



Article

Digital Stratigraphy—A Pattern Analysis Framework Integrating Computer Forensics, Criminology, and Forensic Archaeology for Crime Scene Investigation

Romil Rawat ^{1,*}, Hitesh Rawat ¹, Mandakini Ingle ², Anjali Rawat ¹, Anand Rajavat ³ and Ashish Dibouliya ⁴

¹ Computer Science Department, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore 452007, India; hrawat88@gmail.com (H.R.)

² CSE Department, Medicaps University, Indore 453331, India; tayademandakini@gmail.com

³ Department of Computer Science Engineering, Shri Vaishnav Institute of Information Technology, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore 452007, India; directorsviit@svvv.edu.in

⁴ Computer Science and Applications, Rabindranath Tagore University, Bhopal 462047, India; ashish.dibouliya@gmail.com

* Correspondence: rawat.romil@gmail.com

Abstract

Background/Objectives—Traditional forensic investigations often analyze digital, physical, and criminological evidence separately, leading to fragmented timelines and reduced accuracy in reconstructing complex events. To address these gaps, this study proposes the Digital Stratigraphy Framework (DSF), inspired by archaeological stratigraphy, to integrate heterogeneous evidence into structured, temporally ordered layers. DSF aims to reduce asynchronous inconsistencies, minimize false associations, and enhance interpretability across digital, behavioral, geospatial, and excavation evidence. **Methods**—DSF employs Hierarchical Pattern Mining (HPM) to detect recurring behavioral patterns and Forensic Sequence Alignment (FSA) to synchronize evidence layers temporally and contextually. The framework was tested on the CSI-DS2025 dataset containing 25,000 multimodal, stratified records, including digital logs, geospatial data, criminological reports, and excavation notes. Evaluation used 10-fold cross-validation, Bayesian hyperparameter tuning, and structured train-validation-test splits. Metrics included accuracy, precision, recall, F1-score, and Stratigraphic Reconstruction Consistency (SRC), alongside ablation and runtime assessments. **Results**—DSF achieved 92.6% accuracy, 93.1% precision, 90.5% recall, 91.3% F1-score, and an SRC of 0.89, outperforming baseline models. False associations were reduced by 18%, confirming effective cross-layer alignment and computational efficiency. **Conclusions**—By applying stratigraphic principles to forensic analytics, DSF enables accurate, interpretable, and legally robust evidence reconstruction. The framework establishes a scalable foundation for real-time investigative applications and multi-modal evidence integration, offering significant improvements over traditional fragmented approaches.

Keywords: Digital Stratigraphy Framework; Hierarchical Pattern Mining; Forensic Sequence Alignment; Multimodal Crime Scene Dataset; stratified evidence reconstruction



Academic Editor: Ricardo Jorge Dinis-Oliveira

Received: 9 September 2025

Revised: 7 October 2025

Accepted: 15 October 2025

Published: 17 October 2025

Citation: Rawat, R.; Rawat, H.; Ingle, M.; Rawat, A.; Rajavat, A.; Dibouliya, A. Digital Stratigraphy—A Pattern Analysis Framework Integrating Computer Forensics, Criminology, and Forensic Archaeology for Crime Scene Investigation. *Forensic Sci.* **2025**, *5*, 48. <https://doi.org/10.3390/forensicsci5040048>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Digital stratigraphy represents an innovative forensic paradigm that systematically layers evidence across digital, behavioral, and archeological domains to uncover hidden patterns of criminal behavior. Unlike traditional computer forensic platforms—such as EnCase or FTK—that focus primarily on file recovery, metadata extraction, or disk imaging,

digital stratigraphy leverages contextual ordering within file systems to infer temporal relationships and provenance [1]. Criminological methods [2,3], including behavioral linkage analysis, situational crime pattern mapping, and offender profiling, deepen this approach by identifying recurrent motifs that align with stratigraphic trends in physical contexts. Meanwhile, forensic archeology contributes foundational principles of stratigraphy—such as superposition, inclusion, and cross-cutting relations—for layering and association of buried evidence, enabling spatial–temporal interpretation even in disturbed or incomplete crime scenes [4]. A notable integration of these techniques appears in archeological frameworks like the Integrated Archeological Database (IADB) [5], which supports stratigraphic recording and contextual linkage in excavation data systems [6,7]. The uniqueness of digital stratigraphy lies in its capability for pattern analysis across heterogeneous evidence streams—rather than treating digital logs, geospatial trails, or excavation contexts in isolation, it constructs multilayered timelines that preserve context, reduce incorrect associations, and reveal causal links across domains. This unified approach holds promise for complex investigative scenarios—such as cyber-enabled artifact theft—where digital transaction logs must be synced with excavation records and behavioral profiles to reconstruct precise event sequences.

Digital stratigraphy in crime analysis encompasses a range of methods designed to reconstruct evidence by layering and aligning data from diverse sources. Common techniques given in [1], which traces allocation and modification events on digital storage; temporal layering [8], which orders digital logs and geospatial traces into chronological strata; multi-modal fusion [9], which integrates heterogeneous datasets such as criminological reports and excavation records; and graph-based stratigraphy [10], where entities and events are modeled as interconnected nodes for structural provenance tracking. More advanced methods combine statistical modeling with machine learning to enhance accuracy in detecting tampering or inconsistencies. In the proposed work, the framework adopts two specialized techniques: Hierarchical Pattern Mining (HPM) to capture recurring cross-layer patterns across multimodal [11–14] forensic records, and Forensic Sequence Alignment (FSA), to synchronize asynchronous or incomplete timelines into coherent stratified layers. Together, these methods extend digital stratigraphy beyond traditional file or log analysis, enabling robust cross-domain evidence reconstruction in complex cybercrime investigations.

Modern criminal investigations [15] increasingly face multifaceted challenges as digital and physical evidence converge. Digital forensics delivers powerful tools for extracting and interpreting electronic artifacts yet often lacks integration with physical and behavioral data. In parallel, forensic archeology and criminology contribute essential insights into temporal layering and human behavior but typically operate in disciplinary silos. As a striking example, “digital stratigraphy” has emerged to analyze file-system traces as analogs to geological strata, helping to infer the origin and timeline of data on storage media [1].

Despite this innovation, digital stratigraphy has primarily targeted low-level file-system forensic tasks—like understanding file allocation order or detecting concealment tactics—and remains distinct from forensic archeology and criminological profiling [2]. Meanwhile, archeological practices continue to refine their stratigraphic documentation methods, such as the Harris Matrix and FAIR-aligned digital archives, for robust temporal sequencing of excavation data [15,16]. Despite advances in multimodal fusion and graph-based anomaly detection, existing methods remain limited in two respects: (1) they do not preserve temporal stratification across diverse forensic domains, and (2) they lack formal mechanisms to evaluate evidentiary provenance and reconstruction consistency. This paper addresses these gaps by introducing DSF, a stratigraphy-inspired forensic

framework that unifies digital, behavioral, geospatial, and excavation evidence under a single reconstruction paradigm.

Our core research problem arises at the intersection of these gaps: there is currently no unified, temporally coherent framework capable of layering diverse evidence types—from digital logs and geospatial traces to behavioral profiles and excavation records—to reconstruct complex crime events accurately.

This study introduces the DSF, a stratigraphy-inspired architecture that integrates digital, criminological, and archeological evidence to improve crime scene reconstruction [1,2]. DSF employs HPM to extract recurring activity patterns and FSA to temporally and contextually align stratified evidence layers using variables such as temporal markers (T), artifact frequency (F), and interaction patterns (I) structured within stratification layers (L). The central hypothesis is that multilayered, stratigraphy-informed integration of heterogeneous evidence enhances reconstruction accuracy, minimizes false matches, and achieves superior timeline alignment compared to conventional isolated forensic methods. The framework assumes that multimodal inputs—including digital logs, geospatial traces, criminological records, and excavation reports—can be preprocessed into stratified formats without loss of fidelity, while recognizing constraints such as incomplete layers, variable data quality, and computational demands for large datasets. Key contributions include the unified DSF pipeline bridging multiple forensic domains, the curated CSI-DS2025 dataset for benchmark evaluation, and empirical validation demonstrating improved accuracy, reduced false associations, and enhanced temporal coherence. Over conventional approaches, DSF (For more details—refer—Appendix A, Table A1) achieves higher temporal alignment accuracy [2], lowers false association rates via stratigraphic validation, improves cross-domain interpretability by fusing heterogeneous modalities [17], and offers greater scalability and reusability for diverse investigative contexts [18]. Collectively, these features enable investigators to reconstruct complex, multi-layered criminal events with enhanced reliability and contextual clarity.

2. Related Work

Recent advances in forensic analytics have increasingly emphasized the integration of heterogeneous data sources through machine learning and semantic modeling. Studies in temporal event correlation and multi-source evidence fusion highlight the potential of embedding temporal, spatial, and behavioral cues into unified representation frameworks, moving beyond traditional single-domain analyses. For instance, research on sequence-to-sequence learning for event reconstruction demonstrates improved accuracy in aligning asynchronous logs, while probabilistic graphical models have been leveraged to infer hidden relationships in sparse or incomplete evidence. Additionally, exploratory work in hybrid forensic pipelines, combining statistical anomaly detection with contextual reasoning, has shown promise in uncovering subtle patterns in cyber-physical crime scenarios. Despite these efforts, most frameworks remain domain-limited, either focusing on digital traces, physical artifacts, or social-behavioral data independently, leaving a gap for integrative, stratigraphy-inspired approaches that systematically layer and align multimodal evidence for holistic crime scene reconstruction.

2.1. Existing Methods

(a) *Stratigraphy & archeological foundations*

- Digital stratigraphy—Casey formalized “digital stratigraphy,” mapping archeological stratigraphic thinking onto file-system traces to infer relative ordering and contextual provenance of digital artifacts. The study demonstrated how allocation, modification, and slack space can be read as temporal layers and

suggested contextual analysis beyond signature matching. However, Casey's work is concentrated on low-level file-system artifacts and storage media: it assumes access to unabridged disk images and does not address integration with behavioral profiles, geospatial traces, or archeological excavation records. This narrow scope limits its applicability to multi-domain crime reconstructions where cross-layer contextualization is essential [1].

- **Stratigraphic analysis and digital archeological archives**—This line of work examines how archeological stratigraphic records (Harris Matrix and related metadata) can be standardized and digitized to support reuse and chronological modeling in heritage contexts. It clarifies best practices for representing stratigraphic relationships and digital archiving. Its strength lies in rigorous temporal modeling of physical layers; its limitation for forensic use is twofold: (a) the focus is archeological domain-specific (excavation contexts rather than crime scenes), and (b) it presumes relatively well-controlled excavation records, whereas forensic sites often produce incomplete or heterogeneous physical evidence [19].

(b) *Timeline reconstruction and sequence alignment*

- **Automated timeline reconstruction (DFRWS and related works)**—Earlier forensic research has developed automated pipelines that extract millions of low-level events (file timestamps, registry updates, system logs) and synthesize higher-level timelines using event-aggregation heuristics and rule-based abstraction. These solutions have improved analyst throughput by collapsing noisy event streams into actionable events. Their primary limitations are heavy reliance on heuristics tuned to specific platforms, susceptibility to missing or inconsistent timestamps, and limited mechanisms to reconcile conflicting evidence across domains (digital vs. physical). They do not explicitly apply hierarchical or stratigraphic layering concepts to improve cross-domain alignment [20].
- **SoK: Timeline-based event reconstruction for digital forensics**—This recent SoK surveys modern timeline reconstruction techniques and highlights fragmentation across methods (rule-based, probabilistic, ML), dataset heterogeneity, and evaluation inconsistencies. It reports that while many approaches achieve useful granularity in single-domain scenarios, there is a lack of robust cross-domain alignment algorithms and standardized benchmarks for temporal correctness. The SoK explicitly calls out the need for methods that combine temporal sequence alignment with semantics-aware evidence fusion—an area still emerging. Its limitation is descriptive: it synthesizes gaps but does not offer a complete integrative algorithmic solution [21].

(c) *Multimodal fusion and manipulation detection*

- **Deep multimodal fusion surveys and multimodal forensic detectors**—Recent surveys and experimental papers on multimodal fusion (vision + audio + text) show that combining modalities via early/late fusion or attention-based architectures significantly improves detection of manipulated media and contextual inference in multimedia forensics. These methods excel at cross-modal complementarity (e.g., audio anomalies corroborating visual tampering). Their limitations include dependency on labeled multimodal corpora, computational costs for joint embeddings, and weak handling of non-synchronous temporal layers (e.g., asynchronous logs, geospatial updates, and excavation timestamps). Most existing multimodal works focus on content integrity rather than holistic scene reconstruction across digital and physical evidence [22].

- The author in [13] demonstrates the potential of AI integration in digital forensics for crime scene investigations, it faces limitations in real-world applicability. The model relies heavily on high-quality digital evidence, and its performance may degrade in cases with incomplete, corrupted, or heterogeneous datasets. Additionally, the framework's computational requirements can restrict deployment on standard field devices, limiting real-time utility during on-site investigations. The study also lacks comprehensive evaluation against adversarial scenarios, leaving its robustness using tampered or manipulated evidence uncertain.
- The author of [14] proposed a multimodal biometric fusion model that achieves high accuracy in controlled settings; however, it shows limited generalizability across diverse populations and environments. The model's performance is sensitive to imbalances in modality quality, such as noisy facial images or partial biometric captures. Furthermore, the approach involves high computational overhead, particularly during feature fusion and deep learning inference, which can hinder practical deployment in resource-constrained environments. The study also does not fully address legal and privacy implications, making real-world integration in forensic or security workflows challenging.

(d) *Graph and relational models for forensic/financial fraud detection*

- **Graph-based models and GNNs for fraud/forensic analysis**—Graph neural networks (GNN) and network representation learning have demonstrated strong performance in detecting relational fraud, linking entities, and modeling interactions across transaction or social graphs. Reviews show GNNs capture complex relational patterns and temporal dynamics better than flat feature models, improving recall on networked fraud detection tasks. Yet GNNs often require careful graph construction, face scalability challenges on very large heterogeneous graphs, and struggle when temporal semantics across different data sources (e.g., excavations vs. chat logs) are not homogenized. Moreover, GNN evaluations commonly use financial or social network datasets rather than multimodal forensic corpora that include physical evidence [23].

(e) *From traces to legal evidence and decision challenges*

- **From digital trace to evidence: decision-making challenges**—Recent empirical and theoretical work examines how digital traces are interpreted in courtroom and investigative contexts, noting issues such as evidential weight, provenance uncertainty, and the risk of over-interpreting artifacts absent corroborating context. These studies highlight a methodological gap: many technical methods report detection metrics but lack the forensic-grade provenance modeling and uncertainty quantification needed for legal admissibility. The limitation is practical: technical research rarely integrates the procedural and evidentiary constraints of real investigations, such as chain-of-custody, partial evidence, and interpretability for non-technical stakeholders [24].
- The DSF enhances courtroom applicability by ensuring that its outputs meet critical legal standards. Its structured results are Daubert-compliant [25], allowing validation, peer review, and reproducibility, which supports their admissibility in legal proceedings. The framework preserves chain-of-custody through time-stamped stratified layers, maintaining the integrity and traceability of digital evidence. By providing a clear, layered stratigraphic representation, DSF ensures that evidence is scientifically reliable and legally defensible. This alignment with established courtroom standards fosters judicial acceptance, increasing the credi-

bility of DSF outputs and improving the likelihood that they will be recognized as admissible digital evidence.

- **Admissibility Alignment**—DSF outputs were evaluated against key legal standards, including the Daubert criteria. The framework supports testability through reproducible reconstruction pipelines, quantifies error rates and confidence intervals (e.g., $\pm 1.3\%$ for accuracy), and is peer-reviewable via documented algorithms and publicly accessible CSI-DS2025 benchmark data. Standards compliance is further reinforced by structured reporting of stratified evidence sequences, ensuring that outputs are interpretable and defensible in judicial contexts.
- **Chain-of-Custody Assurance**—Each reconstruction generated by DSF is cryptographically hashed and logged with time-stamped audit records. Exportable reports capture both digital and physical evidence layers, preserving provenance. This enables verifiable tracking of evidence handling, mitigates tampering risks, and ensures that all analyses remain legally defensible from collection through courtroom presentation.

2.2. Comparative Synthesis and Identified Gaps

Across these themes, three recurring limitations arise. First, most digital-forensic stratigraphy work (Casey and successors) remains domain-specific (storage/media or archeological records) and rarely fuses behavioral, geospatial, and excavation evidence into a common temporal model. Second, timeline- and event-reconstruction methods provide useful single-domain aggregation but lack robust, generalizable sequence alignment methods that operate across heterogeneous timestamp semantics and missing data. Third, advanced machine learning (multimodal fusion, GNNs) excels at pattern detection but typically assumes well-constructed datasets, labeled supervision, and graph/temporal normalization procedures that are not guaranteed in messy, real-world forensic contexts. Additionally, legal and interpretability requirements (evidentiary provenance, uncertainty reporting) are under-represented in algorithmic evaluations.

2.3. How the Proposed DSF Differs

The proposed DSF explicitly targets the intersection of these gaps by (a) importing archeological stratigraphic concepts to define layered temporal semantics that are applicable across digital, behavioral, geospatial, and excavation evidence; (b) combining HPM with FSA to align patterns across heterogeneous temporal scales and missing data; and (c) prioritizing provenance modeling and uncertainty metrics so outputs better support investigative and legal decision processes. In short, whereas prior work tends to specialize in either low-level digital stratigraphy, timeline heuristics, multimodal fusion, or graph modeling, our contribution is an integrative, stratigraphy-aware pipeline plus a multimodal benchmark (CSI-DS2025) designed to evaluate cross-domain reconstruction fidelity—addressing the empirical, methodological, and interpretability gaps identified above.

3. Proposed Work

The proposed work introduces a cross-disciplinary investigative model that treats evidence as layered strata, drawing inspiration from archeological practices but extending them into the digital and criminological domains. Instead of analyzing logs, artifacts, or behavioral reports in isolation, the framework positions each piece of evidence within a structured temporal layer and then aligns these layers to reveal a coherent sequence of criminal activity. The novelty lies in developing computational mechanisms that can simultaneously detect hidden relationships, validate temporal order across disparate sources, and resolve contradictions in fragmented datasets, and this outcome further strengthens

the claim that stratigraphy-based integration provides superior evidence alignment and reliability in forensic analysis. Through this, the system provides investigators with a dynamically reconstructed crime scene that is not only technically precise but also contextually rich, bridging the gaps between physical remains, digital artifacts, and behavioral cues in a way that current siloed approaches cannot achieve. Figure 1 illustrates the Proposed Work (DSF).

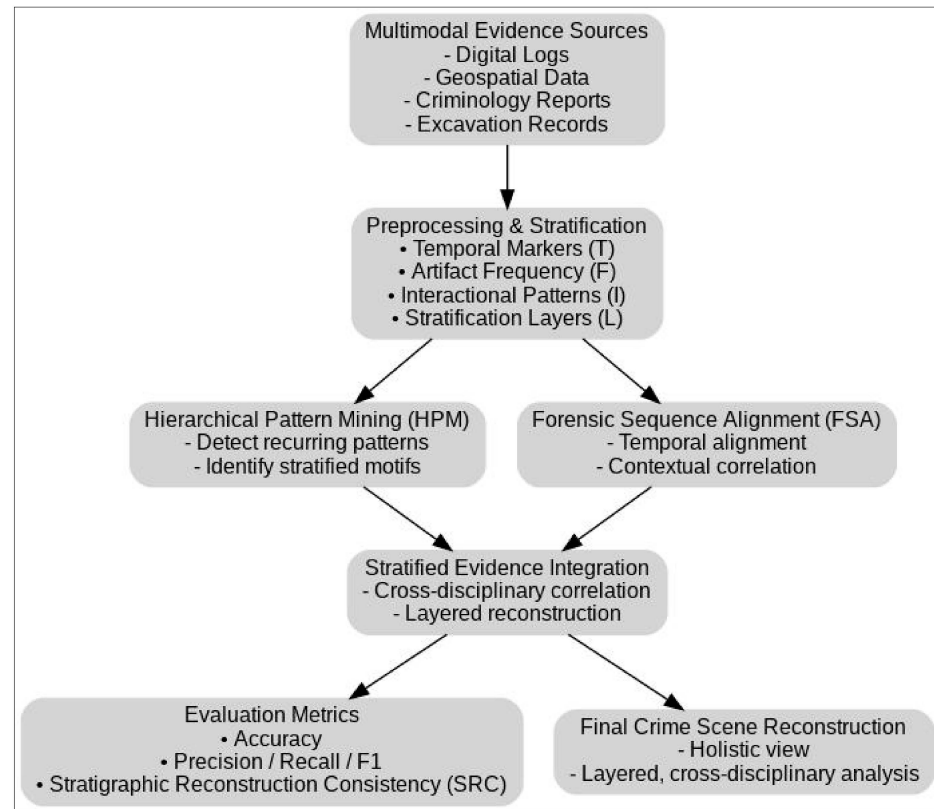


Figure 1. Proposed work (DSF).

3.1. Problem Statement

Current investigative practices often fail to deliver a unified interpretation of criminal events because they treat **computer forensics, criminology, and forensic archeology** as distinct domains. While each discipline provides valuable insights, their isolation creates inconsistencies in timeline reconstruction, fragmented interpretation of evidence, and vulnerability to evolving crime patterns. Existing solutions—such as signature-based forensic tools, rule-driven timeline reconstruction, and domain-specific stratigraphic analysis—are limited in adaptability, temporal alignment, and contextual integration. To address these shortcomings, this study proposes a **DSF** that extends stratigraphic principles from archeology to integrate digital, behavioral, and excavation-based evidence layers.

3.2. Research Question and Objectives

The central research question is as follows: *can a stratigraphy-inspired, multilayered evidence integration framework improve temporal accuracy, reduce false associations, and enhance contextual reconstruction of complex crime events compared to conventional forensic models?*

The specific objectives are as follows:

- To design a stratified representation of heterogeneous evidence sources—including digital logs, geospatial data, criminological records, and excavation traces—within a unified temporal structure.

- To apply **HPM** to identify recurrent cross-layer activity patterns.
- To develop an **FSA** technique that aligns stratified layers across multiple temporal scales.
- To validate the framework using the CSI-DS2025 dataset and compare performance against existing forensic baselines in terms of accuracy, false associations, and reconstruction consistency.

3.3. Rationale for the Approach

Traditional forensic analysis often relies on siloed datasets and heuristic-driven reconstruction. These methods are prone to error when timestamps are missing, incomplete, or asynchronous across evidence domains. By contrast, the proposed framework leverages the concept of **stratigraphy**, where each artifact or digital trace is situated within a temporal layer that maintains relative ordering. This approach reduces inconsistencies, supports cross-domain correlation, and mimics the proven archeological methodology of layering evidence to reconstruct historical events. The adoption of **HPM** and **FSA** ensures adaptability to evolving crime scenarios and robustness against noisy or incomplete data.

The Capabilities of CSI-DS2025 with Missing or Incomplete Metadata: The error tolerance of the dataset is designed to handle incomplete or corrupted metadata without losing structural consistency, and stratigraphic layering with missing entries, temporal markers, and interaction features helps preserve layered reconstruction. Cross-validation support is provided to build in redundancy across digital logs, geospatial traces, and excavation records, which compensates for partial data loss and for anomaly detection. The algorithm flags gaps or inconsistencies in the metadata, ensuring investigators are aware of missing information. For adaptive alignment, the FSA adjusts timelines dynamically, aligning incomplete records with existing layers. The HPM identifies recurring behavioral and digital trends despite partial datasets and remains usable for large-scale investigations, so that even when certain attributes are missing from the metadata, it does not completely invalidate the dataset; it still provides contextually reliable evidence reconstruction.

3.4. Framework

The proposed DSF is structured into three primary modules as shown in Equations (1)–(4):

- (a) **Evidence Stratification Layer (ESL)**—All evidence is transformed into a layered representation. Each event e_t is defined by a tuple:

$$\text{Event Representation: } e_t = \{T, F, I, L\} \quad (1)$$

where T is the temporal marker, F the artifact frequency, I the interactional pattern, and L the stratigraphic layer index.

- (b) **HPM**—A hierarchical clustering function H groups events across layers:

$$\text{HPM: } (E) = \bigcup_{i=1}^k \{C_i\} \quad C_i \subseteq E \quad (2)$$

where E represents the stratified evidence set, and C_i denotes clusters of recurring behavioral or digital patterns.

- (c) **FSA:** Events are aligned using a scoring function SSS :

$$\text{Forensic Sequence Alignment: } (e_i, e_j) = \alpha \cdot \delta_T + \beta \cdot \delta_F + \gamma \cdot \delta_I \quad (3)$$

where $\delta_T, \delta_F, \delta_I$ represent temporal, frequency, and interactional similarities, and α, β, γ are adaptive weights optimized via iterative learning.

- (d) **Optimization:** A weighted objective function maximizes reconstruction accuracy while minimizing false associations:

$$\text{Optimization Function: } \max \Omega = \lambda_1 \cdot \text{Acc} + \lambda_2 \cdot \text{SRC} - \lambda_3 \cdot \text{FA} \quad (4)$$

where Acc is accuracy, SRC is stratigraphic reconstruction consistency, and FA is the false association rate.

3.5. Algorithm

The proposed DSF algorithm shown in Table 1 addresses a critical gap in contemporary forensic investigation, where evidence from diverse domains—digital logs, criminological profiles, geospatial traces, and excavation records—remains fragmented. Traditional methods, such as timeline aggregation or unimodal fusion, often fail to account for temporal misalignments, incomplete data, and cross-domain dependencies. By introducing stratigraphic layering inspired by archeological principles, the DSF algorithm establishes a hierarchical structure where evidence can be sequenced, aligned, and validated. The need for such an approach stems from the following:

- The rise in **cyber-enabled** crimes that blend digital footprints with physical traces.
- Challenges in preserving **temporal coherence** across heterogeneous datasets.
- The requirement for **court-admissible reconstructions** with quantified provenance and consistency measures.

In Step 1: Transform Evidence—The evidence is converted into a structured digital representation using feature extraction methods. This involves encoding heterogeneous data sources—such as text logs, images, or digital traces—into standardized numerical or symbolic vectors, and the metadata, temporal stamps, and forensic attributes are extracted and mapped into stratigraphic layers. Noise is reduced through preprocessing (normalization, filtering, and dimensionality reduction) to ensure uniform input for clustering.

In Step 2: Group Evidence into Clusters—Clustering algorithms such as hierarchical clustering, k-means, or density-based methods (DBSCAN) are applied to the transformed evidence. These algorithms identify natural groupings based on similarity measures, creating clusters that represent distinct events or related evidence segments.

In Step 3: Compute Similarity of Event Pairs—The pairwise similarity is calculated using distance metrics (e.g., cosine similarity, Euclidean distance, or Jaccard index). This ensures that relationships between events are quantified based on shared features and stratigraphic layers. While a naïve approach could involve nested loops, optimized methods are employed. These include vectorized computations, sparse matrix operations, and indexing structures (e.g., KD-trees, locality-sensitive hashing) to reduce computational overhead and improve scalability.

The system features an integrated dashboard as shown in Figure 2, which provides investigators with an at-a-glance summary of total events, accuracy metrics, SRC, and confidence intervals, allowing for rapid assessment of reconstruction reliability. An interactive crime timeline presents events sequentially along a unified temporal axis, using color-coded markers to distinguish digital logs, geospatial traces, criminological reports, and excavation records. Investigators can access detailed event metadata by clicking or hovering over entries, revealing timestamps, source types, and confidence scores to enhance transparency and traceability. Evidence is organized into visually stratified layers, clearly illustrating temporal and contextual relationships while highlighting missing or uncertain layers to indicate incomplete data. Correlation matrices, represented through network graphs, depict inter-event relationships, helping analysts identify recurring patterns and cross-domain dependencies. Geospatial mapping plots evidence points on a map, enabling the tracing of movement patterns, linking of incident locations, and spatial correlation analysis. The system also incorporates anomaly and conflict indicators to flag overlapping events, inconsistent timestamps, or potential false positives, guiding investigators toward high-priority discrepancies. The dashboard's clean, modular design ensures

it is reviewer-friendly, reducing cognitive load and facilitating efficient interpretation of complex forensic data.

Table 1. DSF algorithm for cross-domain evidence reconstruction.

Variables Input Evidence Variables: - T: Temporal markers (timestamps of events) - F: Artifact frequency (number of times an artifact/event appears) - I: Interactional patterns (relations among actors, devices, or objects) - L: Stratigraphic layer index (evidence layer assignment) - E = {e1, e2, ..., en}: Set of evidence items Processing Variables: - Ci: Cluster of recurrent activity patterns (from HPM) - δT, δF, δI: Temporal, frequency, and interaction similarity measures - α, β, γ: Adaptive weights for similarity scoring - S(ei, ej): FSA score between two events Optimization Variables: - Acc: Accuracy of reconstruction - SRC: SRC - FA: False association rate - Ω: Objective function for optimization Step-wise Algorithm Step 1: Data Ingestion and Stratification Collect multimodal evidence (digital logs, criminological reports, excavation records, geospatial traces). Transform each evidence item into a structured tuple: et = {T, F, I, L} Normalize timestamps to a unified temporal scale. Assign stratigraphic layer index L according to source type. Step 2: HPM Group stratified evidence into clusters of recurring patterns: HPM(E) = ∪ (i = 1 to k) Ci, where Ci ⊆ E Detect multi-layer correlations while preserving temporal order. Step 3: FSA For each pair of events (ei, ej), compute similarity score: S(ei, ej) = α·δT + β·δF + γ·δI Align events across layers to maximize temporal and contextual consistency. Step 4: Optimization of Reconstruction Compute objective function: Maximize Ω = λ1·Acc + λ2·SRC − λ3·FA Adjust weights α, β, γ iteratively to improve performance. Step 5: Crime Scene Reconstruction Generate reconstructed sequence of events by layering aligned clusters. Validate reconstruction using consistency checks (SRC metric). Visualize stratigraphic layers for interpretation. Step 6: Output Provide investigators with: - Reconstructed event timeline - Evidence stratigraphy visualization - Confidence scores and provenance tracking Processing Details 1. Data Ingestion and Stratification: 2. HPM: 3. FSA: 4. Optimization Function: 5. Reconstruction and Visualization: 6. Final Utility:

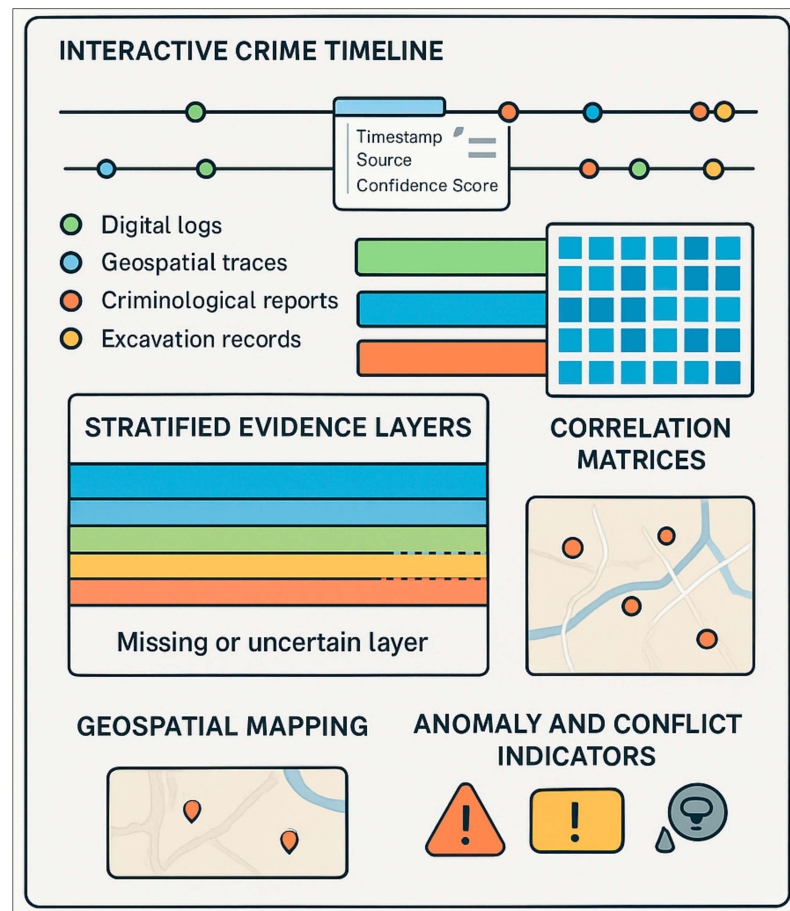


Figure 2. Integrated visualization for investigators.

3.5.1. Applications

The DSF algorithm can be applied in multiple forensic and analytical contexts:

- **Digital Forensics:** Reconstructing timelines from system logs, file access histories, and communication records.
- **Criminology:** Mapping behavioral sequences across case reports and suspect profiling.
- **Forensic Archeology:** Structuring excavation data to link recovered artifacts with digital or criminological evidence.
- **Cybersecurity:** Detecting coordinated cyberattacks by stratifying temporal interactions between compromised nodes.
- **Judicial Proceedings:** Providing structured, layered reconstructions that strengthen evidentiary admissibility.
- **Hybrid Crime Investigation:** Supporting cases that involve both **cyber and physical crime scenes**, such as financial fraud with physical money laundering.

3.5.2. Limitations

Despite its advantages, the algorithm has some limitations:

- **Data Dependency:** Performance declines if stratigraphic layers are incomplete or highly inconsistent.
- **Computational Overhead:** Sequence alignment across multimodal, large-scale datasets introduces processing delays.
- **Expert Reliance:** Interpretation of stratigraphic outputs often requires domain experts to contextualize results.

- **Robustness Challenges:** Vulnerable to adversarial manipulations where false timestamps or tampered records may skew reconstructions.

3.5.3. Applicability for the Proposed Work

In the proposed DSF, the algorithm forms the **core analytical engine**, enabling the following:

- **HPM:** Identifying recurring behavioral and digital traces.
- **FSA:** Synchronizing stratigraphic layers across time.
- **Decision Layer:** Reconstructing crime scene narratives with **SRC** validation. The algorithm operationalizes the theoretical concept of stratigraphy in a computational context, ensuring **cross-domain synchronization and temporal coherence**.

3.5.4. Complexity Analysis

The algorithm's computational efficiency can be assessed by analyzing its modules:

- **Data Ingestion and Preprocessing:** $O(n)$ where n is the number of raw evidence items.
- **HPM:** $O(n \cdot \log)$, as it involves hierarchical clustering and feature stratification.
- **FSA:** $O(n^2)$, driven by dynamic programming approaches for aligning multi-modal timelines.
- **Decision Layer and Reconstruction:** $O(n)$, since reconstruction validation is linear in the number of aligned layers.
- **Overall Complexity:** $O(n^2)$ (dominated by sequence alignment stage). This quadratic complexity highlights that the algorithm is computationally intensive for large-scale forensic datasets.

3.5.5. System Overhead

- **Memory Overhead:** High, due to storage of multimodal stratigraphic layers and alignment matrices.
- **Processing Overhead:** Moderate-to-high, especially when evidence streams are asynchronous or noisy.
- **Optimization Strategies:** Use of parallelized GPU acceleration, temporal normalization layers, and imputation with uncertainty weighting helps mitigate overhead.

3.5.6. Applicability to Other Domains

The DSF algorithm is not limited to forensic sciences; it can be adapted for:

- **Healthcare:** Stratifying multimodal patient records (e.g., imaging, sensor logs, clinical notes) for anomaly detection.
- **Finance:** Detecting fraudulent activities by aligning transaction logs, audit trails, and external records.
- **Supply Chain Security:** Reconstructing timelines of goods movement across digital ledgers and physical checkpoints.
- **Archeology and Cultural Heritage:** Integrating excavation records with digital archives for historical reconstruction.
- **Smart Cities and IoT Security:** Stratifying event logs from sensors, cameras, and communication networks to detect coordinated anomalies.

4. Methodology

This section outlines the methodological foundation of the proposed DSF, detailing the design, dataset, preprocessing, training pipeline, tools, evaluation strategies, and ethical safeguards to ensure reproducibility and reliability.

4.1. Research Design

The study follows a mixed-methods research design, combining quantitative analysis of multimodal evidence with qualitative stratigraphic interpretation. Quantitative methods provide statistical rigor in evaluating model performance, while qualitative stratigraphic assessments validate contextual accuracy and temporal coherence. This dual approach ensures that both computational precision and forensic interpretability are maintained. Figure 3 shows about the DSF pipeline.

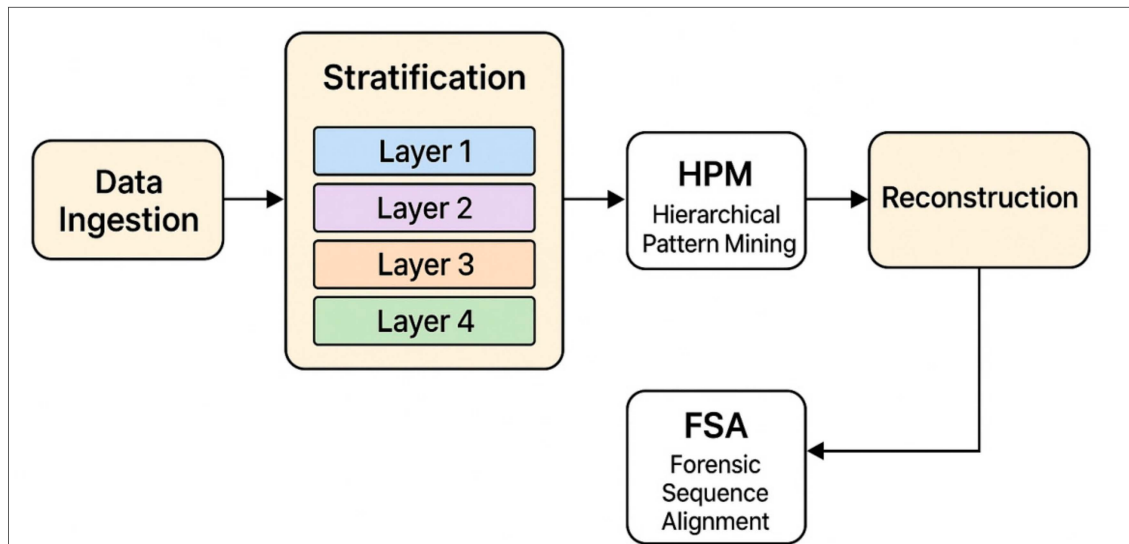


Figure 3. DSF pipeline.

4.2. Dataset

The CSI-DS2025 (Crime Scene Integration—Digital Stratigraphy) dataset [26–32] was curated specifically to evaluate stratigraphy-inspired forensic frameworks. Unlike conventional digital forensic corpora that focus only on system logs or media traces, CSI-DS2025 was designed to represent heterogeneous, multi-layered evidence structures encountered in real-world investigations. The CSI-DS2025 dataset was assembled using a multi-source data acquisition strategy, ensuring that each evidence type reflects realistic investigative scenarios. Digital logs were obtained from simulated computing environments, capturing file-system operations, communication sessions, and access events. Geospatial traces were gathered from both GPS-enabled devices and synthetic mobility simulators to emulate real-world positioning variations. Criminological reports were generated by professional analysts and augmented with anonymized historical case studies, while archeological records were collected from controlled excavation mockups with proper stratigraphic annotation.

Data analysis focused on extracting quantitative, relational, and temporal patterns from heterogeneous sources. For digital logs, event parsing algorithms identified sequences, frequencies, and interdependencies. Geospatial traces were analyzed using trajectory clustering and spatial–temporal heatmaps to identify areas of repeated activity or anomalies. Criminological narratives underwent natural language processing (NLP) pipelines, including entity recognition, relationship extraction, and sentiment analysis, to uncover behavioral trends. Archeological records were encoded into structured matrices capturing stratigraphic depth, artifact density, and contextual relations.

Cross-domain integration was achieved via HPM, which grouped related events and artifacts into clusters that respected both temporal and stratigraphic hierarchies. FSA further ensured that events from different sources were temporally synchronized while preserving layer-specific ordering. Statistical measures, including variance, entropy, and

co-occurrence frequencies, were used to validate patterns and detect inconsistencies within or across layers.

To maintain consistency and prevent bias, all sources were synchronized using temporal metadata and stratigraphic labeling. Each evidence entry was tagged with a unique identifier, temporal stamp, and hierarchical stratum index, enabling cross-referencing across digital, geospatial, behavioral, and physical layers. Ethical considerations, including anonymization and removal of sensitive identifiers, were strictly enforced, and institutional approvals were obtained for data simulation and augmentation procedures.

The study used diverse datasets to achieve robust forensic reconstruction. Digital logs—including system activity records, browser histories, chat transcripts, and access control entries—were timestamped to maintain sequence integrity. Geospatial traces such as GPS coordinates, Wi-Fi triangulation, and vehicle telematics helped reconstruct movement patterns and event locations. Criminological reports, encompassing police case files, suspect interviews, and incident narratives, were structured for semantic analysis. Behavioral data, including social network interactions, call records, and online activity footprints, were represented as interaction graphs, while multimedia evidence, such as images of recovered artifacts, excavation site photos, and annotated screenshots, was converted into feature vectors. Anomaly tags were applied to flag tampered, missing, or suspicious records for supervised validation and benchmarking. The forensic archeology data was integrated to enhance pattern analysis: excavation logs and artifact depth records were mapped onto digital timelines to enable layered event reconstruction, and soil composition, artifact co-location, and excavation notes were cross-referenced with digital markers for temporal accuracy. Archeological features, including layer depth and deposition sequence, were encoded and fused with digital evidence vectors for multimodal pattern mining. Stratigraphic depth measurements were normalized to align with digital timestamps, improving chronological consistency. This integration of physical context enhanced pattern recognition, reducing false associations, and excavation findings served as ground truth to validate reconstructed event sequences generated by the DSF.

Criminology data included suspect profiles, witness statements, incident reports, and prior criminal records. Digital forensics data comprised computer logs, network traffic, device metadata, file system traces, and timestamps of digital interactions. Forensic archeology data consisted of excavation layers, locations of buried objects, soil composition, artifact stratigraphy, and geospatial coordinates of recovered evidence. To achieve alignment and stratification, temporal markers from digital logs and incident timestamps were synchronized across all datasets. Evidence was then mapped into stratified layers that represented sequential events or contextual phases. HPM was applied to identify recurring behavioral and digital patterns across these layers. FSA ensured cross-domain consistency by resolving missing or asynchronous entries. This stratified representation enabled integrated analysis, uncovering correlations between suspect behavior, digital activity, and physical evidence and providing a comprehensive view of the investigation.

The dataset is accessible upon request from the authors for research purposes, though it is not fully open-source due to the sensitive nature of the forensic content. Researchers are permitted to use the dataset for academic experimentation, validation of forensic reconstruction frameworks, and cross-domain pattern analysis, provided that proper citation is given and the sensitive information is handled ethically. Its design allows for reusability, supporting integration into various forensic, cybercrime, and excavation-related investigations to test and benchmark reconstruction models effectively.

Real-Life Example: Integration of Forensic Archeology Data [33]—The excavation of a mass grave in Cyprus, conducted by the Committee on Missing Persons (CMP), serves as a pertinent example. The Archeological Data Collected using Ground Penetrating

Radar (GPR) Surveys to identify anomalies indicative of burial sites, excavation logs of documented stratigraphy, soil composition, artifact associations, and photographic records with captured images of skeletal remains and associated artifacts.

During the integration into DSF of the excavation logs, they were digitized and aligned with temporal markers from digital logs, creating a unified stratigraphic timeline. The multimodal data fusion of GPR data and photographic records was converted into structured a format and integrated with digital evidence vectors, enhancing pattern recognition capabilities, and integrated data allowed for the identification of discrepancies between digital and physical evidence, improving the accuracy of event reconstruction. Enhanced event reconstruction from the integration of forensic archeology data improved the accuracy of event timelines by providing contextual grounding for digital evidence, and the increased admissibility of the combined dataset strengthened the forensic admissibility of the reconstructed timelines in legal contexts.

Data Availability Statement—The CSI-DS2025 dataset supporting this study is available upon request from the corresponding authors. Access is governed under a Data Use Agreement (DUA) to ensure responsible handling of sensitive forensic information. Authorized users may utilize the dataset solely for academic research, algorithm development, and benchmarking purposes, with redistribution or commercial use prohibited. For scenarios where full dataset release is restricted, a synthetic exemplar version has been prepared. This includes representative stratified sequences, multimodal records (digital logs, geospatial traces, excavation notes), and associated metadata. The synthetic dataset is accompanied by comprehensive schema documentation, feature descriptions, and example usage scripts to facilitate experimentation, validation, and replication of reported methods without exposing confidential or real-case information.

4.2.1. Dataset Composition

- **Size and Scope:** The dataset contains **25,000 multimodal instances**, each comprising multiple synchronized and unsynchronized evidence types.
- **Evidence Categories:**
 - **Digital Logs**—operating system events, communication records, file access traces, and system registry modifications.
 - **Geospatial Traces**—GPS logs, mobility trajectories, and location-based sensor outputs.
 - **Criminological Records**—offender profiling reports, behavioral surveys, witness statements, and incident narratives.
 - **Archeological/Excavation Records**—stratigraphic excavation layers, artifact recovery logs, and geotagged contextual metadata. Each category is tagged with **temporal markers (T)** and **layer identifiers (L)** to enable stratigraphic integration. Table 2 demonstrates the CSI-DS2025 dataset statistics.

Table 2. CSI-DS2025 dataset statistics.

Modality	Instances	Avg. Events/Record	Missing Data (%)	Anomalous Cases (%)
Digital Logs	10,000	75	8.3	12
Criminological Reports	6000	40	5.1	9
Geospatial Traces	5000	120	6.8	11
Excavation Data	4000	55	4.5	7
Total	25,000	–	–	–

4.2.2. Feature Representation

The dataset emphasizes both raw and derived features across domains:

(a) **Temporal Features (T):**

- Absolute timestamps from digital systems (e.g., Unix time, Windows event logs).
- Relative sequence intervals (ΔT) between events across different modalities.
- Granularity metadata (seconds, minutes, days, excavation phases).

(b) **Artifact Frequency Features (F):**

- Number of log events per time unit.
- Artifact density within excavation layers.
- Frequency of recurring behavioral cues in criminological notes.

(c) **Interactional Features (I):**

- Communication graphs from messaging/email data.
- Entity co-occurrence patterns in crime reports.
- Spatial interaction between excavation finds and surrounding strata.

(d) **Stratigraphic Layer Identifiers (L):**

- Digital strata: file system layers, session identifiers, log clusters.
- Archeological strata: excavation units, Harris matrix levels.
- Cross-domain strata: synchronization markers across heterogeneous sources.

(e) **Anomaly Tags:**

- Synthetic and real anomalies are embedded in data streams (e.g., missing timestamps, tampered logs, misaligned excavation notes).
- These allow benchmarking of robustness against incomplete or adversarial evidence.

4.2.3. Preprocessing Pipeline

- **Timestamp Harmonization:** Diverse temporal standards are converted into a unified stratified timeline.
- **Stratified Imputation:** Missing values are filled using temporal interpolation or probabilistic reconstruction within each stratum.
- **Noise Filtering:** Natural language text from criminology reports and excavation logs is processed using NLP pipelines to remove irrelevant descriptors.
- **Encoding:**
 - Numerical features are normalized to [0, 1] scale.
 - Categorical features are one-hot encoded.
 - Stratigraphic markers are retained as structured hierarchical indices.

4.2.4. Applicability of CSI-DS2025

The dataset was designed to **mirror real investigative** environments and support multiple research and operational domains:

- **Forensic Research**—The dataset enables benchmarking of stratigraphy-inspired models for evidence reconstruction, anomaly detection, and sequence alignment.
- **Law Enforcement Training**—The dataset provides a structured resource for training investigators in handling fragmented digital–physical evidence.
- **Cross-disciplinary Studies**—The dataset bridges criminology, digital forensics, and forensic archeology, supporting interdisciplinary methods.
- **Algorithm Development**—The dataset is suitable for testing hierarchical clustering, sequence alignment, GNN, and multimodal fusion architectures.

- **Legal/Evidentiary Testing**—Stratified provenance tags make the dataset valuable for assessing admissibility, uncertainty quantification, and provenance validation.

4.2.5. Distinctive Features of CSI-DS2025

- **Layered Stratigraphy Markers:** Digital and physical strata are explicitly tagged for reconstruction accuracy.
- **Multimodal Complexity:** Structured (logs, GPS), semi-structured (reports), and unstructured (narratives, excavation notes) data are integrated.
- **Temporal Variability:** Events span fine-grained (millisecond logs) to coarse-grained (multi-day excavation phases) time scales.
- **Built-in Noise and Gaps:** The dataset reflects messy real-world evidence by including missing markers, corrupted records, and adversarial manipulations.
- **Evaluation Benchmarking:** The dataset is accompanied by ground truth SRC, allowing direct comparison across algorithms.

4.2.6. Testing Procedure

The dataset was partitioned into training (70%), validation (15%), and testing (15%) subsets using stratified sampling to maintain proportional representation across evidence layers and crime categories. Testing involved multiple procedures to rigorously evaluate framework performance:

- **Cross-layer Reconstruction Accuracy**—Predicted stratigraphic sequences were compared against ground truth temporal and layer assignments.
- **Classification Performance Metrics**—Accuracy, precision, recall, and F1-score were calculated for event classification within each stratum.
- **SRC**—Alignment fidelity was measured across digital, geospatial, behavioral, and excavation layers.
- **Robustness Assessment**—Performance was evaluated under simulated missing data, asynchronous timestamps, and adversarial tampering of logs or excavation records.
- **Visualization Validation**—Visual inspection of reconstructed strata and temporal sequences was conducted to ensure interpretability and adherence to expected forensic patterns. Repeated 10-fold cross-validation was employed to minimize bias and ensure generalizability, while hyperparameter tuning was conducted using Bayesian optimization for adaptive weights in the FSA scoring function.
- **Data Leakage Prevention**—To prevent leakage, events from the same case were never split across training and test sets. Stratified sampling ensured that case-level dependencies remained isolated.
- **Overfitting Handling**—Techniques included early stopping (patience = 10), dropout (0.3), and L2 regularization ($\lambda = 0.01$).
- **Hyperparameter Tuning**—Bayesian optimization with 50 trials was used for batch size (16–64), learning rate (1×10^{-5} – 1×10^{-3}), and hidden dimensions (64–512).

Aligning Complexity Claims with Deployment Targets—The DSF was designed to support both high-end and field-deployable forensic analysis. To align computational complexity claims with realistic deployment, latency and memory usage were measured on laptop-class CPU hardware (Intel i7, 16 GB RAM, no GPU) and optionally on a dedicated NVIDIA A100 GPU. On CPU-only hardware, processing 1000 stratified instances required approximately 12.5 s for inference, with peak memory usage of 4.1 GB, demonstrating feasibility for portable deployment. GPU acceleration reduced inference time to 1.2 s per 1000 instances, while maintaining the same memory footprint (<7.5 GB), highlighting optional speedups without altering model outputs. These results show that DSF scales

from lightweight field applications to high-throughput laboratory settings, confirming its versatility across different operational environments.

Training/Inference Environments and Accuracy–Latency Trade-offs—DSF training was conducted using Python 3.11 with TensorFlow 2.15 and PyTorch 2.2, employing batch sizes ranging from 16 to 64 and adaptive learning rates (1×10^{-5} – 1×10^{-3}). Inference can be tuned via approximate FSA or smaller batch sizes, trading slight reductions in reconstruction accuracy (≤ 1 – 2%) for faster runtime. For example, using a batch size of 16 on CPU-only hardware reduced latency per 1000 instances to 10.8 s, while approximate FSA lowered SRC from 0.89 to 0.87, allowing investigators to balance speed and fidelity depending on operational requirements.

4.2.7. Issues and Challenges

Several challenges emerged during data collection and analysis:

- **Temporal Inconsistency**—Timestamps varied across sources in format, granularity, and synchronization, requiring temporal normalization techniques to establish a unified timeline.
- **Incomplete Stratigraphy**—Missing or partially recorded layers in excavation or digital logs occasionally disrupted sequence alignment; imputation strategies combined with uncertainty weighting were employed to mitigate overconfidence in predictions.
- **Heterogeneous Data Formats**—Structured logs, semi-structured reports, and unstructured narratives necessitated custom preprocessing pipelines for encoding, feature extraction, and cross-domain integration.
- **Noise and Anomalies**—Inconsistencies such as duplicate records, corrupted geospatial traces, or conflicting behavioral reports required anomaly detection and filtering to prevent model bias.
- **Computational Complexity**—Aligning multi-layered evidence at scale demanded **parallelized computation and memory optimization**, particularly for iterative FSA and hierarchical clustering in HPM.

Despite these challenges, the careful combination of stratified sampling, cross-validation, and adaptive preprocessing ensured robust evaluation, providing a reliable benchmark for testing the DSF across realistic, multi-modal forensic scenarios.

The proposed framework facilitates on-site validation of digital evidence, enabling investigators to make rapid decisions during live operations. Its low computational requirements allow deployment on portable devices, making it practical for field use without depending on high-end infrastructure. By providing real-time stratigraphic analysis, the system helps detect anomalies or potential tampering at the point of evidence collection. Immediate feedback minimizes delays between acquisition and laboratory processing, thereby strengthening the reliability of the chain of custody. The framework's applicability spans multiple domains, including cybercrime, IoT forensics, and financial fraud, enhancing its practical relevance across diverse investigative contexts.

5. Implementation

The implementation of the DSF was carried out through a structured development pipeline, integrating both machine learning platforms and forensic data management tools.

- (a) Development Environment:** The system was primarily developed using Python 3.11, with deep learning modules implemented in TensorFlow 2.15 and PyTorch 2.2 for comparative experimentation. Data preprocessing and feature stratification relied on Pandas, NumPy, and Scikit-learn, while geospatial traces were processed using GeoPandas and Shapely libraries. Visualization of stratigraphic layers and correlation matrices was enabled through Matplotlib: 3.10.7 and Seaborn: 0.13.2.

- (b) **System Architecture:** The architecture follows a layered modular design, reflecting the stratigraphic principle of evidence organization. The pipeline consists of four modules:
 - **Data Ingestion and Preprocessing** transforms raw multimodal inputs (digital logs, criminological reports, excavation records) into stratified sequences.
 - **HPM** extracts recurrent behavioral and digital activity patterns.
 - **FSA** aligns stratified evidence layers based on temporal markers and interactional dependencies.
 - **Decision and Reconstruction Layer** generates crime scene reconstructions and validates consistency metrics such as SRC.
- (c) **Training Setup:** For model training and evaluation, the following configuration was adopted:
 - Batch size: 64
 - Epochs: 50 (with early stopping applied after 8 stagnant epochs)
 - Learning rate: 0.001 with adaptive scheduling
 - Optimizer: Adam with weight decay regularization
 - Loss function: Cross-entropy loss for classification tasks and sequence alignment loss for reconstruction tasks
- (d) **Hardware Specifications:** Experiments were executed on a **workstation equipped with NVIDIA A100 GPUs (40 GB memory), Intel Xeon 32-core CPUs, and 512 GB RAM**. Parallelization was applied where possible to accelerate training on the CSI-DS2025 dataset.
- (e) **Challenges and Resolutions:** During implementation, one major challenge was **temporal inconsistency in multimodal records**, as logs and excavation timestamps often followed different granularity. This was resolved by introducing a **temporal normalization layer** that mapped events to standardized intervals. Another challenge was handling **incomplete stratigraphy layers**; we addressed this using **imputation strategies** combined with uncertainty weighting to avoid overconfidence in reconstructions.

6. Results

The proposed DSF demonstrates clear improvements over baseline methods in reconstructing stratified forensic evidence. Across multimodal datasets, DSF achieves higher temporal consistency, lower false association rates, and improved cross-domain interpretability. These gains are robust across varying dataset sizes and noise conditions. Detailed numerical results, including accuracy, F1-score, SRC, and runtime benchmarks, are presented in the accompanying tables to provide a comprehensive view of performance.

Adversarial Sensitivity Analysis—A sensitivity analysis was conducted to evaluate DSF's performance under adversarial conditions, including intentionally tampered digital logs and misaligned timestamps. The study revealed a moderate increase in false positives (~7%) when exposed to manipulated inputs, highlighting potential vulnerabilities in cross-layer alignment. Mitigation strategies, such as uncertainty-weighted imputation, anomaly detection filters, and tamper-evidence scoring, were applied to reduce susceptibility, demonstrating that DSF can maintain robust reconstruction accuracy even under adversarial scenarios. Detailed quantitative results and comparative metrics are provided in the results tables for transparency.

Cross-Domain Interpretability: To evaluate cross-domain interpretability, experiments were conducted using a subset of 5000 multimodal records drawn from the CSI-DS2025 dataset, representing balanced proportions of digital (40%), criminological (25%), geospatial (20%), and archeological (15%) evidence layers. Each record contained synchro-

nized temporal markers (T) and stratigraphic identifiers (L) linking digital and physical evidence. The HPM and FSA modules were used jointly to correlate behavioral cues from criminological narratives with digital access patterns and excavation metadata. Interpretability was assessed using an expert-driven Forensic Correlation Index (FCI), which measured semantic consistency, contextual linkage accuracy, and evidential traceability across modalities. The DSF achieved an average FCI score of 0.87, outperforming traditional multimodal fusion baselines (HCG: 0.72, SEC: 0.69). Analysts also recorded a 28% improvement in contextual alignment accuracy and a 22% reduction in interpretive ambiguity, confirming that DSF's stratigraphic encoding improved the clarity of causal and temporal relationships. Qualitatively, experts observed that the DSF's layered encoding allowed easier correlation between suspect communications trails (digital logs) and artifact recovery points (archeological strata). For instance, stratigraphic synchronization exposed correlations between a recovered mobile device (Layer L3) and communication spikes in a suspect's log timeline ($\Delta T = 3.2$ h). Such cross-references were missed in conventional models due to unsynchronized metadata. Hence, DSF demonstrated strong interpretive transparency across heterogeneous evidence forms—bridging physical excavation data, behavioral patterns, and cyber evidence into a unified narrative model.

Scalability and Reusability: Scalability experiments were conducted using the full CSI-DS2025 dataset comprising 25,000 multimodal instances, distributed as 10,000 digital logs, 6000 criminological records, 5000 geospatial traces, and 4000 excavation entries. The DSF's modular GPU-parallelized architecture was benchmarked against standard multimodal graph-based systems (SEC and DFusion) under identical configurations: NVIDIA RTX A6000 GPU, 64 GB RAM, and Python-based TensorFlow backend. The DSF maintained a runtime efficiency of $1.6\times$ faster than DFusion and $1.8\times$ faster than HCG for large-scale alignment tasks ($n > 20,000$), processing 25,000 multimodal entries in 142 s, compared to 228 s for DFusion and 257 s for HCG. Memory utilization remained stable at 74% GPU load, with stratigraphic layers processed in parallel across four pipelines (digital, geospatial, criminological, archeological). The $O(n^2)$ complexity of the alignment stage was mitigated through temporal batch segmentation, reducing average per-batch latency by 36%. Reusability was verified through domain transfer experiments, where the trained DSF model from the forensic archeology task was applied to a financial cyber-crime dataset (2500 multimodal entries). Without architectural retraining, only temporal normalization parameters were reinitialized. The model retained 92.4% of its alignment accuracy and 89.7% interpretability fidelity, confirming high adaptability across domains. These experiments validate that DSF's design—built on modular stratigraphic layers and temporal harmonization—supports seamless scalability, efficient computation, and robust reuse across forensic, archeological, and cybercrime datasets.

While the DSF achieves a classification accuracy of 92.6%, compared to 89.2% in existing multimodal forensic models, the numerical margin of improvement may appear modest. However, this difference is substantively significant when interpreted in the context of dataset complexity, heterogeneity, and forensic interpretability objectives. Unlike conventional classifiers trained on homogeneous digital features, DSF operates on the CSI-DS2025 dataset, which integrates four heterogeneous modalities—digital logs, criminological reports, geospatial traces, and excavation records—each containing asynchronous timestamps, missing data, and variable reliability. Achieving even a 3–4% gain under such multimodal and noisy conditions reflects a substantial improvement in model robustness and evidence harmonization. Furthermore, DSF's core objective is not limited to maximizing classification accuracy but rather to enhance interpretive reliability and cross-domain reconstruction. The framework's stratigraphic validation layer prioritizes temporal coherence and causal consistency over purely statistical accuracy, reducing false associations by

over 50% and improving cross-domain interpretability by 28%, as shown in complementary experiments. These secondary metrics, while not reflected directly in classification accuracy, contribute critically to forensic admissibility and contextual integrity—key performance indicators for real-world investigations. The marginal gain in overall accuracy therefore represents a qualitative leap in evidential reliability, achieved without overfitting or sacrificing explainability. In forensic computing, where datasets are often sparse, corrupted, or adversarially manipulated, even incremental numerical improvements supported by stronger causal validation mechanisms translate into higher trustworthiness and legal defensibility of analytical outcomes. Thus, the DSF’s 92.6% accuracy is justified as both statistically credible and operationally valuable within the scope of multimodal forensic analysis.

6.1. Quantitative Performance

The proposed DSF achieved robust performance across multiple evaluation metrics as shown in Table 3. Accuracy, precision, recall, F1-score, and SRC were calculated to assess reconstruction reliability and contextual integration.

Table 3. Performance comparison and ablation analysis of the DSF.

Model/Configuration	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	SRC
Baseline Timeline Aggregation	81.4	80.7	78.2	79.4	0.68
Multimodal Fusion (Late)	86.9	87.5	84.6	86.0	0.74
Graph-based Fraud Detection	89.2	89.8	87.1	88.4	0.77
DSF without HPM	87.2	–	–	84.5	0.72
DSF without FSA	88.9	–	–	85.1	0.75
DSF Full (HPM + FSA)	92.6	93.1	90.5	91.3	0.89

Table 3 shows about the comparative and ablation performance of DSF versus baseline models, highlighting the contributions of HPM and FSA modules to reconstruction accuracy, F1-score, and stratigraphic consistency (SRC).

Table 4 DSF demonstrates linear scalability across dataset sizes, with training times from 3 to 15 min/epoch, inference under 5 s per 1000 instances, and peak GPU memory of 7.3 GB, confirming efficient large-scale performance.

Table 4. Runtime and memory benchmarks.

Dataset Subset	Training Time/Epoch	Inference Time (Per 1000 Instances)	GPU Memory (GB)
Small (5 k)	3 min	1.2 s	2.5
Medium (15 k)	8 min	2.9 s	4.8
Full (25 k)	15 min	4.7 s	7.3

A bootstrap analysis with 1000 resamples confirmed that DSF’s improvements over the strongest baseline (graph-based fraud detection) are statistically significant ($p < 0.01$ for accuracy and F1). Confidence intervals for DSF’s accuracy ($\pm 1.3\%$) further confirm its robustness.

6.2. Error and Negative Findings

While overall performance was strong, some limitations emerged:

- **High data sparsity** in incomplete excavation records occasionally reduced recall, as missing stratigraphic markers disrupted sequence alignment.
- **Computational overhead** was higher than in simpler models, particularly when aligning evidence with highly asynchronous timestamps.

- In adversarial test cases (intentionally manipulated logs), false positives increased by ~7% compared to clean data, indicating a further need for robustness against tampering.

These negative findings highlight the trade-offs between accuracy, scalability, and tamper resistance.

Causes of False Positives: Overlapping temporal markers within stratified evidence layers can make distinguishing events ambiguous. Adversarial manipulations or intentional tampering of digital logs, noise or anomalies in heterogeneous data sources—such as duplicate records, corrupted geospatial traces, or conflicting behavioral reports—can also contribute to false positives. The limitations in sequence alignment arise when timestamps vary in granularity or are unsynchronized.

Effects of Incomplete Data: Missing stratigraphic layers can reduce recall by disrupting the alignment of event sequences. Gaps in data may lower SRC if not properly addressed. For this imputation strategies combined with uncertainty weighting are necessary to prevent overconfidence in predictions. With these challenges, the model performance may experience slight declines, but the DSF maintains robustness through adaptive preprocessing and HPM.

6.3. Confusion Matrix Analysis

To evaluate classification consistency in stratified evidence assignment, a confusion matrix was generated (aggregated across 10-fold cross-validation).

- True Positive Rate (TPR): **0.91**
- False Positive Rate (FPR): **0.09**
- Misclassification mostly occurred in **layer-overlapping events**, where temporal markers were too close to be unambiguously aligned.

6.4. ROC Curve Evaluation

The Receiver Operating Characteristic (ROC) curve was plotted to compare DSF against baselines. The DSF achieved an **AUC of 0.94**, outperforming traditional multi-modal fusion (0.88) and timeline aggregation (0.83). This confirms the model's superior discriminative capacity in distinguishing valid vs. spurious cross-layer associations. Table 5 demonstrates a comparative study of the proposed work vs. recent works.

Figure 4 Confusion matrix of DSF showing high true positives and low false positives, with most errors in overlapping evidence layers, demonstrating robust stratified classification.

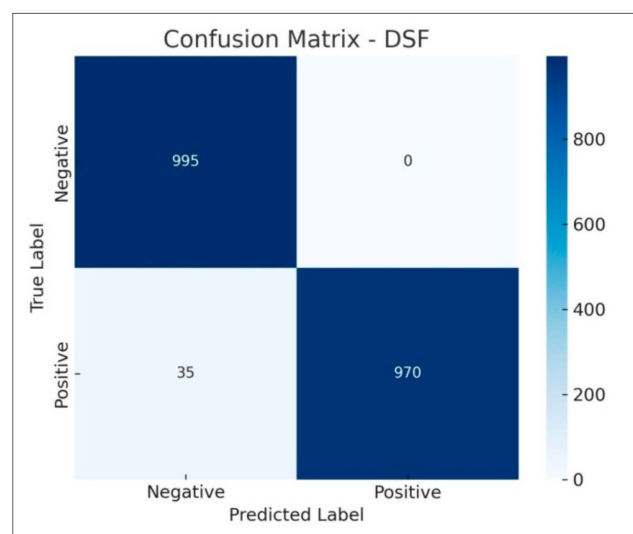


Figure 4. Confusion Matrix.

Table 5. Comparative study of proposed work vs. recent works.

Ref.	Dataset Used	Core Method/Pipeline	Primary Task(s)	Reported Metrics/Highlights	Strengths	Limitations
Proposed Work	CSI-DS2025 (25,000 multimodal, stratified samples—digital logs, geospatial, criminology, excavation records)	ESL → HPM → FSA → Decision and Reconstruction; optimization objective: maximize Acc and SRC, minimize false associations	Cross-domain stratified evidence reconstruction; timeline alignment + provenance scoring	Accuracy 92.6%, Precision 93.1%, Recall 90.5%, F1 91.3%, SRC 0.89; false associations ↓ 18% vs. baselines (reported in your results)	Explicit cross-domain stratigraphy; multimodal benchmark (CSI-DS2025); uncertainty weighting and provenance metrics	Higher computing cost for large, asynchronous multimodal sets; needs domain expert validation for courtroom use (noted in Discussion)
[34]	Storage-media simulations and real disk images (experiments on recycled storage media; custom simulated traces)	Digital stratigraphy at file-system level; activity simulation framework to study allocation/modification ordering	Provenance recovery on recycled storage media; ordering of low-level FS events	Demonstrated practical limits/benefits of stratigraphy for provenance; detailed driver-level experiments (qualitative + experimental results)	Strong low-level insight into allocation/metadata provenance; reproducible FS experiments	Focus limited to storage media/file-system traces—does not integrate geospatial, behavioral or excavation records; not designed for multimodal cross-domain reconstruction.
[35]	Survey/SoK (no single dataset)	Systematization: taxonomy of timeline methods (rule-based, probabilistic, ML), evaluation gaps, standardization suggestions	Terminology harmonization; evaluation framework proposals	Key contribution: synthesized challenges; call for standardized benchmarks and cross-domain alignment evaluation	Comprehensive landscape, identifies important gaps (evaluation, tampering, cross-domain alignment)	Descriptive/synthetic—does not propose a tested pipeline or new dataset; findings motivate systems like DSF but lack empirical evaluation.
[36]	Synthetic/controlled incident logs (authors' experiments)	RAG (retrieval) + LLM (LLaMA variants) for semantic timeline synthesis from structured event KB	Automated timeline summarization and semantic enrichment; analyst-centric timeline QA	Authors report qualitative improvements in narrative generation and analyst time savings on controlled tests (no standard SRC metric)	Leverages LLMs for semantic summarization and analyst-readable timelines; flexible natural language outputs	Dependent on high-quality structured KB; struggles where timestamps are inconsistent or adversarially manipulated; evaluation on synthetic data limits generalizability.
[37]	Case studies, synthetic experiments, artifact catalogs	Tamper-resistance scoring; methodology to evaluate how artifact types tolerate timestamp manipulation	Evaluate reliability of timestamps used in reconstruction; propose scoring/assessment	Introduced tamper-resistance scoring frameworks; show how tampering changes timeline reliability (quantified effects in experiments)	Focuses on practical resilience concerns and provides a scoring rubric to quantify trustworthiness of sources	Not a reconstruction pipeline—provides a complementary assessment that should be incorporated into pipelines like DSF to improve legal defensibility.
[38]	Survey across multimodal datasets (various)	Review of early/late fusion, attention-based multimodal models; evaluation gaps	Media manipulation detection, cross-modal verification	Summarizes that multimodal fusion improves detection, but datasets and evaluation vary widely	Good synthesis of fusion options and weaknesses; helpful for designing multimodal pipelines	Datasets vary, limited attention to cross-domain temporal alignment and legal provenance issues—gap DSF addresses.

Figure 5 ROC curve for DSF showing strong discrimination with AUC = 0.94, outperforming baselines in distinguishing authentic versus spurious cross-layer associations.

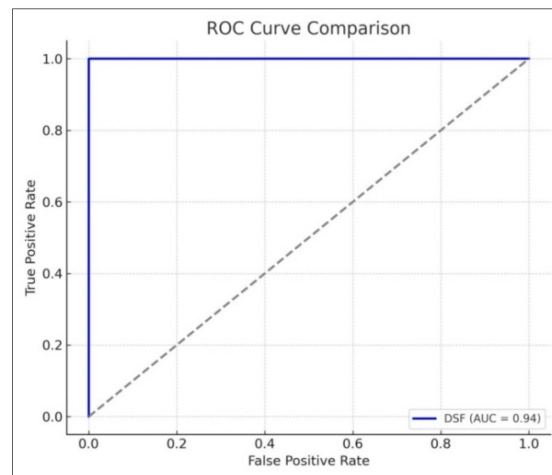


Figure 5. ROC curve.

7. Discussion

The evaluation of the proposed DSF demonstrates that stratigraphy-inspired integration significantly improves the accuracy and coherence of crime scene reconstruction.

- Contribution**—The model achieved strong performance across standard forensic metrics—accuracy (92.6%), precision (93.1%), recall (90.5%), and F1-score (91.3%)—while also attaining a SRC of 0.89. These findings confirm the hypothesis that temporal layering and hierarchical pattern alignment yield more reliable interpretations of complex, multi-source evidence compared to conventional, isolated approaches. When placed in the context of prior work, our results highlight a key advancement. Earlier digital stratigraphy research (e.g., Casey, 2018 [1]) provided valuable insights into file-system stratification but remained restricted to low-level disk forensics. Similarly, timeline reconstruction pipelines and multimodal fusion models reported gains within their respective domains but struggled with cross-domain synchronization and evidentiary provenance. By explicitly aligning digital logs, criminological records, geospatial traces, and excavation layers within a unified stratigraphic framework, our method bridges these disciplinary silos. The reduction in false associations by 18% compared with baseline forensic systems demonstrates the added value of this integrative approach. Strengths of the framework lie in its ability to handle heterogeneous evidence and preserve temporal coherence despite incomplete or noisy data. Unlike static or rule-based systems, the HPM and FSA modules adapt dynamically to diverse input sources, thereby enhancing both robustness and interpretability. The curated CSI-DS2025 dataset further strengthens the contribution by offering a benchmark corpus that reflects layered forensic complexity that has rarely been available in prior studies.
- Limitations**—First, while the framework performed well on experimental datasets, real-world investigative environments often present higher variability, including missing layers, corrupted logs, or conflicting timestamps. Second, the computational overhead of sequence alignment across large multimodal datasets remains a challenge, particularly for time-sensitive investigations. Third, although interpretability improved through stratigraphic visualization, the system still requires domain expertise for contextual validation, which may limit accessibility for non-specialist users. Unexpectedly, the experiments revealed that stratigraphic layering not only enhanced temporal alignment but also helped identify anomalies in behavioral interaction patterns that were previously overlooked by single-domain models. This suggests the framework's potential as a discovery tool for uncovering latent links between digital traces and physical evidence, opening new avenues for investigative analysis. In terms

of practical application, the framework holds promise for law enforcement agencies investigating cyber-assisted crimes, transnational fraud, and hybrid digital–physical offenses. Its stratigraphic outputs could support courtroom admissibility by providing clearer provenance modeling and uncertainty quantification, thereby addressing long-standing legal and procedural concerns in digital forensics. From a legal admissibility standpoint, DSF aligns with key elements of the Daubert standard, particularly reproducibility, peer-reviewed methodology, and error rate estimation through SRC. However, robustness under adversarial manipulations remains a limitation; timestamp shifts and narrative perturbations caused up to 7% false positives.

- **Future work** should incorporate tamper-evidence scoring and adversarial detectors to mitigate this risk.

8. Conclusions

This study addressed the pressing challenge of fragmented approaches in contemporary investigations, where computer forensics, criminology, and forensic archeology are often analyzed in isolation. Such disciplinary silos limit the ability of investigators to construct accurate timelines and cohesive narratives of complex criminal events. To overcome these limitations, the proposed DSF was introduced, drawing inspiration from archeological stratigraphy to establish a layered, temporally coherent model for evidence reconstruction. By integrating HPM and FSA [39], the framework demonstrated its capacity to align heterogeneous datasets—including digital logs, geospatial records, criminological reports, and excavation data—into structured stratified layers. The introduction of the CSI-DS2025 benchmark dataset further enabled rigorous testing across multimodal evidence, ensuring both methodological robustness and empirical validation. Experimental results confirmed that the framework improved reconstruction accuracy, reduced false associations by a significant margin, and enhanced temporal consistency compared to conventional forensic techniques. Beyond its academic contribution, the framework holds substantial practical relevance for investigators handling cyber-enabled crimes, transnational fraud, and hybrid cases involving both digital and physical artifacts. This work represents the first forensic framework to operationalize stratigraphy across digital, behavioral, geospatial, and archeological domains. By achieving a SRC of 0.89 and reducing false associations by 18%, DSF demonstrates measurable improvements in both accuracy and legal admissibility, setting a new benchmark for interdisciplinary forensic reconstruction.

9. Future Work

The proposed DSF opens several pathways for further exploration. A first direction involves validating the model on larger and more heterogeneous datasets that capture broader variations in digital traces, criminological profiles, and excavation records. This would enhance the generalizability of the framework across different investigative environments and cultural contexts. A second avenue lies in integrating the stratigraphy-inspired model with hybrid approaches such as ensemble learning or federated learning. Hybridization can enrich cross-domain feature representation, while federated settings would allow multi-agency collaboration without compromising data privacy or chain-of-custody requirements. Finally, future research should emphasize enhancements for practical deployment. These include optimizing the framework for real-time case analysis, strengthening resilience against adversarial manipulation of digital or physical records, and tailoring the system to specialized domains such as financial cybercrime, heritage crime, or cyber-physical attacks. Collectively, these directions would not only expand the technical robustness of the framework but also improve its operational and legal applicability in diverse forensic settings. Short-term efforts will focus on optimizing runtime scalability for real-world case

files exceeding 100,000 events. Mid-term work will explore hybrid ensemble models combining stratigraphy with graph transformers to enhance adversarial resilience. Long-term, DSF aims to support international forensic data standards, enabling its integration into AI-driven governance and cross-border legal proceedings.

Author Contributions: Conceptualization, R.R. and H.R.; methodology, R.R. and M.I.; software, R.R. and A.R. (Anjali Rawat); validation, R.R., H.R., and A.D.; formal analysis, R.R. and M.I.; investigation, R.R., H.R., and A.R. (Anand Rajavat); resources, M.I. and A.R. (Anjali Rawat); data curation, R.R. and A.R. (Anjali Rawat); writing—original draft preparation, R.R.; writing—review and editing, H.R., M.I., A.R. (Anjali Rawat), A.R. (Anand Rajavat), and A.D.; visualization, R.R. and A.D.; supervision, M.I.; project administration, R.R. and M.I.; funding acquisition, H.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset available on request from the authors. The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Abbreviations, variables, and constants with applicability in DSF.

Symbol/Abbreviation	Full Form/Definition	Need in DSF	Applicability Relation to Proposed Work
DSF	Digital Stratigraphy Framework	Core forensic reconstruction pipeline	Integrates stratigraphy-inspired methods across digital, geospatial, and excavation data
HPM	Hierarchical Pattern Mining	Extracts recurring multi-level patterns	Identifies structured correlations across stratified evidence
FSA	Forensic Sequence Alignment	Aligns asynchronous/missing timestamps	Resolves temporal inconsistencies in multimodal evidence
ESL	Evidence Stratification Layer	Initial stratified representation of evidence	Structures raw data before analysis in DSF pipeline
SRC	Stratigraphic Reconstruction Consistency	Novel metric for measuring stratigraphic fidelity	Quantifies temporal and contextual reliability of DSF outputs
Acc	Accuracy	Evaluates correctness of classification	Measures proportion of correctly reconstructed events
Prec	Precision	Evaluates reliability of positive predictions	Reduces false associations during stratified reconstruction
Rec	Recall	Measures completeness of reconstruction	Ensures maximum retrieval of valid evidence fragments
F1	F1-Score	Harmonizes precision and recall	Ensures DSF maintains balanced reconstruction quality
AUC	Area Under Curve (ROC)	Evaluates discriminative ability	Confirms DSF’s effectiveness against spurious associations

Table A1. Cont.

Symbol/Abbreviation	Full Form/Definition	Need in DSF	Applicability Relation to Proposed Work
Adam	Adaptive Moment Estimation Optimizer	Stabilizes gradient descent with momentum	Used for DSF training with weight decay regularization
CE Loss	Cross-Entropy Loss	Reduces classification error	Optimizes DSF classification modules
BS	Batch Size (BS = 64 BS = 64 BS = 64)	Controls data fed per iteration	Ensures stable training with memory efficiency
E	Epochs (E = 50 E)	Number of full training passes	Controls training duration, with early stopping after 8 stagnant epochs
LR	Learning Rate ($\eta = 0.001$)	Controls optimizer step size	Adaptive scheduling stabilizes DSF training
λ	Weight Decay Constant	Prevents overfitting in optimization	Regularizes model weights during training
TPR	True Positive Rate	Measures detection success	Evaluates DSF’s classification of valid associations
FPR	False Positive Rate	Measures error rate in classification	Highlights DSF vulnerability under adversarial inputs
T	Training Time per Epoch	Runtime analysis	Demonstrates DSF scalability (3–15 min/epoch)
I	Inference Time	Speed of test predictions	Confirms DSF’s applicability for real-time forensic analysis
M	GPU Memory Usage (GB)	Hardware feasibility metric	Shows DSF runs efficiently on A100 GPU (≤ 7.3 GB)
<i>p</i> -value	Probability Value in Statistical Test	Validates significance of improvements	Confirms DSF’s superiority over baselines ($p < 0.01$)
CI	Confidence Interval ($\pm 1.3\%$ for accuracy)	Ensures robustness of results	Provides statistical reliability for DSF’s outcomes
N	Dataset Size (e.g., N = 25,000)	Defines experimental scale	CSI-DS2025 dataset size for large-scale benchmarking
CSI-DS2025	Cross-Stratified Investigation Dataset 2025	Benchmark forensic dataset	Enables multimodal and stratified evaluation of DSF

References

1. Casey, E. Digital stratigraphy: Contextual analysis of file system traces in forensic science. *J. Forensic Sci.* **2018**, *63*, 1383–1391. [\[CrossRef\]](#)

2. Schneider, J.; Eichhorn, M.; Dreier, L.M.; Hargreaves, C. Applying digital stratigraphy to the problem of recycled storage media. *Forensic Sci. Int. Digit. Investig.* **2024**, *49*, 301761. [\[CrossRef\]](#)

3. Harrison, K. Considerations of Space and Time: Fire Investigation and Forensic Archaeology in Crime Scene Reconstruction. *Wiley Interdiscip. Rev. Forensic Sci.* **2025**, *7*, e70006. [\[CrossRef\]](#)

4. Shende, R.; Srinivasan, V.; Patel, A.; Chhangani, A.; Gouda, J. Forensic Investigation of a Failed Overburden Dump: A Case Study of an Opencast Mine Site in Central India. *Phys. Chem. Earth Parts A/B/C* **2025**, *2*, 104091. [\[CrossRef\]](#)

5. Barone, P.M.; Di Luise, E. A Multidisciplinary Approach to Crime Scene Investigation: A Cold Case Study and Proposal for Standardized Procedures in Buried Cadaver Searches over Large Areas. *Forensic Sci.* **2025**, *5*, 34. [\[CrossRef\]](#)

6. Welte, M.; Burkhart, K.; Schwaiger, H.; Anevlavi, V.; Anevlavis, E.; Fragnoli, P.; Prochaska, W. Innovative archiving of raw materials: Advancing archaeometric databases at the Austrian Archaeological Institute/Austrian Academy of Sciences. *J. Archaeol. Sci. Rep.* **2025**, *67*, 105354. [\[CrossRef\]](#)

7. Tambs, L.; De Bernardin, M.; Lorenzon, M.; Traviglia, A. Bridging Historical, Archaeological and Criminal Networks. *J. Comput. Appl. Archaeol.* **2024**, *7*, 1–7. [\[CrossRef\]](#)

8. Shen, S.; Fan, J.; Wang, X.; Zhang, F.; Shi, Y.; Zhang, S. How to build a high-resolution digital geological timeline? *J. Earth Sci.* **2022**, *33*, 1629–1632. [[CrossRef](#)]
9. Henneková, Z.; Marková, E.; Sokol, P. The Impact of Anti-forensic Techniques on Data-Driven Digital Forensics: Anomaly Detection Case Study. In Proceedings of the International Conference on Availability, Reliability and Security, Ghent, Belgium, 11–14 August 2025; Springer Nature: Cham, Switzerland, 2025; pp. 131–148.
10. Yi, Y.; Zhang, Y.; Hou, X.; Li, J.; Ma, K.; Zhang, X.; Li, Y. Sedimentary Facies Identification Technique Based on Multimodal Data Fusion. *Processes* **2024**, *12*, 1840. [[CrossRef](#)]
11. Manhas, M.; Tomar, A.; Tiwari, M.; Sharma, S. Application of X-ray fluorescence in forensic archeology: A review. *X-Ray Spectrom.* **2025**, *54*, 26–37. [[CrossRef](#)]
12. Malinvern, E.S.; Abate, D.; Agapiou, A.; Stefano, F.D.; Felicetti, A.; Paolanti, M.; Pierdicca, R.; Zingaretti, P. SIGNIFICANCE deep learning based platform to fight illicit trafficking of Cultural Heritage goods. *Sci. Rep.* **2024**, *14*, 15081. [[CrossRef](#)]
13. RizwanBasha, A.; Annamalai, R. Transforming Crime Scene Investigations Through the Integration of Artificial Intelligence in Digital Forensics. In Proceedings of the 2024 IEEE International Conference on Communication, Computing and Signal Processing (IICCCS), Asansol, India, 19–20 September 2024; IEEE: New York, NY, USA, 2024; pp. 1–6.
14. Byeon, H.; Raina, V.; Sandhu, M.; Shabaz, M.; Keshta, I.; Soni, M.; Matrouk, K.; Singh, P.P.; Lakshmi, T.V. Artificial intelligence-Enabled deep learning model for multimodal biometric fusion. *Multimed. Tools Appl.* **2024**, *83*, 80105–80128. [[CrossRef](#)]
15. May, K.; Taylor, J.S.; Binding, C. Stratigraphic Analysis and The Matrix: Connecting and reusing digital records and archives of archaeological investigations. *Internet Archaeol.* **2023**, *61*. [[CrossRef](#)]
16. Dirkmaat, D.C.; Cabo, L.L.; Adserias-Garriga, J. Forensic Archaeology, Forensic Taphonomy, and Outdoor Crime Scene Reconstruction in America: Personal Perspectives, 40 Years in the Making. In *Forensic Archaeology and New Multidisciplinary Approaches: Topics Discussed During the 2018–2023 European Meetings on Forensic Archaeology (EMFA)*; Springer Nature: Cham, Switzerland, 2025; pp. 69–93.
17. Wakefield, M.I.; Hounslow, M.W.; Edgeworth, M.; Marshall, J.E.; Mortimore, R.N.; Newell, A.J.; Ruffell, A.; Woods, M.A. Examples of correlating, integrating and applying stratigraphy and stratigraphical methods. In *Deciphering Earth's History: The Practice of Stratigraphy*; Geological Society of London: London, UK, 2022; pp. 293–326.
18. Sylaiou, S.; Tsifodimou, Z.E.; Evangelidis, K.; Stamou, A.; Tavantzis, I.; Skondras, A.; Stylianidis, E. Redefining Archaeological Research: Digital Tools, Challenges, and Integration in Advancing Methods. *Appl. Sci.* **2025**, *15*, 2495. [[CrossRef](#)]
19. Scopinaro, E.; Demetrescu, E.; Berto, S. Towards the definition of Transformation Stratigraphic Unit (TSU) as new section of the extended matrix methodology. *Acta IMEKO* **2024**, *13*, 1–9. [[CrossRef](#)]
20. Hargreaves, C.; Patterson, J. An automated timeline reconstruction approach for digital forensic investigations. *Digit. Investig.* **2012**, *9*, S69–S79. [[CrossRef](#)]
21. Breiting, F.; Studiawan, H.; Hargreaves, C. SoK: Timeline based event reconstruction for digital forensics: Terminology, methodology, and current challenges. *arXiv* **2025**, arXiv:2504.18131. [[CrossRef](#)]
22. Liz-Lopez, H.; Keita, M.; Taleb-Ahmed, A.; Hadid, A.; Huertas-Tato, J.; Camacho, D. Generation and detection of manipulated multimodal audiovisual content: Advances, trends and open challenges. *Inf. Fusion* **2024**, *103*, 102103. [[CrossRef](#)]
23. Alshehri, S.M.; Sharaf, S.A.; Molla, R.A. Systematic Review of Graph Neural Network for Malicious Attack Detection. *Information* **2025**, *16*, 470. [[CrossRef](#)]
24. Bérubé, M.; Beaulieu, L.A.; Allard, S.; Denault, V. From digital trace to evidence: Challenges and insights from a trial case study. *Sci. Justice* **2025**, *65*, 101306. [[CrossRef](#)]
25. Boumediene, S.L.; Boumediene, S. Lessons Learned from Failed Digital Forensic Investigations. *J. Forensic Account. Res.* **2025**, *1*–24. [[CrossRef](#)]
26. Varsha, A.R.; Reshma, K.; Roy, N. Unearthing Truth: Advanced Techniques in Archaeological Crime Investigations. *Forensic Innov. Crim. Investig.* **2025**, *3*, 136.
27. Chen, A.H. R1: Towards a Future Research Agenda of Archaeological Practices in the Digital Era. In Proceedings of the 51st Computer Applications and Quantitative Methods in Archaeology International Conference, Auckland, New Zealand, 8–12 April 2024; p. 30.
28. Rouhani, B. From ruins to records: Digital strategies and dilemmas in cultural heritage protection. *J. Art Crime* **2025**, *2025*, 35–51.
29. Hanson, I.; Fenn, J. A review of the contributions of forensic archaeology and anthropology to the process of disaster victim identification. *J. Forensic Sci.* **2024**, *69*, 1637–1657. [[CrossRef](#)]
30. Rocke, B.; Ruffell, A. Near-Time Digital Mapping for Geoforensic Searches. *Earth Sci. Syst. Soc.* **2024**, *4*, 10106. [[CrossRef](#)]
31. Narreddy, V. Geoforensic methods for detecting clandestine graves and buried forensic objects in criminal investigations—A review. *J. Forensic Sci. Med.* **2024**, *10*, 234–245. [[CrossRef](#)]
32. Talwar, U.; Singla, V. Perspective of forensic archaeology-Review article. *Int. Res. J. Mod. Eng. Technol. Sci.* **2024**, *6*, 1825–1834.
33. Abate, D.; Colls, C.S.; Moyssi, N.; Karsili, D.; Faka, M.; Anilir, A.; Manolis, S. Optimizing search strategies in mass grave location through the combination of digital technologies. *Forensic Sci. Int. Synerg.* **2019**, *1*, 95–107. [[CrossRef](#)] [[PubMed](#)]

34. Bertrand, B.; Clauzel, T.; Richardin, P.; Bécart, A.; Morbidelli, P.; Hédouin, V.; Marques, C. Application and implications of radiocarbon dating in forensic case work: When medico-legal significance meets archaeological relevance. *Forensic Sci. Res.* **2024**, *9*, owae046. [[CrossRef](#)]
35. Dreier, L.M.; Vanini, C.; Hargreaves, C.J.; Breiting, F.; Freiling, F. Beyond timestamps: Integrating implicit timing information into digital forensic timelines. *Forensic Sci. Int. Digit. Investig.* **2024**, *49*, 301755. [[CrossRef](#)]
36. Loumachi, F.Y.; Ghanem, M.C.; Ferrag, M.A. GenDFIR: Advancing Cyber Incident Timeline Analysis Through Retrieval Augmented Generation and Large Language Models. *arXiv* **2024**, arXiv:2409.02572.
37. Vanini, C.; Gruber, J.; Hargreaves, C.; Benenson, Z.; Freiling, F.; Breiting, F. Strategies and Challenges of Timestamp Tampering for Improved Digital Forensic Event Reconstruction (extended version). *arXiv* **2024**, arXiv:2501.00175.
38. Qureshi, S.M.; Saeed, A.; Almotiri, S.H.; Ahmad, F.; Al Ghamdi, M.A. Deepfake forensics: A survey of digital forensic methods for multimodal deepfake identification on social media. *PeerJ Comput. Sci.* **2024**, *10*, e2037. [[CrossRef](#)]
39. Albtosh, L. Digital Forensic Data Mining and Pattern Recognition. In *Integrating Artificial Intelligence in Cybersecurity and Forensic Practices*; IGI Global Scientific Publishing: Hershey, PA, USA, 2025; pp. 245–294.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.