International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Use of Artificial Intelligence in Forensic Medicine and Criminal Investigation: A Critical Analysis

Dr. Vimal Kumar Jha

Associate Professor, Dept of Law, Jagannath Vishwa Law College, Dehradun

Abstract

Applications include (but are not limited to) the use of artificial intelligence to screen post mortem images, classifing histopathology, prioritising digital evidence, identifying a person, detection of manipulated media and to support decision making in investigations. While these applications offer the potential for increased speed, consistency and access to specialist expertise, they also involve unique dangers: forensic products may impact on determinations of cause of and manner of death, and affect a person's liberty and reputation. This paper aims to critically evaluate the technical, ethical, and legal factors of AI in forensic medicine and criminal investigation. It separates assistive measurement from autonomous inference and assesses systems on scientific validity, on the performances of subgroups, on explainability, on provenance, on chain of custody, on privacy and contestability. Several factors are revealed in the analysis, none of which can be resolved by high accuracy of the aggregates: base rates, domain shift, feedback loops, adversarial manipulation, undocumented software changes, and automation bias can turn small technical errors into big miscarriages of justice. A risk leveled assurance model is proposed that may use a range of low-risk tools on a periodic basis, but a system of biometric identification, predictive policing, cause of death suggestion and evidence-authentication will need to be tested independently, authorised by humans, validated, disclosed, monitored and provide actionable opportunities for challenges. AI should be used to complement and inform, not supplant, informed forensic decision-making.

Keywords: artificial intelligence, forensic medicine, criminal investigation, digital forensics, facial recognition, deepfake detection, evidentiary reliability, responsible AI

INTRODUCTION

The business of orensic medicine and the criminal investigation is a process of information intensive business in an extremely high-stakes business. A mistaken prioritization score or a linked face or CT missed fracture may turn an investigation in its entirety and in turn impact freedom. Artificial Intelligence (AI) is expected to significantly lower the number of backlogs, as it can detect patterns in images, signals, documents and networks which are too complex to analyze manually. A forensic application, age estimation, identification, pathology, toxicology and digital evidence are described in the recent literature [1]–[5].

For sure, AI is appealing. Models can filter out thousands of images, prioritize devices to analyze, separate or group similar files, translate or summarize communications, and highlight anomalies for

expert review. These capacities can amplify limited resources for under-sourced medicolegal systems. However there are differences between forensic work and normal business forecasting. But, not only does the model have to find a higher average result but it must also be scientifically valid for the case, repeatable by the other examiner, backed up by preserved evidence, explainable to court, and scrutinizable by the affected person.

There's also the need to distinguish between various forms of using AI in a critical analysis. It's a very different algorithm, improving an image or searching for duplicate files than the one used to identify a suspect or predict an individual will offend. Similarly, a model which measures a lesion for a pathologist is inconsequential to him, whereas a system which suggests a cause of death is more important. Otherwise, all applications are grouped together, preventing differences in autonomy from being observed, differences in the weight of the evidence, scale, reversibility, and differences in the impact on the rights to be seen. It is all right to allow research or triage to proceed with the same model which is not the basis for coercive action.

The paper has four contributions to make. It starts by providing an overview of the present and upcoming areas of forensic medicine and investigation in which AI is being deployed. Second, it assesses their advantages and dangers on the basis of the principle of forensic science and responsible-AI. Third, it contrasts intergovernmental strategies and approaches, such as the European Union AI Act, NIST resources, WHO principles, INTERPOL-UNICRI resources, and the Indian evidentiary context. Finally, it offers an approach to practical assurance based on intended use, independent validation, human oversight, provenance, disclosure, monitoring and contestability.

ANALYTICAL SCOPE AND CONCEPTUAL FRAMEWORK

A. *Review Scope and Sources*

A.Check and review the scope and sources.

This study is a critical narrative study and does not present a thorough review. It brings together best practice research from peer-reviewed forensic-medicine, deepfake, biometric, digital-forensics and predictive-policing literature alongside authoritative standards, and legal instruments. Specially consideration has been afforded to those sources that focus mainly on operational validation, operational demographic impacts, evidence preservation, human rights and court application. The argument is intentionally interdisciplinary because any technically correct model could be the one that is flawed, that could not have been reproduced, or that could not have meaningfully been disputed.

Analysis is targeted at AI that forensic physicians, pathologists, laboratories, police, prosecutors and digital investigators use. Does not include routine Administrative Automation, except that the output may influence the interpretation of the evidence or investigative suspicion. Generative AI is embedded when it writes reports, condenses case content, generates investigative leads, or generates fake media to be verified later in the investigative process. This paper is not testing a particular commercial product but provides a set of criteria that can be used for procurement and deployment by all jurisdictions.

B.Evaluation Criteria

Assessing forensic suitability is done using seven criteria: analytical validity, operational validity, fairness, explainability, integrity, legality, and contestability. Analytical validity is related to the question of whether the method is able to measure what it is designed to measure under controlled circumstances. Operational validity: does it perform well on local devices, populations, image qualities and workflows? Fairness looks at rates of error among subgroups and at the unequal exposure. Explanatory reasons are

reasons that are comprehensible to the appropriate user, not only a visualization. Integrity includes provenance, versioning, chain of custody and reproducibility. Legality is linked to possibilities of authority, need, proportionality and data protection. A contestability requires access to adequate information to be able to independently test.

These criteria have been made more stringent than a benchmark accuracy criterion; they are intentionally so. The long-standing problem with forensic science is the lack of validation and examiner bias, inflated statements, and the lack of quality control systems [13]. While AI can help alleviate some forms of subjective variation, it can also industrialise such variation by propagating the same hidden error. A defensible system should therefore present not only statistical performance, but governance performance as well – who gave permission to use it, what information was gathered, which software version was employed, how uncertainty information was reported, what happened when the output did not match evidence elsewhere.

APPLICATIONS IN FORENSIC MEDICINE

A. *Postmortem Imaging and Pathology*

Deep learning can assist PMCT and postmortem magnetic-resonance imaging by detecting fractures, hemorrhage, gas collections, foreign bodies, and organ abnormalities. It can segment structures and prioritize scans for expert review, potentially reducing workload where specialist radiologists are scarce [3], [4]. Digital pathology models may similarly identify tissue regions, quantify features, or retrieve comparable cases. These applications are promising because they produce localized measurements that a pathologist can verify. Their limits include postmortem artifacts, decomposition, uncommon injury patterns, scanner variation, and small single-center datasets that may not represent routine casework.

TABLE I. REPRESENTATIVE AI APPLICATIONS AND FORENSIC OUTPUTS

Forensic area	Representative AI-supported output
Postmortem imaging	Lesion or fracture flagging, segmentation, scan prioritization
Digital pathology	Region detection, quantitative morphology, case retrieval
Identification	Face, fingerprint, voice, gait, age or skeletal estimation
Toxicology	Chromatogram classification, anomaly and drug-pattern detection
Digital forensics	File triage, malware classification, clustering and entity extraction
Multimedia forensics	Deepfake screening, source attribution, tamper localization
Investigation support	OSINT linking, transcription, translation and draft summarization

B. *Identification, Toxicology, and Injury Analysis*

Other proposed uses include estimation of age, sex, ancestry-related traits, stature, postmortem interval, drug pattern classification, wound characterization, and reconstruction of injury mechanisms [1]. Many are probabilistic and population dependent. A model trained on one demographic, laboratory method, or decomposition environment may not transfer to another. The risk is especially high when a broad biological estimate is converted into a categorical legal conclusion. Forensic reports should therefore preserve the distinction between measured features, model-derived probabilities, and the expert's final opinion.

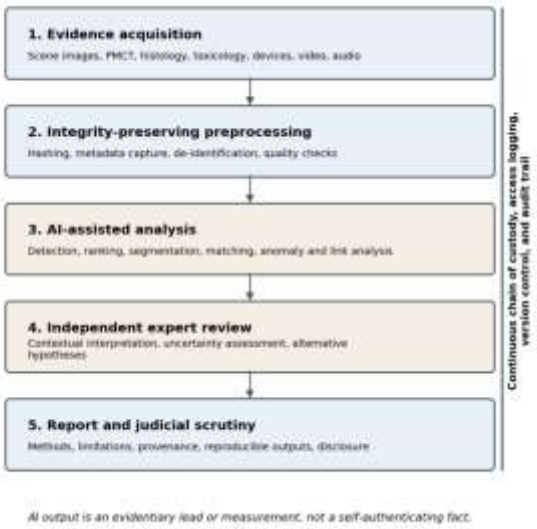

Fig. 1. A defensible AI-enabled forensic workflow. Evidence integrity and expert review remain continuous requirements; the model output is not self-authenticating.

C. *Applications in Criminal Investigation*

AI-supported investigation includes digital-evidence triage, entity and relationship extraction, biometric searching, automated transcription and translation, multimedia authentication, anomaly detection, and open-source intelligence. These tools can reduce the search space, but their apparent precision can be misleading when the target event is rare. For a system with sensitivity Se, specificity Sp, and prevalence π, the positive predictive value is

$$PPV = (Se \times \pi) / [(Se \times \pi) + (1 - Sp)(1 - \pi)], \quad (1)$$

Equation (1) demonstrates the base-rate problem. Even a high-specificity model can generate many false alerts when applied to millions of people or files. Subgroup analysis is equally necessary. A single overall error rate can hide substantial demographic differences, as NIST evaluations of face-recognition systems have shown [7]. A simple disparity measure is

$$\Delta FPR = \max(FPR \text{ across groups}) - \min(FPR \text{ across groups}), \quad (2)$$

The integrity of digital evidence also depends on preserving a verifiable sequence of acquisition and transformation. A model may resize an image, extract frames, or generate embeddings without altering the original, but every derived artifact should be linked to the source, software version, operator, and time. A conceptual chained record is

$$h(i) = H[E(i) \| h(i-1) \| t(i) \| a(i)], \quad (3)$$

Equations (1)–(3) illustrate why "the algorithm matched" is not an adequate forensic statement. The report must describe the search population, decision threshold, subgroup performance, evidence transformations, and the role of the examiner. AI outputs are usually most defensible as investigative leads or measurements. They become more difficult to justify when they are treated as conclusive identification, intent, dangerousness, or credibility.

D. ***Digital Evidence Triage and Cybercrime***

Machine learning can rank files, classify malware, detect explicit or exploitative content, cluster communications, and identify anomalous transactions. Systematic reviews find strong growth in image and multimedia forensics, often using convolutional networks [5]. Such triage may be valuable when investigators face terabytes of data, but it creates a risk of non-review: material ranked low may never be examined. A practical risk-priority score should reflect not just error probability but the severity and scale of its consequences:

$$R = P(\text{error}) \times S(\text{consequence}) \times E(\text{exposure}). \quad (4)$$

High-risk triage requires recall-oriented validation, preservation of unreviewed material, documented search strategies, and random quality-control sampling from the "low relevance" set. Investigators should not infer absence merely because a model did not surface an item. The original evidence must remain available for alternative tools and defense examination. NIST guidance emphasizes reliable acquisition, preservation, analysis, and chain-of-custody practices for digital evidence [6]. AI should operate inside, rather than replace, those controls.

E. ***Biometrics, Multimedia, and Open-Source Intelligence***

Face recognition, voice biometry, gait matching, fingerprint enhancement, license plate matching and across-camera face identification can quickly produce candidate face identities. The value of the investigation is greatest if the reference database is legal and the quality of the image is reasonable and the result has been corroborated from other sources. Many systems have been shown in NIST testing to vary widely by algorithm and/or demographic difference [7]. Once again being detected on a candidate list (or in absence of a list) does not constitute an identification; confirmation should rely on other evidence, documented thresholds and trained examiners who know what to look for to minimise contextual-bias risks.

Synthetic media poses the opposite issue – the AI can manufacture believable evidence, while other AI can try to uncover it. There are useful findings but also some challenges with generalization, interpretability and robustness for deepfake forensics from the literature [14, 15]. Detectors could give good results on established generators, and may fail for a new compression pipeline or manipulation technique. Therefore, an opinion in a courtroom should compete/consist of multiple complementary examinations, such as file structure, file provenance, sensor traces, or inconsistencies in content and detector output, and not be formed from just one probability score.

Generative models can also summarize interviews or prepare reports, translate between languages and suggest links between cases. Their value is to lessen the burden of the writing Staff, and there is no place for hallucinating facts, losing qualifications, and citing things that never happened. All written work must be considered a first draft with a URL, line by line read aloud and until proven, not included in the evidence record. Retention of prompts, model version, retrieval sources, edits may be required if the output was a critical factor in an investigatory decision.

Public records, social media posts, location activity, and communications can all be utilized to expose networks in an effort that ties together open-source intelligence (OSINT) and link analysis systems.Link analysis and open-source intelligence (OSINT) systems can link together public records, social media activity, location and communications to uncover networks. The problem is that information gathered for one purpose can be a lingering intelligence profile, called function creep. Association can be confused with correlation and culpability with association. Investigative graph tools should reveal source records,

level of confidence, temporal setting, other explanations. They should not claim a single, over-arching "risk" label, which conceals the evidence on which it is based.

CRITICAL ANALYSIS OF RISKS

F. *Accuracy, Base Rates, and Domain Shift*

Forensic models frequently encounter data unlike their development sets: different scanners, populations, lighting, languages, compression, decomposition stages, laboratory protocols, or criminal methods. Performance can deteriorate silently because a probability score may remain confident under distribution shift. Validation should therefore include representative local casework, challenging and poor-quality samples, known negatives, rare conditions, and comparison with trained examiners. The result must be reported with uncertainty and limitations appropriate to the decision context, not as a universal accuracy claim.

TABLE II. CRITICAL FAILURE MODES AND MINIMUM CONTROLS

Risk	Typical trigger	Forensic and justice effect	Minimum control
Domain shift	New device, population, or case conditions	Silent accuracy loss; false inclusion or exclusion	Local validation and drift checks
Bias	Skewed labels, proxies, or enforcement feedback	Unequal FPR/FNR; discriminatory suspicion	Subgroup audit, impact assessment, use limits
Opacity	Proprietary model or weak explanation	Result cannot be reproduced or meaningfully challenged	Disclosure, audit access, case-level reasons
Attack / provenance	Manipulation, leakage, or undocumented update	Corrupted output and case contamination	Hashing, version control, security testing, incident response

Benchmark leakage is another concern. Repeated tuning on public datasets can produce impressive results that do not survive independent testing. Vendor claims based on proprietary data are insufficient where the tool influences detention, search, charging, or cause-of-death conclusions. Independent laboratories should be able to test the deployed version under realistic workloads. Material software updates, threshold changes, new reference databases, or changes in data acquisition should trigger documented revalidation rather than being treated as routine maintenance.

G. *Bias, Explainability, and Automation Bias*

Bias enters through labels, sampling, measurement, policing patterns, access to health care, and investigator decisions. Predictive-policing systems are especially vulnerable to feedback loops: recorded incidents guide future deployment, which increases detection in the same areas and is then interpreted as confirmation of risk [16]. In medicine, uneven access to diagnostic testing can similarly distort training labels. Fairness cannot be solved only by deleting protected attributes because location, income, language, and other variables may act as proxies.

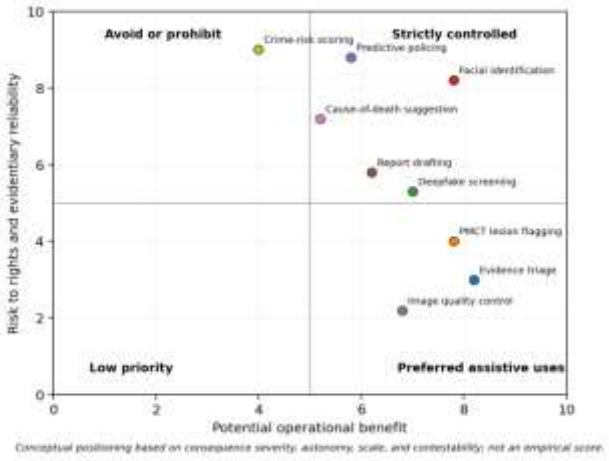

Fig. 2. Conceptual benefit-risk map. Assistive uses with verifiable outputs occupy the preferred region; high-autonomy identification and prediction require strict controls or prohibition.

Explainability should be audience specific. A developer may need feature and error analyses; a forensic examiner needs case-level reasons and failure conditions; a judge or jury needs a clear account of what the system did and did not establish; and the defense needs enough technical information to test the claim. NIST's explainability principles emphasize that systems should provide evidence or reasons, make them meaningful to the user, accurately reflect the process, and operate within knowledge limits [9]. Saliency maps alone rarely satisfy these requirements.

Fig. 3. Continuous assurance lifecycle; monitoring, contest, and retirement are integral to validation.

H. *Privacy, Chain of Custody, and Admissibility*

In unusual circumstances, forensic information can be highly damaging: sex crimes information, medical imaging, biometric information, communications, locations and histories could all suggest fault of the victim, family, or third-party parties that are not involved in the crimes. Collection and re-use should be in compliance with lawful authority, necessity, proportionality, retention and access controls. There are dangers of unauthorised secondary use and risks of breach with large reference databases. De-identification might not be sufficient for case stories, face or genome. PIA must be done for everything beforehand, not after a controversy breaks out.

TABLE III. COMPARATIVE GOVERNANCE APPROACHES

Framework	Scope	Safeguard	Gap
EU AI Act	Biometrics / policing	Bans, high-risk controls, authorization	National implementation
India	Digital evidence / expert opinion	Authenticity, safety, equality, privacy	No unified forensic-AI code
NIST / WHO / INTERPOL	Cross-sector / health / police	Validation, rights, oversight, monitoring	Mostly guidance

I. *Legal and Governance Landscape*

The European Union AI Act uses a risk-based structure and places severe limits on certain biometric and law-enforcement uses. Real-time remote biometric identification in public spaces is generally prohibited for law enforcement, subject to narrowly defined exceptions, necessity and proportionality safeguards, impact assessment, and authorization requirements [12]. The regulation also treats a range of law-enforcement and biometric systems as high risk. Its importance for forensic practice lies less in any single technical standard than in the principle that rights impact determines the required level of control. A simplified proportionality condition is:

$$\text{Deploy only if expected public benefit} > \text{rights harm, and no less intrusive alternative exists.} \quad (5)$$

In India, the Bharatiya Sakshya Adhiniyam, 2023 recognizes electronic and digital records and preserves the relevance of expert opinion [17]. This creates a legal pathway for AI-assisted evidence but does not make algorithmic output automatically reliable. The proponent must still establish integrity, authenticity, and an adequate scientific basis. NITI Aayog's responsible-AI principles emphasize safety, equality, privacy, transparency, and accountability [18], while the data-protection framework adds obligations relevant to sensitive personal data [19]. The I4C and National Cybercrime Forensic Laboratory ecosystem demonstrate growing operational capacity [20].

TABLE IV. FORENSIC AI ASSURANCE CONTROLS

Control	Evidence required	Owner	Release gate
Scientific validity	Independent tests, error bounds, subgroup and stress results	Forensic quality lead plus statistician	Predefined limits met for intended use
Operational integrity	Hashes, version log, disclosure pack, monitoring and incident plan	Laboratory or agency with legal oversight	Reproducible case record and approved controls

International guidance is converging around similar themes. The NIST AI Risk Management Framework organizes governance around mapping, measuring, managing, and governing risk [8]. WHO guidance for AI in health prioritizes autonomy, safety, transparency, responsibility, inclusiveness, and sustainability [11]. The INTERPOL-UNICRI toolkit adapts responsible-AI principles to law enforcement and emphasizes organizational readiness, risk assessment, procurement, human rights, and oversight [10].

These frameworks are not substitutes for law, but they provide operational questions that legislation often leaves unanswered.

Admissibility ultimately depends on jurisdiction-specific evidence rules, but several common requirements recur: relevance, authenticity, reliability, qualified interpretation, and the opportunity for cross-examination. Proprietary secrecy can conflict with these requirements. Full source-code disclosure may not always be necessary, yet the defense should receive meaningful information about the model's purpose, training and validation, thresholds, known limitations, version, error rates, and case-specific processing. A result that cannot be independently tested should receive limited evidentiary weight.

A FORENSIC AI ASSURANCE FRAMEWORK

J. *Risk-Tiered Deployment*

Proposed strategy: classification of use by consequence, autonomy, scale and reversibility. Tier 1 includes low-risk technical support, including duplicate detection, image quality control, or searching. This may be adequately validated audited logs or ordinary validation. Tier 2, which includes substantive analysis that can be reviewed but not-so-essential, such as lesion segmentation, evidence ranking and deepfake screening, requires independent validation and is mandatory for expert confirmation. Tier 3 includes identity, dangerousness, suggestions of causes of death, using predictive policing, or any product that could warrant coercive action – deployment should be subjected to explicit legal permission, rigorous validation, impact assessment, disclosure and independent supervision.

If accuracy is specified, there are some applications that should be excluded, regardless. People should not be criminalized based on their facial appearances nor be presumed to be guilty based on emotion recognition or be judged by a disadvantageous rule due to prediction based on AI. These applications are based on shaky or uncertain principles, they carry the potential for discrimination, and are hard to challenge. The illegitimate objective cannot be overcome by further training data, so a prohibition is a legitimate risk control if the underlying question is not scientifically but has completely contrary fundamental rights consequences.

B.Validation and Documentation

There needs to be some specific use for the validation. A model that has been proven to be suitable for image triage may not be suitable for image identification and a clinic model may not be applicable to a post mortem imaging. The targets to be achieved, types of equipment to be used, quality range, reference standard, exclusion criteria, thresholds, and acceptable error should be predetermined in the protocol. Report sensitivity/specificity, calibration, confidence intervals, failure to process rates, effect of missing or poor data etc. according to the capacity of the data. When there are asymmetric consequences of errors, select metrics and thresholds that capture the asymmetric consequences.

Any independent testing should be done with data that is not available to the developer and should contain adversarial and stress conditions. This includes out-of-sight generators, recompression, cropping, screen recording, audio dubbing, and mixing genuine and synthetic content for deepfake detection. For biometrics, it's about the realistic quality of the images, the scale of the database, the demographics of the photo and lookalike candidates. In forensic medicine, it represents decomposition, interventions, rare diseases and geography. When evaluating computer performance, human-comparison studies should consider whether it was unaided or aided with AI.

Documentation should be enough to recreate the analysis. The base data entries are: Original evidence hash, derived-file hashes, acquisition method, pre-processing, Model and Software Version, configuration,

threshold, reference database, operator, time, output, expert review and override. A model card should include intended uses and uses prohibited, training data provenance, validation populations, known limitations, cyber security controls and change history. Reproducibility means the preservation of "executable environments" or validated "alternatives" besides only screenshots of results.

C.Human Oversight and Contestability

Only if the reviewer is competent, has time, and has means to access the proof of the underlying facts, will human oversight be of any consequence. A rubber-stamp confirmation won't help lower automation bias. In a high-risk situation, interfaces should not show the AI conclusion until an independent observation is made by the examiner. Quality indicators and warning should be displayed on the candidate rank and/or candidate confidence. Any overrides, disagreements, or near misses should be recorded as a safety information and not as operator error.

Contestability involves system and individual results being open to challenge and a process for doing so. Those adversely affected should have the right to seek human consideration for their case, a correction of any error in a database, and an independent review where statutory to statutory. Court or control authorities should have access to validation reports, incidents, and history of software changes. This is a proposition that needs to be expressly defined by the experts, e.g., “the system determined this image was among the candidates under these circumstances;” a statement of “the system proved identity” offers no information regarding when or how it did so, and would not establish a proposition.

D.The procurement effort, monitoring and incident response efforts.

Securing audit rights, version notice, access to technical documentation, data-location and retention controls, cybersecurity obligations, and providing support for legal disclosure, should be features of procurement contracts. Agencies should not endorse products the performance of which can not be evaluated independently or whose licensing terms don't allow the performance to be examined by the defense. Pilot deployment should take place in an observed setting with a predetermined stopping criteria. The cost of the validation, training, storage, disclosures, incidence responses, and re-examination should be part of cost analysis, not just the subscription fee.

Significant changes should be monitored post deployment to alleviate worry of miss-match, misunderstandings of subgroups, volume of false alerts, overrides, complaints and performance of outcomes. Comparisons of the number of cases in the time period with historically established distributions may rely on statistical drift metrics, but a threshold is merely an early-warning indicator. Revalidation should also be asked when error limits are breached, devices or the databases change, or the size and scope of the operation grows.

Having a drift threshold is not an alternative to having an outcome monitored, but it can also be a signal for investigation prior to adverse outcomes in the court proceedings. The agencies should have a data leakage, adversarial attack, false identification, model corruption, or faulty update incident response plan. The plan should designate the person to suspend use, to keep logs, to notify affected cases, and to do a retrospective review. If a defect would have affected prosecution or death certificates in the past, then disclosure and re-examination should be taken into account.

Governance should be kept separate from the ownership of a product. Both developer and operational units are not the only ones who should judge the validity. High-impact systems should have a committee that includes representation from the community, civil rights, cybersecurity, statisticians, data protection officers and forensic practitioners. More robust external internal assessments and/or proficiency tests should be developed to encompass self-assisted workflows for AI. Use Cases, validations, error events

and safeguards can be described in transparent reports that are shared with the public without leaking sensitive methods of investigation.

DISCUSSION

The difference between automation of labour and automation of judgement is the critical one. AI works best when it carries out discrete, testable tasks, like isolating pieces of an image, looking for duplicate images, prioritizing items for review, or measuring a specific feature. It is at its weakest when it transforms historical data into predictions about people, reads in these contested psychological states or arrives at conclusions which are not easily reproducible from the primary sources. The appropriate yardstick is not that AI can replace an expert, but whether human-AI collaboration proves to be more accurate, fair, transparent and accountable than the current human-only process. While responsible adoption can contribute to the enhancement of forensic capacity, this capacity should not be quickly achieved by compromising investigative processes or on the backs of scientific methods.

Conclusion

In forensic medicine and the field of criminal investigation, AI has the potential to make a tangible contribution in areas such as image analysis, information retrieval, digital-evidence triage and/or in the screening of multimedia. But, on the matter of evidence, it is conditional. Inaccuracies in individual pieces of data will swamp errors stemming from poor data quality, from subset biases, from non-clear updates, and from a lack of meaningful challenge. The autonomy and consequences of high-impact applications should be correlated with validation, human authorisation, disclosure, monitoring and oversight within a risk tiered framework. AI can help identify and organize materials and build a case, but the ability to interpret the materials and make any coercive policies in the name of AI must be held by accountable experts and institutions themselves.

REFERENCES

1. L. Tournois, V. Trousset, D. Hatsch, T. Delabarde, B. Ludes, and T. Lefèvre, “Artificial intelligence in the practice of forensic medicine: A scoping review,” Int. J. Legal Med., vol. 138, no. 3, pp. 1023–1037, 2024, doi: 10.1007/s00414-023-03140-9.
2. B. Ondruschka, D. Seifert, N. Rotermund, and A. Fehnker, “A critical review of ‘Artificial intelligence in the practice of forensic medicine: A scoping review,’” Int. J. Legal Med., vol. 138, no. 4, pp. 1667–1668, 2024, doi: 10.1007/s00414-024-03209-z.
3. A. Dobay et al., “Potential use of deep learning techniques for postmortem imaging,” Forensic Sci. Med. Pathol., vol. 16, no. 4, pp. 671–679, 2020, doi: 10.1007/s12024-020-00307-3.
4. F. Dedouit et al., “The current state of forensic imaging—post mortem imaging,” Int. J. Legal Med., vol. 139, no. 3, pp. 1141–1159, 2025, doi: 10.1007/s00414-025-03461-x.
5. T. Nayerifard, H. Amintoosi, A. G. Bafghi, and A. Dehghantanha, “Machine learning in digital forensics: A systematic literature review,” CoRR, abs/2306.04965, 2023, doi: 10.48550/arXiv.2306.04965.
6. B. Guttman, D. R. White, S. Williams, and T. Walraven, Digital Evidence Preservation: Considerations for Evidence Handlers, NISTIR 8387. Gaithersburg, MD, USA: NIST, 2022, doi: 10.6028/NIST.IR.8387.

7. P. Grother, M. Ngan, and K. Hanaoka, Face Recognition Vendor Test Part 3: Demographic Effects, NISTIR 8280. Gaithersburg, MD, USA: NIST, 2019, doi: 10.6028/NIST.IR.8280.
8. E. Tabassi, Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1. Gaithersburg, MD, USA: NIST, 2023, doi: 10.6028/NIST.AI.100-1.
9. P. J. Phillips et al., Four Principles of Explainable Artificial Intelligence, NISTIR 8312. Gaithersburg, MD, USA: NIST, 2021, doi: 10.6028/NIST.IR.8312.
10. INTERPOL and UNICRI, Toolkit for Responsible AI Innovation in Law Enforcement. Lyon, France: INTERPOL, revised 2024.
11. World Health Organization, Ethics and Governance of Artificial Intelligence for Health: WHO Guidance. Geneva, Switzerland: WHO, 2021, ISBN 978-92-4-002920-0.
12. European Parliament and Council, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence," Off. J. Eur. Union, 2024.
13. National Research Council, Strengthening Forensic Science in the United States: A Path Forward. Washington, DC, USA: National Academies Press, 2009, doi: 10.17226/12589.
14. P. M. G. I. Reis and R. O. Ribeiro, "A forensic evaluation method for DeepFake detection using DCNN-based facial similarity scores," Forensic Sci. Int., vol. 358, Art. no. 111747, 2024, doi: 10.1016/j.forsciint.2023.111747.
15. E. Hydara, M. Kikuchi, and T. Ozono, "Empirical assessment of deepfake detection: Advancing judicial evidence verification through artificial intelligence," IEEE Access, vol. 12, pp. 151188–151203, 2024, doi: 10.1109/ACCESS.2024.3480320.
16. T.-W. Hung and C.-P. Yen, "Predictive policing and algorithmic fairness," Synthese, vol. 201, Art. no. 206, 2023, doi: 10.1007/s11229-023-04189-0.
17. Government of India, The Bharatiya Sakshya Adhiniyam, 2023, Act No. 47 of 2023, in force July 1, 2024.
18. NITI Aayog, Approach Document for India: Part 1—Principles for Responsible AI. New Delhi, India: Government of India, 2021.
19. Government of India, Digital Personal Data Protection Act, 2023, and Digital Personal Data Protection Rules, 2025. New Delhi, India: Ministry of Electronics and Information Technology.
20. Ministry of Home Affairs, Government of India, "Indian Cybercrime Coordination Centre (I4C) and National Cybercrime Forensic Laboratory ecosystem," 2026.