www.ijarp.com ISSN 2456-9992 Page: 01-24

International Journal Advanced Research Publications

Review Article

Volume: 02
Issue: 09

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING FOR AUTOMATED DIGITAL EVIDENCE TRIAGE AND PRIORITIZATION: A REVIEW OF METHODS, APPLICATIONS, CHALLENGES, AND RESEARCH GAPS

***[1]Gloria Ebere Amadi, [2]Mba Chioma Juliet**

[1]Department of Computer Science, Enugu State University of Science and Technology (ESUT), Enugu, Nigeria.

[2]Enugu State Polytechnic Iwollo.

Article Received: 13 August 2026, Article Revised: 03 September 2026, Published on: 23 September 2026

***Corresponding Author: Gloria Ebere Amadi**

Department of Computer Science, Enugu State University of Science and Technology (ESUT), Enugu, Nigeria

Doi: https://doi-doi.org/101555/ijarp.1264

ABSTRACT

The rapid proliferation of digital technologies has resulted in an unprecedented growth in the volume, diversity, and complexity of digital evidence available to forensic investigators. Computers, smartphones, cloud platforms, Internet of Things (IoT) devices, messaging applications, social media platforms, and networked systems continuously generate digital artefacts that may become relevant to criminal, civil, corporate, and cybersecurity investigations. Conventional digital forensic processes rely substantially on manual examination and established forensic tools, but the increasing volume of evidence creates significant challenges for timely investigation and contributes to digital forensic backlogs. Artificial intelligence (AI) and machine learning (ML) have consequently emerged as important technologies for supporting automated evidence triage, artefact classification, anomaly detection, relevance assessment, and evidence prioritization. This paper reviews the application of AI and ML to automated digital evidence triage and prioritization, with particular attention to supervised classification, Decision Tree approaches, Bayesian methods, anomaly detection, metadata-based classification, and human-in-the-loop forensic analysis. The review is based on the author's existing literature review and an expanded examination of scholarly publications addressing machine learning and automation in digital forensics. The literature indicates that ML can support the identification and prioritization of potentially relevant artefacts and can reduce the amount of evidence requiring immediate detailed

examination. However, the literature also reveals important limitations involving dataset availability and quality, class imbalance, model generalizability, explainability, false positives and false negatives, encryption, anti-forensic techniques, interoperability, evidence integrity, privacy, legal admissibility, and limited empirical validation in operational environments. The review identifies a research gap in the integration of classification, anomaly detection, relevance assessment, and priority scoring within a unified, investigator-centred evidence-triage process. The paper proposes research directions involving explainable AI, representative forensic datasets, human-in-the-loop validation, cross-domain evaluation, scalable processing, interoperability, and standardized evaluation protocols. The review concludes that AI/ML has considerable potential to support digital forensic triage, but its adoption should complement rather than replace professional forensic judgment.

KEYWORDS: Artificial Intelligence, Machine Learning, Digital Forensics, Digital Evidence, Evidence Triage, Artefact Classification, Evidence Prioritization, Anomaly Detection, Digital Forensic Backlog, Human-in-the-Loop.

1. INTRODUCTION

Digital technologies have fundamentally transformed the nature of evidence available during contemporary investigations. Digital traces may originate from desktop and laptop computers, smartphones, cloud platforms, network devices, IoT systems, removable storage, social media, messaging applications, and enterprise information systems. As the use of these technologies continues to expand, the quantity of digital evidence requiring forensic examination has increased correspondingly.

The growth of digital evidence represents one of the major challenges confronting digital forensic investigators. A single device may contain thousands or millions of files, application records, browser artefacts, system logs, images, videos, metadata records, communications, and deleted information. Investigations involving multiple devices or distributed cloud and IoT environments may generate considerably larger evidence collections. The author's underlying review identifies the increasing complexity of digital evidence as a consequence of IoT, cloud computing, encrypted communication, distributed systems, and other emerging technologies.

The conventional digital forensic process remains essential because evidence must be properly collected, preserved, examined, analyzed, and reported. NIST SP 800-86 describes digital forensics in terms of collection, examination, analysis, and reporting, with

examination involving both automated and manual processing while preserving evidence integrity.

However, automation has become increasingly important because forensic investigators cannot reasonably treat every artefact in a large evidence collection as equally important. Michelet et al. (2023) identify the increasing volume of digital evidence as a central motivation for automation and argue that the field requires clearer terminology and conceptual foundations for different levels of forensic automation.

Artificial intelligence and machine learning provide opportunities to address this challenge. ML algorithms can learn patterns from previously classified evidence and use those patterns to classify or prioritize newly encountered artefacts. Potential applications include artefact classification, relevance determination, anomaly detection, malware identification, image analysis, text classification, behavioural analysis, and evidence prioritization.

The literature reviewed for this study demonstrates that machine learning has already been applied to several digital forensic domains. Nayerifard et al. (2023), in a systematic literature review covering studies published between 2010 and 2021, identified increasing use of machine learning across digital forensic domains and reported substantial use of deep-learning approaches, particularly convolutional neural networks, in image forensics.

A particularly relevant application is automated digital evidence triage. Du and Scanlon (2019) proposed a supervised machine-learning methodology for automatically prioritizing suspicious digital artefacts. Their approach uses features extracted from forensic artefacts and information from previously processed cases to recommend artefacts that are potentially relevant to an investigation. Importantly, the approach is explicitly human-in-the-loop: the model recommends or predicts suspicious artefacts but does not replace the final forensic analysis.

The author's literature review similarly identifies machine learning as an emerging approach for artefact classification, anomaly detection, and evidence prioritization. It also identifies Decision Tree and Bayes Net approaches among the machine-learning methods reported in digital forensic research.

Despite these developments, significant research challenges remain. The reviewed literature identifies limitations involving interoperability, standardization, scalability, training and skills, encryption, anti-forensics, ethical and legal considerations, and limited empirical validation.

This paper therefore reviews AI and ML approaches specifically from the perspective of **automated digital evidence triage and prioritization**.

2. Aim and Objectives of the Review

The aim of this review is to examine the application of artificial intelligence and machine learning to automated digital evidence triage and prioritization and to identify the major research gaps that must be addressed before such approaches can be widely adopted in operational digital forensic environments.

The objectives are to:

1. Examine the evolution of automation in digital forensics;
2. Review applications of ai and ml in digital forensic investigation;
3. Examine machine-learning approaches for digital artefact classification;
4. Review decision tree and bayesian approaches to forensic classification;
5. Examine ai/ml approaches to anomaly detection;
6. Review automated evidence relevance determination and prioritization;
7. Examine human-in-the-loop approaches to ai-assisted forensic analysis;
8. Identify technical, forensic, legal, and ethical limitations;
9. Identify gaps in existing research; and
10. Propose directions for future research on ai-assisted digital evidence triage.

3. Review Questions

The review was guided by the following questions:

RQ1: How is artificial intelligence and machine learning currently being applied in digital forensic investigation?

RQ2: Which machine-learning techniques have been applied to digital artefact classification and evidence triage?

RQ3: How are AI/ML techniques being used for anomaly detection and evidence relevance determination?

RQ4: How can machine learning support digital evidence prioritization and forensic backlog management?

RQ5: What challenges limit the operational adoption of AI/ML-based digital evidence triage?

RQ6: What research gaps remain in AI/ML-assisted digital evidence triage?

4. Review Methodology

4.1 Review Design

This study adopts a structured narrative review approach supported by thematic analysis of the existing literature. The review combines the author's existing literature review with additional scholarly publications addressing AI, ML, automation, evidence classification, evidence relevance, and digital forensic investigation.

The purpose is not to conduct an empirical test of a machine-learning model. Rather, the purpose is to synthesize existing knowledge, compare methodological approaches, identify limitations, and establish areas requiring further investigation.

4.2 Literature Sources

The review considered scholarly literature concerning:

- Artificial intelligence in digital forensics;
- Machine learning in digital forensics;
- Automated forensic investigation;
- Digital evidence triage;
- Artefact classification;
- Metadata-based evidence classification;
- Anomaly detection;
- Evidence relevance determination;
- Evidence prioritization;
- Forensic automation;
- Human-in-the-loop ai;
- Forensic datasets; and
- Digital forensic challenges.

The review was supplemented by authoritative forensic guidance, including NIST's guidance on the digital forensic process. NIST describes collection, examination, analysis, and reporting as core forensic activities and explicitly recognizes the combination of automated and manual methods during examination.

4.3 Selection Criteria

Studies were considered relevant where they addressed one or more of the following:

1. Ai or ml application in digital forensics;
2. Automated classification of digital artefacts;

3. Automated evidence triage;
4. Anomaly detection;
5. Forensic evidence relevance;
6. Evidence prioritization;
7. Forensic automation;
8. Machine-learning-based analysis of digital evidence; or
9. Challenges associated with ai-assisted digital investigations.

Studies focused exclusively on general cybersecurity machine learning without a meaningful digital-forensic application were excluded from the core synthesis.

4.4 Thematic Analysis

The selected literature was synthesized under the following themes:

1. Automation in digital forensics;
2. Ai/ml applications;
3. Artefact classification;
4. Decision tree approaches;
5. Bayesian approaches;
6. Anomaly detection;
7. Evidence relevance;
8. Evidence prioritization;
9. Human-in-the-loop investigation;
10. Forensic backlog management;
11. Challenges and limitations; and
12. Future research.

5. Automation in Digital Forensics

Automation in digital forensics refers broadly to the use of computational mechanisms to perform or support investigative tasks that would otherwise require manual intervention.

Michelet et al. (2023) note that the term automation is not uniformly defined within the digital forensic community. Their study distinguishes different levels of automation and argues that common terminology is necessary because existing approaches range from basic automated operations to highly autonomous processes.

This distinction is important when discussing AI-assisted digital evidence triage.

Automation may occur at several levels:

Level	Description	Example
Basic automation	Repetitive computational task	Hash calculation
Task automation	Automated investigative activity	Keyword searching
Intelligent automation	ML-supported decision/recommendation	Artefact classification
Integrated automation	Multiple automated tasks connected	Evidence triage pipeline
Autonomous operation	Minimal human intervention	Fully automated decision process

For forensic purposes, the transition from conventional automation to intelligent automation is particularly important because AI/ML systems can use learned patterns rather than only predetermined rules.

Nevertheless, full autonomy introduces additional forensic and legal concerns. Consequently, the literature increasingly supports approaches where AI assists investigators rather than independently making final forensic determinations.

6. Artificial Intelligence and Machine Learning in Digital Forensics

AI and ML are increasingly being investigated as mechanisms for dealing with the scale and complexity of digital evidence.

Nayerifard et al. (2023) conducted a systematic literature review of ML in digital forensics and reported growing application of machine-learning methods across different forensic domains. Their review found particularly strong application in image forensics and increasing use of convolutional neural networks.

Fähndrich et al. (2023) similarly conducted a structured literature review of AI in digital forensics and concluded that so-called weak AI and applied AI techniques are already represented in different areas of digital forensics, while strong AI remains largely outside the practical forensic discourse.

The applications of AI/ML identified across the literature include:

- Digital artefact classification;
- Image and video analysis;
- Malware analysis;
- Anomaly detection;
- Network forensic analysis;

- Mobile forensic analysis;
- Text and language analysis;
- Evidence correlation;
- User-behaviour analysis;
- Relevance determination;
- Evidence prioritization; and
- Automated triage.

7. Machine Learning for Digital Artefact Classification

Digital artefact classification involves assigning an artefact to a predefined category based on its characteristics.

Possible categories include:

- Relevant;
- Potentially relevant;
- Irrelevant;
- Suspicious;
- Anomalous;
- High priority;
- Medium priority; and
- Low priority.

Classification can be based on:

- File metadata;
- File type;
- Timestamps;
- Application source;
- Directory location;
- Filesystem characteristics;
- Keywords;
- Communication attributes;
- User activity;
- Content;
- Relationships between artefacts; and
- Investigative context.

The selection of features is highly dependent on the type of investigation.

7.1 Metadata-Based Classification

Metadata provides information about an artefact without necessarily examining the entire content.

Examples include:

- Filename;
- File extension;
- File size;
- Creation timestamp;
- Modification timestamp;
- Access timestamp;
- Directory path;
- Application source; and
- Filesystem information.

Metadata-based approaches are attractive because they can provide relatively inexpensive features for large-scale triage.

Du and Scanlon (2019) proposed an automated methodology using metadata-based classification to prioritize potentially incriminating digital artefacts. Their methodology involves feature extraction, dataset generation, supervised model training, and evaluation.

The approach demonstrates that metadata can provide useful information for prioritization without requiring investigators to manually inspect every artefact.

8. Decision Tree Approaches

Decision Trees are supervised machine-learning models that use a series of feature-based decisions to classify observations.

A simplified forensic classification process may be represented as:

Artefact → Relevant application? → Relevant time period? → Suspicious characteristics? → Priority category

Decision Trees have several potential advantages in digital forensics.

8.1 Interpretability

The decision path can be examined to understand how a classification was produced.

8.2 Relatively Simple Implementation

Decision Trees can be implemented without the computational complexity associated with some deep-learning architectures.

8.3 Feature-Based Analysis

The method can incorporate metadata, temporal, filesystem, and contextual features.

8.4 Potential for Investigator Explanation

Because decision rules can be represented explicitly, they may provide investigators with a more understandable explanation than some black-box models.

However, Decision Trees may suffer from overfitting and may be sensitive to training-data characteristics. Their performance should therefore be established experimentally for the particular forensic dataset under investigation.

9. Bayesian and Bayes Net Approaches

Bayesian approaches provide probabilistic reasoning under uncertainty.

A Bayesian model estimates the probability of a hypothesis or class based on available evidence.

For example, an artefact's relevance may be influenced by:

- File type;
- Application;
- Timestamp;
- User;
- Location;
- Communication activity; and
- Previous classification patterns.

The general Bayesian relationship can be represented as:

P(H|E)=P(E|H)P(H)P(E)P(H|E)=\frac{P(E|H)P(H)}{P(E)}

where:

- HH represents a hypothesis;
- EE represents observed evidence;
- P(H|E)P(H|E) represents the probability of the hypothesis given the evidence;

- P(E|H)P(E|H) represents the probability of observing the evidence if the hypothesis is true; and
- P(H)P(H) represents the prior probability.

Bayesian approaches can therefore provide a useful framework for reasoning under uncertainty.

The author's review identifies Bayes Net among approaches applied to digital forensic and IoT forensic datasets. □

However, reported performance from an individual dataset should not automatically be generalized across all digital forensic environments. Dataset characteristics, feature selection, class distributions, and investigative context can substantially influence model performance.

10. Anomaly Detection in Digital Forensics

Anomaly detection identifies observations that differ from expected patterns.

Unlike supervised classification, anomaly detection can be useful when complete labelled training data are unavailable.

Potential forensic anomalies include:

- Unusual login times;
- Unexpected file access;
- Abnormal communication patterns;
- Unusual network connections;
- Unexpected application activity;
- Unusual data transfer;
- Suspicious file modification;
- Abnormal device behaviour; and
- Deviations from established user behaviour.

An important limitation is that **anomaly does not necessarily mean malicious or criminal behaviour**. An unusual event may have a legitimate explanation.

Therefore, anomaly detection should primarily be regarded as an investigative screening mechanism.

11. Automated Evidence Relevance Determination

Evidence relevance is context-dependent.

An artefact that is highly relevant to one investigation may have little or no relevance to another.

For example:

Investigation	Potentially Relevant Evidence
Financial fraud	Financial records, transactions, spreadsheets
Cyberattack	Logs, network records, malware artefacts
Harassment	Messages, communications, contact records
Missing person	Location records, communications, photographs
Intellectual property theft	Documents, file transfers, removable-media activity

Automated relevance determination therefore requires consideration of investigative context.
Du, Le, and Scanlon investigated automated artefact relevance determination using artefact metadata and associated timeline events. Their work demonstrates the potential for contextual information and temporal relationships to assist in determining artefact relevance.
The important implication is that AI-based forensic triage should not simply ask whether an artefact is suspicious in isolation. It should consider:
"Relevant to what investigation, under what circumstances, and during what period?"

12. Automated Evidence Prioritization

Evidence prioritization extends classification by determining which evidence should receive attention first.

A basic priority structure may be:

1. **Very High Priority**
2. **High Priority**
3. **Medium Priority**
4. **Low Priority**
5. **Unclassified/Requires Review**

The purpose is not to declare evidence legally significant automatically. Rather, prioritization provides an operational mechanism for managing investigator attention.
Du and Scanlon (2019) specifically addressed automated prioritization of suspicious file artefacts to reduce manual investigative effort. Their approach used supervised machine learning and was explicitly designed as a recommendation system operating within a human-in-the-loop workflow.

This distinction between **priority** and **evidentiary significance** is fundamental.

An AI system may recommend that an artefact be examined early. The investigator must subsequently determine its actual forensic relevance through appropriate examination and contextual interpretation.

13. Human-in-the-Loop Digital Forensics

A recurring issue in AI-assisted digital forensics is the degree of human involvement.

Three broad approaches can be distinguished:

Approach	Human Role	Forensic Implication
Manual	Investigator performs analysis	High human workload
AI-assisted	AI recommends; investigator validates	Human oversight retained
Autonomous	AI makes decisions with minimal human involvement	Greater explainability and governance concerns

The human-in-the-loop approach appears particularly appropriate for forensic triage.

Du and Scanlon's methodology explicitly adopts this approach by using ML to predict or recommend potentially suspicious artefacts while leaving final analysis to investigators.

This model can be represented conceptually as:

Digital Evidence → AI/ML Analysis → Recommendation → Investigator Review → Detailed Examination → Forensic Conclusion

Such a workflow preserves the investigator's ability to:

- Accept an AI recommendation;
- Reject an AI recommendation;
- Modify a priority;
- Investigate unexpected findings;
- Request additional evidence; and
- Interpret evidence in its wider context.

14. AI/ML and Digital Forensic Backlogs

Digital forensic backlogs arise when the amount of evidence requiring examination exceeds the capacity of available investigators, tools, and processing infrastructure.

The relationship between evidence volume and investigative workload can be conceptualized as:

Increasing Evidence Volume → Increasing Examination Requirements → Increasing Investigator Workload → Potential Backlog

AI-assisted triage introduces an additional stage:

Increasing Evidence Volume → Automated Triage → Evidence Prioritization → Focused Human Examination

The objective is not necessarily to eliminate forensic examination but to direct human effort toward evidence that is more likely to be useful.

Michelet et al. (2023) identify automation as important for managing increasing volumes of digital evidence. □

15. Comparative Analysis of AI/ML Approaches

Approach	Primary Application	Strength	Important Limitation
Decision Tree	Artefact classification	Interpretability	Overfitting/generalization
Random Forest	Classification	Ensemble robustness	Reduced interpretability
Bayes Net	Probabilistic classification	Handles uncertainty	Requires suitable probability structure
Naïve Bayes	Classification	Computational simplicity	Independence assumption
SVM	Classification	Effective in high-dimensional data	Parameter/kernel selection
Neural Networks	Complex pattern recognition	High modelling capacity	Data and computational requirements
CNN	Image/video forensics	Strong visual pattern recognition	Requires substantial training data
Clustering	Unlabelled evidence	Discovers patterns	Cluster interpretation
Anomaly Detection	Unusual behaviour	Can operate without complete labels	Anomaly does not imply wrongdoing
Metadata-based ML	Evidence triage	Relatively efficient feature extraction	Metadata can be incomplete/manipulated

The systematic literature reviewed by Nayerifard et al. demonstrates the diversity of ML methods being applied across digital forensic domains, while Du and Scanlon provide a particularly relevant example of supervised ML for artefact prioritization.

The literature therefore suggests that algorithm selection should be **problem-specific**, rather than based on the assumption that one algorithm is universally appropriate.

16. Major Challenges

16.1 Dataset Availability

The development of reliable forensic ML systems requires representative datasets.

However, forensic datasets can be difficult to obtain because digital evidence may contain:

- Personally identifiable information;
- Confidential communications;
- Sensitive organizational information;
- Proprietary information; or
- Legally protected material.

16.2 Dataset Quality

A machine-learning model is strongly influenced by the quality of its training data.

Poor-quality datasets may contain:

- Incorrect labels;
- Duplicated artefacts;
- Missing values;
- Unrepresentative cases;
- Biased distributions; or
- Insufficient examples of rare events.

These problems can result in misleading model performance.

16.3 Class Imbalance

Forensic datasets may contain many irrelevant artefacts and comparatively few relevant artefacts.

For example:

Large majority class: irrelevant evidence

Small minority class: highly relevant evidence

A model may achieve high overall accuracy while failing to identify the minority class effectively.

Consequently, forensic ML research should consider precision, recall, F1-score, confusion matrices, and class-specific performance rather than accuracy alone.

16.4 False Positives

A false positive occurs when an artefact is classified as relevant when it is not.

Excessive false positives may increase investigator workload because analysts must review large numbers of incorrectly prioritized artefacts.

16.5 False Negatives

A false negative occurs when relevant evidence is incorrectly classified as irrelevant or low priority.

This is particularly important in forensic triage because a missed artefact could potentially affect the subsequent investigation.

16.6 Explainability

Many AI/ML approaches can be difficult to interpret.

Forensic investigators may need answers to questions such as:

- Why was this artefact classified as relevant?
- Which features influenced the prediction?
- Why was this artefact assigned a high priority?
- What evidence supports the recommendation?
- How reliable is the model in this particular context?

Explainability is therefore an important research requirement.

16.7 Generalizability

A model trained on one dataset may not generalize to:

- Another operating system;
- Another application;
- Another device type;
- Another jurisdiction;
- Another crime type;
- Another evidence source; or
- Another generation of technology.

Cross-dataset and cross-domain validation is therefore necessary.

16.8 Encryption

Encryption can restrict access to potentially relevant evidence.

AI cannot classify information that is completely inaccessible to the forensic process.

Consequently, automated triage must be considered alongside established acquisition and decryption processes where legally and technically appropriate.

16.9 Anti-Forensics

Attackers may deliberately attempt to:

- Delete evidence;
- Manipulate timestamps;

- Alter metadata;
- Hide files;
- Encrypt information;
- Obfuscate communications; or
- Interfere with forensic tools.

Such activities can reduce the reliability of automated classification.

16.10 Evidence Integrity

Automation must operate within established forensic integrity requirements.

NIST guidance emphasizes preservation of data integrity during collection and examination.

AI/ML processing should therefore be conducted in a manner that preserves the original evidence and maintains an auditable record of processing.

16.11 Interoperability

Digital investigations frequently involve multiple forensic tools.

A machine-learning triage system should ideally be capable of exchanging information with established forensic platforms and evidence-processing tools.

16.12 Legal and Ethical Issues

AI-assisted forensic investigation raises questions concerning:

- Privacy;
- Transparency;
- Accountability;
- Explainability;
- Evidentiary admissibility;
- Algorithmic bias;
- Handling of personal data;
- Investigator responsibility; and
- Disclosure of automated processing.

An AI prediction should therefore not automatically be treated as an independent forensic fact.

17. Research Gaps

The review identifies several significant gaps.

Gap 1: Fragmentation of AI/ML Applications

Many studies address specific forensic tasks rather than integrating multiple intelligent functions into a unified triage workflow.

Gap 2: Limited Evidence-Triage Focus

Although ML is widely applied in areas such as image forensics, relatively less attention has been given to broad evidence reduction and prioritization across heterogeneous evidence collections.

Gap 3: Dataset Limitations

There is a continuing need for representative and appropriately governed datasets.

Gap 4: Limited Cross-Domain Validation

Models are often evaluated in specific contexts, creating uncertainty concerning transferability.

Gap 5: Explainability

There is a need for methods that allow investigators to understand and document AI-generated recommendations.

Gap 6: Human-in-the-Loop Evaluation

The literature establishes the concept of human-in-the-loop forensic analysis, but more research is required to measure how investigators interact with AI-generated recommendations in realistic workflows.

Gap 7: Operational Backlog Measurement

There is a need to distinguish **model accuracy** from **actual forensic workflow improvement**.

A model can achieve high classification performance without necessarily reducing investigator workload.

Gap 8: Standardized Evaluation

The literature would benefit from standardized benchmarks for comparing AI/ML approaches using consistent datasets and metrics.

Gap 9: Modern Distributed Evidence

Further research is required for integrated evidence originating from:

- Smartphones;
- Cloud services;
- Iot devices;
- Networks;
- Messaging applications; and

- Distributed platforms.

18. Proposed Conceptual Model for AI-Assisted Evidence Triage

Based on the literature reviewed, an integrated conceptual model can be represented as:

Digital Evidence Sources

↓

Evidence Acquisition and Preservation

↓

Preprocessing and Normalization

↓

Feature Extraction

↓

AI/ML Analysis

↙ ↓ ↘

Classification | Anomaly Detection | Relevance Assessment

↘ ↓ ↙

Evidence Prioritization

↓

Investigator Review

↓

Detailed Forensic Examination

↓

Analysis and Interpretation

↓

Reporting

19. Future Research Directions

19.1 Development of Standardized Datasets

Researchers should develop appropriately governed benchmark datasets containing representative digital forensic artefacts.

19.2 Explainable AI

Future systems should provide interpretable explanations for classifications and priority recommendations.

19.3 Human-in-the-Loop Evaluation

Future research should measure how effectively investigators can work with AI-generated recommendations.

19.4 Multi-Algorithm Comparison

Decision Trees, Random Forests, Bayesian methods, SVM, gradient boosting, and neural networks should be compared under standardized conditions.

19.5 Large-Scale Evidence Processing

Future studies should investigate whether AI/ML systems remain effective when processing very large evidence collections.

19.6 Cloud and IoT Evidence

Because evidence increasingly originates from distributed environments, future models should integrate cloud, mobile, IoT, and network evidence.

19.7 Adversarial and Anti-Forensic Conditions

Models should be evaluated against manipulated metadata, deleted evidence, obfuscation, and other anti-forensic conditions.

19.8 Interoperable Forensic Platforms

AI/ML components should be designed to operate with established forensic tools and evidence formats.

19.9 Standardized Evaluation Metrics

Future research should evaluate not only classification performance but also:

- Evidence reduction;
- Investigator workload;
- Processing time;
- False-negative risk;
- Explainability;
- Reproducibility; and
- Forensic workflow impact.

19.10 Empirical Validation

The most important next stage is implementation and empirical validation. The proposed concepts should be tested using appropriately constructed datasets and controlled experimental conditions.

20. DISCUSSION

The reviewed literature establishes a strong rationale for investigating AI and ML in digital evidence triage. The primary motivation is the growing volume and complexity of evidence and the resulting difficulty of examining all available artefacts manually.

The literature also demonstrates that AI/ML should not be viewed as a single technology or algorithm. Different forensic problems require different computational approaches.

Decision Trees can provide interpretable rule-based classification. Bayesian methods can provide probabilistic reasoning. Neural networks can model complex patterns, particularly in image and multimedia evidence. Anomaly detection can identify unusual behaviours without requiring complete labelled datasets.

However, the usefulness of an AI/ML model depends on the forensic problem, dataset, features, training procedure, validation strategy, and operational environment.

One of the most important findings from the literature is therefore that accuracy alone is not an adequate measure of forensic usefulness.

A model that achieves high accuracy but fails to identify rare relevant artefacts may be unsuitable for forensic triage. Conversely, a model with slightly lower overall accuracy may be more useful if it effectively identifies high-value evidence while providing interpretable recommendations.

The human-in-the-loop principle consequently emerges as an important component of responsible AI-assisted digital forensics.

Du and Scanlon's work provides a useful example: machine learning identifies or recommends suspicious artefacts, while the investigator remains responsible for the final interpretation. □

This approach is consistent with the broader forensic lifecycle described by NIST, in which automated and manual examination methods can coexist while maintaining evidence integrity and supporting subsequent analysis and reporting.

The review therefore supports a model of investigator-centred intelligent automation, rather than unrestricted autonomous forensic decision-making.

21. Implications for Digital Forensic Practice

The reviewed literature has several implications for forensic practitioners.

First, AI/ML can potentially assist investigators in managing large evidence collections.

Second, organizations considering AI-assisted forensic systems should evaluate datasets and models rather than assuming that a commercially available AI system is automatically appropriate for forensic use.

Third, investigators require training in interpreting AI-generated recommendations.

Fourth, forensic laboratories should maintain appropriate audit trails documenting automated processing.

Fifth, organizations should establish governance procedures addressing privacy, security, model validation, and accountability.

Finally, AI-assisted triage should remain integrated with established forensic procedures rather than operating independently.

22. CONCLUSION

The rapid expansion of digital technologies has transformed the quantity and complexity of evidence available to forensic investigators. Computers, mobile devices, cloud systems, IoT platforms, messaging applications, and networks generate extensive digital traces that may need to be examined during investigations. This growing volume of evidence creates substantial pressure on conventional forensic processes and contributes to the need for greater automation.

This review examined the application of artificial intelligence and machine learning to automated digital evidence triage and prioritization. The literature demonstrates that ML can support digital artefact classification, anomaly detection, relevance determination, and prioritization. Decision Tree, Bayesian, ensemble, support-vector, and neural-network approaches have all been investigated in different digital forensic contexts.

A particularly important research direction is metadata-based artefact classification and prioritization. Du and Scanlon's work demonstrates how supervised ML can learn from previously processed forensic cases to recommend potentially suspicious artefacts while maintaining a human-in-the-loop approach.

The literature also indicates that AI/ML adoption remains constrained by dataset limitations, class imbalance, generalizability, explainability, false positives and false negatives, encryption, anti-forensics, interoperability, evidence integrity, and legal and ethical considerations.

The review identifies a particular need for integrated approaches that connect artefact classification, anomaly detection, relevance assessment, and evidence prioritization within an investigator-centred forensic workflow.

The conceptual model developed from this review therefore positions AI/ML as a decision-support and triage mechanism rather than a replacement for forensic investigators.

No original experimental performance results are claimed in this paper. The review instead establishes the theoretical and empirical foundation for subsequent implementation and evaluation. Future research should develop representative datasets, compare alternative ML algorithms, incorporate explainable AI, evaluate human interaction with automated recommendations, and measure the actual impact of AI-assisted triage on evidence-processing time, investigator workload, and digital forensic backlogs.

The eventual objective should not simply be to make digital forensic tools more automated. It should be to develop reliable, explainable, interoperable, scalable, and forensically defensible intelligent systems capable of helping investigators identify and examine the most potentially relevant evidence efficiently while preserving professional human judgment and evidentiary integrity.

REFERENCES

1. Du, X., & Scanlon, M. (2019). Methodology for the automated metadata-based classification of incriminating digital forensic artefacts. In *Proceedings of the 14th International Conference on Availability, Reliability and Security (ARES 2019)*. ACM. https://doi.org/10.1145/3339252.3340517
2. Du, H., Le, N., & Scanlon, M. (2020). Automated artefact relevancy determination from artefact metadata and associated timeline events. *Forensic Science International: Digital Investigation*.
3. Fähndrich, J., Honekamp, W., Povalej, R., Rittelmeier, H., Berner, S., & Labudde, D. (2023). Digital forensics and strong AI: A structured literature review. *Forensic Science International: Digital Investigation, 46*, 301617. https://doi.org/10.1016/j.fsidi.2023.301617
4. Iqbal, S., & Alharbi, S. A. (2019). Advancing automation in digital forensic investigations using machine learning forensics. In *Digital Forensic Science*. IntechOpen. https://doi.org/10.5772/intechopen.90233
5. Javed, A. R., Ahmed, W., Alazab, M., Jalil, Z., Kifayat, K., & Gadekallu, T. R. (2022). A comprehensive survey on computer forensics: State-of-the-art, tools, techniques, challenges, and future directions. *IEEE Access, 10*, 11065–11083.

6. Kent, K., Chevalier, S., Grance, T., & Dang, H. (2006). *Guide to integrating forensic techniques into incident response* (NIST Special Publication 800-86). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-86
7. Michelet, G., Breitinger, F., & Horsman, G. (2023). Automation for digital forensics: Towards a definition for the community. *Forensic Science International, 349*, 111769. https://doi.org/10.1016/j.forsciint.2023.111769
8. Nayerifard, T., Amintoosi, H., Ghaemi Bafghi, A., & Dehghantanha, A. (2023). Machine learning in digital forensics: A systematic literature review. *arXiv*. https://arxiv.org/abs/2306.04965
9. Wu, T., Breitinger, F., & O'Shaughnessy, S. (2020). Digital forensic tools: Recent advances and enhancing the status quo. *Forensic Science International: Digital Investigation, 34*.