Project 3 Part 2

Blake Moore

1. Introduction

Digital forensics is the application of computer science and investigative procedures in order to examine digital evidence derived from digital sources. [1] This includes persevering, collecting, validating, identifying, analyzing, interpreting, documenting, and presenting digital evidence for the reconstruction of criminal events. Artificial intelligence is the science and engineering of making intelligent machines. [2] This means with artificial intelligence (AI), we can engineer computers to perform tasks that traditionally require human intelligence to complete.

The AI field has been growing and evolving very quickly over the last few years and we have seen many different advancements in the capabilities of artificial intelligence. AI can be a tool for many different tasks that humans do, allowing for work to be performed more efficiently and yield the best results possible. This is seen in many different fields, but it is important to study how AI interacts with digital forensics for a few reasons. Artificial intelligence not only can improve the efficiency of digital forensics work, but also ease the workload on the investigators, allowing them to focus on other aspects of their work. By studying AI and its capabilities, we can create new implementations and uses for it, making the digital forensics process easier and more efficient.

2. History and Evolution

Digital forensics formally began in the 1990s. There was some investigation being done before this, however, with the original investigators being “computer forensics technicians”. These original technicians were law enforcement officers who also happened to be computer hobbyists. [3] In 1984 the FBI formed their Computer Analysis and Response Team (CART). In the 1990s it was realized that digital forensics as a field needed standard procedures and protocols to follow. Multiple conferences were held throughout the 1990s, causing the formation of a few different groups to address the issues and lack of standardization within the digital forensics field.

These groups were originally called Technical Working Groups, but in 1999 the name was changed to Scientific Working Groups. [4] This was done to distinguish long-term groups from short-term groups. Short-term groups, or the Technical Working Groups, were only together for a short duration of time, and usually only created a single deliverable, such as a guidebook. Long-term groups, or the Scientific Working Groups, were long-term groups that met at least once a year and had different long-term goals to make improvements in the field of digital forensics. [4] These different kinds of groups have been working over the past few decades to introduce new standards, best practices, and tools for digital forensics investigators to use. Artificial intelligence is another one of these tools that can be utilized by investigators to make their jobs more efficient and easy.

AI started as a concept in science fiction and became something that scientists and researchers would be familiar with through their cultural experiences. In 1950, Alan Turing wrote a paper exploring the possibility of artificial intelligence. [5] Five years later, there was a conference hosted by John McCarthy and Marvin Minsky that aligned researchers with the sentiment that AI was an achievable technology. [5] Over the following years, there were many different successes and setbacks in artificial intelligence, all of which led to a more and more realistic version of the AI we knew from science fiction.

More recently, AI has been recognized as a powerful tool for various tasks, such as pattern recognition and data analysis. Digital forensics investigators face many different challenges in their work, including data overload, data variety, data complexity, and anomaly detection. [6] These challenges were realized as issues that could be mitigated with the use of artificial intelligence. AI can process large amounts of data quickly, analyze it, and identify anomalies automatically. AI has grown to offer many different tools for investigators, such as automation tools, predictive analysis software, and digital evidence analysis tools. [6] Using AI and the many different tools based off of it, investigators can now do their jobs more efficiently.

3. Current Applications of AI in Digital Forensics

Artificial intelligence has become a strong tool for digital forensics investigators to use to aid in their investigations. One such tool is a type of neural network, called a multi-layered network (MLNN). An MLNN is capable of analyzing and classifying images. [7] Image classification is an important part of digital forensics due to a number of reasons. Images recovered as digital evidence can identify criminals, be illegal material themselves, or provide evidence needed for a criminal investigation. Using an MLNN investigators can allow the AI to quickly sift through many images and identify if any of them contain illegal material. However, these image classifiers can sometimes misclassify images. For standard cases, this is not a major issue, but in digital forensics investigations, these misclassifications can cost both time and money if they lead investigators down the wrong trail.

Malware detection is another section of digital forensics where artificial intelligence can now be helpful for investigators. As mentioned earlier, AI is well-suited for pattern recognition. This means that investigators can use AI tools that can scan software for known malicious patterns. This allows investigators to easily identify malware without having to extensively analyze it themselves. Another advantage to using AI to analyze software for malicious content is that the artificial intelligence can identify and classify previously unknown malware. [8] This is due to it being able to learn different types of patterns that indicate malicious software and recognize when there are patterns similar, but not exactly the same as the known patterns. This is very useful for investigators, because the AI may be able to identify malware that the investigator might not have discovered themselves.

Artificial intelligence tasked with malware detection can create its own set of issues, however. A malware detection AI can run into the same false positive issue as the image classifier. In addition to this, the artificial intelligence could return false negatives when scanning software. If an investigator relied on the AI to determine what is malware and what is not, they might run a piece of malware that the AI told them was safe. This could lead to many different issues, including losing the copy of digital evidence the investigator was working with. The other issue is that malware is a rapidly changing field. While the AI may be able to detect malicious content that is similar to known malware, it would not be able to detect completely new methods and types of malware.

Another subsection of digital forensics that is currently being aided by artificial intelligence tools is network traffic analysis. Network traffic analysis is when an investigator monitors and analyzes network traffic for potential malicious activity. Normally this would require the investigators to manually analyze the network traffic at certain intervals. This is not an ideal way to perform this task. Due to not being able to constantly surveil the network traffic, it becomes very easy to miss malicious behavior due to not checking at the correct time. This process would also take a lot of manual investigation time, due to the investigator having to analyze the contents of the traffic themselves. Both of these issues are solved by utilizing an AI solution.

Investigators can use artificial intelligence tools to automatically analyze network packets. [8] This means that the AI can continuously monitor the network traffic, mitigating the interval timing issue that comes with manual analysis. With the AI monitoring the network traffic, no malicious traffic should be missed due to the continuous nature of the analysis. Also due to the automatic scanning of the AI, the investigator spends little to no time manually analyzing the network traffic. This implementation of AI would also be susceptible to returning false positives and negatives. Despite this, this implementation could also benefit from being able to detect previously unknown malicious behavior by finding similarities to known malicious activity, the same as the malware detection implementation.

4. AI-driven Forensics Tool Case Study

Veritone Tracker is an AI-powered tool that helps investigators track individuals across different videos. Veritone is a software company that primarily focuses on artificial intelligence solutions. Veritone is the company that owns and develops this tool, which is built off of their enterprise AI platform aiWARE. This company provides tools that can analyze audio, videos, biometrics, speech, and text. This specific Veritone Tracker tool has the ability to find persons of interest across different videos without any personally identifiable information. [9]

This tool allows the user to identify a person within a video as a “human-like object”. [9] This object is then passed to the AI system, which uses the object to identify and track similar individuals across different videos. This artificial intelligence integration allows the AI to build a profile for the person it needs to track, and automatically search through many different videos searching for this profile.

These different functionalities allow investigators to quickly build timelines detailing events that have transpired on camera for their investigation. This AI tool would also allow investigators to automatically scan camera footage to help find specific persons of interest. If they are scanning for a criminal, this tool can help find and apprehend these individuals quickly. Veritone Tracker can also help investigators in missing-person cases by scanning videos and identifying individuals automatically. This allows investigators to increase the amount of footage that they can intake and process, meaning that they are more likely to find and help more individuals more often.

Veritone Tracker provides many advantages for digital forensics investigators, primarily focused on the automation of the tool. This saves time and makes the investigation process easier for the humans performing it. This tool could also fall into the same issues discussed in section 3 of this paper. It has the possibility of returning both false positives and negatives to the user, potentially causing time loss. Overall, this tool can prove to be useful to investigators due to its ability to automatically scan large amounts of videos for persons of interest.

5. Ethical Implications

Artificial intelligence is at the forefront of the ethics conversation in the computer science and software engineering field at the moment. Utilizing AI in a digital forensics investigation causes many different ethical concerns and considerations. One example of such concerns has already been mentioned multiple times in this paper with examples of false positives and negatives. The accuracy of artificial intelligence is a legitimate concern, especially when it comes to law enforcement scenarios. The idea is that investigators could become too reliant on the results that their AI tools return to them, and treat the AI's results as facts. This becomes an ethical concern because artificial intelligence is not perfect. If an AI tool returns an incorrect result to an investigator who treats the result as fact, the investigator could incorrectly pursue and attempt to prosecute an individual not guilty of any illegal activity.

Another major concern when using AI tools in digital forensics is privacy. This is a prevalent concern when it comes to facial recognition technology or any artificial intelligence tool that attempts to identify people. These tools have to build profiles for the different people it attempts to identify, and these profiles need to be stored in some way somewhere. Data breaches are a serious problem that can have a large effect on innocent people's privacy. In a normal situation, someone may have their credit card information stolen, in which case the affected person can cancel their card and apply for a new one. However, if a data breach happens and someone is able to access these AI profiles of different individuals, it provides attackers with a new avenue to steal people's identities.

Due to these vulnerabilities and even more, it is important for digital forensics investigators to act ethically and responsibly when using artificial intelligence tools in their investigation process. There are a few different standards for investigators to follow to attempt to combat the ethical concerns that come with using AI in their investigations. Digital forensics investigators should strive for integrity, accuracy, and data minimization. [10] Integrity in this context references the validity of information gathered from digital evidence. Accuracy is defined as how closely the gathered evidence represents the real world events. Finally, data minimization is the principle that data should not be used unless it is essential for the reasons stated clearly in advance. [10] Investigators should work to uphold these principles in their work, especially when working with AI-based tools. When using artificial intelligence, investigators need to monitor and validate the results, as well as uphold privacy standards.

6. Future of AI in Digital Forensics

Artificial intelligence already has many uses within digital forensics today, but it is believed that its role in the industry will grow as time progresses. This is due to predicted future innovations in both the technology used in illegal activities and the AI tools used by digital forensics investigators.

One such potential use case would be combating deepfake videos. Deepfake videos are currently being used to create misinformation in social media and will only improve in quality as the technology used to make them improves. One of the ways to combat this is to train AI models and tools to identify if a piece of media is a deepfake video or not. This will also be benefited by AI's ability to automatically process content and analyze it, meaning that posts made to social media sites can be quickly searched and removed if they contain harmful deepfake content.

It is also expected that artificial intelligence will be further utilized in data intake and analysis. As technology progresses computers are able to work with larger and larger sets of data. Eventually, the amount of data investigators need to analyze in search of malicious activity may become too large for them to manually examine. Artificial intelligence tools would be able to automatically intake these large data sets and analyze them for the investigators.

These AI data analysis tools will also become even more useful for their pattern recognition capabilities. As malicious actors learn to hide their behavior, it will become increasingly more difficult for investigators to manually identify harmful traffic or data. AI tools can scan the contents of network traffic packets or other data quickly and identify malicious content, or even discover previously unknown malicious content.

As AI facial recognition or other tools that track and identify people become more advanced and widely used, privacy will become a difficult standard to uphold. Artificial intelligence is exceptional at being able to analyze every detail of the information it is analyzing, meaning every aspect of a person being investigated can be scrutinized and stored by the AI. Due to this, now and moving forward into the future developers and investigators will need to find a balance between the needs of the investigation and the privacy rights of individuals. It is up to developers to not create AI tools that would infringe upon the privacy of individuals. It is the responsibility of the investigators who use the tools to not attempt to use the AI tools in a way that would infringe on individual privacy rights.

7. Conclusion

Artificial intelligence has been a growing and improving technology over the past few years. It has many uses, but its utilization with digital forensics is especially useful. Artificial intelligence is especially useful for pattern recognition and automatically processing large amounts of data. Due to these strengths, AI tools excel in the digital forensics field. There are already AI-powered tools being used that automate portions of investigator's work, and as the technology grows stronger it will become more commonly used.

This increased use of artificial intelligence tools creates some new security and ethical risks, however. AI tools can be used as facial recognition or personal identification tools, raising a new type of privacy concern. Due to these new types of issues and ethical concerns, it will be important for developers and investigators alike to create and use AI tools responsibly and ethically.

Overall, artificial intelligence seems to have a bright future in the digital forensics field. The strength of automation and the potential AI has to detect malicious behavior in data, images, videos, and network traffic is a powerful tool for digital forensics investigators to utilize. Investigators and developers need to ensure that they uphold privacy and security standards when using and creating AI tools. As long as these standards are upheld, and AI tools become more ingrained into the digital forensics investigation process, society and the digital world will become safer landscapes.

8. References

[1] NIST, “Glossary | CSRC,” *Nist.gov*, 2020. https://csrc.nist.gov/glossary

[2] C. Manning, “Artificial Intelligence Definitions,” Stanford University, Sep. 2020. Available: https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

[3] “Digital forensics,” *OpenLearn*, 2012. https://www.open.edu/openlearn/science-maths-technology/digital-forensics/content-section-4.2

[4] Carrie Morgan Whitcomb, “An Historical Perspective of Digital Evidence: A Forensic Scientist’s View,” *International Journal of Digital Evidence Spring*, vol. 1, no. 1, 2002, Available: https://www.utica.edu/academic/institutes/ecii/publications/articles/9C4E695B-0B78-1059-3432402909E27BB4.pdf

[5] R. Anyoha, “The History of Artificial Intelligence,” *Science in the News*, Aug. 28, 2017. https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

[6] S. K. Das, “The Digital Detective’s New Partner: AI in Digital Forensics,” *Medium*, Oct. 03, 2023. https://medium.com/@sourabhkumardas/the-digital-detectives-new-partner-ai-in-digital-forensics-6213f0ecbbd6

[7] A. A. Solanke, “Explainable digital forensics AI: Towards mitigating distrust in AI-based digital forensics analysis using interpretable models,” *Forensic Science International: Digital Investigation*, vol. 42, p. 301403, Jul. 2022, doi: https://doi.org/10.1016/j.fsidi.2022.301403.

[8] S. A. T. Officer, ExterroAugust 29, and 2023, “6 Ways AI Can Revolutionize Digital Forensics,” *Dark Reading*, Aug. 29, 2023. https://www.darkreading.com/dr-tech/6-ways-ai-can-revolutionize-digital-forensics

[9] “An AI-Powered Digital Forensics Solution | Veritone Tracker,” *Veritone*. https://www.veritone.com/applications/tracker/

[10] M. I. Maratsi, O. Popov, C. Alexopoulos, and Y. Charalabidis, “Ethical and Legal Aspects of Digital Forensics Algorithms: The Case of Digital Evidence Acquisition,” *15th International Conference on Theory and Practice of Electronic Governance*, Oct. 2022, doi: https://doi.org/10.1145/3560107.3560114.