John Haggerty and Madjid Merabti, (eds), ACSF 2007: Proceedings of the 2nd Conference on Advances in Computer Security and Forensics, (Liverpool John Moores University, School of Computing & Mathematical Sciences, 2007).

EC-Council Logo. 2022. *Digital Forensics*. [online] Available at: <https://www.eccouncil.org/what-is-digital-forensics/>
[Accessed 3 July 2022].

Exterro. 2022. *The Use Of Artificial Intelligence In Digital Forensics - Exterro*. [online] Available at: <https://www.exterro.com/blog/the-use-of-artificial-intelligence-in-digital-forensics> [Accessed 5 July 2022].

Fahdi, M. Al et al. "Towards An Automated Forensic Examiner (AFE) Based Upon Criminal Profiling & Artificial Intelligence." (2013).

Fahdi, M. & Clarke, Nathan & Li, Fudong & Furnell, Steven. (2016). A suspect-oriented intelligent and automated computer forensic analysis. Digital Investigation. 18. 10.1016/j.diin.2016.08.001.

Fei, B., Eloff, J., Olivier, M. and Venter, H. (2006) The use of self-organising maps for anomalous behaviour detection in a digital investigation, Forensic Science International 162, no. 1-3, 33–37.

Fei, B., Eloff, J., Venter, H., and Olivier, M. (2005) Exploring forensic data with selforganizing maps, Advances in Digital Forensics, Springer, pp. 113–123.

Garfinkel, Simson. (2010). Garfinkel, S.L.: Digital Forensics Research: The Next 10 Years. Digital Investigation 7(suppl.), 64-73. Digital Investigation. 7. 10.1016/j.diin.2010.05.009.

Grance, T. , Chevalier, S. , Scarfone, K. and Dang, H. (2006), Guide to Integrating Forensic Techniques into Incident Response, Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, [online], https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=50875

Irons, Alastair & Lallie, Harjinder. (2014). Digital Forensics to Intelligent Forensics. Future Internet. 6. 584-596. 10.3390/fi6030584.

J. Xiao, S. Li and Q. Xu, "Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation," in IEEE Access, vol. 7, pp. 55432-55442, 2019, doi: 10.1109/ACCESS.2019.2913648.

James, Joshua I & Gladyshev, Pavel. (2013). Challenges with Automation in Digital Forensic Investigations.

Jarrett, A, Choo, K-KR. The impact of automation and artificial intelligence on digital forensics. *WIREs Forensic Sci*. 2021; 3:e1418. https://doi.org/10.1002/wfs2.1418

Kanimozhi, V. and Jacob, T., 2019. *Artificial Intelligence based Network Intrusion Detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing*.

Masombuka, M., Grobler, M., & Watson, B. (2018). *Towards an Artificial Intelligence Framework by Actively Defend Cyberspace,* Reading, England: Academic Conferences International Limited.

Mitchell, F.. "The use of Artificial Intelligence in digital forensics: An introduction." *Digital Evidence and Electronic Signature Law Review* 7 (2014): 35-41.

Philip Turner, 'Unification of digital evidence from disparate sources (Digital Evidence Bags)', Digital Investigation (2005) 2(3), pp 223-228.

Pollitt, M., 2010, January. A history of digital forensics. In *IFIP International Conference on Digital Forensics* (pp. 3-15). Springer, Berlin, Heidelberg.

Solanke, A.A., Biasiotti, M.A. Digital Forensics AI: Evaluating, Standardizing and Optimizing Digital Evidence Mining Techniques. *Künstl Intell* (2022). https://doi.org/10.1007/s13218-022-00763-9

Reiber, L.(2018). How does AI contribute to digital forensics? Forebes.com. Retrieved from
https://www.forbes.com/sites/quora/2019/06/05/how-does-ai-contribute-to-digital-forensics/?sh=5fcb21f7c20a

Acknowledgment

I extend my sincere gratitude to those who helped in the successful completion of this review paper. I thank General Sir John Kotelawala Defence University for providing me with this opportunity to showcase my efforts and improve my research potential

Author Biography

MTA Deen is currently a BSc. Computer Science undergraduate in the Department of Computer Science, Faculty of Computing at General Sir John Kotelawala Defence University.

.

B Hettige is Head of the Department of Computer Engineering, Faculty of Computing, General Sir John Kotelawala Defence University. His research interests include Multi-Agent Systems, Machine Translation, Sinhala Language and Computational Grammar.