Indian Journal of Forensic Medicine and Pathology
Volume 14 Number 2, April - June 2021 (Special Isuue)
DOI: http://dx.doi.org/10.21088/ijfmp.0974.3383.14221.36

REVIEW ARTICLE

Various Applications of Artificial Intelligence in Digital Forensic Investigation

Kiran Singh[1], T Poongodi[2] Shrddha Sagar[3]

Authors' Affiliations:
[1]Assistant Professor,
[2,3]Associate Professor,
School of Computing Science and Engineering, Galgotias University, Greater Noida, Uttar Pradesh 201310, India.

Corresponding Author:
Shrddha Sagar

Email:
sagarshraddha@gmail.com

ABSTRACT

Most data, including books, images, photographs, medical records, and even human genetic material, has moved to digital formats in recent years. Laptops, smartphones, and connected devices are the primary sources of this digital data transformation, and they are quickly becoming an integral part of our everyday lives. We are becoming a soft target for different forms of cybercrime as a result of this transition. Digital forensic investigation allows you to retrieve files from a suspect's laptop that have been accidentally removed or hidden. The available manpower and government resources, however, are insufficient to investigate cybercrime. Unfortunately, current automated forensic protocols and practices necessitate a lot of human contact, which slows down the process and slows down the rate at which digital crimes are committed. Over the last few years, the use of Artificial Intelligence (AI) in Digital Forensic Science (DFS) approaches to cybercrime investigations has gotten a lot of attention. Traditional DFS techniques are no longer applicable to Digital Forensic (DF) investigators, as they frequently require a DF investigator to manually sift through data in order to locate relevant evidence. In order for Digital Forensic Science to keep up with the demands, which are compounded by Big Data, more intelligent Digital Forensic investigation techniques are needed. The main focus of our paper is to conduct a thorough study of the various AI algorithms and their implementation specifically in Digital forensics.

KEYWORDS | ai, machine learning, digital forensics, cybercrime

INTRODUCTION

The terms Digital Forensics (DF) is defined as "The application of scientifically proven and tested strategies in terms of preservation, compilation, authentication, identification, study, interpretation, recording, and presentation of digital evidence obtained from several digital sources in order to ease or further the reconstruction of criminal events, assisting in the detection of unethical practices that have been shown to be adverse to scheduled operations. Demand for DF are becoming increasingly significant in today's world. DF investigation protocols support in collecting critical data from a compromised computer system. And, businesses depend mostly on computers and mobile devices and the Internet nowadays. Moreover, it is essential to gather the necessary evidences from these devices. To promote or refuse any reasoning, an investigator might have information about the incident, where the digital evidences should be obtained from the connected systems. It's crucial to understand how to retrieve digital data that may be useful to investigators. Though, the existing human resources

How to cite this article
Kiran Singh. Various Applications of Artificial Intelligence in Digital Forensic Investigation. Indian J Forensic Med Pathol. 2021;14(3 Special):263-269.

and other available resources are insufficient to adequately prosecute digital crimes committed on digital devices.

Furthermore, current digital forensic techniques and methods necessitate a high level of human involvement, which speeds down the process in comparison to the rate at which digital crimes are committed. The majority of data is now gathered using digital devices and websites confiscated from criminals, or accessible on the Internet or exchanged by telecommunication companies. Forensic investigators should gather and examine information to determine where the crime is committed, who the criminal is, whom the criminal is targeting, whether the criminal is committing the crime, and how the crime happens. Thus, DF is a branch of criminalistics concerned with identifying, acquiring, preserving, analyzing, and presenting the information content of computer systems, or digital devices in general.

It's important to understand how AI is used when investigating AI as a weapon of crime. Investigators can compile and evaluate the dataset, training model, learning model, inference model, and implementation of the AI method used to commit a crime. Investigators should be able to understand the purpose of using AI's function after examining it. It is essential for investigators to be able to distinguish between the developer's objective and the AI outcome. AI programmes, unlike conventional programming, frequently have unintended consequences. Since several AI models exploit random weights in the learning process, AI parameters are often calculated with randomness. As a result, while using the same dataset and learning model, programmes with different parameters and outputs can be developed. It would be difficult to prove if AI was used as a tool, how AI was used, and how much damage AI caused because investigators would be unable to investigate the case.

AN OVERVIEW

By 1970, digital forensics, also known as electronic forensics, had been introduced.[1] The financial crime is found in the first investigation using the criminal's computer. In 1996, the first electronic fraud was reported. In 1996, the initial computer crime was identified in Texas, which ended in a five-year sentence.[2] With the growing popularity of the Internet, computer-based digital crimes began to increase in 1990. At the end of 1990s and beginning of 2000s, computer forensics emerged as a separate discipline. According to CSI surveys, nearly 46% of respondents have been impacted by computer crimes in some way.[3] According to a Gallup poll from 2010, 11% of American adults have been victims of Internet/Computer related crimes. And, the ratio is upto 6–8% high compared from the previous seven years. According to a study conducted by the "Australian Company Crime Survey," financial fraud and data breaches reached A$2,000,000 in 2006.[4] According to the Company Crime Survey, financial fraud and data breaches cost the company A$ 2,000,000 in lost revenue. With the emergence of various digital devices and their growing use for investigative purposes, the term "digital forensic" has become very common.

DF is a legal discipline that deals with the research, identification, conservation, security, extraction, recording, and other forms of digital data treatments. The aim is to transform the data to be assessed as part of a criminal or civil case after it has been processed into an information using appropriate types of reasoning. Consequently, the DF is a division of criminalistics whose practices are aimed at illustrating the presence of digital evidence related to the completion of a prosecution, or even prior to the investigative processes. It is a method of analyzing digital phenomena with the help of science and technology. Theories may also be used legally to address concerns about events that have occurred. DF has rapidly carved out a massive presence in business environments, with the aim of highlighting breaches of industrial secrets and/or security policies, as well as dealing with cyber incidents.

For the past two decades, global use of mobile smart devices has shot up dramatically, and they have become an integral part of our everyday lives. The word "smart computer" refers to a wide range of devices such as smartphones, laptops, GPS, tablets, and so on. Because of their processing power, large storage capacity, and low cost, these smart devices have become increasingly common. As a result, they can store a massive amount of commercial and personal data. These devices play an indispensable role in our everyday lives because they store users' personal and vital information. As emerging technology such as digital devices and the internet become more predominant nowadays, so the number of digital crimes also increases. At last, we're becoming victims of a variety of cybercrime and cyber attacks.

CONCLUSION

In today's world, technology advances at a breakneck pace, and we are constantly exposed to new innovations. AI is one of the most interesting areas of computer science, with a bright future. AI has the ability to make a computer function like a person.

Software forensic tests require the ability to analyze vast volumes of data in a timely manner in order to find relevant evidence during criminal investigations. Time and resource constraints, both computational and human, have a negative effect on the outcomes. As a result, better utilization of available resources is needed, in addition to the capabilities of currently used forensic equipment. The use of Artificial Intelligence in computer forensics is described in this paper. This system is made up of specialized intelligent agents that operate based on technical domain experts' expertise. Their aim is to interpret and compare the data found in an investigation's evidences, then show the most interesting information to the human investigator based on their experience, minimizing the amount of data that must be individually examined. The correlation function aids in the discovery of ties between evidences that a human expert might overlook, particularly given the large amount of data involved. The preservation, collection, and analysis of evidence found in digital storage media are all steps in the forensic investigation of computer systems. In a criminal investigation, digital evidence may be crucial in cases involving child trafficking, document forgery, tax evasion, and even terrorism. The relentless growth in the ability of digital storage media, as well as their widespread presence in everyone's everyday lives, has resulted in an increase in demand for such tests, as well as the amount of data to be examined. Furthermore, the existing collection of forensic methods is insufficiently efficient when it comes to evaluating a large number of evidences and correlating the results. As a result, the work of computer forensic experts is extremely time consuming. Since most forensic instruments lack distributed processing capabilities, the computing resources needed to conduct such investigations are also a concern. To deal with the three issues listed above, a variety of approaches[5] have been proposed: (i) a reduction in the number of evidences to be investigated, (ii) proof correlation, and (iii) forensic examinations computational work distribution. As a branch of the forensic sciences, digital forensics is confronted with new challenges as potential digital evidence grows and expands.[6] Innovative methods for automated investigations are being developed at a rapid pace in the fields of computer science and information technology. Algorithms and techniques for machine learning are used in a wide range of applications. In order to work more effectively, machine learning engineers and forensic analysts must have a thorough understanding of the algorithms in use, how they work, and how they learn from raw data. Digital forensics[7] is becoming a more advanced topic, as well as an important field that often necessitates the analysis and extraction of a large amount of complex data from a crime scene. Examining digital information of the committed crime to be used as legal proof in a court of law is part of a digital forensic investigation. In the process of collecting and analyzing digital data, a variety of machine learning algorithms and techniques can be useful. Machine learning can improve this method by allowing it to deal with large amounts of data in a short period of time while maintaining a high degree of accuracy and producing high-quality results.

Algorithms of AI and ML in Digital Forensic

AI systems have made tremendous progress in solving increasingly complex computational tasks using machine learning as the core technology, making them crucial components of human society's future growth.[8] However, as ML algorithmic models chase prediction precision and become increasingly implicit, explainability becomes a challenge for black-box techniques like ensemble methods and deep neural networks.[9] An examination of current ideas about the application of artificial intelligence technology in forensic science reveals that they are almost always linked to one of the following areas: Traditional peer research capacity building; algorithmizing of the crime investigation process; crime prevention (recognition of signs of imminent crime, etc.).

The described directions are focused on processing big data with machine learning elements that operate within a limited number of parameters to a greater extent. We are currently not considering the use of artificial intelligence in these areas in its entirety, since only a portion of the abilities inherent to the human intellect are used [10]. Simultaneously, technological advancements would inevitably lead to the creation of more advanced approaches that offer new artificial intelligence capabilities comparable to human intelligence.

Inductive reasoning and deductive reasoning are the two key methods used to characterize ML forensics:

i. Inductive Reasoning is based on a broad understanding of particular data. The knowledge gained is fresh and does not preserve the facts. This means that new information will invalidate previously acquired knowledge. There isn't a single well supported hypothesis.There are numerous objectives in this field, including the need to discover general concepts from a limited collection of examples. The examples are referred to as "experience." The foundation for this is to look for common characteristics in different instances. Inductive learning is the foundation of these approaches.

ii. Deductive Reasoning derives insight from well established logic techniques. Deductive reasoning uses well established techniques to derive information from experience. The information is not fresh. However, it is implied in the initial understanding. Established knowledge and its foundation in mathematical logic cannot be invalidated by new knowledge.

AI has been studied in a variety of academic settings, as mentioned in the introduction. From different viewpoints, this section discusses research on the AI security threat and AI-related crime. We also look at cybercrime as identified by the cybersecurity and digital forensics communities. Adopting online personas, known as socialbots, that act like humans is a prime example of malicious AI use.[11]

PRINCIPLE	TRADITIONAL FORENSICS	AI FORENSICS
Meaning	When gathered, the facts and context remain unchanged.	During the learning process, the interpretation of the AI system changes.
Errors	In forensic method, errors may be reported	
Transparency and trustworthiness	Several approaches have been tried and proven.	It's needed for the AI forensic process to be verified.
Reproducibility	It demonstrates a high degree of consistency in quality	Even using the same dataset and learning model, it would be impractical.
Experience	There is a wide range of research. and education available	It should be researched.

Table 1 describes the difference between the conventional forensics and digital forensic using AI techniques. Although the original goal of socialbot was to increase public awareness and collaboration, it has also been used for malicious purposes including phishing, fraud, and political manipulation of online social media campaigns.[12,13] Since it is based on a single user's past activities and public profiles, detecting the malicious socialbot has become a computer security problem. When malicious socialbots are programmed to carry out a political attack, according to social science, the technique has the potential to influence or inflame public opinion.

Some researchers claim that malicious hackers have already begun to use AI to improve their hacking skills and create new forms of cyber attacks. Financial frauds, cyber terrorism, and cyber extortion are all popular cybercrime techniques that have been enhanced with AI. When attempting voice phishing, for example, hackers can fool victims by convincingly imitating the voice of the victims' relatives or friends.[14]

In comparison to previous research, Brundage et al.,[15] concentrate on the potential problems that particular techniques can cause. They looked at three different shifts in the threat landscape: the proliferation of emerging threats, the rise of new threats, and a shift in the threat's traditional cost-cutting approaches (for example, mass spear phishing) allow attackers to invade more targets, resulting in the spread of established threats. The cost of tasks that require human labor can be reduced because of the AI system's scalability. New threats can emerge in order to complete tasks that humans are incapable of completing (for example, imitating individual voices or controlling multiple drones).[16] The typical character of threats may change as highly successful AI attacks become more widespread.

King *et al.,*[17] offered a fresh take on AI defense. They are raising a threat by using the word 'AI crime.' AI crime includes trade, financial markets, and insolvency (e.g. market manipulation, price fixing, collusion), toxic or hazardous substances (e.g. drug distribution, selling, purchasing, possessing banned drugs), and crimes against individuals (e.g. harassment, torture), Theft and fraud, forgery and personation (e.g., spear phishing, credit card fraud). They asked that each of the offences be labelled as having one or more threats associated with it. They focused on human nature when classifying AI security risks: emergence, transparency, surveillance, and control.The psychology danger, for example, means that AI can influence a person's mental state to the point that it promotes or triggers illegal behaviour.

Concerns about AI privacy resulting from the processing of personal data have been the subject of some study. AI applications in healthcare, banking, and education, according to Li and Zhang [18], can lead to privacy concerns. Developers want to obtain as much data as possible because the quantity and quality of training data has a huge effect on AI performance. According to Li *et al.,* the collection of detailed data entails inherent privacy risks.

The previous study has three implications for AI stakeholders. First and foremost, researchers and engineers should be aware that, due to AI's dual-use nature, the device could be used to conduct illegal acts, even though it is intended for legitimate purposes. Since artificial intelligence is a double-edged sword, anyone working in this area must adhere to strict professional ethics. Second, completely new forms of security threats that have never been considered before will arise.

Since AI may perform tasks that were previously thought to be difficult for humans or conventional systems to perform, the risks may be outside the reach of current threats. To avoid AI security threats and respond to AI crime, AI researchers should collaborate with experts from other fields. Finally, the cybersecurity sector's trials and tribulations could help the AI security field. Predictable AI crimes are inextricably related to cybercrime, as previously mentioned in previous studies. This object has two functions. The presence of ICT led to the emergence of cybercrime; the current state of AI security is similar to that of cybersecurity in its early stages.

Comparative study on various AI applications in Digital Forensics

The use of artificial intelligence (AI) improves the chances of detecting and investigating cybercrime. This enables forensic experts to get to the root of the problem rapidly and effectively.

AI aids in the quick resolution of a crime and saves police a lot of money. This will allow them to concentrate more on the areas where cybercrime is most likely to occur. By sifting through unstructured data collected by police, AI can identify suspicious and criminal activity.

Cognitive-Data Analytics is a form of AI that allows you to quickly digest and analyse data. It can also make it easier for investigators to search through criminal records and find possible suspects. [19]

AI will assist in the recognition of specific elements in images and videos that are being investigated. AI may also assist in identifying commonalities in contact, place, and time. This allows authorities to pinpoint the location of the next crime or incident.

Here are some AI techniques that have an effect on digital forensics:

Knowledge Representation

This has to do with what a computer program needs so that it intelligently executes tasks, and also how quantitative methods can feed this information to the program. It can be used to develop better strategies for defending against cyber attacks.

Expert Systems

This clarify why certain processes are carried out and the findings reached during the digital forensic' investigation. It enables an individual to examine and criticize the process and reasoning employed. This can reveal shortcomings in the methods used to draw conclusions. They also make data processing faster.

c) Pattern Recognition

In an investigation, this distinguishes specific types of data clusters. It can be used to identify the contents of pictures, spam emails, and directories on hard drives that contain suspicious files. When used in conjunction with information exploration, it can help analysts spot trends in massive quantities of data.

Some Artificial Intelligence Techniques that Can Aid in Digital Forensics:

a) This recognizes, restricts, and removes threats using live forensics. It's also a good idea to scan criminal history to see who could be responsible.

b) Data recovery is the process of restoring data that has been lost or destroyed. In digital forensics, this is an important technique for retrieving possible evidence.

c) Password recovery is crucial during forensics when password-protected files are to be accessed. It gives you access to these files, which can be presented as evidence.

d) Investigators may use known file filtering to find files that are important to their investigation.

e) Investigators may use a timeline analysis to see the sequence of events that led up to the incident under investigation.

In view of the above, we can see how AI can help forensic analysts in a variety of ways. It removes the need for investigators to sift through various data sources, saving them precious time and effort.

CONCLUSION

After discussing the use of AI and its possible applications in forensic science and criminal investigation to support forensic scientists, police officers, and security personnel, it is obvious that Artificial Intelligence applications or software can assist investigators in greatly reducing the amount of time spent on various tasks at various stages of an analysis and investigation. Saving time will eventually lead to increased efficiencies in the disposition of cases, potentially contributing to the goal of reducing the number of cases that are pending due to sluggish and complicated investigation procedures. More precision, competence, and a lack of prejudice would ultimately lead to proper criminal justice. To make forensic investigation techniques more advanced, scientists and researchers have been designing more AI-based software programs and devices. Our forensic investigation and predictive policing systems as well as our security and defense systems will benefit from new advances in AI. IJFMP

REFERENCES

1. *Pollitt M. A history of digital forensics. In: IFIP International Conference on Digital Forensics. Springer, Berlin, Heidelberg: Springer; 2010. pp. 3-15.*
2. *Dierks MP. Computer network abuse. Harvard Journal of Law & Technology. 1992;6:307*
3. *Richardson R, Director C. CSI computer crime and security survey. Computer Security Institute. 2008;1:1-30*
4. *A.C.E.R.T.A. 2006 Australian Computer Crime and Security Survey. AusCERT& Australian High Tech Crime Center (AHTCC); November 23, 2006.*
5. *Hoelz, B. W., Ralha, C. G., &Geeverghese, R. (2009, March). Artificial intelligence applied to computer forensics. In Proceedings of the 2009 ACM symposium on Applied Computing (pp. 883-888).*
6. *Ganesh, V. (2017). Artificial Intelligence Applied to Computer Forensics. International Journal, 5(5).*
7. *Qadir, A. M., &Varol, A. (2020, June). The role of machine learning in digital forensics. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-5). IEEE.*
8. *West DM. The Future of Work: Robots, AI, and Automation. Washington, D.C: Brookings Institution Press; 2018*
9. *Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, et al. Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion. 2020;58:82-115*
10. *Rubinstein S L 1973 Problems of general psychology (Moscow, Russia: Pedagogy)*
11. *Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu, "Key challenges in defending against malicious socialbots," presented at the 5th USENIX Workshop Large-Scale Exploits Emergent Threats, 2012*
12. *R. W. Gehl and M. Bakardjieva, "Socialbots and their friends," in Socialbots and Their Friends. Evanston, IL, USA: Routledge, 2016, pp. 17–32. S. Rathore, P. K. Sharma, V. Loia, Y.-S. Jeong, and J. H. Park, "Social network security: Issues, challenges, threats, and solutions," Inf. Sci., vol. 421, pp. 43–69, Dec. 2017.*
13. *O. Bendel, "Thesynthetization of human voices," AI & Soc., vol. 34, no. 1, pp. 83–89, 2019.*
14. *M. Brundage et al., "The malicious use of artificial intelligence: Forecasting, prevention, and mitigation," 2018, arXiv:1802.07228. [Online]. Availab le: ttp:// arxiv.org/abs/1802.07228*
15. *G. Allen and T. Chan, Artificial Intelligence and National Security. Belfer Center for Science and International Affairs. Cambridge, MA, USA: Belfer Center for Science and International Affairs, 2017. [Online]. Available: https://www.belfercenter. org/ sites/default/files/files/publication/AI%20 NatSec%20-%20final.pdf*
16. *T. King, N. Aggarwal, M. Taddeo, and L. Floridi, "Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions," SSRN Electron. J., vol. 26, no. 1, pp. 1–32, 2019.*
17. *Li and T. Zhang, "An exploration on artificial intelligence application: From security, privacy and ethic perspective," in Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal. (ICCCBDA), Apr. 2017, pp. 416–420.*
18. *Richmond, K. M. (2020). AI, Machine Learning, and International Criminal Investigations: The lessons from forensic science.*